

Développement et caractérisation de méthodes statistiques de détection et de localisation d'objets dans des images fortement bruitées

THÈSE

présentée et soutenue publiquement le 3 octobre 2001

pour l'obtention du

Doctorat de l'Université d'Aix-Marseille III

(spécialité Traitement des Images)

par

Vincent PAGÉ

Composition du jury

<i>Rapporteurs :</i>	Jean Taboury Pierre Ambs
<i>Directeur de thèse :</i>	Ph. Réfrégier
<i>Co-Directeur de thèse :</i>	F. Goudail
<i>Examineurs :</i>	François Le Chevalier J.Garcia

Mis en page avec la classe thloria.

Remerciements

Je tiens à exprimer ici ma reconnaissance à l'ensemble des personnes qui ont contribué à l'accomplissement de ce travail de thèse.

Je remercie tout d'abord Messieurs Jean Taboury et Pierre Ambs, pour le temps qu'ils ont consacré à rapporter ce manuscrit. Ma gratitude revient également à Monsieur Francois Le Chevalier qui m'a fait l'honneur de présider ce jury, ainsi qu'à Javier Garcia Monreal dont la présence m'aura autant enchanté que notre collaboration passée.

Je suis reconnaissant à Philippe Réfrégier qui a dirigé cette thèse, pour son enthousiasme en matière de sciences ainsi que pour sa patience à l'égard de mes méthodes de rédaction, en espérant qu'il trouve dans ce manuscrit la preuve que ses conseils n'ont pas été inutiles. Je ne peux pas non plus manquer de souligner le travail de François Goudail en l'absence duquel ces travaux n'auraient sans doute pas eu le même aboutissement.

Cette thèse s'est déroulée au *Laboratoire Signal et Image* de l'*Ecole Nationale Supérieure de Physique de Marseille*, devenu depuis *Equipe Physique et Traitement de l'Image* du nouvel *Institut Fresnel*. J'adresse mes remerciements à tous les permanents qui font vivre ce laboratoire : Ceux qui l'ont vu naître, François Goudail, Mireille Guillaume et Thierry Gaidon ainsi que les nouveaux arrivants, Salah Bourennane et Nicolas Bertaux.

Je n'oublie pas Olivier Ruch, Frédéric Guérault et Cristophe Chesnaud, thésards dont les conversations lors du déjeuner m'ont permis de garder le sourire malgré la nourriture servie au Restaurant Universitaire. Merci enfin à mes co-locataires et amis, Celine Defay, Thomas Olivier, Karine Florin, Elise Picherot et Marie Bénion qui ont su faire preuve d'une patience infinie face aux changements d'humeur qui ont eu lieu au cours de ces trois années.

Pour finir, je voudrais remercier Gilles Trossevin, ingénieur système, pour ses nombreuses interventions salvatrices ainsi que Nathalie Goître et Nina Ralijaona, secrétaires successives, pour leur aide précieuse et leur gentillesse.

A toutes celles qui m'auront supporté pendant cette thèse.

Table des matières

Introduction

Chapitre 1

Introduction au filtre GLRT

1.1	Modèle d'images SIR et bruit non recouvrant	6
1.2	Localisation optimale au sens du maximum de vraisemblance	8
1.2.1	Principe	8
1.2.2	Exemple d'application [15]	10
1.3	Détection optimale au sens du rapport de vraisemblance [19]	12
1.3.1	Principe	12
1.3.2	Neyman-Pearson	14
1.3.3	Rapport de vraisemblance généralisé	14
1.3.4	Critère de qualité de la détection	15
1.4	Quelques remarques complémentaires sur ces deux algorithmes	15
1.4.1	Complexité de ces algorithmes	15
1.4.2	Localisation	17
1.4.3	Détection à Taux de Fausse Alarme constant	18
1.5	Conclusion	19

Chapitre 2

Robustesse de l'algorithme GLRT en bruit non homogène

2.1	Influence de la taille de la fenêtre d'analyse en bruit homogène	23
2.1.1	La connaissance des paramètres du fond est une hypothèse forte	23
2.1.2	Utiliser de grandes fenêtres procure une meilleure estimation des paramètres du fond	25
2.1.3	Limite supérieure à la taille des fenêtres	27
2.1.4	Généralisation à différents contrastes	28

2.1.5	Généralisation aux bruits Poissonien et Gaussien	28
2.1.6	Conclusion en bruit homogène	29
2.2	Bruit non homogène	30
2.2.1	Modèle d'images inhomogènes	31
2.2.2	Echec des grandes fenêtres	31
2.2.3	Inhomogénéités trop importantes	33
2.3	Conclusion	33

Chapitre 3

Cas de cibles de tailles inconnues

3.1	Chute des performances lorsque la taille de la cible varie	37
3.2	Rappels de la théorie de l'estimation et de la décision	38
3.2.1	Hypothèses	38
3.2.2	Estimation de la taille au sens du maximum de vraisemblance . . .	39
3.2.3	Estimation de la taille au sens du maximum a posteriori	40
3.2.4	Approche du Marginal Bayésien	40
3.2.5	Conclusion	41
3.3	Choix d'une méthode de gestion du paramètre de taille	41
3.3.1	Cas de lois Gamma d'ordre L	42
3.3.2	Influence du prior sur les performances de l'AMB	42
3.4	Première illustration des performances de cet algorithme	44
3.5	Choix d'un seuil de détection	44
3.6	Exemples d'utilisation de l'algorithme $GLRT_{mv}$	47
3.7	Conclusion	48

Chapitre 4

Exemples d'application : Imagerie SAR et Imagerie Polarimétrique

4.1	Imagerie SAR	54
4.2	Imageur Polarimétrique	56
4.2.1	Principe	56
4.2.2	Modèle d'images	57
4.2.3	Bruit homogène	59
4.2.4	Bruit non homogène	60
4.2.5	Représentation naturelle des images ICEO	62
4.2.6	Applications aux images réelles d'un imageur polarimétrique	63

4.3 Conclusion	65
--------------------------	----

Chapitre 5

Filtre GLRT multi régions

5.1 Modèle d'image pour cibles multirégions	70
5.2 Détection optimale de cibles multi-régions	72
5.2.1 Rapport de vraisemblance pour des objets multi-régions	73
5.2.2 Illustration des performances sur des images synthétiques	73
5.2.3 Illustration sur des images réelles	75
5.3 Limite du nombre de régions utilisables	78
5.4 Estimation du nombre de régions de la cible	82
5.4.1 Pénalisation du nombre de régions	84
5.4.2 Choix du paramètre de pénalisation	86
5.4.3 Complexité de cet algorithme	88
5.5 Estimation simplifiée du nombre de régions	89
5.6 Performances de ces algorithmes	90
5.7 Conclusion	94

Chapitre 6

Corrélateurs optiques

6.1 Structures de corrélateurs optiques	96
6.1.1 Différents types de codages de l'image d'entrée.	99
6.2 Comparaison entre le GLRT et le NLJTC	101
6.3 Choix d'un codage	103
6.3.1 Position du problème	103
6.3.2 Exemple	104
6.3.3 Etude de l'influence du codage en phase en présence de bruit non recouvrant	107
6.3.4 Illustration dans le cas d'images réalistes	113
6.4 Conclusion	114

Conclusion

1 Taille de la fenêtre d'analyse	118
2 Cibles de tailles inconnues	118
3 Détection de cibles dans des images issues de capteurs réels	119

4	Cibles composées de plusieurs régions	120
5	Implémentation optique des filtres	121
6	Perspectives	122

Annexe A
Principe de la description de longueur minimale (MDL)

Annexe B
Calculs

B.1	Détection a Taux de Fausse Alarme Constant en bruit gaussien	129
B.2	Algorithme du Maximum de Vraisemblance pour des bruits de Poisson . .	130

Bibliographie

Bibliographie	133
----------------------	------------

Introduction

Les progrès constants des systèmes d'imagerie et des calculateurs numériques ont permis l'élaboration de nouveaux systèmes de traitement d'images. Les applications de ces systèmes sont nombreuses et la mise au point de tels systèmes fait l'objet d'une forte demande de la part des industriels, du domaine médical et du domaine militaire.

Cependant, nous sommes encore loin du scénario des auteurs de science-fiction, dans lequel des calculateurs surpuissants permettent de remplacer systématiquement les opérateurs humains dans une activité de reconnaissance d'objet, de localisation, ou d'interprétation du contenu des images. De fait, en matière de traitement, le cerveau humain est encore bien souvent une référence inégalée. L'existence d'un algorithme universel, adapté à tout type d'image, et tout type de tâche est encore pour le moment une chimère.

C'est pourquoi un gros effort de structuration a été fait au cours des trentes dernières années, permettant de définir des catégories de problèmes bien distincts, auxquels on s'est efforcé d'apporter des solutions précises, en clarifiant à la fois contraintes et objectifs. En restreignant les problèmes, on a pu voir émerger des solutions efficaces, et utilisables comme les briques de bases de systèmes plus complexes. On distingue ainsi des objectifs, comme par exemple la localisation, la détection, l'identification ou la segmentation, chacun de ces objectifs étant associé à un contexte précis.

Nous nous sommes intéressés à la localisation et à la détection dans une image, d'objets dont on connaît une référence. Une telle opération, effectuée couramment en traitement du signal radar, est relativement bien maîtrisée dans ce domaine. En traitement d'images, un grand nombre de problèmes n'ont pas encore trouvé de solutions définitives.

Au nombre de ces problèmes, on peut notamment citer le cas d'images fortement bruitées, dans lesquelles l'objet n'apparaît pas clairement. La présence de fonds structurés dans les images est également source de nombreux problèmes, à tel point que le fond est souvent lui-même considéré comme du bruit en soi. Par ailleurs, si les niveaux de gris de l'objet cherché peuvent fluctuer, il convient de procéder à une détection ou une localisation qui ne se base que sur des caractéristiques spatiales de l'objet. En restreignant encore les connaissances dont on dispose, peut également se poser le problème de la détection de cibles de tailles inconnues.

On le voit, un même problème, comme celui de la détection, se décline ainsi en une multitude de sous-problèmes, de difficulté grandissante, et dont le traitement nécessite des algorithmes robustes et adaptés.

De façon très générale, lorsque l'on cherche à tirer des conclusions en dépit de faibles connaissances, en disposant d'un ensemble diversifié de données, une approche statistique semble indiquée. Des solutions adaptées à la localisation d'objets dont on ne connaît

que la forme ont ainsi été développées au laboratoire, s'appuyant essentiellement sur la théorie statistique de l'estimation. Ces algorithmes peuvent être développés en supposant connus ou inconnus les niveaux de gris de l'objet cherché [22]. Le cas où ces niveaux de gris sont inconnus correspond à beaucoup d'applications pratiques. En conséquence, nous nous sommes intéressés à l'approfondissement de techniques dédiées à ces cas. Cet approfondissement fera l'objet des cinq premiers chapitres de ce manuscrit.

Dans le **chapitre 1**, nous reprendrons les modèles d'images et les algorithmes issus de travaux antérieurs à la réalisation de cette thèse. Un modèle d'image est toujours sujet à caution, ne pouvant prétendre, on l'a dit, correspondre efficacement à tous les cas possibles. Le modèle qui a été retenu est simple, et général. Il suppose l'image composée de zones homogènes disjointes, la cible et le fond. C'est un modèle de bruit non recouvrant, différent du modèle de bruit additif le plus classique. Nous verrons sur quelques exemples qu'il permet de construire un algorithme de localisation robuste et rapide, basé sur le calcul de vraisemblances. Il s'agit de l'algorithme que nous nommerons algorithme du ML (Maximum Likelihood). Cependant, nous montrerons plus loin que dans le cas de fonds non homogènes, il accuse des pertes de performances sévères. L'autre algorithme, que nous noterons algorithme GLRT (Generalized Likelihood Ratio Test), est un algorithme de détection utilisé en localisation de façon empirique, mais efficace, et qui permet dans certains cas de dépasser les limites du premier dans le cas de fonds inhomogènes.

L'algorithme GLRT utilise une fenêtre qui se déplace sur l'image pour analyser son contenu. L'influence de la taille de cette fenêtre d'analyse sur les performances du GLRT en localisation fera l'objet du **chapitre 2**. Nous montrerons que le choix d'une taille pour cette fenêtre est guidé par la recherche d'un compromis, entre la qualité de l'estimation des paramètres lorsque le fond est homogène, et la robustesse aux inhomogénéités pouvant apparaître dans des images ne correspondant pas au modèle. Ce compromis pouvant être atteint assez facilement, et permettant de dépasser les limites de l'algorithme du maximum de vraisemblance, nous utiliserons préférentiellement l'algorithme GLRT par la suite.

Dans le **chapitre 3**, nous généraliserons ensuite cet algorithme au cas de cibles de tailles inconnues, en utilisant un banc de filtres représentant la même cible à différentes échelles. Nous montrerons que cette généralisation, utilisée couramment de façon empirique, conserve un caractère théorique rigoureux dans le cadre de l'estimation statistique lorsque le filtre de base est le GLRT. Dans ce cadre, nous confronterons deux approches classiques, l'estimation de taille au sens du maximum de vraisemblance, et l'approche marginale bayésienne, qui consiste à intégrer le paramètre de taille, c'est à dire globalement à sommer les sorties de chaque filtre. Cette dernière approche est souvent supposée plus efficace que la première. Nos observations montreront que ce n'est pas le cas dans nos applications, ce que nous analyserons brièvement.

Dès lors, nous disposerons d'algorithmes théoriquement efficaces, pour des modèles d'images simples. Encore faut-il que leur application à des images réelles soit possible. Dans les systèmes actuels d'imagerie active, le processus physique de formation des images est relativement bien connu. La statistique des intensités des images y est correctement modélisée. Le GLRT peut donc être appliqué efficacement sur de telles images. Si dans

le cas de l'imagerie radar de type SAR, ces algorithmes peuvent s'appliquer sans trop de problèmes, dans le cas d'un imageur polarimétrique actif comme celui développé par Thales optronique SA (ex Thomson-CSF Optronique), il peut être nécessaire de procéder à des transformations des données avant traitement. Nous discuterons de ceci tout au long du **chapitre 4**.

Enfin, le **cinquième et dernier chapitre** sur ces algorithmes optimaux sera consacré à une deuxième généralisation du GLRT, permettant de combler l'écart actuellement existant entre le traitement de cibles de niveaux de gris connus, et celui de cibles homogènes de niveaux de gris inconnus. Nous supposerons la cible constituée d'un ensemble de régions dont on connaît la forme, mais dont l'intensité moyenne est inconnue. Nous obtiendrons ainsi un algorithme GLRT multi-régions, particulièrement indiqué dans le cas de la localisation d'objet 3D, lorsque les conditions d'illuminations peuvent changer. Nous mènerons une analyse théorique des pertes de performances d'une modélisation utilisant de trop nombreuses régions. Nous nous servirons de cette analyse pour proposer deux modifications successives de cet algorithme visant à sélectionner automatiquement le nombre de régions de la référence à utiliser.

Si la plupart des systèmes de traitement sont des calculateurs numériques, la recherche du temps réel dans le traitement des données nécessite souvent de recourir à des solutions alternatives. Le **chapitre 6** concernera ainsi la mise en oeuvre d'algorithmes de localisation par des systèmes optiques. Ces systèmes permettent en effet, au moins théoriquement, d'obtenir des cadences de traitement beaucoup plus élevées que des systèmes numériques. La mise en oeuvre des algorithmes présentés dans la première partie du manuscrit dans un système purement optique ou mixte optique-numérique, est actuellement à l'étude, mais ne peut être considérée comme opérationnelle. Il est alors nécessaire de recourir à des techniques sous-optimales. Nous prendrons ici l'exemple d'un Corrélateur Optique Non Linéaire à Transformation Jointe. Ces corrélateurs nécessitent l'introduction de l'image à traiter dans le système. Pour cela, on utilise un modulateur spatial de lumière, qui peut coder cette image sous la forme d'une modulation de phase, ou d'une modulation d'amplitude. Le choix d'une modulation particulière pouvant modifier les performances, nous avons voulu caractériser de façon phénoménologique son influence. Nous montrerons ainsi que dans le cas de bruits non recouvrant, le codage en phase permet d'augmenter la robustesse du système aux faibles contrastes existants entre la cible et le fond.

Enfin, nous concluerons sur ces études et proposerons des perspectives sur la continuation de ces travaux.

Chapitre 1

Introduction au filtre GLRT

Sommaire

1.1	Modèle d'images SIR et bruit non recouvrant	6
1.2	Localisation optimale au sens du maximum de vraisemblance	8
1.2.1	Principe	8
1.2.2	Exemple d'application [15]	10
1.3	Détection optimale au sens du rapport de vraisemblance [19]	12
1.3.1	Principe	12
1.3.2	Neyman-Pearson	14
1.3.3	Rapport de vraisemblance généralisé	14
1.3.4	Critère de qualité de la détection	15
1.4	Quelques remarques complémentaires sur ces deux algorithmes	15
1.4.1	Complexité de ces algorithmes	15
1.4.2	Localisation	17
1.4.3	Détection à Taux de Fausse Alarme constant	18
1.5	Conclusion	19

Tout au long de cette thèse, nous nous sommes placés dans le cadre de la théorie statistique de l'estimation. Nous considérons donc qu'une tâche de localisation consiste à estimer la position de la cible dans l'image. Une opération de détection consistera à effectuer la discrimination entre deux hypothèses, celle que la cible est présente et celle que la cible est absente. Pour réaliser ces deux opérations nous utiliserons une analyse des propriétés statistiques de l'image. Ce chapitre d'introduction a pour objet de reprendre les bases théoriques nécessaires à l'élaboration de filtres de détection et de localisation efficaces dans ce contexte.

Ces deux tâches sont théoriquement très différentes, bien qu'elles soient effectuées de façon conjointe dans beaucoup de cas. En effet, si l'on souhaite effectuer la détection d'une cible dans une image, on souhaite également savoir où elle se trouve dans cette image. De fait, la détection ne devrait consister qu'à dire : une cible est présente ou une cible est absente dans cette image. La volonté d'effectuer ces deux tâches en même temps entraîne souvent une confusion entre les moyens permettant d'effectuer l'une ou l'autre au mieux.

Commençons par expliquer ce qu'est la localisation dans le cadre de l'estimation statistique. Localiser la cible consiste à estimer le paramètre de position de la cible. Il s'agit donc de définir une probabilité que la cible se trouve à un emplacement donné, connaissant les données observées (la scène). On pourra choisir comme estimation de position celle qui maximise cette probabilité. Pour définir cette probabilité, il est nécessaire de définir un modèle probabiliste pour les scènes à traiter.

La détection quand à elle, consiste à discriminer deux hypothèses, celle que la cible est présente, et celle qu'elle est absente. De la même façon que précédemment, il faudra mesurer la probabilité que chaque hypothèse soit exacte, pour pouvoir choisir quelle décision doit être retenue. Encore une fois, pour mesurer ces probabilités, il est nécessaire de recourir à un modèle probabiliste des données observées, c'est à dire d'un modèle pour les scènes sur lesquelles nous allons travailler. Quelle que soit la tâche à laquelle nous voudrions nous atteler, il est donc nécessaire de définir un modèle d'image avant toute chose.

1.1 Modèle d'images SIR et bruit non recouvrant

Le modèle d'image le plus simple consiste à supposer que la scène n'est perturbée que par du bruit additif [58]. Ce modèle, issu des études portant sur le radar a été le premier repris en traitement d'images. Il peut en effet modéliser le bruit d'un détecteur. De nombreux filtres de localisation ou de détection ont été développés afin de traiter le cas de scènes dans lesquelles la cible apparaît sur un fond d'intensité moyenne faible, et perturbées par un bruit couvrant à la fois cible et fond. Néanmoins dans beaucoup d'applications, le fond de la scène est beaucoup plus complexe. Si l'on considère une image réaliste, d'un avion survolant une ville, le fond de la scène, constitué d'arbres, d'immeubles, peut présenter des contrastes plus importants, et des bords plus saillants que la cible. De ce fait, la principale cause d'échec, en localisation ou en détection, n'est pas le bruit additif dû au système d'acquisition mais la totalité du fond de la scène. De plus, ce bruit n'affecte pas l'intensité des pixels de la cible. On l'appelera donc bruit non recouvrant [30]. Depuis le début des années 90, les chercheurs ont remarqué les performances médiocres des filtres classiques en présence de bruit non recouvrant [12, 30]. Les raisons de cette relative inefficacité ont été étudiées [16] et des algorithmes prenant en compte la présence de bruit non-recouvrant sur les images ont été proposés [46, 29, 31, 18, 44, 36, 11]. La figure 1.1 illustre la différence majeure existant entre ces deux types de bruits. Le bruit non recouvrant étant plus représentatif de nombreux problèmes de traitement d'image, c'est vers ce type de bruits que nous avons tourné notre choix.

Dès lors, nous allons considérer le modèle d'image suivant permettant de modéliser un bruit non recouvrant [30] : dans ce qui suit, nous noterons $\mathbf{s} = [s(i), i \in [1, N]]$ le vecteur contenant l'intensité des N pixels de l'image observée. La notation de l'image sous une forme monodimensionnelle nous permettra de simplifier les notations sans perte de généralité, si l'on suppose que les pixels de cette images sont rangés dans un vecteur dans l'ordre lexicographique. L'hypothèse de bruit non recouvrant signifie que l'image est composée de deux régions disjointes : une région correspondant à l'objet à détecter ou à localiser et la région complémentaire, appelée région fond. Si l'on identifie la région objet

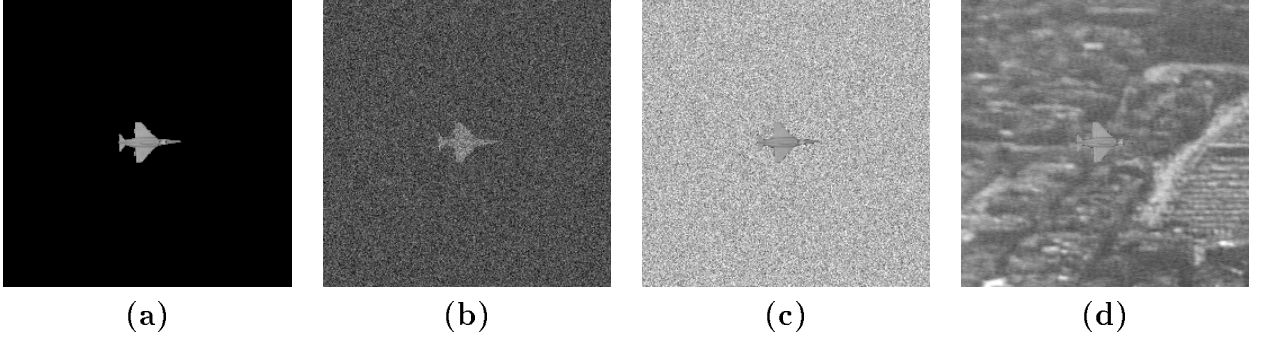


FIG. 1.1: Illustration des différents bruits pouvant affecter une image.

(a) image de référence. (b) scène avec du bruit additif. (c) scène avec du bruit non-recouvrant. (d) scène réaliste. Nous pouvons noter que le bruit prépondérant dans la scène réaliste est non recouvrant.

centrée à la position j de la scène par son support \mathbf{w}_j tel que :

$$\begin{cases} w_j(i) = 1 & \text{si le pixel d'indice } i \text{ appartient à la cible} \\ w_j(i) = 0 & \text{sinon} \end{cases} \quad (1.1)$$

nous noterons $\mathbf{1} - \mathbf{w}_j$ ou $\bar{\mathbf{w}}$ le complémentaire de \mathbf{w}_j dans \mathbf{s} .

Alors l'intensité du pixel i de la scène observée peut s'écrire

$$s(i) = a(i)w_j(i) + b(i)(1 - w_j(i)) \quad (1.2)$$

Nous supposons ensuite que l'intensité de chaque pixel de la cible (et respectivement du fond) est une réalisation d'un champ aléatoire \mathbf{a} (respectivement \mathbf{b}). Les champs \mathbf{a} et \mathbf{b} sont statistiquement indépendants. C'est pourquoi ce modèle d'image est habituellement appelé modèle SIR pour *Statistically Independant Regions* [18, 21, 15]. C'est ce modèle que nous adopterons pour toute la suite de ce mémoire.

Il nous faut maintenant définir les propriétés statistiques des champs \mathbf{a} et \mathbf{b} . Nous supposons que les intensités des pixels sont décorréliées. Par ailleurs, nous nous donnerons donc un modèle paramétrique de la densité de probabilité (DDP) de l'intensité des pixels de chaque région. Cette densité de probabilité sera notée P_{θ_a} sur la cible, avec θ_a l'ensemble des paramètres de la densité de probabilité choisie. A titre d'exemple, si on suppose que l'intensité de chaque pixel de la cible suit une loi gaussienne de moyenne m_a et de variance σ_a^2 , alors $\theta_a = (m_a, \sigma_a^2)$. On définit de même P_{θ_b} la densité de probabilité de l'intensité de chaque pixel du fond.

En supposant connus le support de la cible \mathbf{w}_j , et les paramètres θ_a et θ_b , la probabilité d'observer la scène \mathbf{s} , que l'on nomme *vraisemblance* de la scène peut maintenant s'écrire

$$P(\mathbf{s}|j, \theta_a, \theta_b) = P(\mathbf{s} \cap \mathbf{w}_j | \theta_a) \times P(\mathbf{s} \cap (\mathbf{1} - \mathbf{w}_j) | \theta_b) \quad (1.3)$$

Le choix du modèle paramétrique de la densité de probabilité pour l'intensité des pixels de chaque région sera dicté par des considérations physiques sur la formation de l'image acquise.

A titre d'exemple, on pourra noter les choix classiques suivants, de quelques DDP appartenant à la famille des lois exponentielles.

Système d'acquisition	Densité de probabilité choisie
Imagerie optique pretraitée	Gaussienne
Imagerie cohérente (image d'intensité)	Gamma
Imagerie sonore (image d'amplitude)	Rayleigh
Imagerie à faible flux de photons	Poisson
Imagerie binaire	Bernouilli

Les équations de densités de probabilités [10] que nous avons effectivement utilisées dans ce mémoire sont rappelées dans le tableau 1.1.

Lois	DDP : $P_\theta(x)$	Paramètre : θ
Gaussienne	$\frac{1}{\sqrt{2\pi}\sigma} \exp[-\frac{(x-m)^2}{2\sigma^2}]$	m, σ
Gamma	$\frac{L^L}{m\Gamma(L)} \left(\frac{x}{m}\right)^{L-1} \exp[-L\frac{x}{m}]$	m, L
Poisson	$\sum_{n \in \mathbb{N}} \delta(x - n) \frac{m^n}{n!} \exp[-m]$	m
Bernouilli	$(1 - p)\delta(x) + p\delta(x - 1)$	p

TAB. 1.1: DDP des lois de probabilité considérées dans ce manuscrit. $\delta(x)$ correspond à la distribution de Dirac, \mathbb{N} représente l'ensemble des entiers naturels et $n! = \prod_{k=1}^n k$.

Si l'on considère des densités de probabilités de la famille exponentielle, il est commode dans les calculs de s'intéresser au logarithme de la vraisemblance, que nous nommerons *logvraisemblance*.

$$\mathcal{L}(\mathbf{s}|\mathbf{j}, \theta_a, \theta_b) = \mathcal{L}(\mathbf{s} \cap \mathbf{w}_j|\theta_a) + \mathcal{L}(\mathbf{s} \cap \mathbf{1} - \mathbf{w}_j|\theta_b) \quad (1.4)$$

La figure 1.2 représente un exemple de modèle d'image utilisant une densité de probabilité spatialement décorrelée et Γ d'ordre 1 pour décrire l'intensité des pixels de chaque région. Ce modèle de DDP est classiquement utilisé en imagerie radar.

Nous disposons dès lors d'un modèle statistique précis des images à traiter, et pouvons maintenant définir à la fois localisation et détection de façon rigoureuse.

1.2 Localisation optimale au sens du maximum de vraisemblance

1.2.1 Principe

Dans une tâche de localisation, nous souhaitons déterminer "au mieux" la position de la cible dans une scène, compte tenu des données dont on dispose. Ces données sont la

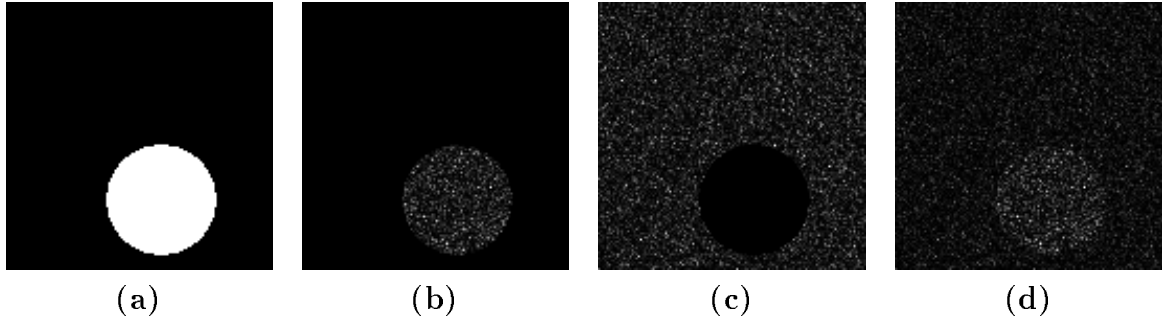


FIG. 1.2: Illustration du modèle d'image SIR : (a) masque \mathbf{w}_j de l'objet présent dans la scène. (b) niveaux de gris de l'objet : champ aléatoire \mathbf{a} (c) niveaux de gris du fond : champ aléatoire \mathbf{b} (d) scène observée. Ici, l'intensité de chaque pixel suit une loi Γ d'ordre 1 et de moyenne $m_a = 1$ sur la cible et $m_b = 2$ sur le fond.

scène observée, la référence de la cible cherchée et les éventuelles informations *a priori* dont on peut disposer sur le paramètre. Dans le cas de la localisation, ce peut être par exemple une connaissance approximative de la zone de l'image où la cible peut se trouver.

Il existe plusieurs méthodes permettant de déterminer le paramètre d'intérêt à partir de ces données. Parmi elles, la théorie statistique de l'estimation présente l'avantage de proposer une définition précise du terme "au mieux", à l'aide du modèle d'image présenté dans la section précédente. Si le vecteur \mathbf{s} représente la scène observée et j le paramètre recherché (en l'occurrence, la position de la cible), on cherchera à maximiser la grandeur $P[j|\mathbf{s}]$, la probabilité que la référence de la cible se trouve à la position j , connaissant la scène \mathbf{s} . Cette grandeur est nommée probabilité *a posteriori*.

L'utilisation des lois de Bayes [10] permet d'exprimer cette probabilité en fonction de la vraisemblance $P[\mathbf{s}|j]$ d'obtenir la scène \mathbf{s} :

$$P[j|\mathbf{s}] = P[\mathbf{s}|j]P[j]/P[\mathbf{s}]. \quad (1.5)$$

$P[\mathbf{s}]$ étant indépendant de j , et en l'absence de connaissance *a priori* sur la position de la cible, nous avons

$$P[j|\mathbf{s}] \propto P[\mathbf{s}|j] \quad (1.6)$$

où le symbole \propto signifie "proportionnel à".

La fonction $P[\mathbf{s}|j]$ s'appelle *fonction de vraisemblance*, car elle représente la probabilité d'observer \mathbf{s} compte tenu de la valeur de j . En d'autres termes, elle exprime la vraisemblance des mesures \mathbf{s} connaissant le paramètre j . La solution optimale pour la localisation au sens du *maximum de vraisemblance* consiste à attribuer à j la valeur qui maximise $P(\mathbf{s}|j)$. Le principe du maximum de vraisemblance signifie qu'il est raisonnable de choisir la valeur de j qui rend les données mesurées \mathbf{s} très vraisemblables. La position de la cible, estimée au sens du maximum de vraisemblance, sera alors choisie comme suit :

$$\hat{j} = \arg \max_{j \in [0:N-1]} P[\mathbf{s}|j] \quad (1.7)$$

Du fait de la monotonie de la fonction logarithme, ceci est équivalent à maximiser la logvraisemblance.

$$\mathcal{L}[\mathbf{s}|j] = \log P[\mathbf{s}|j] \quad (1.8)$$

$$\hat{j} = \arg \max_{j \in [0:N-1]} \mathcal{L}[\mathbf{s}|j] \quad (1.9)$$

1.2.2 Exemple d'application [15]

Supposons que l'intensité des pixels de la cible et du fond suivent une loi gamma d'ordre 1 et que tous les pixels sont indépendants entre eux (bruit décorrélé). Le seul paramètre définissant les DDP de ces lois sont leurs moyennes (cf. tableau 1.1).

On a :

$$\begin{aligned} P_{\mathbf{a}}^{m_a}(\mathbf{a}) &= \prod_{i=0}^{N_a} \frac{1}{m_a} \exp\left[-\frac{a_i}{m_a}\right] \\ P_{\mathbf{b}}^{m_b}(\mathbf{b}) &= \prod_{i=0}^{N_b} \frac{1}{m_b} \exp\left[-\frac{b_i}{m_b}\right] \end{aligned} \quad (1.10)$$

où N_a est le nombre de pixels du support de la cible (cardinal de \mathbf{w}_j) et N_b le nombre de pixels du fond (cardinal de $\bar{\mathbf{w}}_j$).

Pour localiser la cible, nous devons disposer d'une référence de cette cible. Cette référence regroupe les informations dont nous disposons sur la cible. Ces informations consistent en :

- La forme de la cible, un masque binaire que nous noterons simplement \mathbf{w}
- Ses niveaux de gris.

Modéliser les niveaux de gris de la cible sous forme de champ aléatoire (ici un champ gamma d'ordre 1) permet de regrouper la connaissance des niveaux de gris en un seul paramètre : la moyenne de ces niveaux de gris (m_a).

D'après ce qui précède, pour localiser la cible, nous devons déplacer le masque \mathbf{w} en chaque position δ de l'image, pour calculer la vraisemblance de chaque hypothèse de position. Le masque \mathbf{w} , déplacé par translation à la position δ de la scène sera noté \mathbf{w}_δ . Alors, d'après les équations 1.4 et 1.10, l'expression de la logvraisemblance peut s'écrire :

$$\mathcal{L}(\delta, m_a, m_b) = -N_a \log(m_a) - N_b \log(m_b) - \frac{1}{m_a} \sum_{\mathbf{w}_\delta} s_i - \frac{1}{m_b} \sum_{\bar{\mathbf{w}}_\delta} s_i \quad (1.11)$$

On peut remarquer que :

$$\begin{aligned} \sum_{\mathbf{w}_\delta} s_i &= \sum_{i=0}^{N-1} s_i w_{i-\delta} = [\mathbf{s} \otimes \mathbf{w}]_\delta \\ \sum_{\bar{\mathbf{w}}_\delta} s_i &= \sum_{i=0}^{N-1} s_i (1 - w_{i-\delta}) = \sum_{i=0}^{N-1} s_i - [\mathbf{s} \otimes \mathbf{w}]_\delta \end{aligned}$$

où \otimes représente l'opérateur de corrélation. L'équation 1.11 peut donc s'écrire :

$$\mathcal{L}(\delta, m_a, m_b) = -N_a \log(m_a) - N_b \log(m_b) - \frac{1}{m_a} [\mathbf{s} \otimes \mathbf{w}]_\delta - \frac{1}{m_b} \left(\sum_{i=0}^{N-1} s_i - [\mathbf{s} \otimes \mathbf{w}]_\delta \right) \quad (1.12)$$

Cette expression laisse supposer que les paramètres des lois (dans notre exemple, les moyennes de la cible et du fond) sont connus *a priori*. En pratique, il est courant que cette connaissance ne soit pas disponible. Ces paramètres sont alors qualifiés de paramètres de *nuisance*. Leurs valeurs ne nous intéressent pas mais leur méconnaissance empêche le calcul de 1.12. Un moyen simple de gérer ce problème consiste à remplacer m_a et m_b dans l'équation 1.12 par leurs estimées au sens du maximum de vraisemblance. On choisira comme valeur des paramètres de nuisance celle qui donne la vraisemblance la plus grande :

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial m_a} = 0 & \iff \hat{m}_a = \frac{1}{N_a} \sum_{\mathbf{w}_\delta} s_i = \frac{1}{N_a} [\mathbf{s} \otimes \mathbf{w}]_\delta \\ \frac{\partial \mathcal{L}}{\partial m_b} = 0 & \iff \hat{m}_b = \frac{1}{N_b} \sum_{\bar{\mathbf{w}}_\delta} s_i = \frac{1}{N_b} \left(\sum_{i=0}^{N-1} s_i - [\mathbf{s} \otimes \mathbf{w}]_\delta \right) \end{aligned} \quad (1.13)$$

Le filtre optimal en localisation pour ce modèle lorsque les paramètres sont inconnus est alors :

$$\mathcal{L}_\delta^{inc} = -N_a \log \left(\frac{1}{N_a} [\mathbf{s} \otimes \mathbf{w}]_\delta \right) - N_b \log \frac{1}{N_b} \left(\sum_{i=0}^N s_i - [\mathbf{s} \otimes \mathbf{w}]_\delta \right) \quad (1.14)$$

Il est important de noter que la grandeur \mathcal{L}_δ^{inc} , du fait de l'estimation des paramètres n'est plus le logarithme d'une vraisemblance à proprement parler. En toute rigueur, il faut utiliser le terme de pseudo vraisemblance, ou de *Vraisemblance Généralisée* (Generalized Likelihood, en anglais).

La figure 1.3 illustre un exemple de localisation d'objet par le filtre de l'équation 1.14, dans une scène correspondant au modèle de scène SIR. Ce filtre permet de localiser la cible sans difficulté.

La référence dont nous disposons pour localiser la cible n'est qu'un masque binaire, représentant la forme de la cible. Lorsque nous travaillons avec le filtre de l'équation 1.14, estimant les paramètres de la cible, cela signifie que cette forme est la seule connaissance *a priori* dont nous disposons.

Ceci permet de définir des algorithmes capables de localiser des cibles dans un ensemble d'images sur lequel les niveaux de gris de la cible sont susceptibles de fluctuer fortement d'une image à l'autre.

Le prix à payer est une description très simple de la cible, sous la forme d'une région homogène.

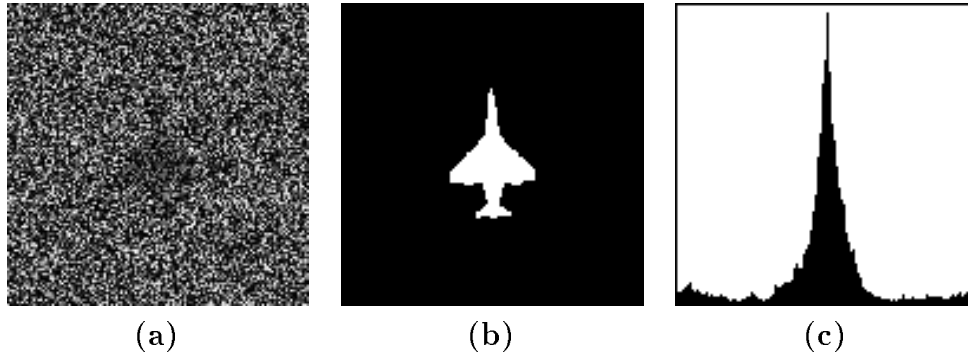


FIG. 1.3: **(a)** : scène \mathbf{s} correspondant au modèle de l'équation 1.14. **(b)** : objet de référence \mathbf{w} utilisé pour construire les filtres. **(c)** : maximum de chaque colonne du plan de sortie résultant du traitement de l'image (a) avec le filtre de l'équation 1.14.

1.3 Détection optimale au sens du rapport de vraisemblance [19]

Pour une opération de détection, il s'agit, dans un échantillon donné, de déterminer si la cible est présente ou non. En d'autres termes, en chaque point de la scène, il s'agira de déterminer s'il existe une cible à cette position ou non.

1.3.1 Principe

L'approche optimale donnée par la théorie de la décision est la suivante : Il faut définir une fenêtre d'analyse \mathbf{F}_δ , centrée sur le δ^{eme} pixel de la scène. Dans cette fenêtre d'analyse, nous cherchons à effectuer un test permettant de discriminer les deux hypothèses suivantes :

$$\begin{cases} \text{Hypothèse } \mathcal{H}_{1,\delta} & \text{La cible est présente au centre de la fenêtre } F_\delta \\ \text{Hypothèse } \mathcal{H}_{0,\delta} & \text{La fenêtre } F_\delta \text{ ne contient que du bruit} \end{cases} \quad (1.15)$$

Ces hypothèses sont résumées dans la figure 1.4 :

Comme dans le cas de la localisation, les grandeurs qui nous intéressent sont les probabilités suivantes :

$$\begin{cases} P(\mathcal{H}_{1,\delta}|\mathbf{s}) & \text{Probabilité de l'hypothèse } \mathcal{H}_{1,\delta}, \text{ connaissant la scène } \mathbf{s} \\ P(\mathcal{H}_{0,\delta}|\mathbf{s}) & \text{Probabilité de l'hypothèse } \mathcal{H}_{0,\delta}, \text{ connaissant la scène } \mathbf{s} \end{cases} \quad (1.16)$$

De ces grandeurs, on tire le rapport des probabilités suivant :

$$R_\delta = \frac{P(\mathcal{H}_{1,\delta}|\mathbf{s})}{P(\mathcal{H}_{0,\delta}|\mathbf{s})} = \frac{P(\mathbf{s}|\mathcal{H}_{1,\delta})}{P(\mathbf{s}|\mathcal{H}_{0,\delta})} \times \frac{P(\mathcal{H}_{1,\delta})}{P(\mathcal{H}_{0,\delta})} \quad (1.17)$$

La règle qui minimise la probabilité d'erreur dans la prise de décision est donnée par

$$\begin{cases} R_\delta > 1 \rightarrow & \text{détection d'une cible} \\ R_\delta < 1 \rightarrow & \text{non détection d'une cible} \end{cases} \quad (1.18)$$

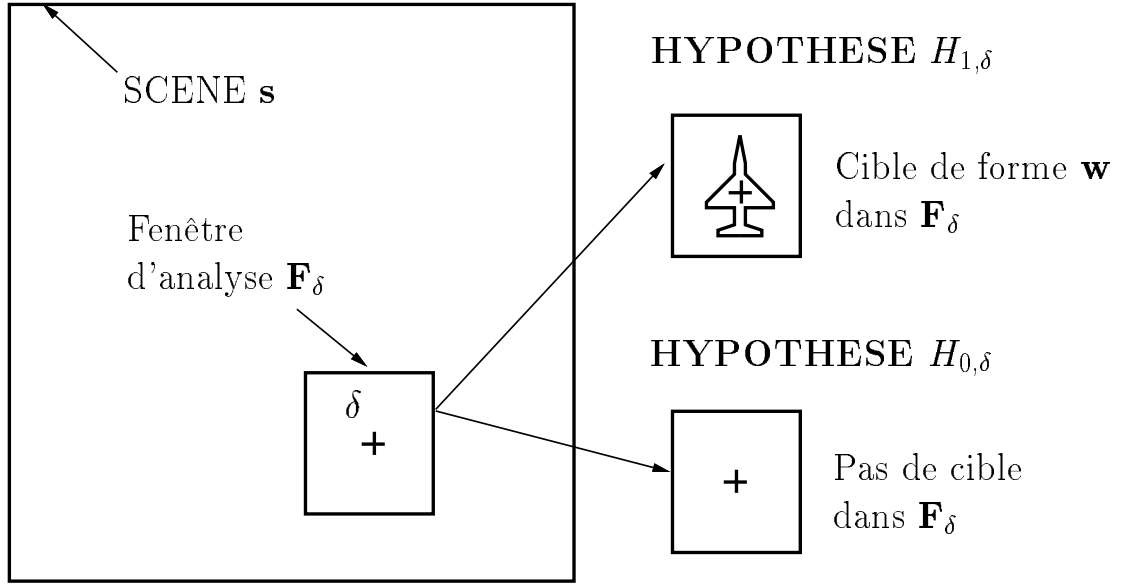


FIG. 1.4: Principe de la détection d'une cible dans une image en utilisant le rapport de vraisemblance.

La grandeur $P(\mathcal{H}_{0,\delta}|\mathbf{s})$ (respectivement $P(\mathcal{H}_{1,\delta}|\mathbf{s})$) est nommée vraisemblance de l'hypothèse $\mathcal{H}_{0,\delta}$ (resp. $\mathcal{H}_{1,\delta}$).

Définissons alors le rapport de vraisemblance Λ_δ :

$$\Lambda_\delta = \frac{P(\mathbf{s}|\mathcal{H}_{1,\delta})}{P(\mathbf{s}|\mathcal{H}_{0,\delta})} \quad (1.19)$$

La règle de décision donnée par l'équation 1.18 est équivalente à la suivante :

$$\begin{cases} \Lambda_\delta > \frac{P(\mathcal{H}_{0,\delta})}{P(\mathcal{H}_{1,\delta})} \rightarrow \text{détection d'une cible} \\ \Lambda_\delta < \frac{P(\mathcal{H}_{0,\delta})}{P(\mathcal{H}_{1,\delta})} \rightarrow \text{non détection d'une cible} \end{cases} \quad (1.20)$$

Notons que nous utilisons des probabilités $P(\mathcal{H}_{0,\delta})$ et $P(\mathcal{H}_{1,\delta})$, qui sont respectivement les probabilités que la cible soit à la position δ , et que rien ne se trouve à la position δ , indépendamment de la scène observée. Donner une mesure de probabilité pour ces événements est au cœur même des théories bayésiennes. Nous ne discuterons pas ici ce point, qui tient aujourd'hui encore plus du débat philosophique que des sciences appliquées. Il n'en reste pas moins que la quantité $\frac{P(\mathcal{H}_{0,\delta})}{P(\mathcal{H}_{1,\delta})}$ est difficile à évaluer. Lorsque l'on travaille sur une image isolée, nous pourrions faire l'hypothèse qu'elle est indépendante de la position testée, puisqu'aucune position n'est *a priori* plus susceptible d'être le centre de la cible¹.

Néanmoins, nous ne pouvons pas la déterminer plus précisément, puisque la fréquence d'apparition d'une cible est inconnue. Nous ne pouvons donc que la fixer arbitrairement, en lui attribuant une valeur τ . Dès lors, la décision de détection est prise en considérant la

¹Ceci est faux dans un contexte de poursuite dans une séquence d'image. Dans ce cas, il est intéressant de relier la position de la cible dans l'image au temps t , à celle estimée dans l'image au temps $t - 1$.

valeur de ce rapport de vraisemblance par rapport à τ . Ceci revient à choisir de détecter une cible lorsque la vraisemblance qu'une cible soit présente est τ fois supérieure à la vraisemblance de l'absence de cible.

$$\begin{cases} \Lambda_\delta > \tau \rightarrow & \text{déttection d'une cible} \\ \Lambda_\delta < \tau \rightarrow & \text{non détection d'une cible} \end{cases} \quad (1.21)$$

Enfin, prendre le logarithme du rapport de vraisemblance est équivalent à prendre la différence des logvraisemblances, nous aurons donc recours à la grandeur \mathcal{R}_δ définie par :

$$\mathcal{R}_\delta = \log \Lambda_\delta \quad (1.22)$$

Des résultats équivalents sont obtenus en considérant l'approche de Neyman-Pearson, qui s'interprète plus simplement.

1.3.2 Neyman-Pearson

La stratégie de Neyman-Pearson consiste à fixer un taux d'erreur acceptable dans la prise de décision. Par exemple, la probabilité de fausse alarme (PFA) est la probabilité de décider qu'une cible est présente alors que la fenêtre d'analyse ne contient que du fond. On fixe alors un seuil correspondant à la PFA tolérée, et l'on cherche à maximiser la probabilité de détection pour cette PFA. On peut montrer que le rapport de vraisemblance donné par l'équation 1.19 réalise cette optimisation.

La règle de décision ainsi obtenue est identique à celle de l'équation 1.21, mais le seuil τ est ici déterminé à partir d'une erreur acceptable, et non plus d'une connaissance a priori sur les probabilités de présence ou d'absence de cible.

Nous dirons donc que le rapport de vraisemblance est optimal en détection, car il fournit la meilleure probabilité de détection pour une PFA fixée.

1.3.3 Rapport de vraisemblance généralisé

Comme précédemment, lorsque les paramètres des lois sont inconnus, pour pouvoir calculer 1.19, il est possible de les estimer au sens du maximum de vraisemblance. Dès lors, on ne peut plus parler de Rapport de vraisemblance mais de Rapport de Vraisemblance Généralisé, que nous noterons GLRT pour *Generalized Likelihood Ratio Test*.

Reprenons l'exemple des champs aléatoires de statistiques gamma d'ordre 1 utilisé dans la section 1.2.2 en localisation. Le calcul de la vraisemblance de l'hypothèse $\mathcal{H}_{1,\delta}$ est très semblable à celui présenté précédemment, si ce n'est que le fond est maintenant défini comme le complémentaire du support de la référence dans la fenêtre d'analyse F_δ . Le calcul de la vraisemblance dans l'hypothèse $\mathcal{H}_{0,\delta}$ consiste simplement à calculer la vraisemblance d'une région homogène dans cette fenêtre.

Si les paramètres de ces lois sont connus, on obtient facilement :

$$\begin{aligned} \mathcal{R}_\delta(m_a, m_b) &= -N_a \log m_a - N_b \log m_b + N_F \log m_b \\ &\quad - \frac{1}{m_a} [\mathbf{s} \otimes \mathbf{w}]_\delta - \frac{1}{m_b} ([\mathbf{s} \otimes \bar{\mathbf{w}}]_\delta) \\ &\quad + \frac{1}{m_b} ([\mathbf{s} \otimes \mathbf{F}]_\delta) \end{aligned} \quad (1.23)$$

Si les paramètres sont inconnus, nous reprenons la solution qui consiste à les estimer au sens du maximum de vraisemblance. Lorsque les paramètres doivent être estimés, il est nécessaire d'établir une différence entre la définition des paramètres du fond estimés sur la fenêtre, et la définition des paramètres du fond estimés sur le complémentaire de la cible dans la fenêtre, car il s'agit d'estimation de variables aléatoires effectuées dans des hypothèses différentes, donc sans lien. Le logarithme du rapport de vraisemblance généralisé s'exprime alors par :

$$\mathcal{R}_\delta^{inc} = -N_a \log \left(\frac{1}{N_a} [\mathbf{s} \otimes \mathbf{w}]_\delta \right) - N_b \log \left(\frac{1}{N_b} [\mathbf{s} \otimes \bar{\mathbf{w}}]_\delta \right) + N_F \log \left(\frac{1}{N_F} [\mathbf{s} \otimes \mathbf{F}]_\delta \right) \quad (1.24)$$

Avant de présenter les performances en détection de ce type de filtres, nous devons définir un critère permettant de mesurer ces performances.

1.3.4 Critère de qualité de la détection

Dans une opération de détection, le seul critère valide devrait définir la capacité de l'algorithme à discriminer les deux hypothèses : présence ou absence de cible. A chaque valeur du seuil de détection τ correspond la probabilité de détection (probabilité de prendre la décision qu'une cible est présente, lorsque cette cible est effectivement présente), que nous noterons PD. Mais ce seuil correspond également à une probabilité de Fausse Alarme, comme nous l'avons signalé en présentant l'approche de Neyman-Pearson, qui est la probabilité de prendre la décision qu'une cible est présente alors que la fenêtre d'analyse ne contient que du fond.

En factorisant ces deux grandeurs, on obtient la courbe COR (Courbes Opérationnelles du Récepteur) des algorithmes de détection, qui exprime la probabilité de détection en fonction de la probabilité de Fausse alarme. Ces courbes sont paramétrées par le seuil de détection choisi. Un seuil de détection nul correspond à une probabilité de détection de 1, mais à une probabilité de fausse alarme également de valeur 1. Un seuil infini permet d'obtenir une PFA nulle, au prix d'une PD nulle également. Ce que signifie l'approche de Neyman-Pearson est qu'aucun algorithme ne devrait pouvoir fournir de courbe supérieure à l'algorithme du Rapport de vraisemblance. Il faut bien noter que dans le cas du rapport de vraisemblance généralisé (GLRT) ceci n'est pas assuré de façon théorique.

La figure 1.5 présente une courbe COR représentative des performances des algorithmes présentés précédemment, ainsi que quelques images représentatives de ce qu'expriment ces courbes.

1.4 Quelques remarques complémentaires sur ces deux algorithmes

1.4.1 Complexité de ces algorithmes

Quel que soit l'algorithme utilisé, un regard sur les équations 1.14 et 1.24 permet de remarquer que les opérations les plus coûteuses en temps de calcul sont de simples

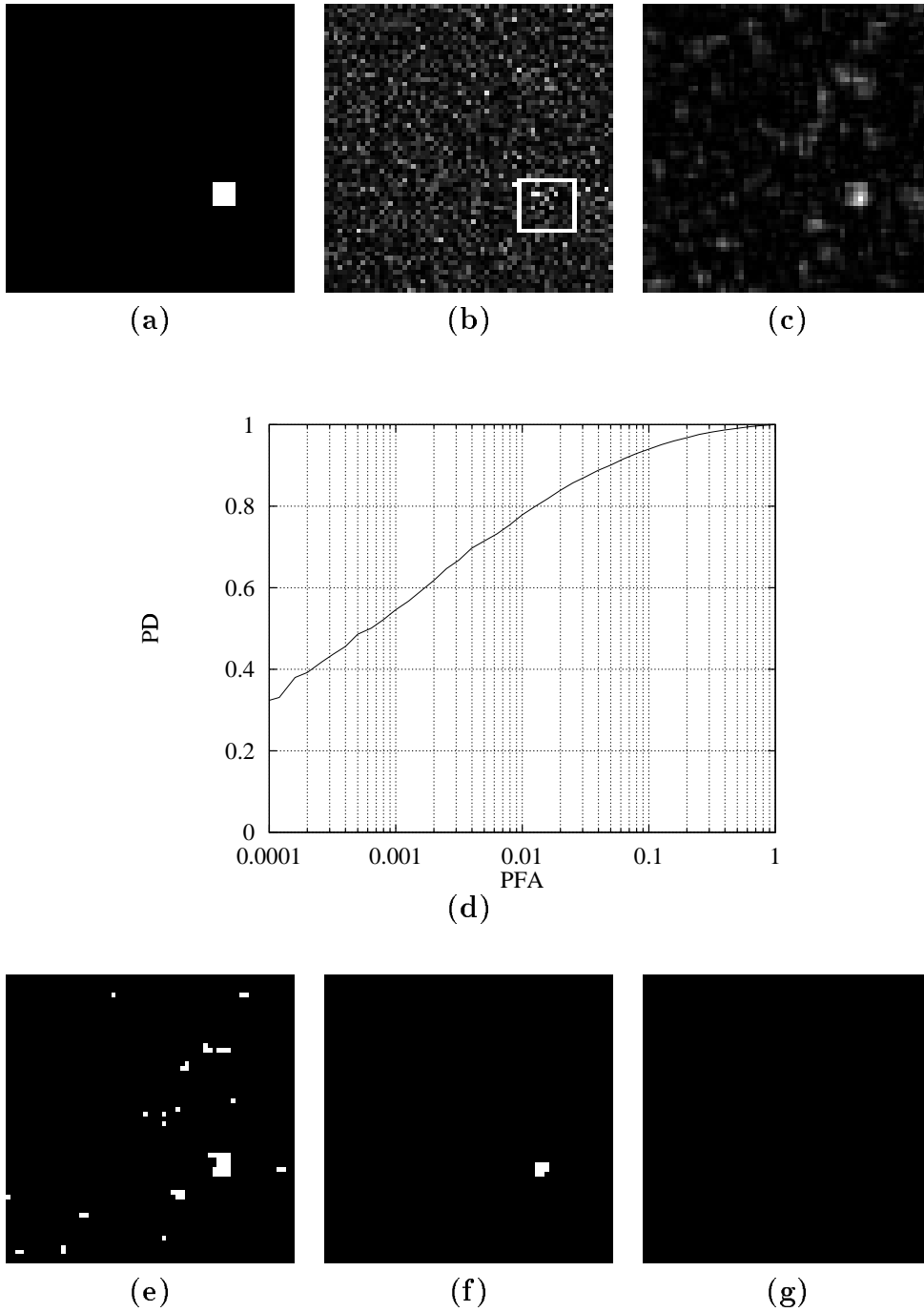


FIG. 1.5: Courbe COR, et images équivalentes. **(a)** Cible carrée de 5×5 pixels apparaissant dans **(b)**. **(b)** Scène composée de bruit gamma d'ordre 1. La cible a une moyenne de 2, le fond une moyenne de 1. **(c)** Plan de sortie de l'algorithme \mathcal{R}_δ^{inc} (cf eq. 1.24). **(d)** Courbe COR obtenue avec cet algorithme **(e)** Plan de sortie (c) seuillé pour obtenir une PFA de 10^{-2} **(f)** Plan de sortie (c) seuillé pour obtenir une PFA de 10^{-3} **(g)** Plan de sortie (c) seuillé pour obtenir une PFA de 10^{-4}

corrélations. On en compte une dans le cas de l'algorithme du ML et deux pour le GLRT, dans le cas de bruit gamma. Si N est le nombre de pixels de l'image à traiter, la complexité d'une corrélation est en $O(N \log N)$, si l'on utilise une transformée de Fourier Rapide.

Les autres opérations sont peuvent être menées comme un post-traitement en chaque pixel, donc d'une complexité en $O(N)$.

Ces deux algorithmes sont donc d'une complexité comparable à celle de filtres linéaires, comme le filtre adapté, sans présenter les mêmes inconvénients face à du bruit non recouvrant.

1.4.2 Localisation

Dans une tâche de localisation, l'algorithme à utiliser est l'algorithme du maximum de vraisemblance (ML). Utiliser l'algorithme du GLRT pour ces cas correspond à une méthode empirique. Néanmoins, il a été montré [19] qu'il permet parfois de dépasser les performances de l'algorithme ML lorsque les scènes ne correspondent pas au modèle SIR, qui suppose les fonds homogènes. Nous reviendrons sur ce point dans le chapitre 2.

Avant cela, il est important d'observer les liens existants entre ces deux approches. Ceci est obtenu très rapidement en comparant les équations respectives de la vraisemblance et du rapport de vraisemblance.

Lorsque les paramètres θ_a de la cible et les paramètres θ_b du fond sont connus, nous pouvons reprendre l'expression de la logvraisemblance de l'équation 1.4, pour définir la logvraisemblance \mathcal{L}_δ .

$$\mathcal{L}_\delta = \log P(\mathbf{s} \cap \mathbf{w}_\delta | \theta_a) + \log P(\mathbf{s} \cap (1 - \mathbf{w}_\delta) | \theta_b) \quad (1.25)$$

Du fait de l'indépendance statistique entre les régions \mathbf{w}_δ et $(1 - \mathbf{w}_\delta)$, on a :

$$\log P(\mathbf{s} \cap (1 - \mathbf{w}_\delta) | \theta_b) = -\log P(\mathbf{s} \cap \mathbf{w}_\delta | \theta_b) + \log P(\mathbf{s} | \theta_b) \quad (1.26)$$

Le dernier terme de cette équation est indépendant de δ , et donc sans intérêt en localisation, ce qui amène l'expression simplifiée suivante :

$$\mathcal{L}_\delta = \log P(\mathbf{s} \cap \mathbf{w}_\delta | \theta_a) - \log P(\mathbf{s} \cap \mathbf{w}_\delta | \theta_b) \quad (1.27)$$

De même, en partant de l'expression du rapport de vraisemblance \mathcal{R}_δ :

$$\mathcal{R}_\delta = \log P(\mathbf{s} \cap \mathbf{w}_\delta | \theta_a) + \log P(\mathbf{s} \cap (1 - \mathbf{w}_\delta) \cap \mathbf{F}_\delta | \theta_b) - \log P(\mathbf{s} \cap \mathbf{F}_\delta | \theta_b) \quad (1.28)$$

En remarquant que $\mathbf{F}_\delta = \mathbf{w}_\delta + (1 - \mathbf{w}_\delta)$, on arrive simplement à

$$\mathcal{R}_\delta = \log P(\mathbf{s} \cap \mathbf{w}_\delta | \theta_a) - \log P(\mathbf{s} \cap \mathbf{w}_\delta | \theta_b) \quad (1.29)$$

ce qui montre l'équivalence entre les approches du rapport de vraisemblance et du maximum de vraisemblance lorsque les paramètres sont connus. En d'autre termes, lorsque les paramètres sont connus, la taille de la fenêtre F n'intervient donc pas dans le calcul du rapport de vraisemblance.

Cette situation change lorsque les paramètres sont inconnus et estimés au sens du Maximum de vraisemblance. Dans ce cas, l'estimation au sens du maximum de vraisemblance des paramètres θ_a et θ_b nécessite la connaissance des densités de probabilité paramétriques des champs \mathbf{a} et \mathbf{b} . Dans ce cas, le lien peut être fait entre le GLRT et le ML, lorsque la taille de la fenêtre est la plus grande possible, dans notre cas, de la taille de la fenêtre. Dans ce cas, ces deux algorithmes sont encore équivalents. Illustrons ce qui vient d'être dit.

Insistons sur le fait que lorsque les paramètres doivent être estimés, il est nécessaire d'établir une différence entre la définition des paramètres du fond estimés sur la fenêtre, et la définition des paramètres du fond estimés sur le complémentaire de la cible dans la fenêtre, car il s'agit d'estimation de variables aléatoires effectuées dans des hypothèses différentes, donc sans lien.

Les paramètres estimés dans chaque région seront notés $\hat{\theta}_a$, $\hat{\theta}_b$, $\hat{\theta}_c$, pour les régions cibles, complémentaire de la cible et fenêtre. Le GLRT s'écrit alors :

$$\mathcal{R}_\delta^{inc} = \log P(\mathbf{s} \cap \mathbf{w}_\delta | \hat{\theta}_a) + \log P(\mathbf{s} \cap (1 - \mathbf{w}_\delta) \cap \mathbf{F}_\delta | \hat{\theta}_b) - \log P(\mathbf{s} \cap \mathbf{F}_\delta | \hat{\theta}_c) \quad (1.30)$$

et le ML :

$$\mathcal{L}_\delta^{inc} = \log P(\mathbf{s} \cap \mathbf{w}_\delta | \hat{\theta}_a) + \log P(\mathbf{s} \cap (1 - \mathbf{w}_\delta) | \hat{\theta}_b) \quad (1.31)$$

Lorsque la fenêtre F utilisée est de la taille de l'image toute entière, on a instantanément $\mathbf{s} \cap (1 - \mathbf{w}_\delta) \cap \mathbf{F}_\delta = \mathbf{s} \cap (1 - \mathbf{w}_\delta)$ et $\mathbf{s} \cap \mathbf{F}_\delta = \mathbf{s}$. Le calcul du second terme de l'équation 1.30 est le même que celui de l'équation 1.31, et le dernier terme de l'équation 1.30 devient indépendant de la position δ testée. Maximiser \mathcal{R}_δ^{inc} ou maximiser \mathcal{L}_δ^{inc} est alors strictement équivalent.

1.4.3 Détection à Taux de Fausse Alarme constant

La principale différence entre détection et localisation, comme évoquée dans l'introduction, se situe dans la définition d'un seuil de détection. Nous nous concentrerons ici sur le cas d'algorithmes supposant les paramètres des régions inconnus. Comme cela a été dit précédemment, l'approche la plus commune en détection consiste à définir un seuil de détection, que l'on relie à la Probabilité de Fausse Alarme acceptée.

Lorsque aucune cible n'est présente à un emplacement donné, nous sommes en droit d'attendre que la probabilité de Fausse Alarme ne soit pas dépendante des paramètres statistiques du fond. Ceci permet de fixer un seuil de détection qui soit indépendant des paramètres du fond. On dit alors que le détecteur est à Taux de Fausse Alarme Constant (TFAC).

La démonstration du fait que le GLRT est un détecteur à taux de fausse alarme constant est quasi immédiate d'après ce qui précède : nous ne l'explicitons que pour le cas des lois Gamma, bien que ce résultat puisse être étendu au cas de lois gaussiennes (cf. Annexe B.1).

Soit un échantillon homogène \mathbf{s} , de moyenne m_b . Appliquons lui la transformation suivante :

$$\mathbf{s}' = \lambda \mathbf{s}$$

\mathbf{s}' est alors un échantillon homogène, obéissant à une loi Gamma de même ordre que \mathbf{s} , et de moyenne λm_b .

Observons l'expression du GLRT à paramètres inconnus appliqué à la scène \mathbf{s}' :

$$\begin{aligned}
 \mathcal{R}_\delta^{inc}(\mathbf{s}') &= -N_a \log \left(\frac{1}{N_a} [\mathbf{s}' \otimes \mathbf{w}]_\delta \right) - N_b \log \left(\frac{1}{N_b} [\mathbf{s}' \otimes \bar{\mathbf{w}}]_\delta \right) + N_F \log \left(\frac{1}{N_F} [\mathbf{s}' \otimes \mathbf{F}]_\delta \right) \\
 &= -N_a \log \left(\lambda \frac{1}{N_a} [\mathbf{s} \otimes \mathbf{w}]_\delta \right) - N_b \log \left(\lambda \frac{1}{N_b} [\mathbf{s} \otimes \bar{\mathbf{w}}]_\delta \right) + N_F \log \left(\lambda \frac{1}{N_F} [\mathbf{s} \otimes \mathbf{F}]_\delta \right) \\
 &= \mathcal{R}_\delta^{inc}(\mathbf{s}) + (N_F - N_a - N_b) \log(\lambda) \\
 &= \mathcal{R}_\delta^{inc}(\mathbf{s})
 \end{aligned} \tag{1.32}$$

La réponse du GLRT à \mathbf{s}' est donc la même que sa réponse à \mathbf{s} , ce qui assure un taux de fausse alarme constant.

Une démonstration similaire permet de montrer que l'algorithme ML n'est pas TFAC. On peut donc voir l'estimation de la vraisemblance de l'hypothèse \mathcal{H}_0 comme une normalisation de la sortie de l'algorithme par rapport à la moyenne du fond.

Qu'il s'agisse de détection ou de localisation, ces deux algorithmes sont proches, bien que le ML présente l'avantage d'offrir un cadre théorique plus sûr dans le cas de la localisation, et que le fait que le GLRT présente un taux de Fausse Alarme indépendant de la moyenne du fond soit un atout majeur en détection.

1.5 Conclusion

Dans ce chapitre d'introduction, nous avons repris les bases théoriques permettant de définir des algorithmes ayant un caractère optimal dans le contexte de l'estimation statistique. Il nous a en effet semblé important de disposer de tels algorithmes, dont les performances sont relativement faciles à caractériser. Définir si un algorithme est meilleur qu'un autre est en effet difficile, dans un contexte aussi vaste que celui du traitement d'image. La variété des images à traiter est en effet telle qu'il est hors de question de définir un algorithme performant dans tous les cas. Dans le cadre que nous avons choisi, le choix d'un algorithme est un peu simplifié : Si les images correspondent au modèle, nous sommes assurés qu'aucun algorithme de localisation n'obtiendra de résultats meilleurs que les algorithmes optimisant la probabilité de localisation correcte.

Nous avons donc choisi de reprendre un modèle d'image (SIR) précis [18], correspondant à du bruit non recouvrant, qui nous a semblé plus adapté que le modèle de bruit additif classique aux images les plus courantes.

Dans le cas de cibles définies uniquement par leur forme, sans connaissance sur l'intensité de leurs niveaux de gris, ce modèle a permis de définir des algorithmes que nous avons divisés en deux classes, algorithmes de localisation (de type ML) et algorithmes de détection (de type GLRT). Nous insistons sur le fait que ces algorithmes ont été développés préalablement aux travaux présentés dans ce mémoire.

Nous avons ensuite brièvement noté les différences entre ces algorithmes. Notamment, l'algorithme GLRT offre une propriété très pratique : celle de fournir une réponse indépendante de la moyenne du fond. Ceci est obtenu au prix de l'apparition d'un paramètre

supplémentaire : la définition de la fenêtre d'analyse \mathbf{F} . L'un de nos objectifs a donc été d'analyser l'influence de cette fenêtre sur les caractéristiques du GLRT.

En ayant souligné la différence entre les domaines d'optimalité respectifs du ML et du GLRT , nous pouvons maintenant définir le GLRT comme un algorithme sous optimal en localisation, qui tends vers l'algorithme optimal du ML, lorsque la taille de la fenêtre d'analyse tends à occuper toute l'image.

Ceci simplifie considérablement l'étude qui fait l'objet du chapitre suivant. Celle-ci consistera simplement à montrer que l'on peut définir des algorithmes GLRT présentant des performances en localisation proches de celles des algorithmes du ML, sans pour autant devoir utiliser des fenêtres trop grandes.

Chapitre 2

Robustesse de l’algorithme GLRT en bruit non homogène

Sommaire

2.1	Influence de la taille de la fenêtre d’analyse en bruit homogène	23
2.1.1	La connaissance des paramètres du fond est une hypothèse forte	23
2.1.2	Utiliser de grandes fenêtres procure une meilleure estimation des paramètres du fond	25
2.1.3	Limite supérieure à la taille des fenêtres	27
2.1.4	Généralisation à différents contrastes	28
2.1.5	Généralisation aux bruits Poissonien et Gaussien	28
2.1.6	Conclusion en bruit homogène	29
2.2	Bruit non homogène	30
2.2.1	Modèle d’images inhomogènes	31
2.2.2	Echec des grandes fenêtres	31
2.2.3	Inhomogénéités trop importantes	33
2.3	Conclusion	33

Nous avons précédemment présenté l’algorithme GLRT, utilisable en localisation comme une alternative possible à l’algorithme du Maximum de Vraisemblance. Comme nous l’avons vu, l’algorithme GLRT nécessite l’utilisation d’une fenêtre d’analyse. La taille de cette fenêtre est pour le moment indéfinie.

Dans son utilisation originale en détection de contours, Oliver [38] utilise ce filtre pour détecter non pas une région de forme connue incluse dans un fond, mais une rupture de statistique entre deux régions de tailles et de formes inconnues. La taille de la fenêtre est définie par la taille caractéristique des régions présentes dans les images, afin que la fenêtre à une position donnée ne recouvre au maximum que deux régions. De plus, comme il recherche une rupture de statistique, aucune région ne doit être privilégiée en taille. Sa fenêtre est donc divisée en deux parties égales. Ceci, ramené à notre problème où la taille d’une région (la cible) est fixée, consisterait à prendre une taille de fenêtre égale au double de la taille de la cible. Ce choix n’est pas forcément judicieux pour la détection de cibles.

Une autre utilisation du GLRT a été proposée par *Guérault et al* [19], dans le cadre de la localisation d'objets apparaissant sur des fonds réalistes. Dans ce cas, le GLRT est présenté comme un moyen d'améliorer les performances de l'algorithme du Maximum de Vraisemblance, lorsque les scènes à traiter ne correspondent plus au modèle d'image SIR présenté dans la section 1.1. La taille de la fenêtre est alors choisie en fonction des inhomogénéités du fond, mais constitue un paramètre libre de l'algorithme. Ajoutons que ces travaux ont été menés en considérant le problème de la localisation de cibles de grandes tailles, pour lesquels les problèmes d'estimations de paramètres sont moins cruciaux que dans notre cas.

Notre travail a consisté à approfondir cette dernière approche, pour expliciter de façon plus précise les motivations possibles pour le choix de cette taille dans le cas de petites cibles.

Nous montrerons que le choix de cette taille relève d'un compromis à effectuer entre estimation correcte des paramètres du fond dans le cas de fond homogène, et robustesse de l'algorithme lorsque le fond est inhomogène. Nous montrerons de plus que ce compromis est réalisé de façon efficace pour des fenêtres à peine plus grandes que la cible. Nous établirons ce résultat pour différents bruits de la famille exponentielle (Gamma, Poisson, gaussien). Pour de telles fenêtres, les performances de l'algorithme GLRT sont équivalentes à celles du Maximum de Vraisemblance lorsque le bruit est homogène. En revanche, lorsque le fond est inhomogène, les performances de l'algorithme GLRT seront bien meilleures.

Dans la première partie, nous nous concentrerons sur le cas de bruit homogène. Nous comparerons les performances de 4 algorithmes GLRT se différenciant par la connaissance ou l'estimation des paramètres dans chacune des régions. Nous montrerons ainsi que la connaissance du fond est la plus déterminante pour ces performances. Dès lors, lorsque ces paramètres ne sont pas connus, il sera primordial de les estimer avec précision, c'est à dire d'augmenter la taille de la fenêtre d'analyse. Nous montrons de plus qu'à partir d'une certaine taille de la fenêtre, le gain obtenu grâce à la meilleure estimation du fond tend à stagner. Nous pourrions ainsi limiter la taille de la fenêtre sans que les performances en soient affectées.

Dans la seconde partie, nous étudierons le cas d'images de bruit non homogènes, donc ne correspondant pas au cas d'optimalité des algorithmes du Maximum de Vraisemblance et du GLRT. Nous présenterons un modèle utilisable pour synthétiser de telles images. Nous montrerons que dans ce cas, le GLRT est très supérieur au maximum de vraisemblance. Il est en effet préférable de limiter l'influence de l'inhomogénéité du fond, ce qui peut être obtenu en utilisant des fenêtres de petites tailles.

2.1 Influence de la taille de la fenêtre d'analyse en bruit homogène

2.1.1 La connaissance des paramètres du fond est une hypothèse forte

Dans cette section, nous évaluerons l'influence de l'estimation des différents paramètres de nuisance de l'algorithme GLRT. Ces paramètres de nuisance sont les paramètres des lois décrivant le bruit présent dans les images. Pour cela, nous comparerons les résultats obtenus lorsque l'algorithme GLRT estime ces paramètres au sens du MV, et les résultats obtenus dans le cas limite d'une estimation parfaite, correspondant au cas où ces paramètres sont connus.

Nous définissons ainsi 4 versions différentes de l'algorithme GLRT :

- $GLRT_C^C$: Les paramètres de la cible et du fond sont connus.
- $GLRT_C^E$: Les paramètres de la cible sont connus, ceux du fond sont estimés.
- $GLRT_E^C$: Les paramètres de la cible sont estimés, ceux du fond sont connus.
- $GLRT_E^E$: Les paramètres de la cible et du fond sont estimés.

Nous prenons ici l'exemple où les niveaux de gris de la cible et du fond sont décorrélés et suivent une distribution Gamma d'ordre L . Dans ce cas, les paramètres de nuisance se résument à m_c (moyenne de la cible) et m_f (moyenne du fond). S'ils ne sont pas connus, nous les remplaçons par leur estimation au sens du maximum de vraisemblance, qui se résume dans ce cas au calcul d'une moyenne arithmétique.

La figure 2.1 rappelle les principales notations utilisées dans les équations qui suivent.

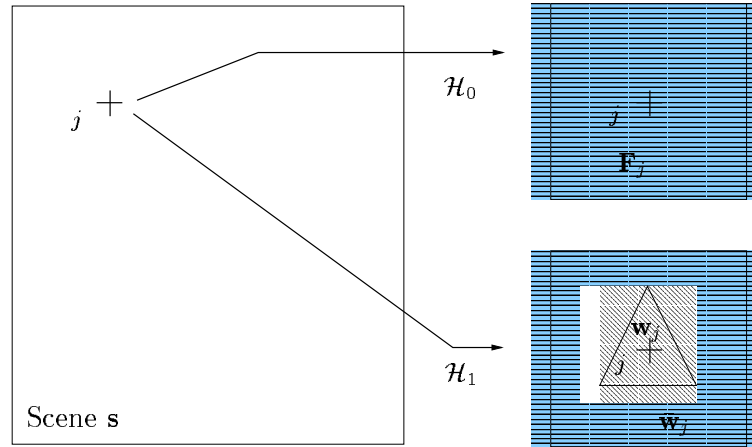


FIG. 2.1: Principales notations utilisées : En chaque position j de la scène s , on teste à l'intérieur de la fenêtre F_j les deux hypothèses : H_0 : fenêtre ne contenant que du bruit ; H_1 : cible de forme w présente au centre de F_j

Le logarithme du rapport de vraisemblance associé à chaque version de l'algorithme pour une position j dans l'image est rappelé dans le tableau ci-dessous.

$$\begin{aligned}
 r_C^C(j) &= -N_w \log(m_c) - N_{\bar{w}} \log(m_f) + N_F \log(m_f) \\
 &\quad - N_w \left(\sum_{i \in \mathbf{w}_j} \frac{s_i}{N_w} \right) \frac{1}{m_c} - N_{\bar{w}} \left(\sum_{i \in \bar{\mathbf{w}}_j} \frac{s_i}{N_{\bar{w}}} \right) \frac{1}{m_f} + N_F \left(\sum_{i \in \mathbf{F}_j} \frac{s_i}{N_F} \right) \frac{1}{m_f}
 \end{aligned} \tag{2.1}$$

$$\begin{aligned}
 r_C^E(j) &= -N_w \log(m_c) - N_{\bar{w}} \log \left(\sum_{i \in \bar{\mathbf{w}}_j} \frac{s_i}{N_{\bar{w}}} \right) + N_F \log \left(\sum_{i \in \mathbf{F}_j} \frac{s_i}{N_F} \right) \\
 &\quad - N_w \left(\sum_{i \in \mathbf{w}_j} \frac{s_i}{N_w} \right) \frac{1}{m_c} - N_{\bar{w}} + N_F
 \end{aligned} \tag{2.2}$$

$$\begin{aligned}
 r_E^C(j) &= -N_w \log \left(\sum_{i \in \mathbf{w}_j} \frac{s_i}{N_w} \right) - N_{\bar{w}} \log(m_f) + N_F \log(m_f) \\
 &\quad - N_w - N_{\bar{w}} \left(\sum_{i \in \bar{\mathbf{w}}_j} \frac{s_i}{N_{\bar{w}}} \right) \frac{1}{m_f} + N_F \left(\sum_{i \in \mathbf{F}_j} \frac{s_i}{N_F} \right) \frac{1}{m_f}
 \end{aligned} \tag{2.3}$$

$$r_E^E(j) = -N_w \log \left(\sum_{i \in \mathbf{w}_j} \frac{s_i}{N_w} \right) - N_{\bar{w}} \log \left(\sum_{i \in \bar{\mathbf{w}}_j} \frac{s_i}{N_{\bar{w}}} \right) + N_F \log \left(\sum_{i \in \mathbf{F}_j} \frac{s_i}{N_F} \right) \tag{2.4}$$

La première de nos expériences a consisté à comparer ces algorithmes, pour des régions \mathbf{w} et $\bar{\mathbf{w}}$ de tailles identiques, dans des scènes de bruit Gamma d'ordre 1. Nous avons souligné dans le chapitre 1 les liens existants entre les algorithmes de type GLRT et les algorithmes de type ML. En localisation, ces algorithmes sont équivalents lorsque la fenêtre est de la taille de l'image complète. Nous avons donc choisi pour caractériser l'influence de la taille de la fenêtre extérieure, d'évaluer les performances de ces algorithmes en localisation.

Tout d'abord, nous commençons par étudier le cas où les régions \mathbf{w} et $\bar{\mathbf{w}}$ ont la même taille. La figure 2.2 montre l'évolution de la probabilité de localisation correcte, en fonction du contraste, pour une cible composée de 8 pixels, et une fenêtre d'analyse en contenant 16. La probabilité de localisation correcte est la probabilité que le maximum du plan de sortie se situe exactement à l'emplacement de la cible. Les images sont de taille 128×128 pixels, et sont composées de bruit gamma d'ordre 1.

On peut mesurer sur cette courbe l'influence de la connaissance ou de l'estimation des différents paramètres. Bien évidemment, les performances les pires sont obtenues pour l'algorithme $GLRT_E^E$ qui estime à la fois m_c et m_f : il s'agit du cas où la connaissance *a priori* est minimale. Le corolaire de ceci est que l'algorithme $GLRT_C^C$ présente les meilleures performances, quel que soit le contraste, car il utilise le maximum de connaissance *a priori*. De la même façon, les performances des deux autres algorithmes viennent se ranger entre les deux cas précédents.

Il est plus intéressant de comparer les gains obtenus en injectant telle ou telle connaissance *a priori* : Bien que les régions \mathbf{w} et $\bar{\mathbf{w}}$ aient la même surface, la connaissance la plus importante est celle des paramètres du fond (m_f). De fait, ne connaître que les paramètres du fond se révèle presque aussi efficace que la connaissance complète des paramètres. En revanche, la seule connaissance de m_c (algorithme $GLRT_C^E$) ne permet de réaliser qu'un gain plus modeste.

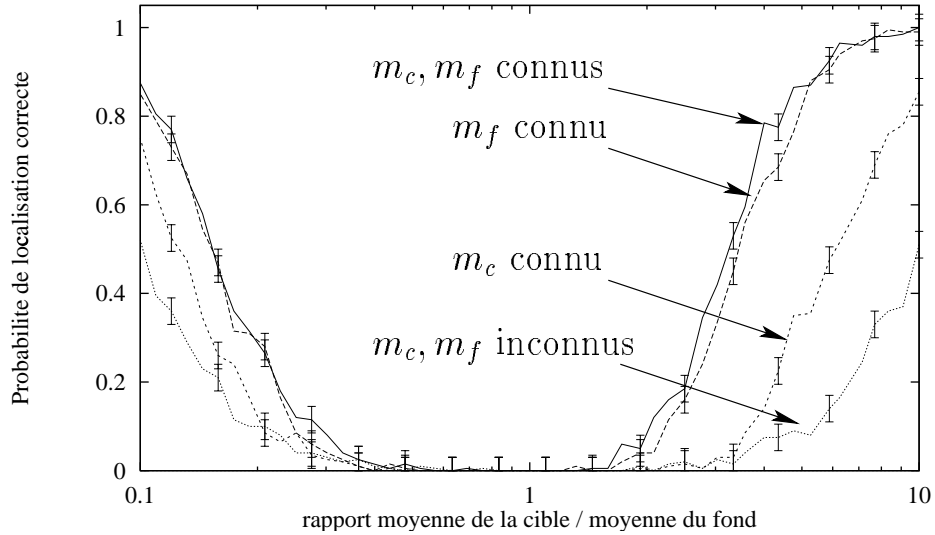


FIG. 2.2: Influence de l'estimation des différents paramètres de nuisance de l'algorithme GLRT : Probabilité de localisation correcte en fonction du contraste pour les 4 versions possibles de l'algorithme GLRT. Ces versions diffèrent par la connaissance a priori des paramètres des lois de probabilité de la cible et du fond. Ces paramètres sont connus ou estimés au sens du maximum de vraisemblance (cf. équations 2.1 à 2.4). La cible a une surface de 8 pixels, la fenêtre d'analyse 16. Les images sont composées de bruit Gamma d'ordre 1. Pour chaque valeur du contraste, la probabilité de localisation correcte a été estimée sur 200 réalisations de ces bruits. La demi-hauteur des barres d'erreur correspond à l'écart-type de l'estimateur de la probabilité de localisation.

Dans beaucoup de problèmes de localisation et de détection, les paramètres de la cible et du fond sont supposés inconnus. Nous venons de voir que pour de petites cibles, la connaissance des paramètres du fond permet cependant d'obtenir des performances bien meilleures. Dès lors, lorsque les paramètres du fond sont inconnus, il est primordial que leur estimation soit aussi précise que possible.

2.1.2 Utiliser de grandes fenêtres procure une meilleure estimation des paramètres du fond

L'influence de la taille de la fenêtre sur le comportement de l'algorithme reste à démontrer. L'algorithme GLRT procède à l'estimation des paramètres de chaque région avec l'estimateur du maximum de vraisemblance. Dans le cas du bruit Gamma, le seul paramètre à estimer est sa moyenne. L'estimateur du MV de la moyenne est dans ce cas la moyenne arithmétique.

Sur un échantillon s homogène de taille N composé de bruit (décorrélé) Gamma d'ordre 1, et de moyenne m , la moyenne arithmétique :

$$\hat{m} = \frac{1}{N} \left[\sum_{i=1}^N s_i \right] \quad (2.5)$$

obéit à une loi Gamma d'ordre N :

$$P(\hat{m}|m) = \frac{1}{\Gamma(N)} \left(\frac{N}{m}\right)^N \exp\left[-N\frac{\hat{m}}{m}\right] \hat{m}^{N-1} \quad (2.6)$$

L'espérance et la variance de l'estimateur \hat{m} sont alors donnés par:

$$E[\hat{m}] = m \quad (2.7)$$

$$Var[\hat{m}] = \frac{1}{N}m^2 \quad (2.8)$$

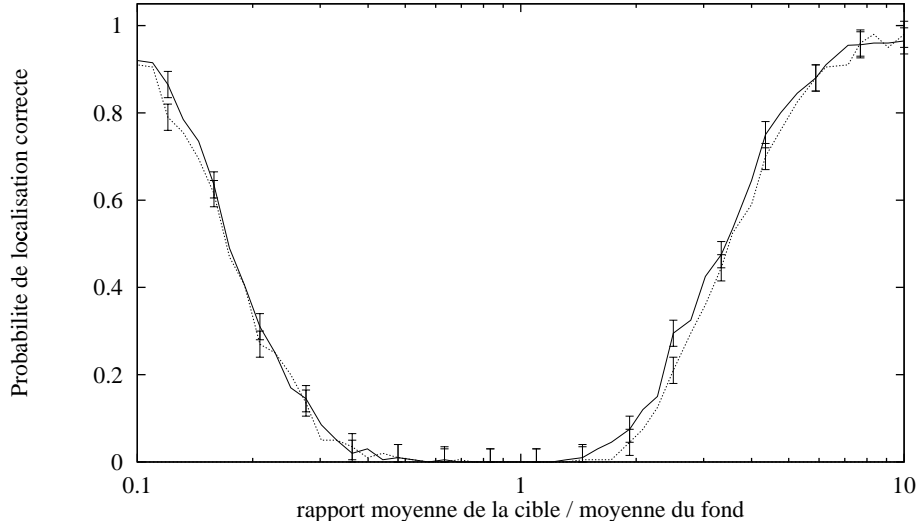


FIG. 2.3: Influence des grandes fenêtres d'analyse sur l'estimation des paramètres : Probabilité de localisation correcte en fonction du contraste.

— : $GLRT_C^C$: m_c et m_f connus.

..... : $GLRT_E^E$: m_c et m_f estimés sur une grande fenêtre d'analyse.

La cible est un carré de taille 3×3 pixels. La fenêtre de l'algorithme $GLRT_E^E$ est un carré de taille 121×121 . Pour chaque valeur du contraste, la probabilité de localisation est estimée sur 200 réalisations de bruit. La demi-hauteur des barres d'erreur correspond à l'écart-type de l'estimateur de la probabilité de localisation.

L'estimateur du maximum de vraisemblance est non biaisé et sa variance décroît en $\frac{1}{N}$. Dès lors, plus le nombre d'échantillons sera grand, plus l'estimation de cette moyenne sera précise. On voit donc tout l'intérêt qu'on peut trouver à utiliser de grandes fenêtres d'analyses, permettant d'estimer les paramètres du fond avec précision.

Dans la figure 2.3, nous montrons une manifestation de ce phénomène, en comparant les performances de l'algorithme $GLRT_C^C$ (paramètres connus), aux performances de l'algorithme $GLRT_E^E$ utilisant une grande fenêtre d'analyse. Ces performances sont mesurées par la probabilité de localisation correcte, en fonction du contraste présent dans la scène.

Quel que soit ce contraste, les performances de ces algorithmes sont similaires, ce qui confirme que la non-connaissance des paramètres du fond peut être compensée par une bonne estimation.

2.1.3 Limite supérieure à la taille des fenêtres

Nous avons vu dans la section 1.4 que lorsque les paramètres sont inconnus, le GLRT utilisant la fenêtre la plus grande possible (de la taille de la scène) est strictement équivalent en localisation à l'algorithme du Maximum de Vraisemblance, lequel maximise la probabilité de localisation correcte. Ceci concorde avec les conclusions de la section précédente, qui tendent à faire choisir des fenêtres de la taille la plus grande possible. Dès lors, les performances de l'algorithme du Maximum de Vraisemblance constitueront une borne supérieure pour les performances du GLRT.

Il est alors intéressant d'observer l'évolution des performances du GLRT en fonction de la taille de la fenêtre d'analyse, et de les comparer aux performances obtenues par le Maximum de Vraisemblance. Dans la figure 2.4, nous présentons sur un exemple l'évolution de la probabilité de localisation correcte en fonction de la taille de la fenêtre d'analyse, pour différentes tailles de cibles.

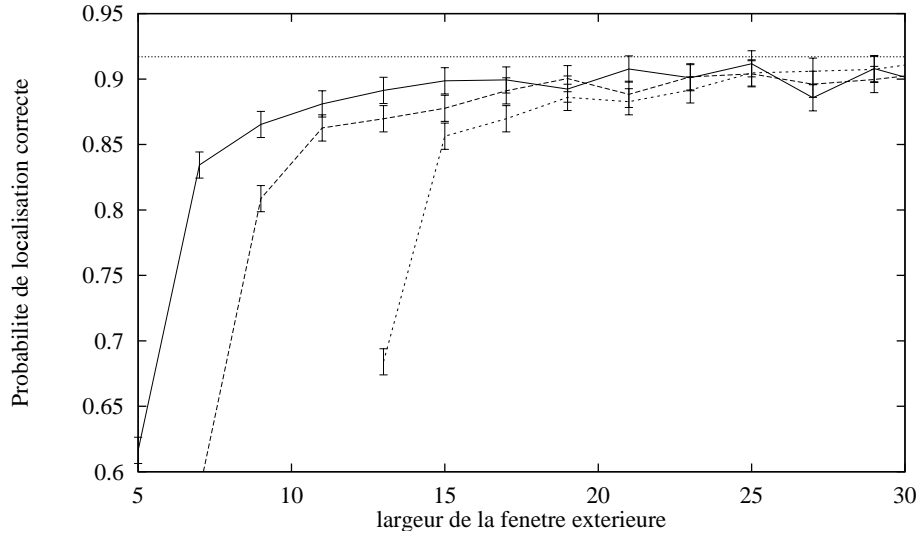


FIG. 2.4: Probabilité de localisation correcte en fonction de la taille de la fenêtre d'analyse, pour différentes tailles de cibles.

— : cible de taille 3×3 pixels
 - - - : cible de taille 5×5 pixels
 : cible de taille 11×11 pixels
 : courbe obtenue pour l'algorithme du ML.

Le contraste est fixé pour que le filtre ML présente une probabilité de localisation correcte de 0.9 pour ces trois tailles. Pour chaque largeur de la fenêtre, la probabilité de localisation est estimée sur 3000 réalisations de bruit. La demi-hauteur des barres d'erreur correspond à l'écart-type de l'estimateur de la probabilité de localisation.

Nous avons considéré une cible carrée de tailles 3×3 apparaissant avec un contraste \mathcal{C} donné dans une image de 128×128 pixels perturbée par du bruit gamma d'ordre 1. Pour chaque taille de fenêtre d'analyse, nous avons estimé la probabilité de localisation correcte, sur 3000 réalisation de ces scènes. Cette courbe est comparée à la probabilité de localisation correcte obtenue en utilisant l'algorithme du Maximum de Vraisemblance.

La même expérience a été refaite pour des cibles de 5×5 , et 11×11 pixels. Bien évidemment, pour une valeur donnée du contraste \mathcal{C} , la probabilité de localisation correcte augmente avec la taille de la cible. Pour étalonner nos courbes, nous avons fixé la valeur du contraste \mathcal{C} de façon à ce que la probabilité de localisation de l'algorithme du Maximum de Vraisemblance soit de l'ordre de 0.9 quelle que soit la taille de la cible. Ceci correspond pour la cible 3×3 (resp 5×5 et 11×11) à un contraste de 7 (resp 3.8 et 2.5)

Cette figure montre que les performances de l'algorithme augmentent avec la taille de la fenêtre, du fait d'une meilleure estimation des paramètres du fond. De plus, ces performances tendent à égaler les performances de l'algorithme du Maximum de Vraisemblance.

Le point primordial de cette expérience est que le gain obtenu en augmentant la taille de la fenêtre stagne très rapidement. De ce fait, il est possible d'utiliser une fenêtre de taille réduite pour le GLRT, tout en conservant des performances très proches de celles du ML. Par exemple, si la fenêtre du GLRT a une largeur de dix pixels de plus que la cibles, les performances ne décroissent que de 5 pour cent, par rapport à celles du ML.

D'un point de vue pratique, il est alors possible de tirer de ces expériences une détermination quasi automatique de la fenêtre d'analyse. Celle ci peut être obtenue en effectuant quelques dilatations morphologiques [51] de la référence, par un élément structurant carré centré de taille 3×3 pixels. Par exemple, une fenêtre plus large de dix pixels que la cible est obtenue par 5 dilatations de la référence, lorsque celle ci est un carré.

2.1.4 Généralisation à différents contrastes

Pour que ces résultats soient utilisables en pratique, il est important de vérifier que le choix de la taille de cette fenêtre d'analyse est valide quel que soit le contraste entre la cible et le fond. La figure 2.5 montre les courbes de probabilité de localisation correcte en fonction du contraste, pour différentes tailles de fenêtres. La cible est ici un carré de taille 3×3 . les fenêtres d'analyse utilisées par le GLRT sont des carrés de tailles respectives 5×5 , 11×11 et 23×23 pixels Ceci correspond à des fenêtres obtenues par 1, 3 et 10 dilatations de la forme initiale de la référence. Nous faisons également figurer sur cette figure les résultats obtenus avec l'algorithme du ML, comme borne supérieure des performances. On peut constater qu'un nombre de dilatations compris entre 3 et 10 suffit pour obtenir des performances très proches de l'algorithme du ML.

2.1.5 Généralisation aux bruits Poissonien et Gaussien

Nous voulons ici illustrer le fait que les observations précédentes sont toujours valides dans le cas où le bruit ne suit plus une loi gamma, mais d'autres lois de la famille exponentielle.

Comme dans la section 2.1.3, observons l'évolution de la probabilité de localisation correcte en fonction de la taille de la fenêtre, pour différentes tailles de cibles. Nous présentons dans la figure 2.6 les résultats obtenus dans le cas de bruits de Poisson, et de bruits Gaussien. Les paramètres sont inconnus et estimés au sens du Maximum de

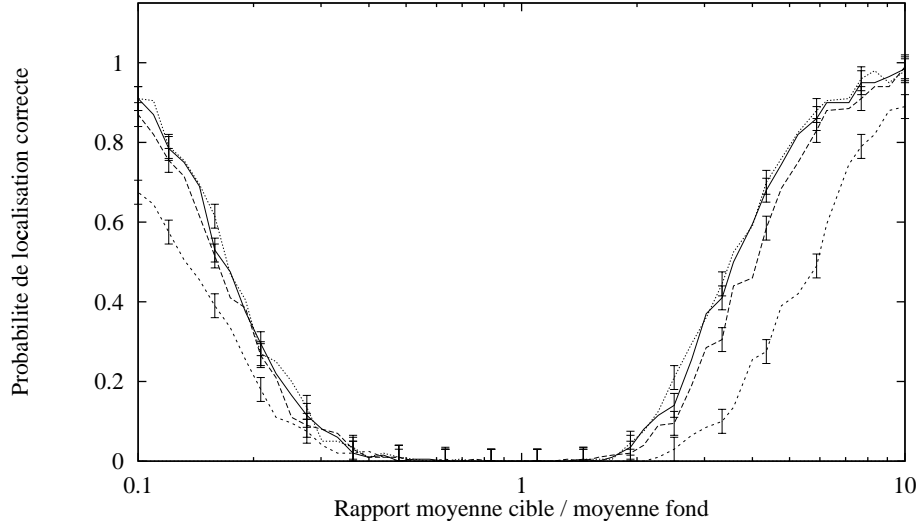


FIG. 2.5: Probabilité de localisation correcte en fonction du contraste, pour différentes tailles de fenêtres. La cible est un carré de taille 3×3 pixels. La fenêtre de l'algorithme GLRT est obtenue par N dilatations successives de la référence de la cible.

..... : algorithme ML

———— : algorithme GLRT, $N = 10$

— — — : algorithme GLRT, $N = 3$

- - - - - : algorithme GLRT, $N = 1$.

Pour chaque valeur du contraste, la probabilité de localisation est estimée sur 200 réalisations de bruit. La demi-hauteur des barres d'erreur correspond à l'écart-type de l'estimateur de la probabilité de localisation.

vraisemblance. Les algorithmes correspondants sont donnés par les équations suivantes :

$$r_E^E(j) = -N_w \hat{m}_w(j) \log \hat{m}_w(j) - N_{\bar{w}} \hat{m}_{\bar{w}}(j) \log \hat{m}_{\bar{w}}(j) + N_F \hat{m}_F(j) \log \hat{m}_F(j) \quad \text{Bruit de Poisson} \quad (2.9)$$

$$r_E^E(j) = -N_w \log \hat{\sigma}_w^2(j) - N_{\bar{w}} \log \hat{\sigma}_{\bar{w}}^2(j) + N_F \log \hat{\sigma}_F^2(j) \quad \text{Bruit Gaussien}$$

Encore une fois, lorsque la taille de la fenêtre augmente, les performances de l'algorithme GLRT atteignent très vite un maximum, donné par l'algorithme ML.

2.1.6 Conclusion en bruit homogène

Conformément à ce qui a été énoncé dans le chapitre 1, en localisation, l'algorithme optimal est l'algorithme du ML. Ces performances sont donc toujours supérieures à celles de l'algorithme du GLRT. De plus, lorsque la taille de la fenêtre d'analyse du GLRT augmente, ces deux algorithmes deviennent équivalents. Nous avons interprété cela en

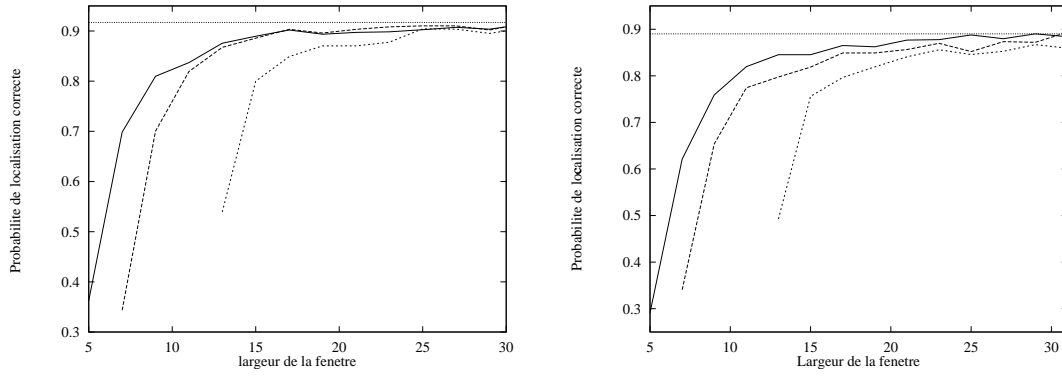


FIG. 2.6: Probabilité de localisation correcte en fonction de la taille de la fenêtre d'analyse, pour différentes tailles de cibles. **à gauche** : Cas du bruit de poisson, **à droite** : cas du bruit gaussien.

— : cible de taille 3×3 pixels
 - - - : cible de taille 5×5 pixels
 : cible de taille 11×11 pixels
 : courbe obtenue pour l'algorithme du ML.

Le contraste est fixé pour que le filtre ML présente une probabilité de localisation correcte identique pour ces trois tailles. La fenêtre est un carré de largeur variable. Pour chaque largeur de la fenêtre, la probabilité de localisation est estimée sur 3000 réalisations de bruit.

nous basant sur la qualité de l'estimation des paramètres de nuisance, qui est d'autant meilleure que la fenêtre d'analyse du GLRT est grande.

Cependant, en comparant les cas où les paramètres sont connus ou estimés, on peut remarquer qu'à partir d'une certaine taille de fenêtre, le gain obtenu est négligeable. Nous avons validé cette conclusion pour différentes tailles de petites cibles, différents contrastes, et différents types de bruits de la famille exponentielle. Dans chacun de ces cas, une fenêtre obtenue par quelques dilatations morphologiques de la références suffisent pour obtenir des performances proches de celles du ML. Si cela est nécessaire, on pourra donc utiliser un algorithme de type GLRT construit avec une fenêtre de petite taille, tout en conservant des performances proches de l'algorithme optimal. Dans la section suivante, nous allons voir un exemple de bruits pour lequel ceci est intéressant.

2.2 Bruit non homogène

Le modèle SIR qui sous tends la construction des algorithmes du Maximum de Vraisemblance et du GLRT supposent la propriété d'homogénéité des champs aléatoires sur la cible et sur le fond. Un champ aléatoire homogène est défini par le fait que ses propriétés statistiques ne varient pas d'un point à un autre.

Prenons l'exemple d'images dans lesquelles les histogrammes locaux du fond diffèrent d'une région à une autre de l'image. Dans ce cas, l'hypothèse que le fond de l'image est un champ homogène ne permet pas en pratique de modéliser de façon efficace les images qui nous intéressent. L'algorithme du ML devient donc sous-optimal. En revanche l'algorithme GLRT n'utilise l'hypothèse d'homogénéité du fond que localement (dans la

fenêtre d'analyse). Dès lors, l'utilisation de fenêtres de tailles réduite permet de conserver une modélisation efficace des images à traiter.

Pour vérifier que la non homogénéité du fond est la cause des différences de performances entre l'algorithme GLRT et l'algorithme du MV, nous les testons sur des images synthétiques, inspirées d'un modèle standard [54].

2.2.1 Modèle d'images inhomogènes

Ce modèle consiste à commencer par générer une image de bruit gaussien \mathbf{n}^c corrélé, dont la fonction d'autocorrélation est une exponentielle de largeur l . Ce bruit a pour moyenne m_b , et pour écart-type σ .

La scène à traiter est alors générée de la façon suivante :

- Si le pixel i appartient au fond, s_i est une réalisation de variable exponentielle de moyenne $m_f = n_i^c$.
- Si le pixel i appartient à la cible, s_i est une réalisation de variable exponentielle de moyenne m_c . (La cible est homogène)

Ainsi, nous disposons d'une scène dans laquelle l'intensité de chaque pixel du fond suit une loi gamma d'ordre 1, mais le bruit est non homogène, puisque la moyenne de cette loi varie d'un point à un autre de l'image.

La figure 2.7 présente un exemple de notre modèle d'image. L'image 2.7(a) est celle du support de la cible cherchée, qui est un carré de taille 11×11 pixels. Dans la figure 2.7(b) se trouve l'image du champ aléatoire gaussien corrélé qui simule les inhomogénéités. L'image 2.7(c) est celle du fond aléatoire gamma inhomogène et enfin dans la figure 2.7(d) apparaît la scène de bruit gamma non homogène, de longueur de corrélation $l = 20$, d'écart-type $\sigma = 0.1$.

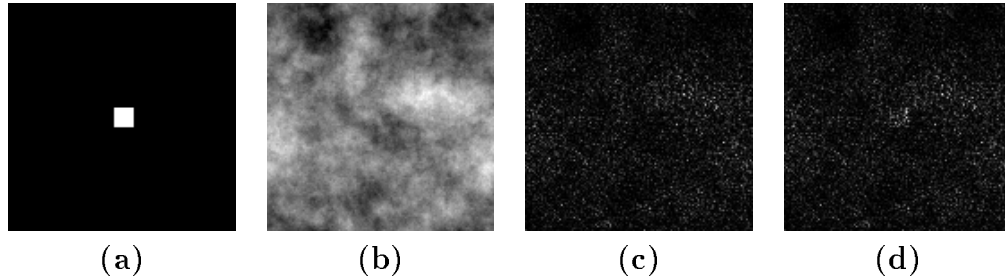


FIG. 2.7: Exemple de scène synthétique non homogène : **(a)** : cible cherchée : carré de taille 11×11 pixels **(b)** : champ aléatoire gaussien corrélé **(c)** : fond non homogène **(d)** : scène complète. La longueur de corrélation du champ aléatoire n^c est fixée à $l = 20$. L'écart-type du bruit corrélé est $\sigma = 0.1$.

2.2.2 Echec des grandes fenêtres

Nous essayons de localiser la cible qui se trouve au centre de la scène de bruit non homogène présentée dans la figure 2.7.d . Nous appliquons sur cette scène les algorithmes

GLRT construits avec différentes tailles de fenêtres. Les fenêtres sont obtenues par dilata-tions successives de la référence. La figure 2.8 montre le maximum sur chaque colonne de chacun des plans de sortie ainsi obtenus. Nous présentons également les résultats obtenus par l'algorithme du ML.

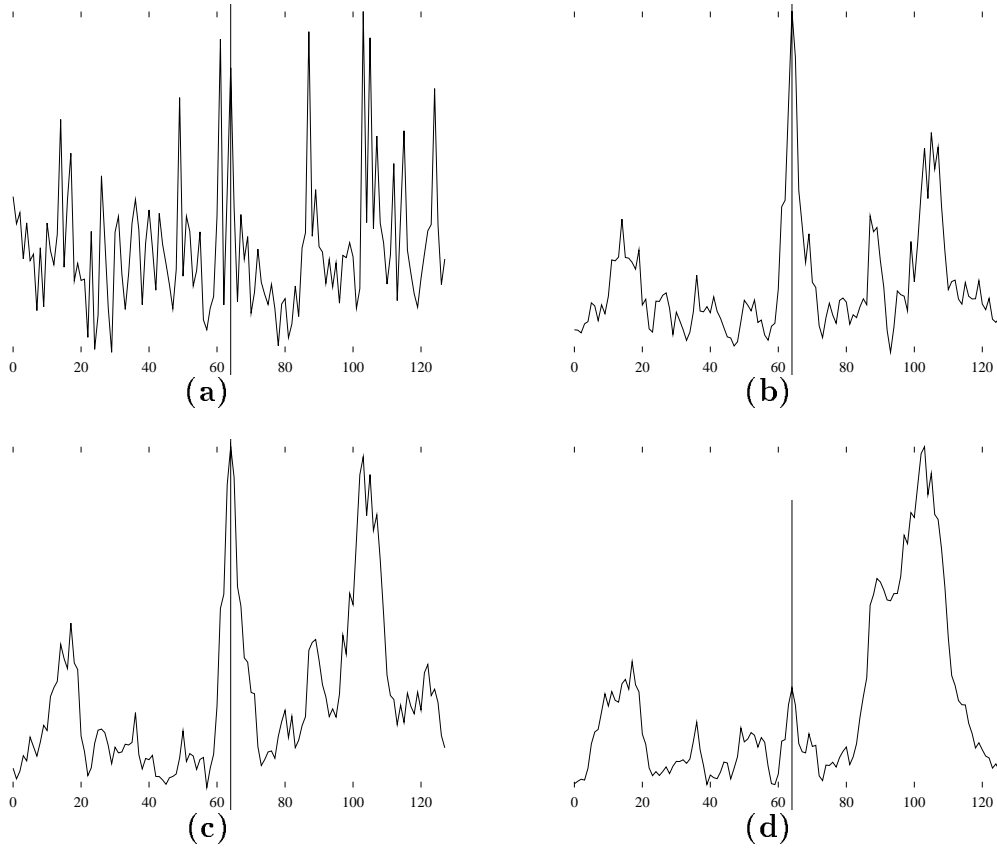


FIG. 2.8: Maximum sur chaque colonne du plan de sortie des différents algorithmes appliqués à l'image présentée dans la figure 2.8 (a) : GLRT : fenêtre obtenue par 1 dilatation, (b) : GLRT : fenêtre obtenue par 5 dilatation, (c) : GLRT : fenêtre obtenue par 10 dilatation, (d) : ML. La position de la cible est indiquée par un trait vertical.

La courbe **2.8.a** confirme que l'utilisation d'une fenêtre trop petite conduit à un échec de la localisation. Les paramètres du fond sont alors mal estimés. Des fenêtres trop grandes **2.8.c** et **2.8.d** conduisent également à un échec, du fait de l'inhomogénéité du fond. En revanche, l'utilisation d'une fenêtre de taille modeste permet de réaliser un bon compromis entre estimation des paramètres du fond et robustesse aux inhomogénéités, et conduit à la localisation de la cible. Notons que la longueur de corrélation (ie : la taille des inhomogénéités) est ici très proche de la largeur de la cible.

Cette conclusion se retrouve sur la figure 2.9 où nous estimons la probabilité de localisation correcte du GLRT, en fonction de la taille de la fenêtre, dans le cas de bruits corrélés. Une fenêtre obtenue par quelques dilata-tions de la référence (entre 5 et 10) limite l'influence des inhomogénéités sur les performances de l'algorithme. Lorsque la longueur de corrélation du bruit est proche de celle de la cible, la localisation devient plus difficile.

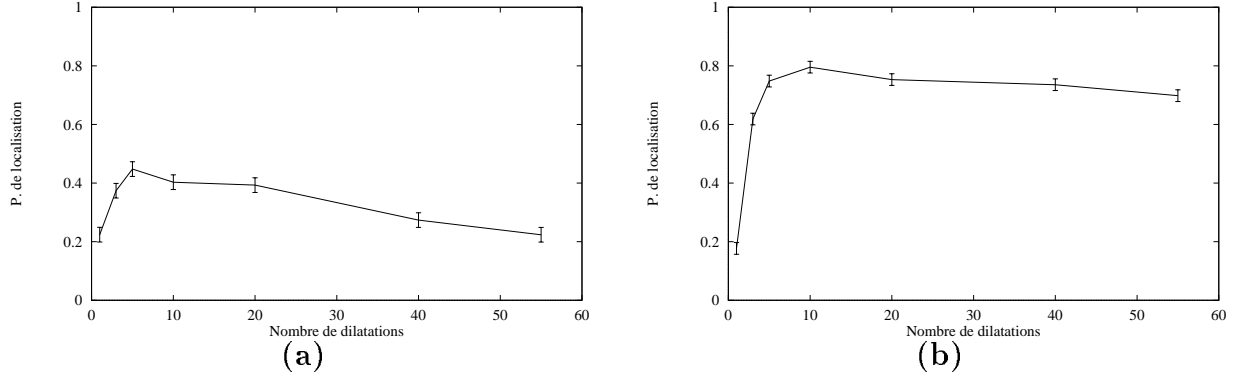


FIG. 2.9: Probabilité de localisation correcte en fonction de la taille de la fenêtre. Cette taille est exprimée en fonction du nombre de dilatations appliquées à la référence. La cible est un carré de taille 11×11 , de moyenne 10, placée dans un bruit corrélé de moyenne 1, de variance 0.7 et de longueur de corrélation variable. **(a)** : Longueur de corrélation 20. **(b)** : Longueur de corrélation 50. Les courbes sont estimées sur 400 réalisations des scènes. La demi-hauteur des barres d'erreur correspond à l'écart-type de l'estimateur de la probabilité de localisation.

2.2.3 Inhomogénéités trop importantes

Nous avons établi que le choix d'une taille de fenêtre relevait d'un compromis entre estimation des paramètres, et inhomogénéités du fond. Bien évidemment, ce compromis peut s'avérer impossible à réaliser, lorsque les inhomogénéités deviennent trop importantes, c'est à dire lorsque la variance des inhomogénéités devient trop grande devant le contraste entre la cible et le fond $\sigma \simeq (m_c - m_b)$. Nous illustrons ceci dans la figure 2.10.

Pour réussir à localiser la cible dans de telles images, il conviendrait de disposer de plus d'information *a priori* sur les intensités des pixels de la cible ou celles du fond. Dans le contexte où seule la forme de la cible est connue, les inhomogénéités constituent un véritable problème dont il faudra essayer de s'affranchir avant tout. Nous verrons dans le chapitre 4 un exemple de cas où nous renoncerons au caractère optimal de l'algorithme pour réduire au maximum les inhomogénéités présentes dans les scènes.

2.3 Conclusion

Dans ce chapitre, nous avons mis en évidence l'influence de l'estimation des paramètres de nuisance des algorithmes statistiques. Cette estimation peut devenir problématique si les échantillons utilisés pour l'estimation sont trop petits. La taille de la cible étant fixée, le seul paramètre libre est la taille de la fenêtre. En augmentant cette taille, nous avons montré que l'estimation des paramètres du fond devient très rapidement correcte. De plus, nous avons montré que les performances de l'algorithme GLRT, pourtant sous-optimal en localisation, étaient équivalentes à celles de l'algorithme ML dédié à cette tâche, dès lors que la fenêtre est obtenue par quelques dilatations du support de la référence. Dès lors, en bruit homogène, l'algorithme GLRT peut être considéré comme équivalent à l'algorithme ML.

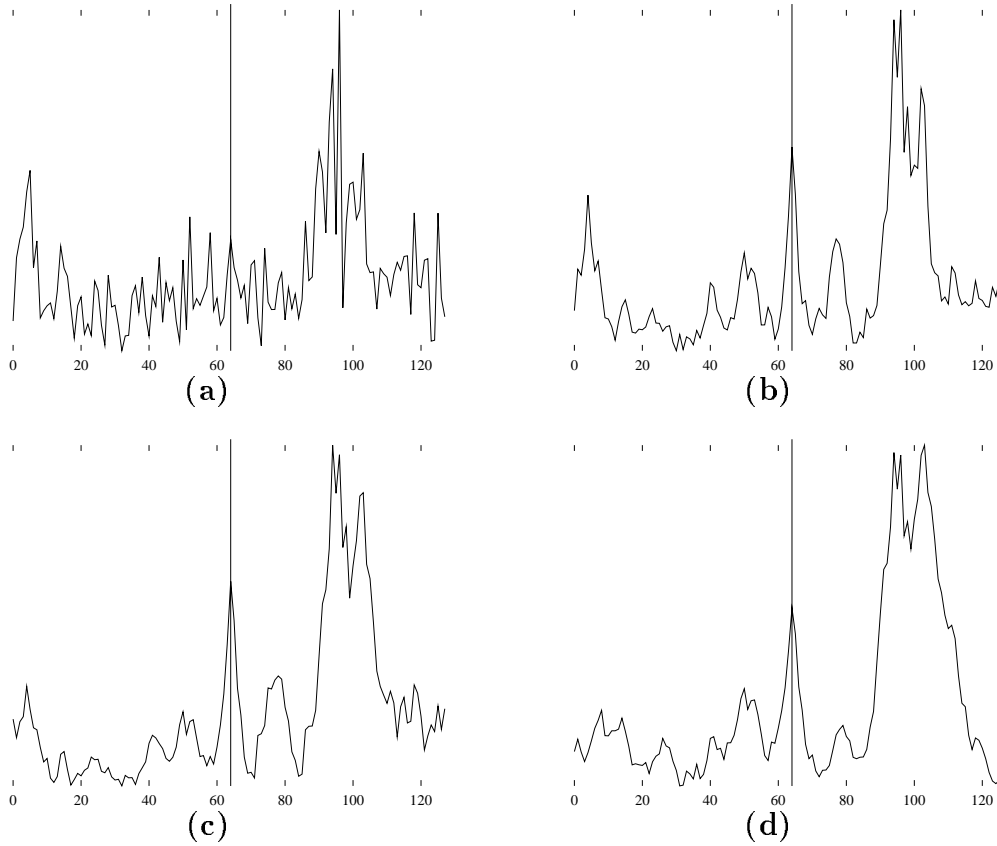


FIG. 2.10: Inhomogénéités de forte variance : Maximum sur chaque colonne du plan de sortie des différents algorithmes appliqués à l'image présentée dans la figure 2.8 **(a)** : GLRT : fenêtre obtenue par 1 dilatation, **(b)** : GLRT : fenêtre obtenue par 5 dilatation, **(c)** : GLRT : fenêtre obtenue par 10 dilatation, **(d)** : ML.

$\sigma = 0.3, m_c = 2, m_b = 1$ La position de la cible est indiquée par un trait vertical.

En bruit inhomogène, l'algorithme du ML n'est plus efficace car le modèle sur lequel il est construit n'est pas adapté à ce type de bruit. En revanche, l'algorithme GLRT, ne faisant l'hypothèse d'homogénéité du fond que localement, permet dans certains cas de dépasser ces limites. Evidemment, dans le cas d'inhomogénéités trop fortes, des algorithmes ne supposant connue que la forme de la cible connaissent des difficultés. Dans ce type de cas, il sera nécessaire de considérer la réduction des inhomogénéités comme une préoccupation majeure pour la localisation ou la détection. C'est sur cette base que nous nous appuierons dans le chapitre 4 pour choisir un mode de représentation des images polarimétriques.

Cependant, avant cela, nous avons souhaité nous intéresser à un cas également très commun en traitement d'images, celui de cibles de tailles inconnues. Nous montrerons qu'il est possible de définir des algorithmes de façon rigoureuse dans le contexte de l'estimation statistique, pour ce type de problèmes. Ces algorithmes présentent des propriétés très intéressantes dans la gestion des fausses alarmes. La présentation de ces algorithmes fera l'objet du chapitre suivant.

Chapitre 3

Cas de cibles de tailles inconnues

Sommaire

3.1	Chute des performances lorsque la taille de la cible varie . .	37
3.2	Rappels de la théorie de l'estimation et de la décision	38
3.2.1	Hypothèses	38
3.2.2	Estimation de la taille au sens du maximum de vraisemblance .	39
3.2.3	Estimation de la taille au sens du maximum a posteriori	40
3.2.4	Approche du Marginal Bayésien	40
3.2.5	Conclusion	41
3.3	Choix d'une méthode de gestion du paramètre de taille . . .	41
3.3.1	Cas de lois Gamma d'ordre L	42
3.3.2	Influence du prior sur les performances de l'AMB	42
3.4	Première illustration des performances de cet algorithme . .	44
3.5	Choix d'un seuil de détection	44
3.6	Exemples d'utilisation de l'algorithme $GLRT_{mv}$	47
3.7	Conclusion	48

L'algorithme GLRT des chapitres précédents suppose que la taille de la cible est connue. Cependant, dans de nombreuses applications, cette taille est inconnue, soit du fait qu'on ne sache pas réellement quelle est la taille de la cible, soit que le bruit présent tende à faire disparaître une partie de la cible.

Pourtant lorsque la référence est inadaptée à la taille de la cible réellement présente, les performances des algorithmes chutent sévèrement. Dès lors, il est nécessaire d'élargir notre modèle pour lui faire prendre en compte la possibilité de traiter des cibles de différentes tailles. De nombreuses études ont porté sur ce problème, et, de façon générale, on peut considérer trois groupes de méthodes.

La première approche consiste à construire un filtre invariant aux changements d'échelle de la cible. On pourra s'inspirer de la méthode des filtres SDF [41], pour construire un filtre invariant par changement d'échelle. Supposons que l'on dispose d'un ensemble d'images constituant une base d'apprentissage dans lesquelles la cible apparaît à différentes échelles.

Il est possible avec cette méthode de construire un filtre optimisant par exemple la robustesse au bruit, en imposant que les réponses du filtre aux différentes images de la base d'apprentissage soient identiques. Dans ce cas, le filtre obtenu sera plus ou moins invariant aux changements d'échelle. Il se pose néanmoins le problème de la définition d'une base d'apprentissage adaptée au problème à résoudre. Par ailleurs, ces filtres sont dits "à structure imposée", puisque les filtres recherchés sont exclusivement des filtres linéaires, donc éventuellement peu adaptés au problème du bruit non recouvrant.

Le deuxième groupe de méthodes consiste à travailler sur des transformations de l'image, comme par exemple la transformation de Mellin [6][5], dans laquelle on peut éliminer le terme dépendant du facteur d'échelle avec lequel la cible apparaît dans l'image. Cette méthode pose le problème du choix du centre de la transformation et induit une perte de l'invariance en translation². Cette dernière propriété est primordiale dans des applications de localisation et de détection et nous considérons sa conservation comme une priorité pour construire un algorithme permettant de traiter le cas de cibles de tailles inconnues. Des solutions ont été proposées afin de disposer d'une invariance en translation et aux changements d'échelle [9], mais ces solutions n'offrent cependant aucune garantie d'optimalité. Il en est de même pour l'approche multirésolution, utilisant une décomposition en ondelettes, telle que celles présentées par Mallat [33]. Des améliorations de celle-ci permettent de satisfaire la condition d'invariance par translation [53][1]. Notons cependant que la décomposition en ondelettes s'apparente à la corrélation de l'image avec une fonction d'échelle. Ces fonctions d'échelle ne permettent cependant pas de prendre en compte la forme spécifique de l'objet cherché. Elles sont choisies pour leurs propriétés mathématiques (continuité, respect de l'équation de dilatation), plus que pour leur adéquation avec l'objet cherché. Cette approche, très utile pour la détection d'objets de forme mal définie reste heuristique dans le cas de cibles dont la forme est connue.

Le troisième grand groupe de solutions consiste à considérer toutes les références aux différentes tailles puis à fusionner les résultats d'une façon ou d'une autre. Cette méthode est coûteuse en temps de calcul, mais semble la plus efficace. En effet, une telle approche laisse toute latitude dans le choix des filtres à appliquer, qui pourront ainsi être défini de façon optimale pour un modèle d'image. On dispose également de la même latitude dans le choix de la méthode de fusion.

Dans le cadre de la théorie de la décision, nous montrerons que cette dernière approche est une extension naturelle de l'algorithme GLRT, le paramètre de taille apparaissant comme un nouveau paramètre de nuisance, au même titre que les paramètres des lois dont nous avons parlé précédemment. Les algorithmes ainsi définis conservent donc l'assise théorique du GLRT conventionnel.

Pour gérer les paramètres de nuisance, la théorie statistique dans laquelle nous nous sommes placés pour cette thèse propose deux grands groupes de solutions :

- Estimation de ces paramètres (maximum de vraisemblance ou maximum a posteriori). Ces méthodes offrent l'avantage de fournir une estimation d'un paramètre qui peut être utilisé plus tard. Par exemple, l'estimation de taille peut servir de base pour initialiser une segmentation plus précise.

²ie : la réponse R_δ du filtre en présence d'une cible placée à la position δ est identique à celle obtenue par une cible à la position δ' , à la translation $\delta' - \delta$ près.

- Elimination des paramètres par intégration, par la méthode du Marginal Bayésien.

La deuxième méthode est en général supposée plus efficace que la première. Nous montrerons ici que dans les conditions les plus générales, où nous disposons d'informations *a priori* minimales sur la taille de la cible, la méthode la plus adaptée est celle du maximum de vraisemblance. L'introduction d'une information *a priori* qui ne correspond pas à la réalité des images présentées nuit en effet aux performances des algorithmes.

De plus, nous montrerons qu'une propriété très intéressante du GLRT peut être utilisée pleinement lorsque la taille de la cible est estimée. On peut en effet considérer avec une bonne approximation que le GLRT donne une réponse identiquement distribuée quelle que soit la taille de la référence et celle de la fenêtre, lorsqu'on l'applique sur des zones homogènes. Dès lors, le choix d'un seuil de détection ne dépendra pas de l'estimation de taille. Ceci simplifie considérablement le nombre de paramètres à contrôler lors de l'utilisation du GLRT en détection.

Enfin nous illustrerons les performances de cet algorithme sur quelques exemples de scènes synthétiques et réalistes, montrant son efficacité.

3.1 Chute des performances lorsque la taille de la cible varie

On l'a vu, l'algorithme GLRT suppose la taille de la cible connue. Dans de multiples opérations de détection de petites cibles, la taille apparente des cibles dans les scènes peut varier. Dès lors, la taille de la référence choisie pour construire le filtre GLRT est erronée.

Nous illustrons ici sur un exemple simple les pertes de performances en détection que cette erreur sur la taille de la référence peut entraîner.

La cible à détecter est un carré de taille 3×3 pixels. Nous avons appliqué deux algorithmes GLRT pour tenter de détecter cette cible : Le premier a été construit avec une référence dont la taille est celle de la cible. Le second a été construit avec une référence de taille 7×7 pixels. Pour ces deux algorithmes, la fenêtre d'analyse était carrée, de taille 19×19 .

Nous comparons les performances de ces deux algorithmes en détection, lorsque la cible est noyée dans du bruit gamma d'ordre 1. Le contraste entre la cible et le fond est fixé à 4. Les courbes COR obtenues sont présentées dans la figure 3.1.

La perte de performance observable sur ces courbes montre l'intérêt que l'on peut trouver à développer des algorithmes efficaces pour la détection de cibles de taille inconnue.

Les algorithmes de type GLRT ayant démontré leur efficacité lorsque la taille de la cible est connue, il nous a semblé intéressant de vérifier si cette approche pouvait être généralisée aux cas de cibles dont la taille est un paramètre inconnu. Pour cela, nous commencerons par faire quelques rappels sur la théorie de l'estimation. Ceux-ci feront l'objet de la prochaine section.

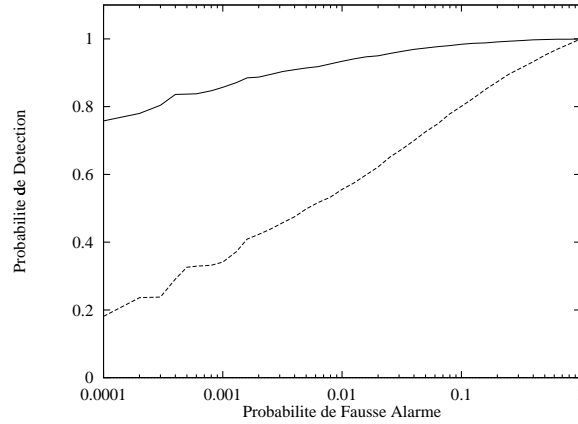


FIG. 3.1: Pertes de performances pour des tailles de références inadaptées à celles de la cible : Courbes COR obtenues pour la détection d'une cible de taille 3×3 pixels, lorsque la taille de la référence est correcte, ou incorrecte.

— : Algorithme *GLRT* construit pour une référence de taille 3×3 pixels.

— — : Algorithme *GLRT* construit pour une référence de taille 7×7 pixels.

La fenêtre d'analyse est un carré de taille 19×19 . Les scènes sont composées de bruit Gamma d'ordre 1. Le contraste entre la cible et le fond vaut 4. Ces courbes COR sont mesurées sur 50000 réalisations.

3.2 Rappels de la théorie de l'estimation et de la décision

3.2.1 Hypothèses

Etant donné un échantillon X (une portion d'image), il s'agit ici de détecter un objet dans cet échantillon, donc de prendre une décision entre deux hypothèses :

- $\mathcal{H}1$: L'objet est présent dans l'échantillon considéré.
- $\mathcal{H}0$: L'objet est absent dans l'échantillon considéré.

Nous supposons que si un objet est présent, il est placé au centre de l'échantillon. En revanche, nous supposons inconnue la taille de l'objet. Cette taille sera notée T . Nous supposons de plus qu'elle ne pourra prendre qu'un nombre fini N de valeurs. Nous noterons Tj , $j \in [1..N]$, l'ensemble des différentes tailles de cibles possibles. $\mathcal{H}1$ se décompose ainsi en N hypothèses $\mathcal{H}1_j$, où $\mathcal{H}1_j$ est l'hypothèse qu'une cible de taille Tj se trouve au centre de l'échantillon. Notons que le calcul des vraisemblances de chaque hypothèse de taille doit être effectué dans la même fenêtre d'analyse qui est également celle utilisée pour calculer la vraisemblance de l'hypothèse $\mathcal{H}0$, sous peine de perdre le cadre théorique qui sous-tend cette technique.

Ces différentes hypothèses sont résumées dans la figure 3.2.

Comme précédemment, nous nous intéresserons au rapport des vraisemblances Λ entre l'hypothèse $\mathcal{H}1$ et l'hypothèse $\mathcal{H}0$, qui permet d'effectuer la prise de décision en optimisant la probabilité de détection pour une probabilité de fausse alarme donnée (voir la section 1.3).

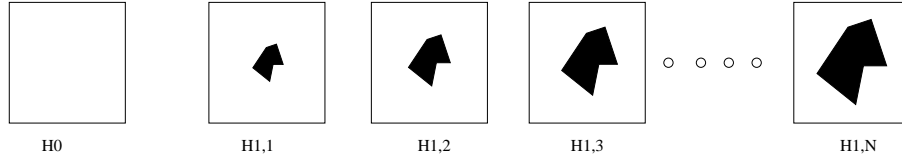


FIG. 3.2: Recensement des différentes hypothèses pour une détection à taille de cible inconnue.

$$\Lambda = \frac{P(X/\mathcal{H}1)}{P(X/\mathcal{H}0)} \quad (3.1)$$

Nous envisageons trois méthodes pour gérer le paramètre de taille de la cible :

- L'estimation de ce paramètre au sens du Maximum de Vraisemblance (MV).
- L'estimation de ce paramètre au sens du Maximum A Postérieur (MAP).
- Elimination par intégration par l'Approche du Marginal Bayésien (AMB).

L'hypothèse $\mathcal{H}0$ ne dépend pas d'une quelconque taille de cible. La seule grandeur qui change dans l'expression de Λ pour ces trois méthodes est la vraisemblance de $\mathcal{H}1$.

3.2.2 Estimation de la taille au sens du maximum de vraisemblance

Dans ce cas, on choisira comme taille de cible estimée, celle qui maximise la vraisemblance d'observer l'échantillon X , en sachant qu'une cible est présente, et en connaissant sa taille.

La taille Tj estimée est donc donnée par :

$$\hat{T}j_{mv} = \arg \max_{j \in [1 \dots N]} P(X|\mathcal{H}1_j) \quad (3.2)$$

L'expression de la vraisemblance de l'hypothèse $\mathcal{H}1$ et du rapport de vraisemblance Λ est obtenue en réinjectant cette taille estimée dans l'équation 3.1, ce qui revient à choisir comme expression du rapport de vraisemblance le maximum des rapports de vraisemblances obtenu pour chacune des tailles :

$$\Lambda_{mv} = \frac{\arg \max_{j \in [1 \dots N]} P(X|\mathcal{H}1_j)}{P(X|\mathcal{H}0)} \quad (3.3)$$

$$\Lambda_{mv} = \arg \max_{j \in [1 \dots N]} \frac{P(X|\mathcal{H}1_j)}{P(X|\mathcal{H}0)} \quad (3.4)$$

Compte tenu de la monotonie de la fonction logarithme, maximiser l'expression précédente revient à maximiser son logarithme. Ceci simplifiant les calculs qui suivront, nous nous intéresserons pour la suite au logarithme du rapport de vraisemblance. Celui ci peut être exprimé en fonction du logarithme du rapport de vraisemblance de chaque hypothèse de taille, que nous notons \mathcal{R}_j .

$$\mathcal{R}_j = \log \left[\frac{P(X|\mathcal{H}1_j)}{P(X|\mathcal{H}0)} \right] \quad (3.5)$$

Il vient alors :

$$\mathcal{R}_{mv} = \arg \max_{j \in [1 \dots N]} \mathcal{R}_j \quad (3.6)$$

L'équation 3.6 montre que la solution du MV au problème de la gestion du paramètre de taille revient simplement à appliquer les filtres GLRT obtenus pour chaque taille possible de cible, et à sélectionner le filtre qui a donné la plus grande réponse. La décision de détection ou de non-détection sera prise en fonction de la sortie de ce filtre. Si la détection est effectuée dans une image, il faudra effectuer cette sélection en chaque pixel de l'image.

3.2.3 Estimation de la taille au sens du maximum a posteriori

Dans ce cas, on suppose connue la densité de probabilité $P(Tj)$ du paramètre de taille. Cette connaissance *a priori* peut être obtenue si on dispose d'une base d'apprentissage représentative des scènes où une cible apparaît.

La probabilité conjointe $P(X, Tj|\mathcal{H}1)$ de l'échantillon X observé et de la taille de la cible s'exprime en utilisant les lois de Bayes par :

$$P(X, Tj|\mathcal{H}1) = P(X|Tj, \mathcal{H}1) P(Tj|\mathcal{H}1) \quad (3.7)$$

L'estimation du paramètre de taille ne dépend plus seulement de l'estimation de la vraisemblance de chacune des hypothèses de taille, mais aussi de la probabilité qu'une cible de taille donnée apparaisse dans l'échantillon.

$$\hat{T}j_{map} = \arg \max_{j \in [1 \dots N]} P(X|Tj, \mathcal{H}1) P(Tj|\mathcal{H}1) \quad (3.8)$$

Cette estimation de la taille est alors ré-injectée pour le calcul du rapport de vraisemblance qui devient :

$$\Lambda_{map} = \frac{P(X|\hat{T}j_{map})}{P(X|\mathcal{H}0)} \quad (3.9)$$

Dans le cas le plus général où l'on fait l'hypothèse que toutes les tailles de cible sont équiprobables, l'estimation du Maximum a Posteriori est équivalente à celle du Maximum de vraisemblance. Nous ne développerons donc pas d'étude sur cette méthode, car nous ne considérons que des cas où la taille des cibles est absolument inconnue. On pourrait lui voir néanmoins une application en poursuite de cible, où la taille de la cible à un instant donné peut être liée à la taille de la cible l'instant précédent.

3.2.4 Approche du Marginal Bayésien

Dans une opération de détection, nous ne cherchons pas à déterminer la taille de la cible, mais à discriminer les hypothèses $\mathcal{H}1$ et $\mathcal{H}0$. L'estimation de la taille de la cible n'a été effectuée que dans le but de pouvoir estimer les vraisemblances respectives de ces deux hypothèses.

Toutefois, l'estimation de la taille de la cible n'est pas indispensable. En effet, l'approche du marginal bayésien permet d'estimer ces vraisemblances sans estimer les paramètres de nuisance.

Cette approche se base sur la formule des probabilités totales [60] : celle-ci nous dit que la probabilité qu'une cible soit présente dans la scène est égale à la somme pondérée des probabilités qu'une cible de taille j soit présente dans la scène (équation 3.10).

$$P(X|H1) = \sum_{j=1}^N P(X|\mathcal{H}1, T_j) P(T_j) \quad (3.10)$$

En réinjectant cette expression dans 3.1, le rapport de vraisemblance s'écrit alors :

$$\Lambda_{amb} = \frac{\sum_{j=1}^N P(X|\mathcal{H}1, j) P(T_j)}{P(X|\mathcal{H}0)} = \sum_{j=1}^N \frac{P(X|\mathcal{H}1_j) P(T_j)}{P(X|\mathcal{H}0)} \quad (3.11)$$

Dans le cas le plus général où les tailles des cibles sont supposées équiprobables, nous obtenons alors :

$$\Lambda_{amb} = \sum_{j=1}^N \frac{P(X|\mathcal{H}1_j)}{P(X|\mathcal{H}0)} \quad (3.12)$$

Comme précédemment, nous désirons exprimer le logarithme de ce rapport en fonction de \mathcal{R}_j , le logarithme du rapport de vraisemblance de chaque hypothèse de taille. Il vient alors instantanément :

$$\mathcal{R}_{amb} = \log \left[\sum_{j=1}^N e^{\mathcal{R}_j} \right] \quad (3.13)$$

Ceci revient à sommer l'exponentielle de l'intensité en sortie des filtres GLRT obtenus pour chaque taille de cible possible, avant d'effectuer la prise de décision. Encore une fois, pour faire de la détection sur toute une image, nous effectuerons la même opération en chaque pixel.

3.2.5 Conclusion

Nous nous sommes intéressés ici au cas de cibles de tailles variables. Dans ce cas, nous avons montré que le rapport de vraisemblance utilisé pour la détection d'une cible de taille connue se généralise de façon naturelle au cas de cibles de tailles inconnues, en considérant la taille comme un paramètre de nuisance. Nous avons retenu deux possibilités pour gérer ce paramètre de nuisance : l'estimation de la taille au sens du maximum de vraisemblance, et l'élimination de ce paramètre par utilisation de l'approche du Marginal Bayésien.

L'ensemble de ces résultats sera utilisé dans la section suivante pour déterminer quelle approche est préférable, en terme de robustesse à l'hypothèse que les tailles de cibles sont équiréparties.

3.3 Choix d'une méthode de gestion du paramètre de taille

Nous avons précédemment évoqué deux façons de gérer le paramètre de nuisance que constitue la taille de la cible. Ces deux méthodes sont :

- L'estimation au sens du MV.
- L'intégration de ce paramètre par l'Approche du Marginal Bayésien

Nous en avons tiré deux algorithmes de détection, correspondant à chacune de ces deux approches. Nous allons montrer ici que l'algorithme conçu sur la base d'une estimation de la taille au sens du MV est plus efficace que celui obtenu par l'Approche du Marginal Bayésien. En effet, l'approche du marginal bayésien nécessite la définition de la densité de probabilité de la taille des cibles présentes. Lorsque cette densité de probabilité est inadéquate, l'AMB se révélera moins efficace que le maximum de vraisemblance.

Pour cela, nous effectuerons une simple comparaison des performances en détection de chacune des deux méthodes. Ces performances sont évaluées en observant les courbes COR de chacun des deux algorithmes.

3.3.1 Cas de lois Gamma d'ordre L

En partant de la densité de probabilité d'une loi Gamma de moyenne m et d'ordre L

$$P^{m,L}(X = x) = \frac{1}{\Gamma(L)} \left(\frac{L}{m} \right)^L x^{L-1} e^{-\frac{Lx}{m}} \quad (3.14)$$

on peut facilement définir le rapport de vraisemblance généralisé \mathcal{R}_j associé à une hypothèse de taille de cible T_j , lorsque l'ordre des lois Gamma est connu et les moyennes de la cible et du fond sont inconnues :

$$\mathcal{R}_j(\mathbf{s}) = L \left[-N_{w_j} \log\left(\frac{1}{N_{w_j}}[\mathbf{s} \otimes \mathbf{w}_j]\right) - N_{\bar{w}_j} \log\left(\frac{1}{N_{\bar{w}_j}}[\mathbf{s} \otimes \bar{\mathbf{w}}_j]\right) + N_F \log\left(\frac{1}{N_F}[\mathbf{s} \otimes \mathbf{F}]\right) \right] \quad (3.15)$$

Les expressions du GLRT pour des cibles de tailles inconnues selon que l'on estime la taille de la cible, ou que l'on utilise l'approche du Marginal Bayésien sont facilement obtenues en ré-injectant ce rapport de vraisemblance dans les équations 3.6 et 3.13. Les deux algorithmes correspondants seront notés respectivement $GLRT_{mv}$ et $GLRT_{amb}$.

Ici nous souhaitons détecter une cible carrée dont la taille était supposée varier d'une scène à l'autre (3, 5, 7 ou 9 pixels de largeur). La fenêtre d'analyse était un carré de taille 19×19 .

3.3.2 Influence du prior sur les performances de l'AMB

Lors du calcul de la logvraisemblance de l'algorithme $GLRT_{amb}$, pour nous placer dans les conditions les plus générales possibles, nous avons choisi une densité de probabilité uniforme sur l'ensemble des tailles de cibles possibles. Afin de mesurer l'influence de cette hypothèse, nous avons étudié deux cas :

- La taille des cibles apparaissant est équirépartie sur l'ensemble des scènes testées.
- La taille des cibles apparaissant est fixée.

Nous avons effectué les deux expériences suivantes : Nous avons estimé les courbes COR obtenues par les deux algorithmes, dans le cas où la taille de la cible apparaissant est tirée de façon aléatoire entre les différentes valeurs possibles (3, 5, 7, 9 pixels de largeur).

Le tirage aléatoire de la taille suivait une loi uniforme. Ces courbes COR sont présentées dans la figure 3.3.a.

La deuxième expérience consistait à effectuer cette opération de détection dans le cas d'une cible de taille fixe (5×5 pixels). Les courbes COR correspondantes sont présentées dans la figure 3.3.b.

Dans chacun de ces cas, les scènes étaient composées de bruit Gamma d'ordre 50. La cible apparaissait avec un contraste de 1.1 sur le fond.

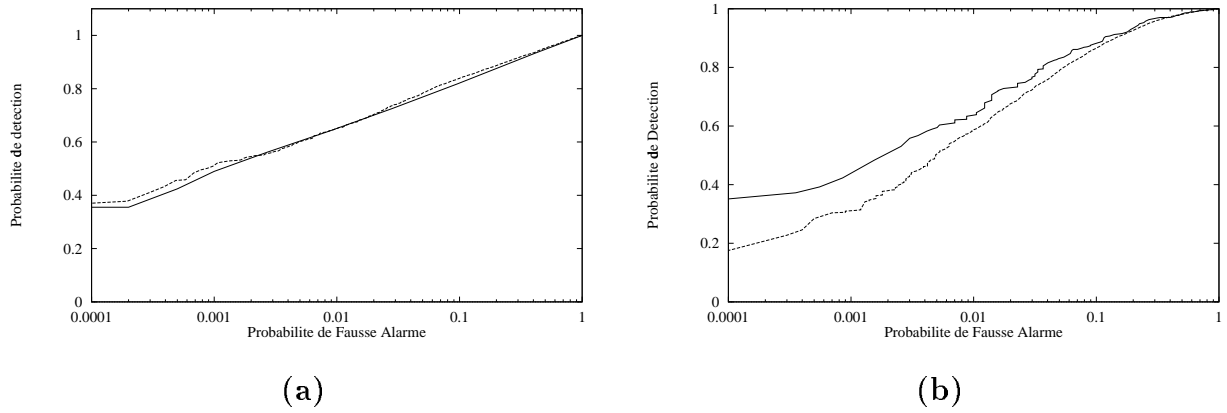


FIG. 3.3: Robustesse face aux connaissances a priori sur la répartition des tailles des cibles : Courbes COR obtenues pour la détection d'une taille de cible inconnue pour les deux méthodes de gestion du paramètre de taille.

— : Algorithme $GLRT_{mv}$.

— — : Algorithme $GLRT_{amb}$.

(a) : Répartition uniforme des tailles des cibles. (b) : Répartition non uniforme des tailles des cibles. Les cibles possibles sont des carrés de largeur 3, 5, 7 ou 9 pixels. La fenêtre d'analyse est un carré de taille 19×19 . Les scènes sont composées de bruit Gamma d'ordre 50. Le contraste entre la cible et le fond vaut 1.1. Ces courbes COR sont mesurées sur 50000 réalisations.

Ces courbes montrent que l'Approche du Marginal Bayésien n'apporte qu'un gain très modeste par rapport à la méthode du MV lorsque le prior utilisé est adéquat. Le cas où les cibles ont toutes la même taille constitue un exemple pour lequel notre prior est inadapté. L'estimation au sens du MV se révèle alors supérieure à l'Approche du Marginal Bayésien. En d'autres termes, l'estimation de taille au sens du MV permet de concevoir des algorithmes plus robustes en détection.

Nous avons ensuite vérifié ce résultat expérimental pour différents ordres des lois Gamma. Les résultats sont présentés dans la figure 3.4 et mènent aux mêmes conclusions. Pour traiter le problème des cibles de taille inconnue, nous choisirons donc d'estimer ces tailles au sens du Maximum de Vraisemblance.

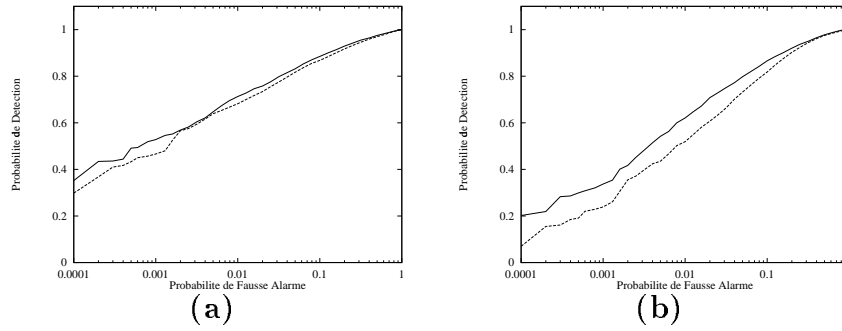


FIG. 3.4: Robustesse face aux connaissances a priori sur la répartition des tailles des cibles : Courbes COR obtenues pour la détection d'une taille de cible inconnue pour les deux méthodes de gestion du paramètre de taille.

— : Algorithme $GLRT_{mv}$.

- - : Algorithme $GLRT_{amb}$.

(a) : Ordre des lois Gamma : 1 (contraste 3). (b) : Ordre des lois Gamma : 10 (contraste 1.5). Les références possibles sont des carrés de largeurs 3, 5, 7 ou 9 pixels. La cible est un carré de taille 3×3 . La fenêtre d'analyse est un carré de taille 19×19 .

3.4 Première illustration des performances de cet algorithme

Après avoir déterminé quelle gestion du paramètre de taille choisir, nous pouvons illustrer les performances de l'algorithme $GLRT_{mv}$ qui suppose inconnue la taille de la cible. Pour illustrer cela, nous avons effectué la même expérience que dans la section 3.1.

La cible à détecter est un carré de taille 3×3 , noyé dans du bruit exponentiel non recouvrant. Cette cible apparaît avec un contraste de 4 sur le fond. Nous présentons dans la figure 3.5 les courbes COR de notre algorithme $GLRT_{mv}$. Cette figure présente également les courbes précédemment obtenues pour les algorithmes utilisant une taille fixe, juste ou fausse (3×3 et 7×7 pixels).

Bien entendu, l'estimation de taille conduit à des pertes de performances par rapport à un algorithme supposant la taille de la cible connue. Cependant, cette courbe montre que l'algorithme $GLRT_{mv}$ permet de compenser correctement une erreur de taille de référence.

3.5 Choix d'un seuil de détection

Les courbes COR présentent les performances d'un algorithme de détection pour toutes les valeurs de seuils possibles. En pratique, la détermination d'un seuil est néanmoins un problème important. Comme on l'a vu, l'algorithme $GLRT_{mv}$ à taille de cible inconnue choisit le filtre donnant la réponse maximale parmi l'ensemble des filtres possibles. Dès lors, il importe de savoir si le seuil de détection doit être fixé en fonction du filtre sélectionné.

Selon la théorie de Neyman-Pearson, le seuil de détection fixe la probabilité de fausse alarme. Nous étudierons donc la réponse de chacun des filtres dans le cas où l'échantillon

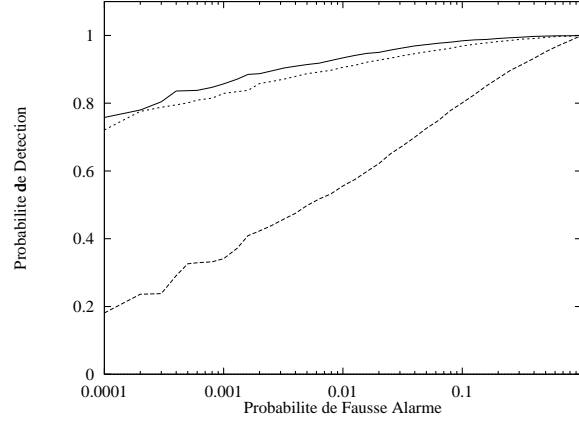


FIG. 3.5: Performances en détection pour une cible de taille inconnue : Courbes COR obtenues pour la détection d'une cible de taille 3×3 pixels, lorsque la taille de la référence est correcte, incorrecte, ou estimée au sens du MV .

— : Algorithme *GLRT* construit pour une référence de taille 3×3 pixels.

— — : Algorithme *GLRT* construit pour une référence de taille 7×7 pixels.

..... : Algorithme *GLRT_{mv}*, estimant la taille de la cible.

La fenêtre d'analyse est un carré de taille 19×19 . Les scènes sont composées de bruit Gamma d'ordre 1. Le contraste entre la cible et le fond vaut 4. Ces courbes COR sont mesurées sur 50000 réalisations de scènes.

considéré ne contient pas de cible. Nous montrerons alors que la densité de probabilité du logarithme du rapport de vraisemblance est indépendante de la taille de la référence, ainsi que de la taille de la fenêtre. Ce résultat est primordial en pratique, car il simplifie considérablement la tâche de détection, en limitant le nombre de paramètres sur lesquels la décision doit être prise. En d'autres termes, la probabilité de fausse alarme ne dépendra que du seuil de détection, et non plus de la taille des régions.

Pour démontrer cela, nous nous appuyons sur un résultat connu dans le domaine des probabilités et des processus stochastiques. On trouve ainsi dans [39, 50] le théorème qui suit :

Si N est le nombre de pixels de l'échantillon, n_1 le nombre de paramètres à estimer de l'hypothèse \mathcal{H}_1 et n_0 le nombre de paramètres de l'hypothèse \mathcal{H}_0 alors, si l'hypothèse \mathcal{H}_0 est vraie :

$$\lim_{N \rightarrow \infty} 2\mathcal{R}_j^{inc} = X \quad (3.16)$$

Avec X distribué selon une loi du χ^2 d'ordre $d = n_1 - n_0$. Une telle loi a pour densité de probabilité [39]

$$P_d(x) = \frac{1}{2^{d/2}\Gamma(d/2)} x^{d/2-1} \exp[-x/2] \quad (3.17)$$

Ainsi, la distribution de $2\mathcal{R}_j^{inc}$ en l'absence de cible est asymptotiquement indépendante du nombre de pixels de la cible et de la fenêtre. Ce résultat est intéressant en soi, car il montre que l'on ne peut pas déterminer sans ambiguïté qu'un échantillon est homogène, quel que soit sa taille, du fait des fluctuations statistiques.

Dans le cas de lois Gamma, il suffit d'estimer un paramètre par région. Le théorème précédent implique donc que lorsque la taille de la référence T_j tends vers l'infini, la distribution de \mathcal{R}_j^{inc} tend vers une loi limite, proportionnelle à une loi du χ^2 . Ce résultat théorique n'a d'intérêt pour nous que si ce comportement asymptotique est atteint très rapidement.

Pour le vérifier, la figure 3.6 présente la densité de probabilité estimée de $2\mathcal{R}_j^{inc}$ pour une cible de taille 3×3 et une fenêtre de taille 9×9 , ainsi que la densité de probabilité théorique d'une loi du χ^2 d'ordre 1 (cf. équation 3.17).

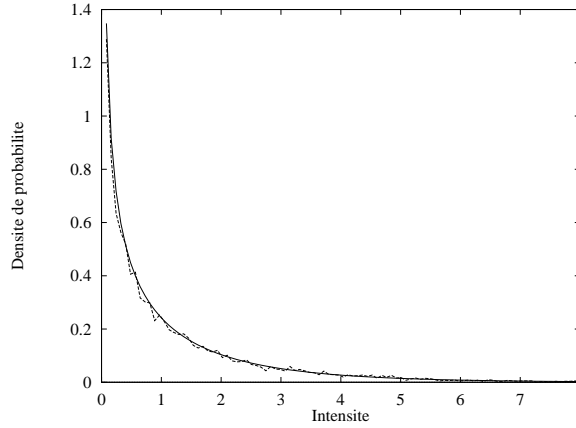


FIG. 3.6: Comparaison de la DDP du GLRT et d'une loi du χ^2 d'ordre 1, en l'absence de cible :
 — — — : DDP de $2lr_j$ pour une référence de 3×3 pixels, et une fenêtre 9×9 .
 — : Densité de probabilité d'une loi du χ^2 d'ordre 1 (cf. équation 3.17).
 Ces densités de probabilités sont estimées sur 10000 réalisations de bruit gamma.

La superposition de ces courbes nous permet d'utiliser la propriété 3.16 même dans le cas de cibles très petites. Il est donc possible d'utiliser un seuil de détection identique pour toutes les tailles de cibles possibles. Dès lors l'utilisation du GLRT estimant la taille de la cible est justifiée, tant par ses performances en détection que par sa simplicité de mise en œuvre. Dans le cas de l'algorithme que nous avons présenté, la taille de la fenêtre est la même, quelle que soit la taille de cible considérée. Dans notre contexte, il suffira donc de définir la PFA souhaitée simplement en fonction du seuil de détection, sans tenir compte de la taille de la cible.

La propriété 3.16 implique également que cette PFA est indépendante de la taille de la fenêtre. Pour le montrer, nous avons généré 10000 échantillons homogènes d'une loi Gamma d'ordre 1 et calculé le logarithme du rapport de vraisemblance de l'équation 3.15 en considérant des cibles et des fenêtres de tailles très différentes. Pour chaque taille, nous représentons dans la figure 3.7 la densité de probabilité estimée du GLRT. Encore une fois, la superposition de ces courbes montre que le seuil de détection peut effectivement être choisi indépendamment de la taille des cibles et de la taille des fenêtres. Dans le contexte de la théorie statistique de l'estimation que nous avons choisi, nous avons vu que les fenêtres d'analyse doivent être identiques. Cependant, la propriété d'indépendance entre le seuil de détection et la taille de la fenêtre nous a paru suffisamment importante pour être démontrée ici.

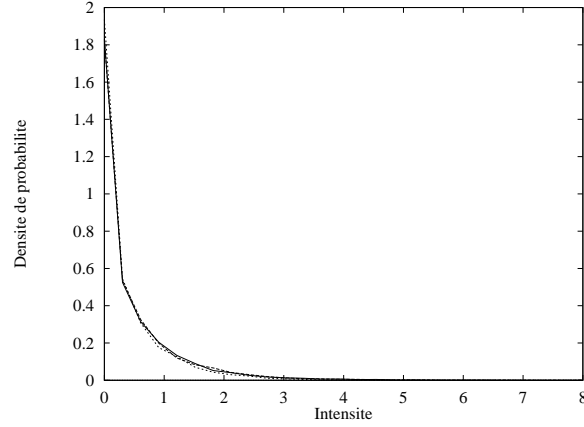


FIG. 3.7: Densité de probabilité du logarithme du rapport de vraisemblance en l'absence de cible, pour différentes tailles de références et de fenêtres. Les images sont composées de bruit Gamma d'ordre 1.

— : $GLRT$ pour une référence de 3×3 pixels, et une fenêtre 11×11 .

- - - : $GLRT$ pour une référence de 3×3 pixels, et une fenêtre 33×33 .

..... : $GLRT$ pour une référence de 11×11 pixels, et une fenêtre 33×33 .

Les scènes sont composées de bruit Gamma d'ordre 1. Ces densités de probabilités sont estimées sur 10000 réalisations de bruit.

Ce résultat est généralisable à différents ordres de lois Gamma, comme illustré dans la figure 3.8.

Dans le cas d'images composées de bruit Gamma, l'algorithme $GLRT$ supposant la taille de la cible connue présente donc une probabilité de fausse alarme constante, quelle que soit la moyenne du fond (cf. section 1.4.3), mais aussi quelle que soit la taille de la référence et de la fenêtre. Outre les avantages que cela procure dans une application de détection de cibles de tailles inconnues, ceci permettrait dans une application plus industrielle de simplifier considérablement la taille des abaques à construire pour la détermination d'un seuil de détection.

On aboutit aux mêmes conclusions lorsque les scènes sont composées de bruit gaussien, et que l'on utilise l'algorithme $GLRT$ adapté à ce bruit. Dans ce cas, il faut noter que le nombre de paramètres à estimer est doublé, puisque l'on estime à la fois moyenne et variance. La loi correspondante est donc une loi du χ^2 d'ordre 2. La figure 3.9 montre la densité de probabilité estimée en sortie de l'algorithme $GLRT$ prévu pour du bruit gaussien (cf. équation 2.9), appliqué à une image de bruit homogène gaussien décorrélé.

Avant de conclure sur ce chapitre, nous présentons dans la section suivante quelques exemples d'utilisation de l'algorithme $GLRT$ supposant la taille de la cible inconnue, en localisation/détection.

3.6 Exemples d'utilisation de l'algorithme $GLRT_{mv}$

Nous avons pu montrer l'intérêt théorique d'un algorithme de type $GLRT_{mv}$ pour la résolution du problème de la détection de cibles de taille inconnue. Nous avons également

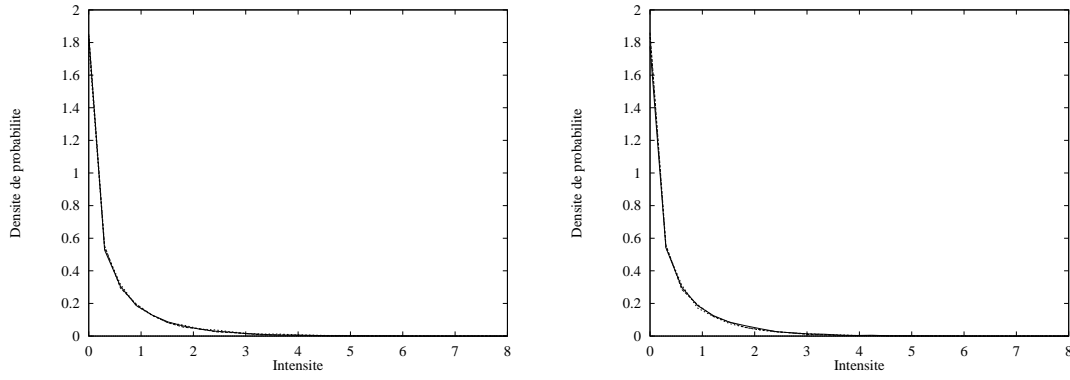


FIG. 3.8: Ordre des lois Gamma et densité de probabilité du logarithme du rapport de vraisemblance en l'absence de cible. **à gauche** : Ordre 3; **à droite** : Ordre 10. Les histogrammes sont obtenus pour différentes tailles de fenêtres et de référence.

— : *GLRT* pour une référence de 3×3 pixels, et une fenêtre 11×11 .

- - - : *GLRT* pour une référence de 3×3 pixels, et une fenêtre 33×33 .

. . . : *GLRT* pour une référence de 11×11 pixels, et une fenêtre 33×33 .

Ces densités de probabilités sont estimées sur 10000 réalisations de bruit.

montré qu'il permettait d'obtenir des performances en détection comparables à celles d'un algorithme supposant connue la taille de la cible. Ici, nous voulons illustrer visuellement ses performances en l'appliquant sur des images contenant effectivement trois cibles de tailles différentes dans le cas de scènes synthétiques et dans le cas de scènes plus réalistes.

La scène à traiter est une scène de bruits Gamma d'ordre 1 qui contient 3 cibles de 3, 5 et 7 pixels de large. La scène est une image 128×128 pixels. Le contraste sur la cible 3×3 est de 2. Le contraste sur la cible 5×5 est de 1.2. Le contraste sur la cible 7×7 est de 1.1.

On présente en figure 3.10 la scène à traiter, le plan de sortie de l'algorithme supposant inconnue la taille des cibles et le plan de sortie seuillé permettant la détection des cibles présentes.

Pour démontrer la robustesse de nos algorithmes lorsque les distributions des niveaux de gris de la scène ne suivent plus des lois gamma décorréelées, nous testons nos algorithmes sur une scène réaliste. La seule connaissance dont nous disposons ici est la forme des cibles à différentes tailles. En dépit de la difficulté de ces scènes, les algorithmes parviennent à détecter correctement chacune des cibles présentes.

3.7 Conclusion

Dans ce chapitre, nous nous sommes intéressés à la détection de cibles de tailles inconnues. Nous avons montré que la non-connaissance de la taille de la cible pouvait induire une baisse de performance importante pour les algorithmes existants. Nous avons donc envisagé de traiter ce problème comme une extension des algorithmes *GLRT* présentés précédemment. Nous avons montré que sur la base d'une modélisation rigoureuse du problème, il était possible de définir des algorithmes optimisant la probabilité de détection

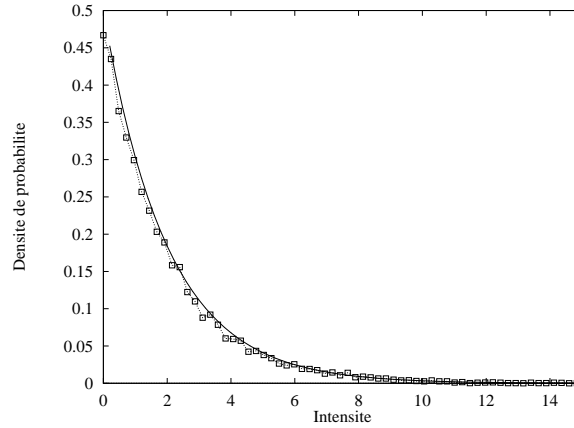


FIG. 3.9: Comparaison de la DDP du GLRT adapté à du bruit gaussien et d'une loi du χ^2 d'ordre 2, en l'absence de cible :

— — — : DDP de $2lr_j$ pour une référence de 3×3 pixels, et une fenêtre 9×9 .

— : Densité de probabilité d'une loi du χ^2 d'ordre 2 (cf. équation 3.17).

Ces densités de probabilités sont estimées sur 10000 réalisations de bruit.

pour une probabilité de fausse alarme donnée, lorsque la taille des cibles varie. Ceci revient à effectuer le filtrage de l'image par une batterie de filtres GLRT correspondant à des tailles de cibles différentes, avant de fusionner les résultats. Cette fusion des résultats s'apparente à une gestion de paramètres de nuisance, problème auquel la théorie de l'estimation apporte des solutions bien connues.

Parmi les deux classes de solutions possibles pour gérer le paramètre de taille, nous avons montré expérimentalement dans le cas d'images composées de bruit gamma, que l'estimation au sens du Maximum de Vraisemblance est plus robuste que l'Approche du Marginal Bayésien. Plus précisément, si la distribution *a priori* des tailles de cible ne correspond pas à la réalité des images présentées, le maximum de vraisemblance est préférable. Cet algorithme, le $GLRT_{mv}$ se définit de la façon suivante : En chaque pixel du plan de sortie de l'algorithme, on choisira le filtre ayant donné la sortie la plus forte.

Il se pose alors la question du choix d'un seuil de détection. En approfondissant les propriétés des différents filtres composant l'algorithme $GLRT_{mv}$, nous avons pu montrer que, dans le cas de bruit Gamma et gaussien, la probabilité de fausse alarme ne dépend que du seuil de détection. En particulier, elle est indépendante de la taille de la cible et de la fenêtre utilisée. Ce résultat est en soi important, dans le cas de cibles de tailles connues, pour réduire le nombre de paramètres à étudier dans la construction d'un abaque.

Dans le cas de la détection de cibles de taille inconnue, cela nous permet d'utiliser le même seuil de détection quelle que soit la taille sélectionnée. Ceci permet de limiter le nombre de paramètres de l'algorithme, ce qui d'un point de vue pratique est primordial.

Enfin, nous avons illustré sur quelques exemples les performances de ces algorithmes, dans le cas d'images synthétiques correspondant au modèle d'image pour lequel ils ont été construits, ainsi que dans le cas d'images réalistes, dans lesquelles ni la cible ni le fond ne peuvent être modélisés efficacement par des bruits homogènes.

Nous disposons donc désormais d'un algorithme rapide et robuste face aux inhomogé-

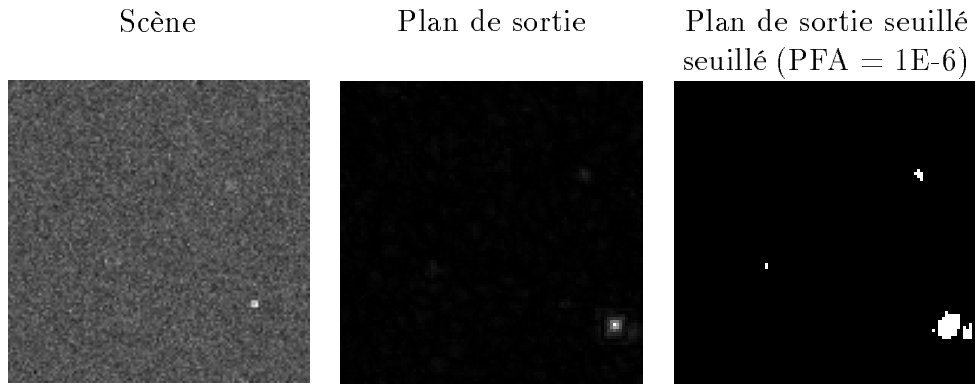


FIG. 3.10: Détection de cibles de tailles inconnues, scènes correspondant au modèle d'image pour du bruit Gamma

néités du fond et capable de traiter le cas de cibles de tailles inconnues. Dans le chapitre suivant, nous analyserons deux applications dans lesquelles ces deux problèmes sont particulièrement importants.

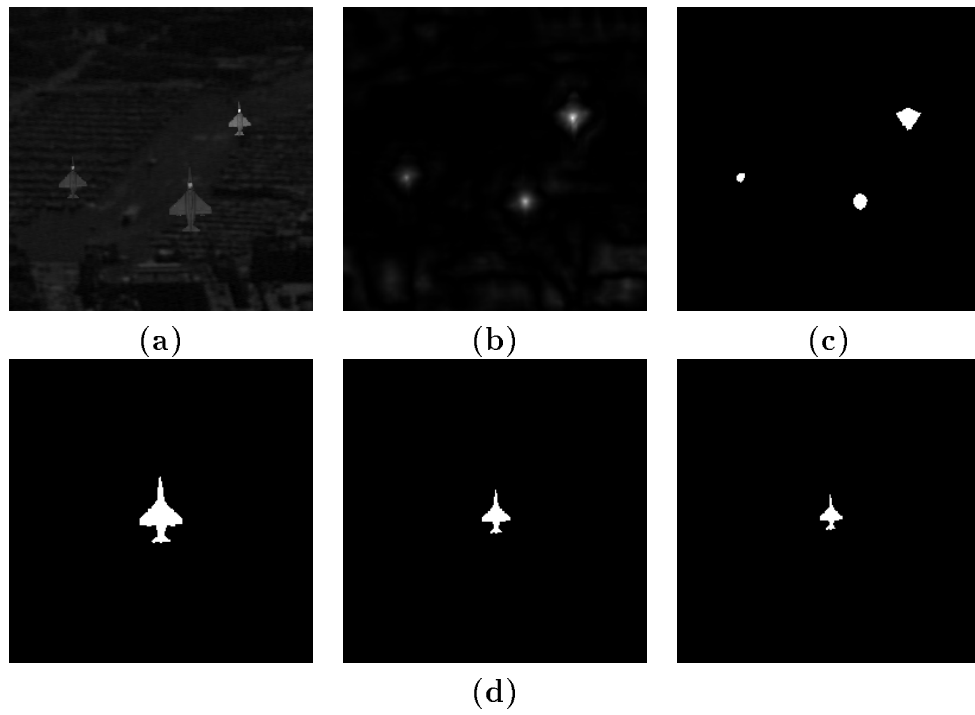


FIG. 3.11: Détection de cibles de tailles inconnues, application à des scènes réalistes :

(a) : Scène à traiter

(b) : Plan de sortie de l'algorithme GLRT estimant la taille des cibles

(c) : Plan de sortie seuillé, indiquant la détection des cibles

(d) : Références des différentes cibles.

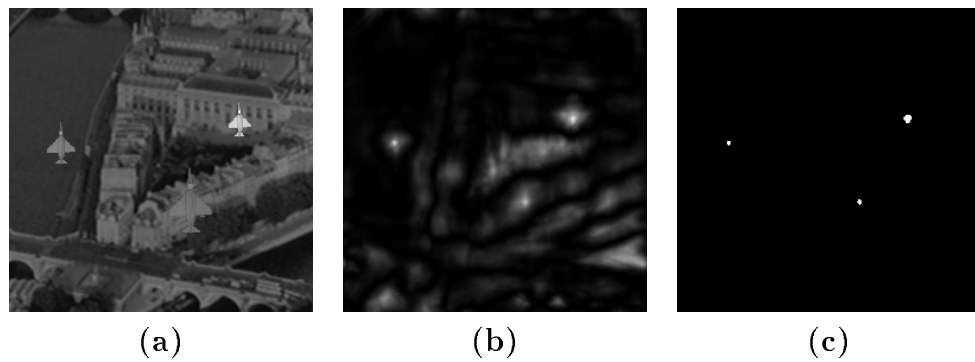


FIG. 3.12: Détection de cibles de tailles inconnues, scènes réalistes. (a) : Scène à traiter, (b) : plan de sortie et (c) : plan de sortie seuillé. Les références utilisées sont celles de la figure 3.11.

Chapitre 4

Exemples d'application : Imagerie SAR et Imagerie Polarimétrique

Sommaire

4.1	Imagerie SAR	54
4.2	Imageur Polarimétrique	56
4.2.1	Principe	56
4.2.2	Modèle d'images	57
4.2.3	Bruit homogène	59
4.2.4	Bruit non homogène	60
4.2.5	Représentation naturelle des images ICEO	62
4.2.6	Applications aux images réelles d'un imageur polarimétrique . .	63
4.3	Conclusion	65

La détection, la localisation, et l'identification automatique d'objets dans des images ou des séquences d'images sont des opérations essentielles dans de nombreux domaines, tels que la télédétection, le contrôle industriel, ou la surveillance. L'imagerie passive, qui utilise le rayonnement électromagnétique des objets pour en former l'image présente des limitations qu'il est souvent souhaitable de dépasser. Par exemple, un imageur fonctionnant dans le visible n'est pas adapté à l'observation nocturne. L'utilisation de capteurs infra-rouge, sensibles à des différences d'émission thermique, permet de s'affranchir partiellement de ce problème mais nécessite des systèmes optiques assez complexes.

L'imagerie active est une technique en plein essor, dans laquelle le système fournit lui-même la source d'illumination de la scène à imager. L'image est alors créée à partir de la rétro-diffusion du rayonnement émis par la source d'illumination. Le contrôle de cette source d'illumination offre une grande marge de manoeuvre sur les paramètres des systèmes. En particulier, le choix de la longueur d'onde utilisée permet d'améliorer la résolution des systèmes, ou d'imager des scènes inaccessibles habituellement. C'est le cas de l'imagerie SAR (Synthetic Aperture Radar), qui utilise un rayonnement micro-ondes qui traverse les nuages, permettant l'acquisition d'images de zones géographiques placées en permanence sous une couverture nuageuse.

L'imagerie active cohérente permet l'acquisition de paramètres de la scène plus complets que le seul facteur de réflexion en intensité des objets. On peut ainsi utiliser une résolution en temps de la lumière réfléchie pour dresser une image de distance de la scène observée, ou détecter et estimer le décalage Doppler introduit par les éléments de la scène qui sont en mouvement. On dispose ainsi de systèmes permettant d'imager les propriétés des objets présents, sous des formes très différentes.

L'imageur polarimétrique actif fait partie de cette classe de systèmes d'acquisition. Il est sensible à la polarisation de l'onde rétro-diffusée. La connaissance de cette polarisation permet de quantifier la dépolarisation induite par la diffusion des différents éléments de la scène. Celle-ci dépendant de la rugosité des objets, on peut ainsi distinguer un objet du fond, même si ses propriétés de réflexion en intensité sont identiques.

Cependant, les imageurs actifs présentent également quelques inconvénients. Les images obtenues en lumière cohérente sont perturbées par du bruit de speckle, qui dégrade la qualité des images de façon conséquente. Pour obtenir des algorithmes efficaces, il est nécessaire de prendre en compte les caractéristiques de ce type de bruit. Comme nous l'avons vu, les algorithmes statistiques sont bien adaptés à ce problème.

Nous donnerons un exemple d'application des algorithmes travaillant avec des cibles de taille inconnue pour des images SAR réelles, ainsi que pour des images obtenues à l'aide d'un imageur polarimétrique. Dans le second cas, les images obtenues présentent un fond très inhomogène, ce qui nous conduira à effectuer une transformation préalable des données avant de leur appliquer nos algorithmes.

4.1 Imagerie SAR

La localisation et la détection dans des images SAR est un problème complexe et très étudié. Le lecteur pourra se référer à [38] pour une vision globale du problème. De ce fait, notre propos n'est pas ici de présenter une solution à la localisation ou à la détection de petites cibles dans des images de type SAR. Nous nous contenterons de montrer l'intérêt de l'algorithme GLRT estimant la taille des cibles en tant qu'élément constituant d'un système plus complexe.

Il est en effet bien connu que le bruit de speckle, omniprésent dans les images SAR [38], peut être modélisé comme un bruit gamma d'ordre 1, correspondant à l'exemple d'application présenté dans la section 3.6.

Nous appliquons cet algorithme à deux vignettes extraites d'une image réelle d'une observation datant de 1991 par le satellite ERS-1 d'une région agricole de l'Ukraine. Ce type d'image est délivrée par l'ESA et nous a été fournie par le CNES. Chaque vignette à une taille de 128×128 pixels. Pour chaque vignette, nous présentons dans la figure 4.1 le plan de sortie de l'algorithme, ainsi que la ou les zones susceptibles d'être détectées. Ces zones sont encadrées dans l'image d'origine.

Ces résultats montrent que l'algorithme réussit à détecter indifféremment des cibles plus claires ou plus sombres que le fond qui les entoure. Les algorithmes de type GLRT qui supposent la taille de la cible connue sont régulièrement mis en oeuvre dans le domaine de l'imagerie SAR, notamment en détection de contour. Ce fut même leur premier domaine d'application [13, 37].

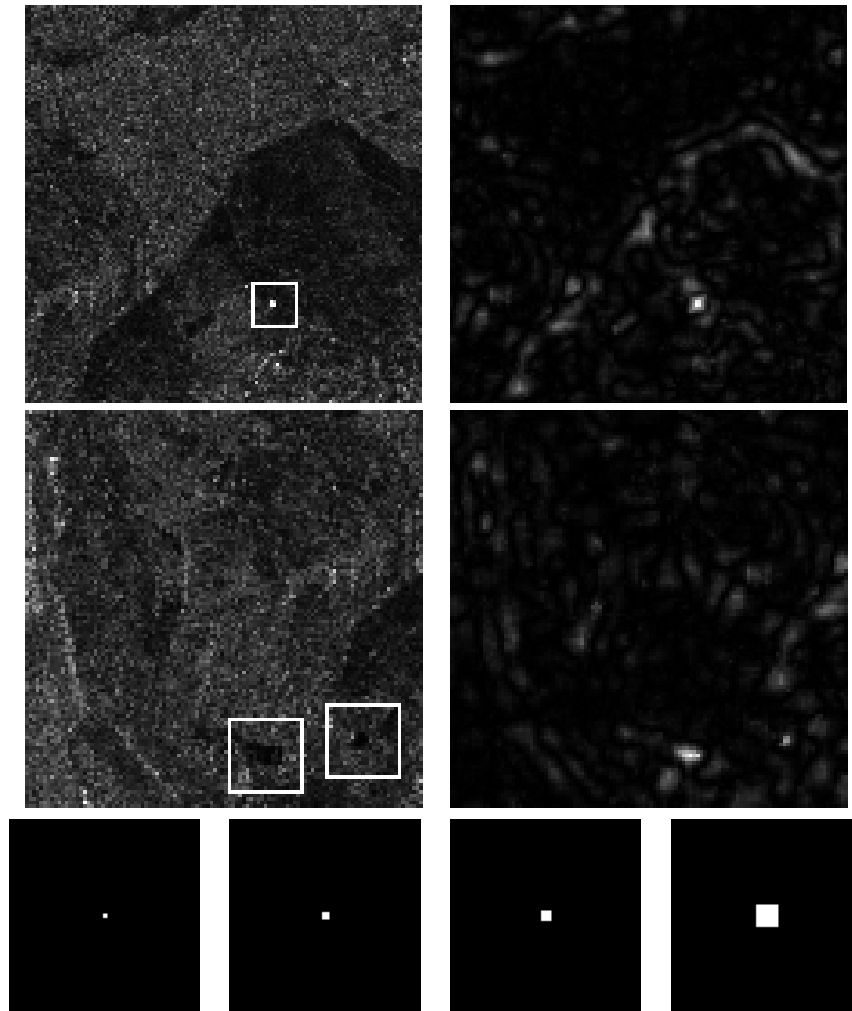


FIG. 4.1: Exemple d'application au cas d'images SAR, vignettes extraites d'une observation de l'Ukraine, acquise en 1991 par le satellite ERS-1 (Copyright CNES/ESA). Première ligne : image originale et plan de sortie de l'algorithme GLRT multi tailles. Deuxième ligne : idem. troisième ligne : références et fenêtre utilisées pour la construction de l'algorithme

Il faut néanmoins noter la présence de nombreuses fausses alarmes dans les plans de sortie présentés. En particulier, les "frontières" délimitant deux régions homogènes apparaissent fortement dans les plans de sortie. En effet, l'algorithme GLRT n'est pas très discriminant et tend à fournir des réponses fortes en présence d'une différence de statistique dans la fenêtre d'analyse, même lorsque les régions n'ont pas la forme exacte de l'objet cherché.

Pour résoudre ce problème, Oliver propose de recourir à une segmentation préalable de l'image, permettant de regrouper les zones où le fond est de moyenne constante [38]. Cette segmentation ne prend pas en compte les cibles de petites tailles, qui peuvent être détectées en effectuant une recherche uniquement sur les zones de réflectivité constante.

Enfin, une autre piste pour l'élimination de ce problème, est l'utilisation de données plus complexes. En effet, nous n'effectuons la détection que par l'analyse des niveaux de

gris de la scène. Les techniques d'imagerie hyperspectrales, qui sont apparues ces toutes dernières années [26], permettent d'obtenir pour chaque scène, un ensemble d'images correspondant chacune à une longueur d'onde d'illumination précise. Ainsi, on peut disposer d'une signature spectrale de la cible cherchée, qui permet une discrimination plus efficace. Notons que le GLRT est souvent retenu pour les applications de détection hyperspectrale, pour ses qualités de simplicité et d'efficacité [27].

4.2 Imageur Polarimétrique

Contrairement au cas de l'imagerie SAR, pour laquelle une application directe du GLRT est envisageable, dans le cas de l'imagerie polarimétrique active, nous verrons que les données nécessitent un travail préalable et la définition d'un mode de représentation adapté. Pour cela, commençons par définir le principe d'un imageur polarimétrique simple tel que celui utilisé par Thales Optronique.

4.2.1 Principe

L'imageur polarimétrique de Thales Optronique [4] illumine la scène avec un faisceau de lumière cohérente issu d'un groupe de diodes lasers. Cette lumière possède un état de polarisation bien défini, modifié à volonté à l'aide d'un générateur d'état de polarisation. La lumière réfléchiée par la scène est décomposée en deux parties grâce à un analyseur d'état de polarisation, pour former deux images. La première image $s_1(i)$ est formée à partir de la lumière réfléchiée ayant le même état de polarisation que l'onde incidente. La seconde $s_2(i)$ est formée à partir de la lumière dont l'état de polarisation est orthogonal à celui de l'onde incidente. Il s'agit de lumière qui a été dépolarisée. Ces deux composantes forment une représentation bi-canal de la scène.

Le schéma optique de l'imageur est présenté dans la figure 4.2

La somme de ces deux images est simplement l'image d'intensité réfléchiée par la scène. On peut aussi former l'Image de Contraste d'Etats Orthogonaux (ICEO) [17] $\rho(i)$:

$$\rho(i) = \frac{s_1(i) - s_2(i)}{s_1(i) + s_2(i)} \quad (4.1)$$

Cette image permet de quantifier la manière dont la lumière incidente est dépolarisée par la zone de la scène correspondant au pixel (i) . Elle permet donc de faire apparaître un contraste entre les zones plus ou moins dépolarisantes de la scène.

Un exemple des différentes images acquises avec cet imageur est présenté dans la figure 4.3.

L'imageur polarimétrique illumine la scène avec une lumière cohérente. Les images s_1 et s_2 sont donc perturbées par un bruit de speckle. Une étude préalable de ces images a permis de montrer que la statistique des niveaux de gris de ces images suit une loi Gamma d'ordre élevé.

La première question qui se pose est de déterminer quelles images utiliser pour des tâches de détection avec ce type d'imageur. Nous pouvons envisager d'utiliser l'image

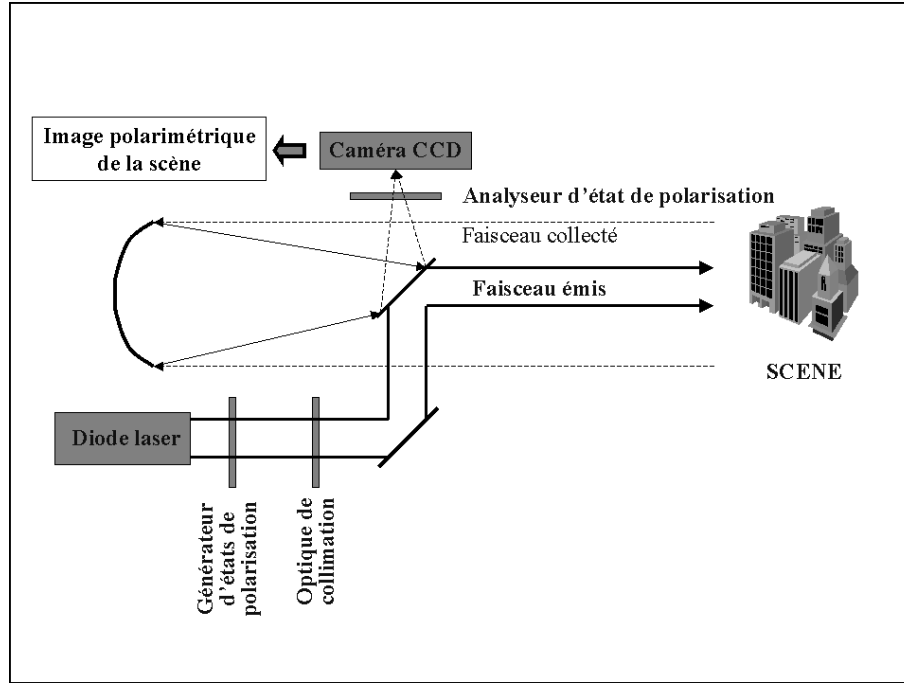


FIG. 4.2: Principe de fonctionnement de l'imageur actif polarimétrique

bi-canal (IBC) [17] composée du couple (s_1, s_2) , ou de travailler sur l'image ICEO directement.

Nous allons démontrer la supériorité d'une approche utilisant l'image bi-canal en présence d'un éclairage uniforme. En revanche, un éclairage non uniforme implique des pertes de performances que nous avons explicitées dans le chapitre 2. Dès lors, nous montrerons que l'image ICEO permet de s'affranchir en partie de ce problème, et constitue donc un choix préférable d'images sur lesquelles utiliser nos algorithmes.

4.2.2 Modèle d'images

L'imageur polarimétrique fournit en premier lieu l'image bi-canal (IBC) :

$$s = (s_1, s_2) \quad (4.2)$$

Le modèle d'image choisi est le même que celui que nous avons présenté dans la section 1.1. Nous supposons donc que l'image est composée d'une zone cible \mathbf{w} et d'une zone fond $\bar{\mathbf{w}}$. De plus, nous considérons que ces deux zones sont indépendantes l'une de l'autre (modèle SIR). Dans chacun des canaux, les niveaux de gris de chaque zone sont distribués selon une loi Gamma de même ordre et de moyenne variable. Nous ferons de plus l'hypothèse que les champs aléatoires sont décorrélés.

Ainsi, le modèle de scène choisi est donné par :

$$\forall k \in [1, 2] \quad s_k(i) = a_k(i) w(i) + b_k(i)(1 - w(i)) \quad (4.3)$$

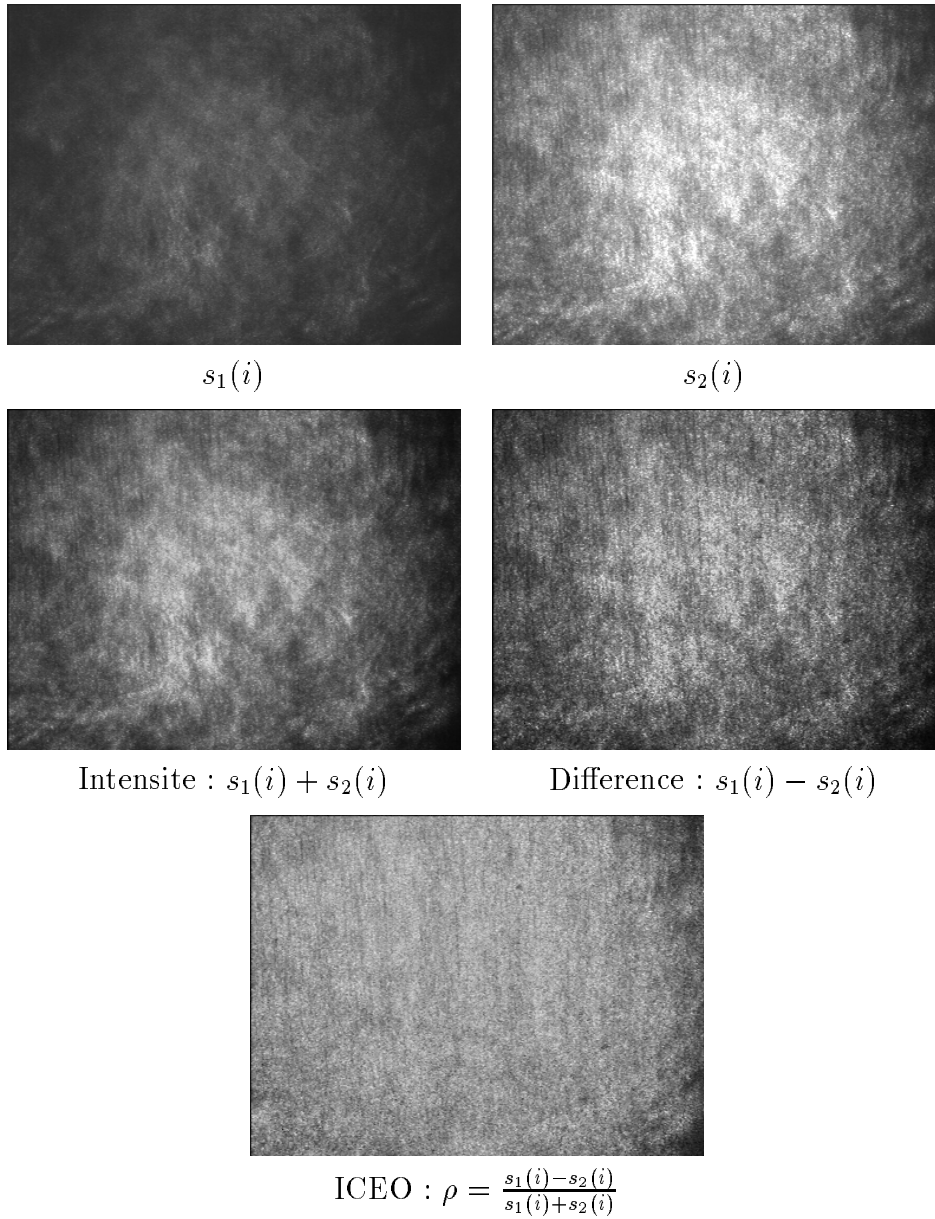


FIG. 4.3: Exemple d'image acquises à l'aide d'un imageur polarimétrique sur un fond de bois

où \mathbf{a}_k (resp. \mathbf{b}_k) sont des réalisations de champs aléatoires spatialement décorrélés dont la densité de probabilité paramétrique est celle d'une loi Gamma d'ordre L et de moyenne μ_k^a (resp. μ_k^b).

Expression du GLRT en bi-canal avec des statistiques Gamma

Comme nous l'avons dit, les fluctuations des niveaux de gris de l'IBC suivent approximativement des lois Gamma. Dans ce cas, l'expression du GLRT lorsque les paramètres des densités de probabilité sont inconnus et estimés au sens du maximum de vraisemblance se déduit aisément de celle de son équivalent en monocanal (cf. équation (3.15)),

qui a été présentée dans le chapitre 3. L'expression du GLRT correspondant, que nous noterons IBC-GLRT est la suivante [17] :

$$\begin{aligned} \mathcal{R}(i) = & -L N_w \left[\sum_{k=1}^2 \log \hat{\mu}_k^a(i) \right] - L N_{\bar{w}} \left[\sum_{k=1}^2 \log \hat{\mu}_k^b(i) \right] \\ & + L N_F \left[\sum_{k=1}^2 \log \left[\frac{N_w}{N_F} \hat{\mu}_k^a(i) + \frac{N_{\bar{w}}}{N_F} \hat{\mu}_k^b(i) \right] \right] \end{aligned} \quad (4.4)$$

avec

$$\hat{\mu}_k^a(i) = \frac{1}{N_w} \sum_{(m) \in \mathbf{w}^{(i)}} s_k(m) \quad (4.5)$$

$$\hat{\mu}_k^b(i) = \frac{1}{N_{\bar{w}}} \sum_{(m) \in \bar{\mathbf{w}}^{(i)}} s_k(m) \quad (4.6)$$

où $\mathbf{w}^{(i)}$ représente une version du masque binaire \mathbf{w} qui est centrée sur le pixel (i) . $\bar{\mathbf{w}}^{(i)}$ est le complémentaire de $\mathbf{w}^{(i)}$ dans la fenêtre d'analyse.

Expression du GLRT en mono-canal avec des statistiques gaussiennes

L'ICEO [17] est une image mono-canal, dont la densité de probabilité peut être approchée par une gaussienne. Nous considérons donc une image mono-canal dont les fluctuations de niveaux de gris suivent des lois gaussiennes de moyennes et de variances inconnues. Lorsque ces paramètres sont estimés au sens du maximum de vraisemblance, nous obtenons l'algorithme que nous noterons ICEO-GLRT. L'expression de $\mathcal{R}(i)$ est alors la suivante :

$$\mathcal{R}(i) = -N_w \log [\hat{\sigma}^a(i)]^2 - N_{\bar{w}} \log [\hat{\sigma}^b(i)]^2 + N_F \log [\hat{\sigma}^c(i)]^2 \quad (4.7)$$

où

$$[\hat{\sigma}^a(i)]^2 = \frac{1}{N_w} \sum_{(m) \in \mathbf{w}^{(i)}} [\rho(m)]^2 - \left[\frac{1}{N_w} \sum_{(m) \in \mathbf{w}^{(i)}} \rho(m) \right]^2 \quad (4.8)$$

$$[\hat{\sigma}^b(i)]^2 = \frac{1}{N_{\bar{w}}} \sum_{(m) \in \bar{\mathbf{w}}^{(i)}} [\rho(m)]^2 - \left[\frac{1}{N_{\bar{w}}} \sum_{(m) \in \bar{\mathbf{w}}^{(i)}} \rho(m) \right]^2 \quad (4.9)$$

$$[\hat{\sigma}^c(i)]^2 = \frac{1}{N_F} \sum_{(m) \in \mathbf{F}^{(i)}} [\rho(m)]^2 - \left[\frac{1}{N_F} \sum_{(m) \in \mathbf{F}^{(i)}} \rho(m) \right]^2 \quad (4.10)$$

4.2.3 Bruit homogène

L'IBC-GLRT est l'algorithme de détection optimal au sens du maximum de vraisemblance pour une image décrite par un modèle SIR et où la densité de probabilité

des niveaux de gris est de type Gamma. Dans le cadre de ce modèle, il optimise donc la probabilité de détection à taux de fausse alarme donné. En revanche, l'ICEO-GLRT correspond à un algorithme sous-optimal, puisque nous avons fait l'hypothèse d'une distribution gaussienne. De plus, en calculant l'ICEO à partir de l'IBC, on perd de l'information par rapport à cette dernière. Les performances des deux algorithmes dépendent des paramètres $\gamma_1 = \mu_1^a/\mu_1^b$ et $\gamma_2 = \mu_2^a/\mu_2^b$, qui représentent les contrastes de niveaux de gris moyens entre la cible et le fond dans chacun des deux canaux.

Il est clair que quelles que soient les valeurs de γ_1 et de γ_2 , l'IBC-GLRT conduira à de meilleures performances que l'ICEO-GLRT. En effet, examinons tout d'abord l'ensemble des couples (γ_1, γ_2) pour lesquels ces deux algorithmes sont inefficaces, c'est-à-dire pour lesquels ils donnent une probabilité de détection égale à la probabilité de fausse alarme.

L'IBC-GLRT est inefficace lorsque $\gamma_1 = \gamma_2 = 1$, c'est à dire lorsque le contraste est nul dans chaque canal. Cette configuration est représentée par le point P dans le plan (γ_1, γ_2) (cf. figure 4.4).

Considérons maintenant l'ICEO-GLRT. Sur la cible, dans le canal 1, on peut écrire : $s_1(i) = \mu_1^a s'_1(i)$, ou s' suit une loi Gamma de moyenne unitaire. Des notations similaires sont adoptées pour chaque région et pour chaque canal. Notons alors $\gamma_a = \mu_1^a/\mu_2^a$, et $\gamma_b = \mu_1^b/\mu_2^b$. On peut ré-écrire l'équation 4.1 sous la forme

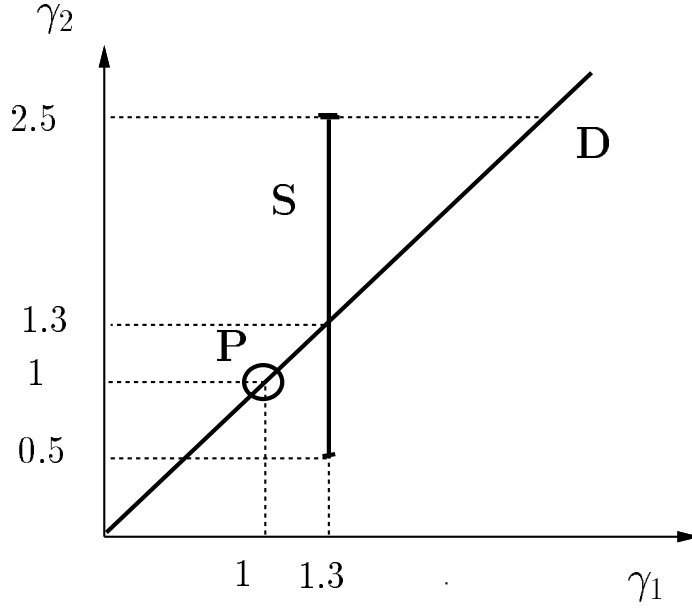
$$\begin{aligned} \rho(i) &= \frac{\gamma_a s'_1(i) - s'_2(i)}{\gamma_a s'_1(i) + s'_2(i)} \text{ sur la cible} \\ \rho(i) &= \frac{\gamma_b s'_1(i) - s'_2(i)}{\gamma_b s'_1(i) + s'_2(i)} \text{ sur le fond} \end{aligned}$$

La distribution des variables aléatoires $s'_k(i)$ étant connue et ne dépendant d'aucun paramètre, la distribution des niveaux de gris de l'image ICEO ne dépend donc que de γ_a sur la cible et de γ_b dans le fond. Lorsque ces paramètres sont égaux, les niveaux de gris de la cible et du fond sont identiquement distribués et rendent la discrimination entre ces régions impossible. L'algorithme ICEO-GLRT est donc inefficace lorsque $\gamma_a = \gamma_b$ ce qui, après quelques calculs, revient à $\gamma_1 = \gamma_2$. Le lieu de ces points dans le plan (γ_1, γ_2) est la droite D de pente unitaire passant par l'origine (cf. figure 4.4).

Pour illustrer ce résultat, considérons la détection d'une cible carrée de 3×3 pixels en utilisant une fenêtre de recherche de 5×5 pixels. Nous avons fixé $\gamma_1 = 1.3$ et fait varier γ_2 de 0.5 à 2.5, ce qui correspond à un déplacement sur le segment S représenté sur la figure 4.4. Nous avons représenté la variation de P_D à $P_{FA} = 10^{-3}$ en fonction de γ_2 sur la figure 4.5. On peut noter que la probabilité de détection obtenue en utilisant l'IBC-MLRT est toujours supérieure à celle obtenue avec l'ICEO-MLRT. La performance minimale est obtenue pour l'IBC-MLRT lorsque $\gamma_2 = 1$ (le contraste est nul sur le second canal) et pour l'ICEO-MLRT lorsque $\gamma_2 = \gamma_1 = 1.3$.

4.2.4 Bruit non homogène

Supposons qu'une image IBC (s_1, s_2) puisse être obtenue avec une illumination uniforme unitaire. Lorsque cette illumination n'est pas uniforme, nous pouvons modéliser ses variations spatiales par un facteur multiplicatif θ . La scène IBC (s'_1, s'_2) observée sous une

FIG. 4.4: Graphe représentant le plan (γ_1, γ_2) .

telle illumination s'écrit

$$\begin{aligned} s'_1(i) &= \theta(i) s_1(i) \\ s'_2(i) &= \theta(i) s_2(i) \end{aligned} \tag{4.11}$$

$$\tag{4.12}$$

Les deux canaux de l'IBC sont alors inhomogènes, et comme on l'a vu précédemment, l'inhomogénéité du fond conduit à des pertes importantes en détection et en localisation.

En revanche, on peut vérifier très simplement que l'ICEO calculée à partir de s'_1 et s'_2 est indépendante du facteur d'illumination θ . En effet :

$$ICEO'(i) = \frac{s'_1(i) - s'_2(i)}{s'_1(i) + s'_2(i)} = \frac{s_1(i) - s_2(i)}{s_1(i) + s_2(i)} \tag{4.13}$$

L'image ICEO, dans l'hypothèse de l'équation 4.12, est donc insensible aux variations de l'illumination. Ceci est observable sur les images de la figure 4.3. L'inhomogénéité d'éclairement très apparente dans chaque canal est fortement atténuée sur l'image ICEO. On peut néanmoins remarquer sur cette image quelques inhomogénéités résiduelles, sans doute dues à la texture du matériau.

Pour montrer l'intérêt de l'ICEO dans le cas d'éclairement non homogène, nous avons repris le modèle de bruit inhomogène de la section 2.2.1. Ce modèle dépend de deux paramètres : la variance de l'illumination σ^2 et la longueur de corrélation l des inhomogénéités. Nous avons ensuite estimé les performances en détection des algorithmes IBC-GLRT et ICEO-GLRT en présence de ce bruit, pour différentes longueurs de corrélation. La variance de l'illumination est fixée à 0.05. Notons que lorsque la longueur de corrélation tends vers l'infini, le bruit devient homogène.

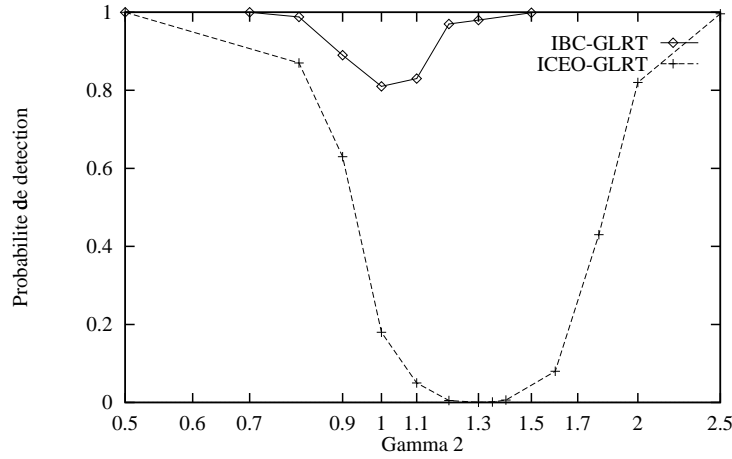


FIG. 4.5: Probabilité de détection (P_D) en fonction de γ_2 pour $P_{FA} = 10^{-3}$ en utilisant l'IBC-GLRT et l'ICEO-GLRT. La cible \mathbf{w} est un carré de 3×3 pixels et la fenêtre de recherche F un carré de 5×5 pixels. On a fixé $\gamma_1 = 1.3$, et P_D et P_{FA} ont été estimées sur 10^4 réalisations.

Ces performances sont mesurées en termes de courbes COR (cf section 1.3.4), et présentées dans la figure 4.6. La taille de la cible est ici connue, il s'agit d'un carré de taille 3×3 . La fenêtre d'analyse est un carré de taille 11×11 . On constate que les performances de l'IBC-GLRT diminuent de façon sensible avec la longueur de corrélation du bruit l , par rapport aux performances observées en bruit homogène. En revanche, les performances de l'ICEO-GLRT sont, par définition insensibles à ce type d'inhomogénéités, de telle sorte qu'elles deviennent rapidement meilleures que celles de l'IBC-GLRT pour les longueurs de corrélation inférieures à 30. Dans cette simulation, les bruits Γ représentant le speckle sont d'ordre 50. Le rapport des moyennes était fixé à 1.6 dans le premier canal et à 1 dans le second canal.

Les images obtenues par l'imageur polarimétrique étant fortement inhomogènes, l'utilisation d'images de type ICEO sera préférable. Par ailleurs, nous avons précédemment montré l'influence de l'estimation des paramètres des lois sur les performances des algorithmes. L'algorithme ICEO-GLRT estime des variances. Or l'estimateur de variance fluctue fortement sur des cibles aussi petites que celles que nous avons considéré. la prochaine section montrera comment une transformation de l'ICEO permet de limiter l'influence de ce problème.

4.2.5 Représentation naturelle des images ICEO

Nous avons montré que les images ICEO, du fait de leur insensibilité à la figure d'éclairément, permettent d'obtenir des performances supérieures à celle obtenues en bi-canal. Il en sera de même pour toute tranformation bijective des images ICEO, conservant cette propriété. Dans le cadre de la recherche de paramètres réduits pour caractériser les propriétés des images ICEO, *Goudail et al* ont proposé [17] d'utiliser à la place de l'image ICEO une image dite DLOG, obtenue de la façon suivante :

$$DLOG(i) = \log(s_1(i)) - \log(s_2(i)) \quad (4.14)$$

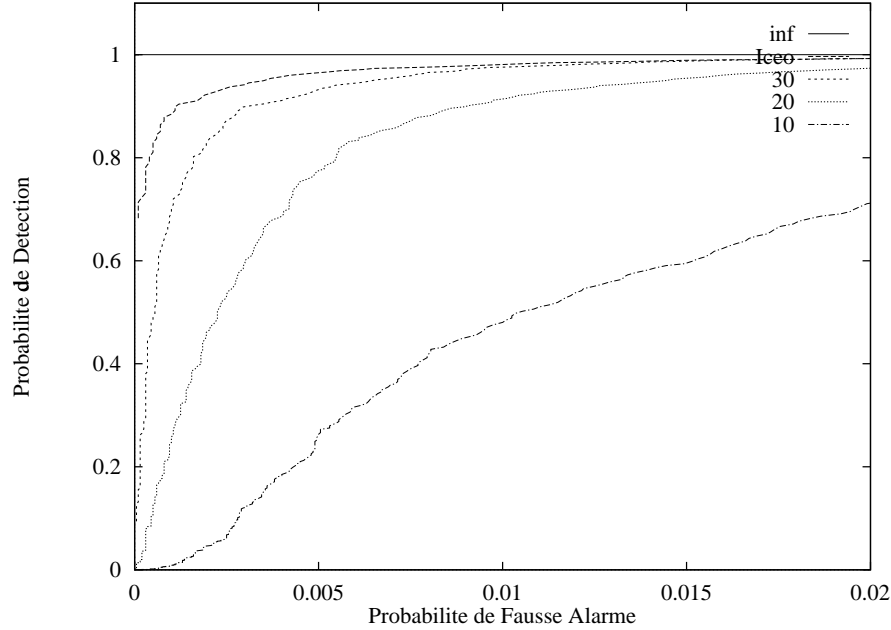


FIG. 4.6: Influence de l'inhomogénéité du fond : Courbes COR obtenues pour la détection d'une cible de taille 3×3 pixels pour différentes configurations : De haut en bas : IBC-GLRT en présence d'illumination uniforme (ligne horizontale) ICEO-GLRT (indépendant de l'illumination) IBC-GLRT en présence d'illumination non uniforme avec $l = 30$; $l = 20$; $l = 10$; Les courbes COR sont estimées sur 50000 réalisations.

On peut montrer simplement que ces images offrent la même propriété d'annulation des inhomogénéités que les images ICEO. De plus, la densité de probabilité des intensités dans des régions homogènes est toujours correctement modélisée par des lois gaussiennes, mais deux régions homogènes ne se différencient plus désormais que par une différence de moyenne, les variances étant identiques. L'algorithme optimal en détection dans ce cas est donné par :

$$DLOG - GLRT(i) = \left(\left[\frac{1}{N_a} \mathbf{s} \otimes \mathbf{w} \right](i) - \frac{1}{N_b} \mathbf{s} \otimes \bar{\mathbf{w}}(i) \right)^2 \quad (4.15)$$

Cette équation ne fait intervenir que la moyenne de chaque région. L'estimation des paramètres dans ce cas est donc plus précise que celle de l'algorithme GLRT estimant les variances. Ceci permet d'obtenir des résultats meilleurs en détection, comme illustré dans la figure 4.7.

Cette représentation est donc plus adaptée pour nos algorithmes.

4.2.6 Applications aux images réelles d'un imageur polarimétrique

Le calcul de l'algorithme GLRT estimant la taille de la cible est immédiat, puisqu'il consiste à prendre le maximum du plan de sortie de différents algorithmes supposant la taille des cibles connue.

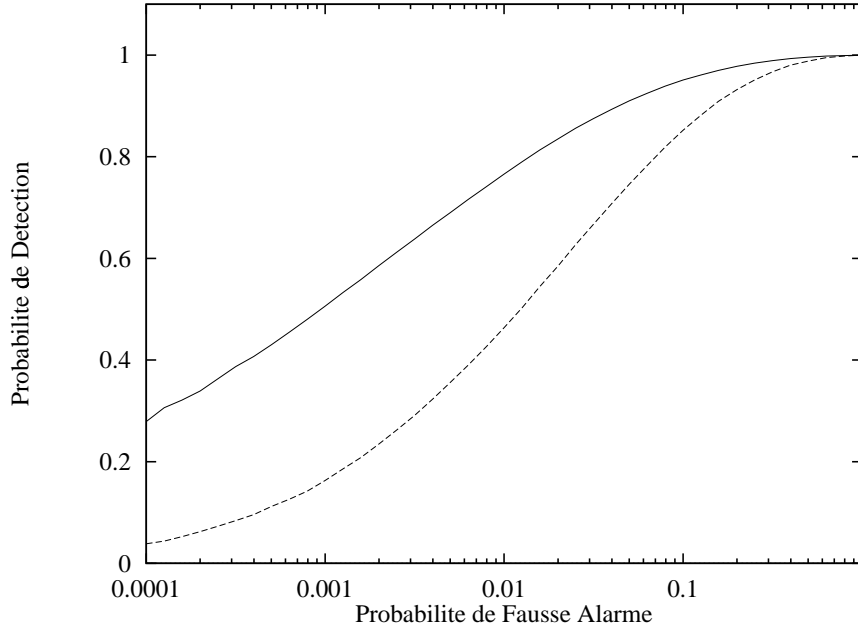


FIG. 4.7: Courbes COR obtenues par les algorithmes ICEO-GLRT et DLOG-GLRT, sur des scènes de bruit Gamma d'ordre 50 pour la détection d'une cible de 3×3 pixels. Le contraste dans le canal 1 était de 1.2, le contraste dans le canal 2 de 1.1. Le DLOG-GLRT n'utilise qu'une estimation des moyennes de chaque région, alors que l'ICEO-GLRT estime les variances de chaque région.

— : Algorithme ICEO-GLRT
 - - - : Algorithme DLOG-GLRT

Dans les images réelles dont nous disposons, les cibles sont mal résolues, et nous ne disposons d'aucune référence de leur forme, ni aucune indication de leur taille. Nous avons donc choisi des cibles carrées, de surface 3×3 , 7×7 et 15×15 pixels, définissant ainsi des régions d'intérêt.

Nous présentons dans la figure 4.8 un exemple d'image IBC, d'une cible en métal posée sur un support de bois, ainsi que l'image DLOG associée. Le résultat du filtrage par l'algorithme DLOG-GLRT travaillant à taille de cible inconnue est également fourni. Cet algorithme a permis dans le cas présent, de localiser la cible correctement.

On peut observer que la texture du fond tend à produire des pics parasites. Spécifiquement dans le cas de cibles de tailles inconnues, ces inhomogénéités sont de la taille de la référence la plus petite, et peuvent conduire à l'apparition de fausses alarmes.

La figure 4.9 montre la même expérience conduite pour une cible de papier posée sur un fond de bois.

Dans les deux cas présentés, les algorithmes permettent de définir une région d'intérêt centrée sur la cible. Dans le dernier couple d'images présenté, on peut voir que la présence d'une fausse cible (en bas à droite des images) entraînerait l'échec de la localisation, le maximum du plan de sortie n'étant pas placé sur la cible. Cette fausse cible constitue néanmoins une région d'intérêt, dont la détection peut permettre de servir de base à

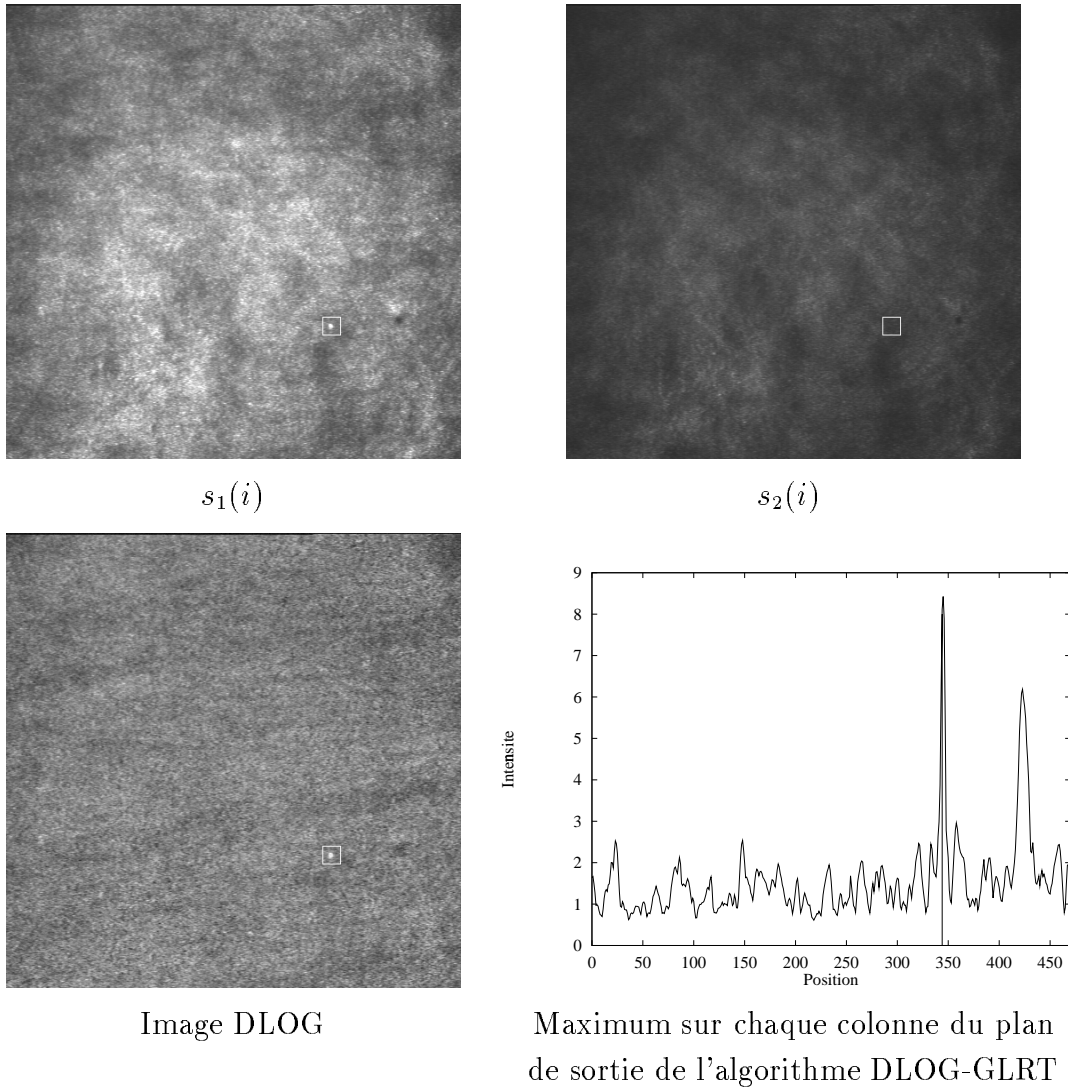


FIG. 4.8: Exemple de résultat obtenu sur des images polarimétrique en recherche de régions d'intérêt avec l'algorithme DLOG-GLRT estimant la taille de la cible. La cible est soulignée dans chaque image par un carré blanc. Il s'agit d'une pièce en métal sur un fond de bois

une analyse postérieure de ces images. Dans notre cas, où les cibles sont petites, de niveaux de gris et de taille inconnus, les résultats précédents peuvent être considérés comme satisfaisants.

4.3 Conclusion

Dans ce chapitre, nous nous sommes intéressés à l'application des algorithmes développés dans les chapitres 2 et 3. Ces algorithmes permettent d'effectuer une détection ou une localisation de façon optimale au sens de la théorie de la décision. Il tiennent compte des lois physiques sous-tendant la formation des images, et sont donc particulièrement adaptés au cas d'images obtenues en imagerie active, où la modélisation des statistiques

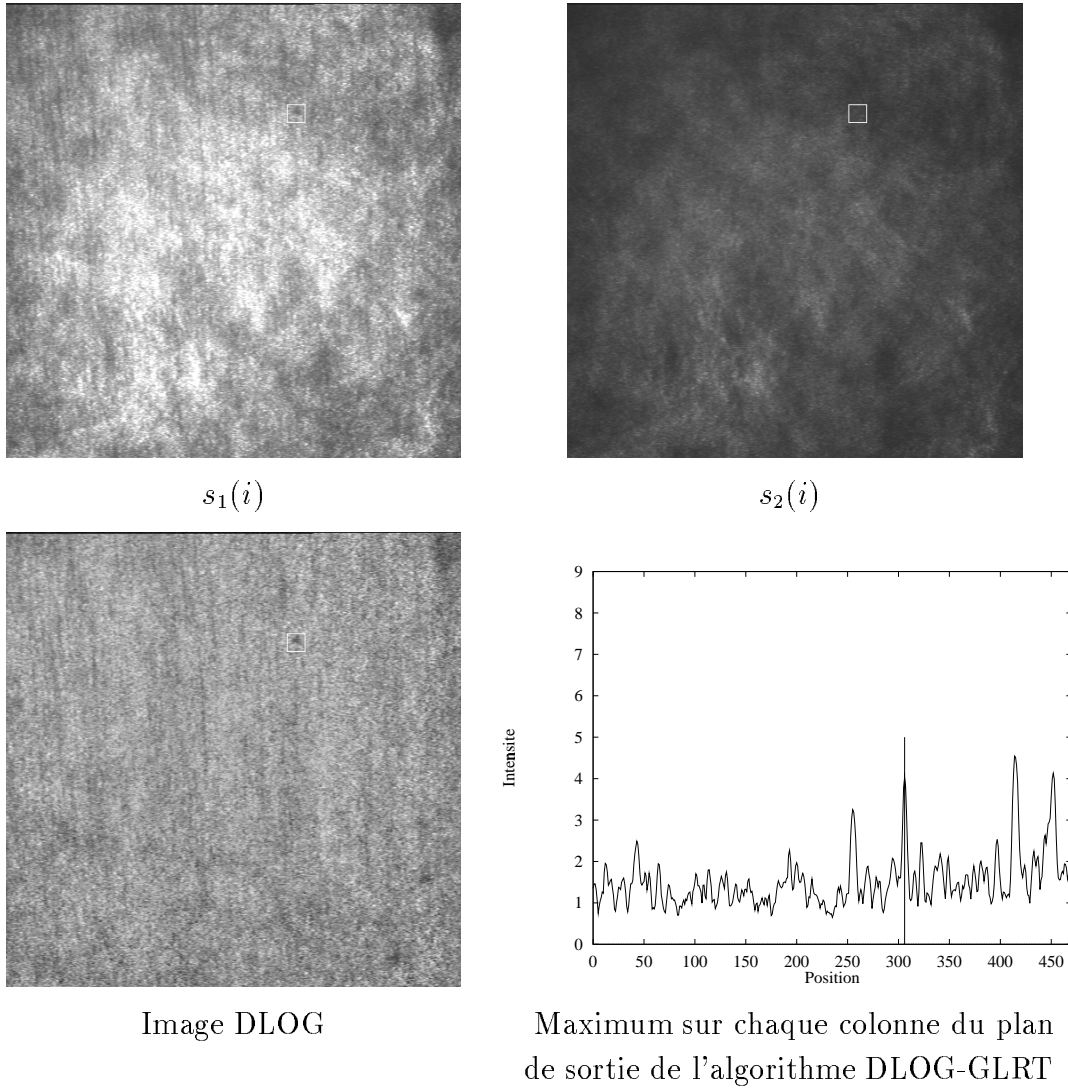


FIG. 4.9: Exemple de résultat obtenu sur des images polarimétriques en recherche de régions d'intérêt avec l'algorithme DLOG-GLRT estimant la taille de la cible. La cible est soulignée dans chaque image par un carré blanc. Il s'agit d'un morceau de papier sur un fond de bois.

des images est relativement bien connue. Nous avons montré dans le chapitre 2 le problème induit par les inhomogénéités du fond. Ce problème se pose de façon plus pressante dans le cas d'images texturées comme celles que nous avons présentées.

L'imageur actif polarimétrique fournit une image bi-canal (IBC), fortement perturbée par des inhomogénéités de la figure d'éclairage. Ces inhomogénéités sont telles que les algorithmes IBC-GLRT que nous avons élaborés présentent de piètres performances. Nous avons alors proposé de recourir à des méthodes sous-optimales consistant à considérer la réduction des inhomogénéités comme une priorité. Ceci peut être obtenu en combinant les données bicanal de l'imageur en une image monocanal, dite ICEO, dont les statistiques peuvent être modélisées par des gaussiennes. La perte d'information qui en découle est en effet négligeable devant le gain réalisé en conservant des images à peu près homogènes.

Ceci définit un algorithme ICEO-GLRT, dont les performances sont déjà plus acceptables.

L'estimation des paramètres de nuisance étant également un problème, nous avons utilisé une transformation récemment proposée [17] de l'image ICEO en une image DLOG, possédant les mêmes propriétés face aux inhomogénéités. Dans les images DLOG, la distribution des niveaux de gris est également gaussienne, mais de variance fixe. L'estimation des moyennes étant plus précise qu'une estimation de variance, pour un même nombre d'échantillons, ceci a permis de définir un algorithme plus efficace, le DLOG-GLRT.

De plus, nous avons montré que sur de telles images, l'utilisation d'un algorithme ne supposant pas connue la taille de la cible permettait de définir des régions d'intérêts. Il semble difficile de procéder à une détection plus discriminante en l'absence de connaissance a priori sur les cibles. En effet, dans tout ce qui a été présenté jusque là, nous avons supposé inconnus les niveaux de gris de la cible. Ceci nous a conduit à utiliser un modèle de cible homogène. Dans le chapitre suivant, nous nous intéresserons à une généralisation des algorithmes GLRT, permettant de dépasser ce modèle, définissant l'objet comme une composition de régions homogènes mais dont les paramètres peuvent varier.

Chapitre 5

Filtre GLRT multi régions

Sommaire

5.1	Modèle d'image pour cibles multirégions	70
5.2	Détection optimale de cibles multi-régions	72
5.2.1	Rapport de vraisemblance pour des objets multi-régions	73
5.2.2	Illustration des performances sur des images synthétiques . . .	73
5.2.3	Illustration sur des images réelles	75
5.3	Limite du nombre de régions utilisables	78
5.4	Estimation du nombre de régions de la cible	82
5.4.1	Pénalisation du nombre de régions	84
5.4.2	Choix du paramètre de pénalisation	86
5.4.3	Complexité de cet algorithme	88
5.5	Estimation simplifiée du nombre de régions	89
5.6	Performances de ces algorithmes	90
5.7	Conclusion	94

Dans tout ce que nous avons présenté auparavant, nous avons toujours supposé que les niveaux de gris de la cibles étaient inconnus. Ceci conduit à effectuer la localisation ou la détection des objets seulement à l'aide de leur forme extérieure. Cette hypothèse nous a paru limitative dans de nombreux cas. Avant d'exposer l'approche que nous proposons, nous devons considérer divers travaux existants lorsque ces niveaux de gris sont connus, ou inconnus.

Le premier filtre reconnu par toute la communauté du traitement d'image a été le filtre adapté [58]. Ce filtre est optimal au sens du rapport signal sur bruit en sortie du filtre, lorsque le bruit est additif et lorsque les niveaux de gris de la cible sont connus. Il a ensuite été démontré que l'hypothèse de bruit additif était souvent limitative [30] et des méthodes heuristiques ont été introduites pour traiter le cas de bruits non recouvrant [31].

Dans le cas de bruit non recouvrant, on a ensuite vu émerger des techniques basées sur la théorie statistique de l'estimation, pour définir des filtres optimisant la probabilité de localisation correcte. Nous avons présenté de telles techniques dans les chapitres

précédents. Ces approches nécessitent un modèle d'image simple, et ont principalement été appliquées pour des objets dont les niveaux de gris sont supposés homogènes [18]. La connaissance des niveaux de gris de la cible à un facteur d'illumination près a également été considérée [22][29].

En l'état actuel des choses, on pourra donc diviser les approches statistiques construites sur un modèle de bruit non recouvrant en deux catégories :

- Celles qui supposent les niveaux de gris de la cibles inconnus, et qui recourent à une modélisation de ces niveaux de gris sous la forme d'un champ aléatoire homogène.
- Celles qui supposent la structure des niveaux de gris de la cible connus, à quelques paramètres prêts.

De fait, la connaissance des niveaux de gris de la cible prend deux formes : connaissance de l'intensité de chaque pixel de la cible, et connaissance des liens existants entre deux pixels de la cible. Il existe donc un écart important entre cette connaissance, et la méconnaissance totale des niveaux de gris de la cible. Nous nous proposons ici de combler cet écart en modélisant la cible sous la forme d'un assemblage de régions homogènes, dont les niveaux de gris suivent des lois de densités de probabilité dont les paramètres sont inconnus. Cette approche permet de conserver le caractère d'optimalité théorique des algorithmes présentés auparavant sans pour autant nécessiter une trop grande simplification des modèles des cibles à détecter, localiser ou reconnaître.

Nous comptons donc développer ici un algorithme adapté à des images dans lesquelles apparaît un objet dont les caractéristiques spatiales sont connues, mais dont l'intensité d'une région peut fluctuer fortement, par rapport au fond environnant, ou par rapport aux autres régions composant l'objet.

Notons que ce problème est habituellement traité en considérant des points caractéristiques de l'objet, le plus souvent des frontières entre ses différentes régions. Une extraction de bord permet d'atténuer l'influence du changement de moyennes éventuel entre les régions. Dans le cas des images que nous présenterons, une telle approche est inemployable. Le niveau de bruit est en effet tel que la notion de contour est alors très mal définie.

L'algorithme multi régions que nous présentons utilise une description plus précise de la cible qu'un algorithme construit pour une cible homogène, donc plus efficace. Cependant, en estimant les statistiques sur des régions plus petites, on peut s'attendre à une augmentation des fluctuations en sortie du filtre. En particulier, nous montrerons que cette augmentation se produit lorsqu'aucune cible n'est présente, ce qui augmente la PFA du système, pour un seuil de détection donné. Nous livrerons une étude théorique des statistiques en sortie de ce type d'algorithmes, qui permettra de définir les pertes éventuelles liées à son utilisation.

Cette étude met en avant l'influence du nombre de régions sur ces pertes. Nous proposerons donc ensuite deux algorithmes permettant de choisir le nombre de régions le plus adapté pour décrire la cible afin de limiter ces pertes.

5.1 Modèle d'image pour cibles multirégions

Supposons que notre cible soit découpée en L régions indépendantes. Nous conservons ici le modèle d'image SIR présenté dans la section 1.1 qui sous-tend la construction de

tous nos algorithmes. La principale différence dans le cas de cibles multi-régions est que le champ aléatoire \mathbf{a} est maintenant découpé en L champs aléatoires indépendants $\mathbf{a}_k, k \in 1..L$, de densités de probabilité $P_{\theta_a^k}$. θ_a^k est l'ensemble des paramètres définissant la densité de probabilité du champ \mathbf{a}_k .

Si le fond est supposé homogène, il est représenté par un champ aléatoire \mathbf{b} de densité de probabilité P_{θ_b} . La figure 5.1 représente ce modèle d'image.

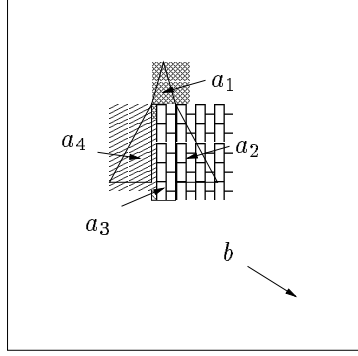


FIG. 5.1: Modèle d'image pour une cible composée de 4 régions. Les niveaux de gris de chaque région de la cible sont représentés par un champ aléatoire $\mathbf{a}_k, k \in 1..4$. Les niveaux de gris du fond sont représentés par un champ aléatoire \mathbf{b} .

La référence de l'objet à localiser ou à détecter est maintenant un ensemble de supports \mathbf{w}_k , prenant la valeur 1 à l'intérieur de la région i et 0 en dehors. L'union de ces supports constitue la silhouette \mathbf{w} de l'objet complet, comme indiqué dans la figure 5.2.

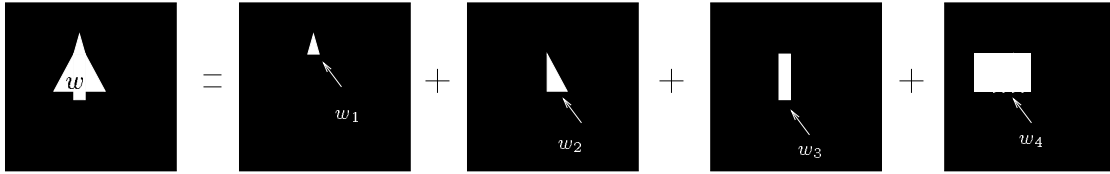


FIG. 5.2: Masques définissant la cible dans le modèle multi-régions

Pour faciliter la visualisation des différents masques correspondant à nos cibles, par la suite, nous les représenterons dans une seule image, avec des niveaux de gris différents. Insistons bien sur le fait que ces niveaux de gris ne sont pas représentatifs des paramètres des champs aléatoires représentant chaque région. Un exemple de cette visualisation est présenté dans la figure 5.3.

Une scène \mathbf{s} correspondant à ce modèle peut être représenté par l'équation 5.1 :

$$s(i) = \sum_{k=1}^L a^k(i)w_k(i) + b(i)(1 - w(i)) \quad (5.1)$$

Nous représentons dans la figure 5.4 un exemple très simple de ce type de scènes, dans lequel les niveaux de gris de chaque région sont des réalisations de variables aléatoires indépendantes distribuées suivant une PDF gaussienne.

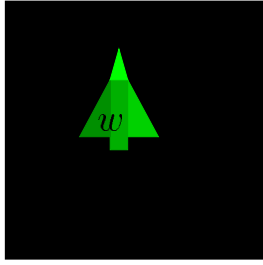


FIG. 5.3: Visualisation simplifiée des masques définissant la cible dans le modèle multi-régions

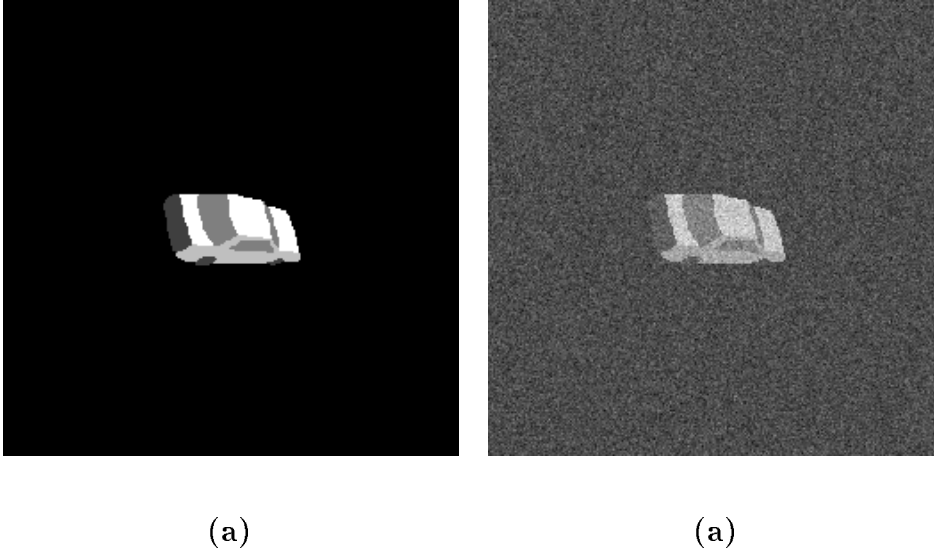


FIG. 5.4: Illustration du modèle d'image de l'équation 5.1. **(a)** : Masque de la cible. **(b)** : Scène s , composée de bruits gaussiens décorrélés.

5.2 Détection optimale de cibles multi-régions

Dans le but de localiser la cible, le filtre optimal serait construit sur la base d'une estimation de position au sens du Maximum de Vraisemblance, comme cela a été présenté au chapitre 1.2. Nous avons cependant vu au chapitre 2 que les performances en localisation des filtres de type GLRT sont équivalentes à celles des filtres optimaux au sens du Maximum de Vraisemblance lorsque le bruit est homogène, et très supérieures dans le cas de bruit non homogène. Il suffit pour cela d'utiliser une fenêtre d'analyse de taille adaptée.

De plus, en détection, les filtres GLRT optimisent la probabilité de détection pour une probabilité de Fausse Alarme donnée, alors que le filtre obtenu par le Maximum de Vraisemblance ne possède aucune propriété spécifique pour ces tâches.

Dès lors, nous nous orientons naturellement vers la définition d'un filtre de type GLRT, adapté à des scènes contenant des cibles composées de plusieurs régions. Notons qu'un algorithme similaire, de type ML a été utilisé dans [20], dans le cas d'images prétraitées. Dans ce cas, l'image prétraitée était supposée contenir simplement trois zones, la cible,

le fond et une région frontière issue du prétraitement. La taille de cette dernière zone pouvait varier d'une image à l'autre, ne faisant pas partie du modèle de la cible.

5.2.1 Rapport de vraisemblance pour des objets multi-régions

Comme au chapitre 1, pour chaque position de la fenêtre d'analyse \mathbf{F} sur la scène, nous cherchons à discriminer deux hypothèses :

$$\left\{ \begin{array}{ll} \text{Hypothèse } \mathcal{H}_1 & \text{La cible, composée de } L \text{ régions est présente au centre de la fenêtre } F \\ \text{Hypothèse } \mathcal{H}_0 & \text{La fenêtre } F \text{ ne contient que du bruit} \end{array} \right. \quad (5.2)$$

Le filtre optimal pour ce type de modèle sera noté $GLRT^{(L)}$. Le calcul de ce type de filtre se déduit aisément en généralisant ce qui a été fait précédemment, et nous nous contenterons de donner ici l'expression du filtre $GLRT^{(L)}$ dans le cas de bruits gamma d'ordre 1 décorrélés, lorsque les moyennes θ de chaque région (cible et fond) sont connues ou inconnues.

Le développement des calculs permet d'obtenir, lorsque les paramètres sont connus :

$$\begin{aligned} \mathcal{R}_{connu}^{(L)} = & - \sum_{k=1}^L N^k \log \theta_{\mathbf{a}^k} - N_b \log \theta_{\mathbf{b}} + N_F \log \theta_{\mathbf{F}} \\ & - \sum_{k=0}^{L-1} \left[\frac{1}{\theta_{\mathbf{a}^k}} \sum_{i \in \mathbf{w}^k} s_i \right] - \frac{1}{\theta_{\mathbf{b}}} \sum_{i \in \mathbf{b}} s_i + \frac{1}{\theta_{\mathbf{F}}} \sum_{i \in \mathbf{F}} s_i \end{aligned} \quad (5.3)$$

Où N^k , N_b , and N_F sont les nombres de pixels de chaque région de la cible, du complémentaire de la cible dans la fenêtre d'analyse, et de la fenêtre d'analyse.

Comme dans les chapitres précédents, lorsque ces paramètres sont inconnus, nous les estimons au sens du maximum de vraisemblance. En remplaçant les paramètres par leurs estimations, on obtient :

$$\mathcal{R}_{inc}^{(L)} = - \sum_{k=1}^L N^k \log \left[\frac{1}{N^k} \sum_{i \in \mathbf{w}^k} s_i \right] - N_b \log \left(\frac{1}{N_b} \sum_{i \in \mathbf{b}} s_i \right) + N_F \log \left(\frac{1}{N_F} \sum_{i \in \mathbf{F}} s_i \right) \quad (5.4)$$

Les filtres obtenus dans les équations 5.3 et 5.4 permettent la détection d'un objet composé d'un nombre L de régions, lorsque les moyennes de chacune des régions sont respectivement connues, ou inconnues.

Le filtre $\mathcal{R}_{connu}^{(L)}$ sera représentatif d'un filtre supposant la connaissance des niveaux de gris de la cible, de type filtre adapté.

5.2.2 Illustration des performances sur des images synthétiques

La scène de la figure 5.5.a contient trois versions du même objet, mais dans lesquelles les moyennes de chaque région changent. Cet objet contient 2568 pixels au total, et est décomposé en 4 régions de tailles similaires. Pour la première version de cet objet (à

gauche), les moyennes de chaque région sont identiques et fixées à 35. Pour les deux autres versions de l'objet (au milieu et à droite), les moyennes de chaque région de l'objet sont respectivement 5, 10, 30 et 35, de la plus sombre à la plus claire. La moyenne du fond vaut 20. On pourra noter que pour les deux dernières versions de l'objet, le contraste entre la cible totale et le fond est faible. La figure 5.5.b présente cette même scène, perturbée par du bruit exponentiel (multiplicatif).

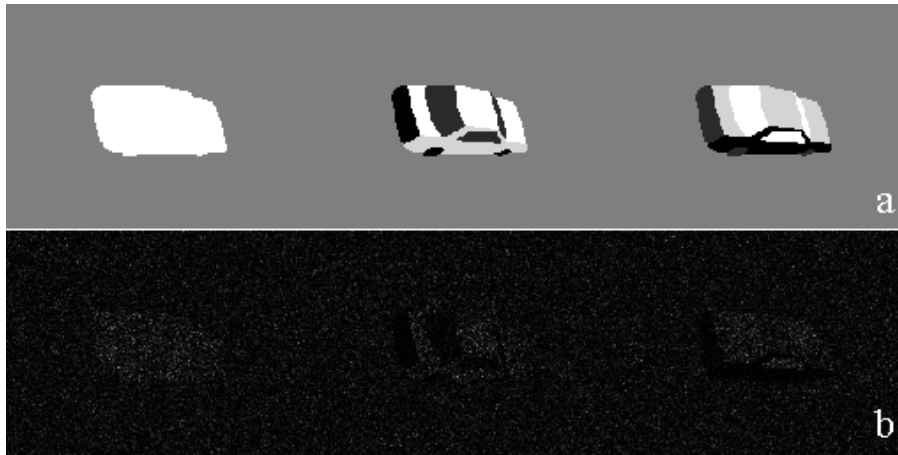


FIG. 5.5: Scène synthétique : **a** : image sans bruit, **b** : Image perturbée par du bruit exponentiel.

Nous testons sur cette la scène (figure 5.5.b) trois versions de l'algorithme GLRT :

- $GLRT^{(1)}$ conçu pour une cible 1 région, paramètres inconnus.
- $GLRT^{(4)c}$ conçu pour une cible 4 régions, paramètres connus.
- $GLRT^{(4)}$ conçu pour une cible 4 régions, paramètres inconnus.

L'algorithme supposant les paramètres connus a été conçu pour la cible placée au centre de la figure 5.5.a.

La figure 5.6 illustre le comportement de ces trois algorithmes, sous la forme du maximum sur chaque colonne de leur plan de sortie. Nous pouvons classer leurs performances en comparant l'information a priori que avons considérée pour chacun de ces filtres, et l'information réellement présente dans les scènes.

L'algorithme GLRT 1 région présente de meilleures performances que les autres pour l'objet composé d'une seule région, au prix d'un échec lorsque l'objet est composé de plusieurs régions (figure 5.6.a). On peut remarquer qu'il a tendance à dédoubler le pic pour donner une réponse forte sur les régions de la cible les plus contrastées par rapport au fond, "biaisant" ainsi la localisation.

L'algorithme GLRT 4 régions à paramètres connus est efficace lorsque les caractéristiques de l'objet à détecter coïncident avec la connaissance a priori pour laquelle il a été conçu. Il échoue lorsque l'objet ne présente pas les caractéristiques correctes, soit en nombre de régions, soit dans la répartition des moyennes de ces régions (figure 5.6.b).

L'algorithme GLRT 4 régions à paramètres inconnus, permet de détecter correctement chaque version de l'objet, qu'il soit composé de 4 régions ou d'une seule région. Dans le cas d'un objet composé d'une seule région, il se comporte moins bien que le GLRT 1 région, lequel dispose de plus d'information a priori.

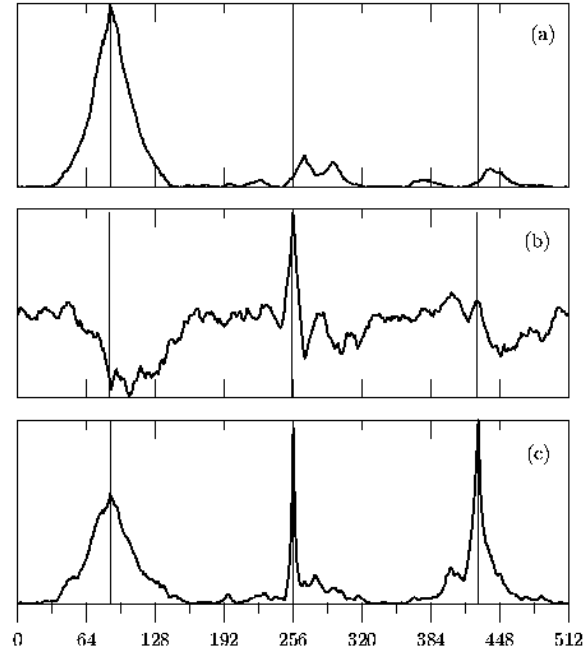


FIG. 5.6: Maximum suivant chaque colonne du plan de sortie de trois algorithmes appliqués à la scène 5.5.b :

(a) : $GLRT^{(1)}$, paramètres inconnus; (b) : $GLRT_c^{(4)}$, paramètres connus; (c) : $GLRT^{(4)}$, paramètres inconnus. Les traits verticaux indiquent la position de la cible.

Nous pouvons conclure en disant que la connaissance a priori des niveaux de gris de la cible est une connaissance trop forte dans beaucoup de cas. Par ailleurs, la simplification du modèle de cible en une seule région uniforme peut également poser problème lorsque la cible présente une structure interne très marquée. Entre ces deux extrêmes, un algorithme ne supposant connu que la répartition spatiale des niveaux de gris permet bien souvent d'obtenir des résultats très satisfaisants.

L'algorithme $GLRT^{(L)}$ est donc une solution simple et efficace aux problèmes de localisation ou de détection de cibles contrastées, mais dont les niveaux de gris peuvent fluctuer. Nous montrons dans la section suivante un exemple de situation réelle pour laquelle ce problème peut se poser.

5.2.3 Illustration sur des images réelles

Pour démontrer le vaste champ d'application de ces algorithmes, nous présentons ici les résultats d'une seconde expérience, de localisation, effectuée sur des images réelles.

L'image présentée dans la figure 5.7.a est un ensemble de prises de vue d'un objet 3D par une caméra CCD, dans différentes conditions d'illuminations. Cet objet présente 5 facettes, qui constituent les différentes régions de l'objet (figure 5.7.b).

Dans les conditions d'acquisition, la lumière illuminant l'objet était proche du niveau de détection minimal de la caméra CCD. On est alors en présence d'un faible flux de

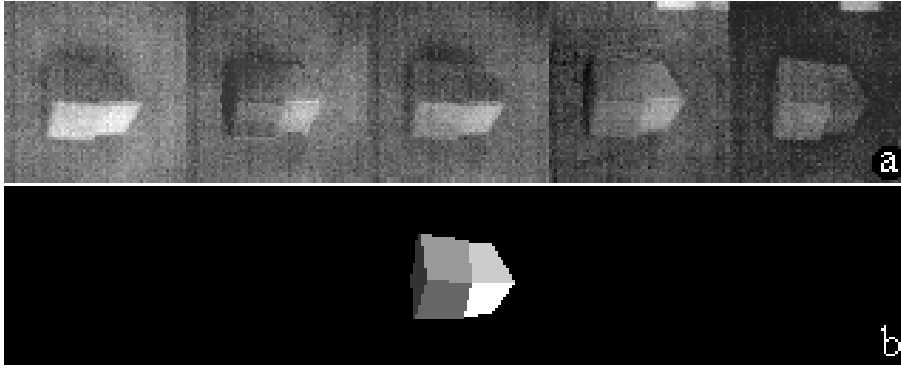


FIG. 5.7: exemple de scène réelle un même objet placé dans des conditions d'illumination différentes. Chaque facette de l'objet constitue ainsi une région de l'objet. **(a)** : Images prises par une caméra CCD à faible flux lumineux, pour un même objet 3D sous 5 illuminations différentes. **(b)** : Description des différentes régions de l'objet

lumière. La distribution des niveaux de gris dans ce cas est correctement modélisée par un bruit de Poisson, dont les paramètres se réduisent à sa moyenne.

Le filtre GLRT pour du bruit de poisson lorsque les moyennes de chaque région m_a^k (pour la cible) et m_b (pour le fond) sont inconnues est donné par l'équation 5.5 (cf. annexe B.2).

$$r'_{un} = \sum_{k=0}^{L-1} N^k \hat{m}_{\mathbf{w}^k} \log(\hat{m}_{\mathbf{w}^k}) + N_b \hat{m}_b \log(\hat{m}_b) - N_F \hat{m}_F \log(\hat{m}_F) \quad (5.5)$$

où \hat{m}_a^k est la moyenne empirique estimée sur la région k de la cible, \hat{m}_b est la moyenne du fond estimée sur le complémentaire de la cible dans la fenêtre et \hat{m}_F est la moyenne du fond estimée sur la totalité de la fenêtre d'analyse.

La figure 5.8.a présente les plans de sortie obtenus par le filtre GLRT 1 région. Les deux derniers objets, dont les niveaux de gris sont relativement homogènes, peuvent être correctement localisés et reconnus par ce filtre. En revanche, le pic est beaucoup moins marqué pour les trois autres objets qui présentent de plus fortes variations internes et un contraste plus faible avec le fond. Encore une fois, on peut remarquer que les pics sont dédoublés et déplacés par rapport à la vraie position de l'objet, pour se positionner sur les zones les plus homogènes de l'objet.

Les plans de sortie de l'algorithme GLRT 5 régions (figure 5.8.b) montrent que celui-ci permet de localiser correctement chaque version de l'objet. Il se comporte donc de façon plus robuste aux changements d'illumination de la scène.

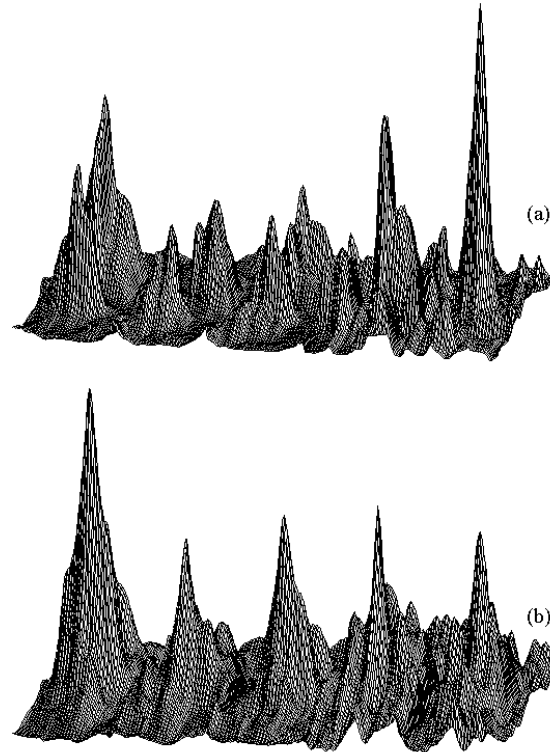


FIG. 5.8: Résultats obtenus sur les images réelles de la figure 5.7.a : **(a)** : Plan de sortie de l'algorithme GLRT 1 région. **(b)** : Plan de sortie de l'algorithme GLRT 5 région. Ces filtres sont conçus pour du bruit de poisson (cf. équation 5.5).

5.3 Limite du nombre de régions utilisables

Nous avons vu que l'augmentation du nombre de régions utilisées pour décrire les cibles permet souvent d'accroître les performances des algorithmes. Cependant, l'augmentation du nombre de régions conduit à une augmentation des fluctuations du rapport de vraisemblance en l'absence de cible. En effet, si on doit estimer les paramètres statistiques sur un grand nombre de petites régions, la variance de l'estimation est importante. En conséquence, ceci accroît la Probabilité de Fausse Alarme. Si la Probabilité de Détection n'augmente pas en conséquence, les performances des algorithmes multi-régions chutent. En d'autres termes, utiliser une description de la cible avec L régions alors que celle-ci est homogène implique des pertes. Ces pertes croissent avec le nombre de régions non nécessaires.

Pour s'en convaincre, nous avons estimé les courbes COR obtenues par des algorithmes utilisant un nombre L croissant de régions, sur des scènes composées d'une cible homogène. La cible est un carré de $S = 400$ pixels de surface, décrite par S/L régions de même tailles. Les scènes sont perturbées par du bruit exponentiel et le contraste a été fixé à 1.5.

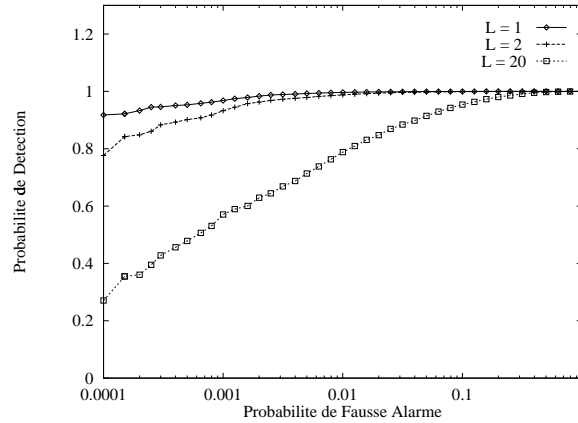


FIG. 5.9: Pertes de performances lors de l'utilisation de régions non nécessaires : Courbes COR obtenues pour des algorithmes GLRT utilisant L régions pour décrire une cible homogène. La cible est un carré de 400 pixels de surface, divisé en L régions identiques. La fenêtre d'analyse contient 800 pixels. Le contraste est fixé à 1.5. Les courbes COR sont obtenues pour 50000 réalisations.

Quantification des pertes

Pour étudier les pertes induites par l'introduction de régions trop nombreuses, nous nous intéressons tout d'abord à la modification subie par le rapport de vraisemblance lorsque l'on divise une région de la référence en plusieurs parties. Nous montrons ensuite que cette division, lorsqu'elle n'est pas nécessaire pour décrire la cible, élargit la distribution statistique des valeurs en sortie du filtre, à la fois sur la cible et sur le fond. Discriminer cible et fond devient donc plus difficile. Nous calculons ensuite une métrique simple, le rapport de Fisher, permettant de quantifier ces pertes.

Nombre de régions et rapport de vraisemblance

Pour commencer, nous pouvons donner l'équation du filtre GLRT défini pour un objet homogène composé de N_I pixels dans le cas de bruit exponentiel :

$$\mathcal{R}_{inc}^{(1)} = -N_I \log \left[\frac{1}{N_I} \sum_{i \in \mathbf{w}} s_i \right] - N_b \log \left[\frac{1}{N_b} \sum_{i \in \mathbf{b}} s_i \right] + N_F \log \left[\frac{1}{N_F} \sum_{i \in \mathbf{F}} s_i \right] \quad (5.6)$$

L'équation de ce filtre pour le même objet, décomposé en deux régions de N_I^1 et N_I^2 pixels est :

$$\begin{aligned} \mathcal{R}_{inc}^{(2)} = & -N_I^1 \log \left[\frac{1}{N_I^1} \sum_{i \in \mathbf{w}^1} s_i \right] - N_I^2 \log \left[\frac{1}{N_I^2} \sum_{i \in \mathbf{w}^2} s_i \right] \\ & - N_b \log \left[\frac{1}{N_b} \sum_{i \in \mathbf{b}} s_i \right] + N_F \log \left[\frac{1}{N_F} \sum_{i \in \mathbf{F}} s_i \right] \end{aligned} \quad (5.7)$$

La différence entre ces deux filtres est alors :

$$\delta = \mathcal{R}_{inc}^{(2)} - \mathcal{R}_{inc}^{(1)} = -N_I^1 \log \left[\frac{1}{N_I^1} \sum_{i \in \mathbf{w}^1} s_i \right] - N_I^2 \log \left[\frac{1}{N_I^2} \sum_{i \in \mathbf{w}^2} s_i \right] + N_I \log \left[\frac{1}{N_I} \sum_{i \in \mathbf{w}} s_i \right] \quad (5.8)$$

Dans cette équation, nous reconnaissons l'expression d'un filtre GLRT conçu pour détecter un objet de support \mathbf{w}^1 avec une fenêtre $\mathbf{F} = \mathbf{w}^1 + \mathbf{w}^2 = \mathbf{w}$.

Nous avons montré dans la section 3.5 que sur des zones homogènes, la distribution de l'intensité en sortie de ces filtres est indépendante de N_I^1 et N_I^2 , car elle tend très rapidement vers une loi limite. Plus précisément, la distribution de 2δ suit une loi du χ^2 d'ordre 1. La moyenne de δ vaut 0.5 et sa variance 0.5.

Diviser sans nécessité une région en deux parties revient à ajouter au logarithme du rapport de vraisemblance une variable aléatoire de moyenne 0.5 et de variance 0.5. La distribution du logarithme du rapport de vraisemblance subit donc une translation et un élargissement. Ce phénomène se produisant dans l'image à la fois sur la cible et sur le fond, discriminer cible et fond devient plus difficile, comme nous l'avons représenté schématiquement sur la figure 5.10.

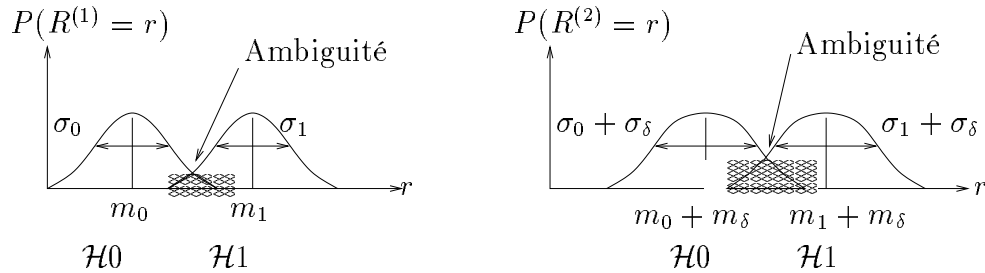


FIG. 5.10: Influence de l'ajout d'une variable aléatoire sur la capacité de discrimination.

Ceci se généralise à la division de la référence en L régions non nécessaires. On peut considérer l'apparition de chaque nouvelle région comme le résultat de la division d'une

région plus grande en deux parties. Il faudra $L - 1$ divisions pour obtenir la description de l'objet en L régions. D'après ce qui précède, chaque division d'une région homogène en deux parties produit l'ajout d'une variable aléatoire $\delta^{(k)}$ distribuée de la même façon que δ . Nous illustrons cela dans la figure 5.11.

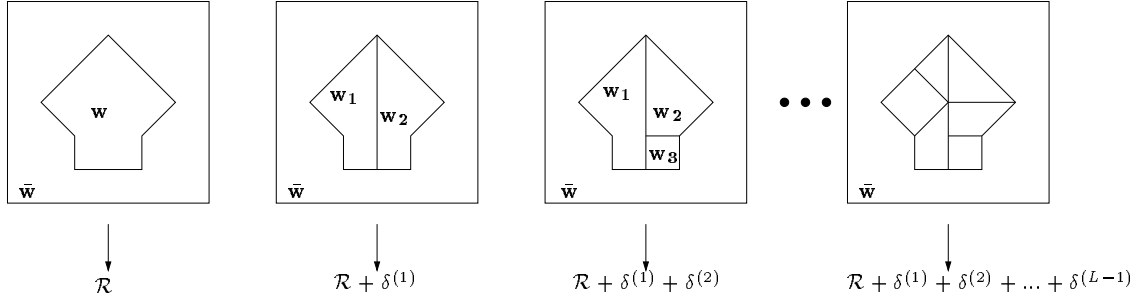


FIG. 5.11: Influence de la division d'une région en plusieurs parties sur le logarithme du rapport de vraisemblance. Les variables aléatoires $\delta^{(k)}$ sont identiquement distribuées et telles que $2\delta^{(k)}$ suit une loi du χ^2 d'ordre 1.

Notons $\zeta^{(L-1)} = \sum_{k=1}^{L-1} \delta^{(k)}$. En faisant l'hypothèse que les variables $\delta^{(k)}$ sont indépendantes, il apparaît que $2\zeta^{(L-1)}$ suit une loi du χ^2 d'ordre $L - 1$ (cf. équation 3.17). On a alors

$$\mathcal{R}_{inc}^{(L)} - \mathcal{R}_{inc}^{(1)} = \zeta^{(L-1)} \quad (5.9)$$

avec $\zeta^{(L-1)}$ une variable aléatoire de densité de probabilité connue.

Pour illustrer cette propriété, nous avons appliqué les filtres $GLRT^{(4)}$ et $GLRT^{(1)}$ sur des échantillons homogènes, et calculé la différence $\zeta^{(3)}$ entre leurs réponses respectives. Nous présentons (figure 5.12) l'histogramme de $X = 2\zeta^{(3)} = 2(\mathcal{R}_{inc}^{(4)} - \mathcal{R}_{inc}^{(1)})$, comparé à la densité de probabilité d'une loi du χ^2 d'ordre 3. Ces expériences ont été menées dans le cas où l'objet total contient $4N$ pixels et où chaque région a une taille $N = 10$ ou 20 pixels.

Ceci nous permet de déduire le comportement statistique de $\mathcal{R}_{inc}^{(L)}$ en fonction de celui du $\mathcal{R}_{inc}^{(1)}$, lorsque la cible est homogène. En particulier, si l'on dispose de réalisations de $\mathcal{R}_{inc}^{(1)}$, on pourra simuler les réalisations de $\mathcal{R}_{inc}^{(L)}$ grâce à la relation donnée par l'équation 5.9. Ces réalisations simulées de $\mathcal{R}_{inc}^{(L)}$ permettent d'estimer la courbe COR obtenue en appliquant le filtre $GLRT^{(L)}$ sur des scènes où la cible éventuellement présente est homogène.

Nous illustrons cela dans la figure 5.13 en comparant les courbes COR expérimentales obtenues par les algorithmes GLRT prévus pour 2, 10 et 20 régions et les courbes COR simulées par la méthode précédente, à partir de la sortie de l'algorithme GLRT 1 région. La cible est homogène et contient 300 pixels, le contraste vaut 1.3.

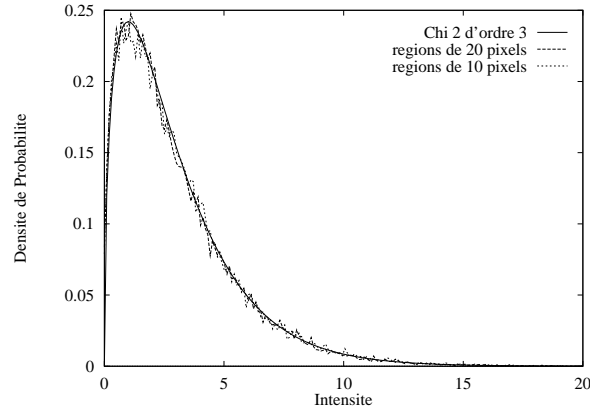


FIG. 5.12: Densité de probabilité la différence du log du rapport de vraisemblance entre le $GLRT^{(1)}$ et le $GLRT^{(4)}$ lors de l'introduction de 3 régions non-nécessaires : histogramme expérimental de $2\zeta^{(3)} = 2(\mathcal{R}_{inc}^{(4)} - \mathcal{R}_{inc}^{(1)})$ sur une région homogène.

..... : Chaque région à une surface de 20 pixels

— — — : Chaque région à une surface de 10 pixels

———— : DDP d'une loi du χ^2 d'ordre 3.

Les histogrammes sont estimés sur 10000 réalisations.

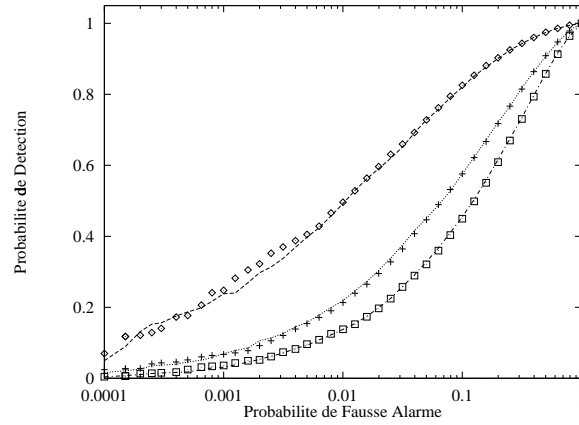


FIG. 5.13: Simulation des pertes dues à l'introduction de régions non nécessaires :

◇ : Courbe COR expérimentale, $L = 2$

⊕ : Courbe COR expérimentale, $L = 10$

□ : Courbe COR expérimentale, $L = 20$

Les lignes pleines représentent les courbes COR obtenues par notre simulation.

La cible contient 300 pixels au total, le contraste vaut 1.3. La fenêtre d'analyse contient 600 pixels. Chaque courbe est obtenue pour 50000 réalisations.

Rapport de Fisher

Nous avons précédemment expliqué la raison pour laquelle l'utilisation de trop nombreuses régions entraînait une baisse des courbes COR.

Il peut être utile de disposer d'une grandeur numérique permettant de caractériser ces pertes. Un moyen simple, bien qu'incomplet d'exprimer les pertes en détection est d'observer le rapport de Fisher en sortie du filtre. Soit m_1, σ_1^2 , la moyenne et la variance de l'intensité en sortie du filtre lorsqu'une cible est présente. Soit m_2, σ_2^2 , la moyenne et la variance de l'intensité en sortie en l'absence de filtre. Le rapport de Fisher, qui donne une indication de la qualité de la discrimination entre la cible et le fond vaut :

$$F = (m_2 - m_1)^2 / (\sigma_1^2 + \sigma_2^2)$$

Supposons maintenant que nous divisons une région de la référence en L parties, sans nécessité. Le logarithme du rapport de vraisemblance à l'emplacement de la cible et sur le fond se voit ajouter une variable aléatoire X telle que $2X$ suit une loi du χ^2 d'ordre $L - 1$. La moyenne et la variance de X sont respectivement $(L - 1)/2$ et $0.25L$.

En conséquence, m_1 et m_2 se trouvent translatés de la même quantité $0.5(L - 1)$, correspondant à la moitié de la moyenne d'un χ^2 d'ordre $(L-1)$. De la même façon, chaque variance se voit ajouter la valeur $0.25L$.

Le rapport de Fisher devient donc :

$$F' = (m_2 - m_1)^2 / (\sigma_1^2 + \sigma_2^2 + 0.5(L))$$

Une augmentation trop importante du nombre de régions induit donc des pertes caractérisables qualitativement grâce à l'équation précédente. Pour pallier à ce problème, nous proposons maintenant une approche permettant d'estimer le nombre de régions utilisé par le filtre.

5.4 Estimation du nombre de régions de la cible

Dans la première partie de ce chapitre, nous avons introduit l'algorithme GLRT multi-région dans le but de traiter le cas d'images dans lesquelles l'intensité des différentes régions de la cible peut fluctuer fortement d'une image à l'autre. Par exemple, en imagerie infrarouge, un même véhicule apparaîtra de façon très différente si son moteur est allumé ou éteint, définissant ainsi des régions utiles ou non pour la description de ce véhicule. Le même phénomène se produit pour un objet 3D placé dans des conditions d'illumination différentes.

Nous avons ensuite montré que l'augmentation outrancière du nombre de régions utilisé pour décrire l'objet cherché pouvait faire décroître les performances de l'algorithme. En effet, l'augmentation du nombre de régions entraîne un recouvrement plus grand des distributions du rapport de vraisemblance, en présence ou en l'absence de cible.

Il peut alors être judicieux de mettre au point un algorithme permettant d'estimer le nombre de régions de la cible. Ceci devrait permettre de tirer pleinement parti d'une description multi-région des objets, tout en limitant les problèmes liés à l'introduction de régions inutiles.

Ce type de problème est classique en segmentation. Dans ce cas, sans connaissance a priori du contenu de la scène, il n'est pas envisageable d'explorer l'ensemble des formes possibles d'objet. On recourt alors à des techniques de type fusion/division, permettant de limiter l'ensemble des solutions possibles. Ce type d'approche nécessite la définition d'une image présegmentée, sur laquelle on effectue l'étape de fusion/division. On choisira de fusionner deux régions ou de diviser une région en deux parties en fonction d'un critère de similarité, dont les plus simples sont le critère de Fisher [57], le test du χ^2 [2], ou un critère de type *Description de Longueur Minimale*, noté MDL (pour *minimum description length* en anglais) [62].

Dans cette section, nous nous intéressons à la construction d'un algorithme permettant de réaliser cette sélection d'un modèle décrivant la cible cherchée, pour des applications de localisation/détection. Ici, contrairement au cas de la segmentation, nous supposons connu l'ensemble des descriptions possibles de la cible.

En ne faisant aucune hypothèse sur la probabilité qu'une cible présente ait un nombre de régions donné, la détection et l'estimation du nombre de régions sont alors deux tâches différentes et indépendantes. L'algorithme doit donc estimer en chaque point de l'image le nombre P de régions d'une cible éventuellement présente. Sur la base de cette estimation, la détection doit discriminer l'hypothèse de présence d'une cible composée de P régions et l'hypothèse d'absence de cible. En d'autres termes, nous devons effectuer une sélection de l'ordre du modèle de cible choisi, avant d'effectuer la détection de la cible. Ce processus est résumé dans la figure 5.14.

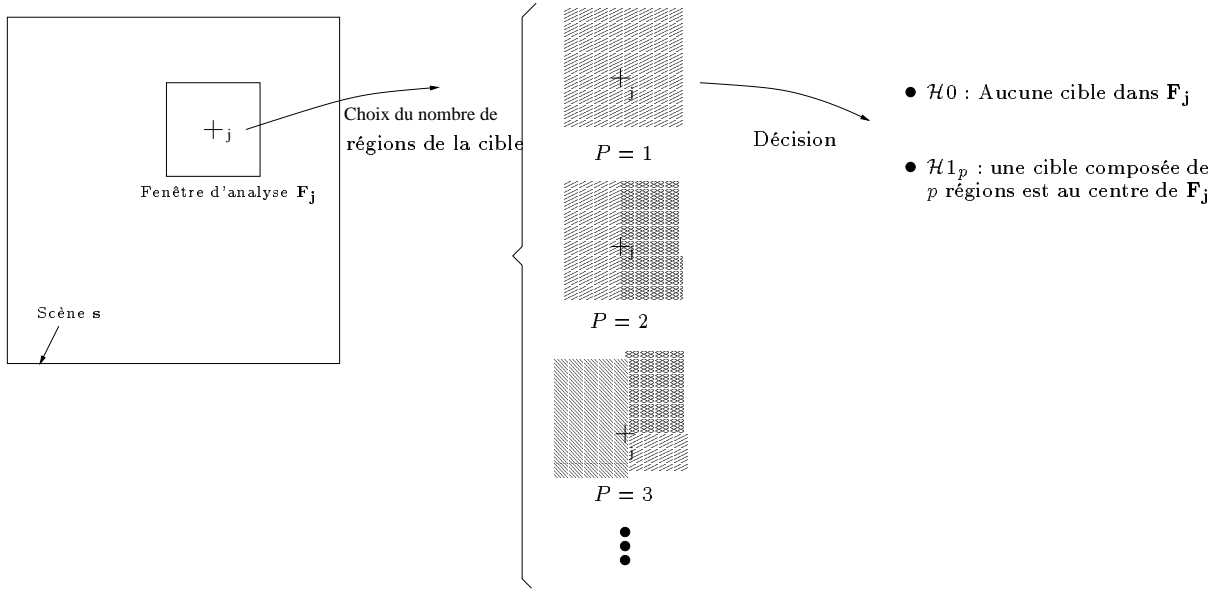


FIG. 5.14: Principe général de l'algorithme de détection/localisation estimant le nombre de régions. En chaque pixel j de l'image, le nombre de régions de la cible éventuelle est estimé. On utilise alors le filtre GLRT correspondant à cette cible pour effectuer la localisation ou la détection.

Pour l'estimation du nombre de régions, une première idée est d'estimer ce nombre au sens du maximum de vraisemblance, ainsi que nous l'avons fait pour l'estimation de la

taille des cibles dans le chapitre 3. Cela reviendrait à calculer en chaque pixel de l'image la vraisemblance de l'hypothèse $\mathcal{H}1$, pour chaque cible possible. Ceci peut être effectué rapidement à l'aide d'un banc de filtre. La description choisie pour la phase de détection serait alors celle ayant donné la plus grande vraisemblance.

Cependant, le nombre de régions composant la cible peut être vu comme représentatif de l'ordre de notre modèle. Or il est connu que le maximum de vraisemblance ne permet pas de sélectionner l'ordre d'un modèle. Plus explicitement, l'étude menée dans la section précédente a permis de montrer que la logvraisemblance de l'hypothèse que la cible est composée de $L + 1$ régions est toujours plus grande que la logvraisemblance d'une hypothèse de L régions. En effet, la différence entre ces deux logvraisemblances est proportionnelle à une variable aléatoire du χ^2 , toujours positive. Une sélection du nombre de région au sens du maximum de vraisemblance donne donc toujours l'estimation du nombre de régions maximal.

5.4.1 Pénalisation du nombre de régions

Pour s'affranchir de ce problème, une solution pragmatique consiste à sélectionner le nombre de régions en fonction d'un critère correspondant à la fonction de vraisemblance, à laquelle nous ajoutons un terme pénalisant l'augmentation de la complexité du modèle (ie : le nombre de régions utilisé pour décrire la cible). Nous supposons connu le nombre maximal L de régions composant la cible. Nous choisirons donc le nombre P_j de régions choisi pour définir la cible cherchée dans la fenêtre \mathbf{F}_j , placée à la position j dans la scène s de la façon suivante :

$$P_j = \operatorname{argmax}_{k \in [1..L]} [\mathcal{L}^k(\mathbf{s}, j) - f(k)] \quad (5.10)$$

Avec $\mathcal{L}^k(\mathbf{s}, j)$, la logvraisemblance de l'hypothèse d'une cible composée de k régions et f , une fonction croissante.

La fenêtre d'analyse utilisée pour la détection étant la même quel que soit le nombre de régions choisies, ceci est équivalent à effectuer la sélection en considérant le logarithme du rapport de vraisemblance entre les hypothèses : $\mathcal{H}0$ et $\mathcal{H}1_k$, d'absence de cible et de présence d'une cible composée de k régions.

$$P_j = \operatorname{argmax}_{k \in [1..L]} [\mathcal{R}^{(k)}(\mathbf{s}, j) - g(k)] \quad (5.11)$$

Il s'agit donc maintenant de définir la fonction de pénalisation g . On trouve dans la littérature un grand nombre de points de vues pour définir ce terme. Selon le critère d'information d'Akaike [49], ce terme de pénalisation se justifie par la suppression d'un biais dans l'estimation de la logvraisemblance. Ce biais est lié au nombre de paramètres libres du modèle. Plus précisément, ce biais est le nombre de paramètres du modèle divisé par deux. Dans le cas du modèle de bruit Gamma que nous avons utilisé, seules les moyennes doivent être estimées. La fonction de pénalisation serait donc :

$$g(k) = \frac{1}{2}k \quad (5.12)$$

Cette correction correspond exactement à la moyenne de la variable aléatoire ajoutée au logarithme du rapport de vraisemblance que nous avons évoqué dans la section précédente.

Une approche différente est celle de la description de longueur minimale, notée MDL (pour Minimum Description Length), que nous explicitons plus en détail dans l'annexe A. La pénalisation ne s'explique alors pas par la correction d'un biais. Le principe sous-jacent est de n'utiliser un modèle complexe que si celui-ci porte effectivement de l'information. Pour définir cela, le modèle retenu est celui qui minimise la longueur du code nécessaire pour coder l'information présente dans l'image. En d'autres termes, le terme de pénalisation correspond au nombre de bits nécessaires pour coder la description de l'objet [48]. Le calcul de ce nombre de bits n'est pas des plus évident, bien qu'en recourant à certaines approximations on puisse disposer d'une valeur très générale. Encore une fois, la fonction de pénalisation prend une forme linéaire.

C'est donc la forme générale que nous retiendrons également pour pénaliser le nombre de régions décrivant la cible

$$g(k) = \alpha k \quad (5.13)$$

où α est un paramètre permettant de régler l'importance du terme de pénalisation.

Dans ce qui précède, nous n'avons considéré que le nombre de régions utilisé pour modéliser la cible. Cependant, supposons qu'un objet décrit par L régions au maximum, ne soit composé que de P régions effectives, c'est à dire de régions de moyennes différentes. Il existe dans ce cas plusieurs possibilités de configurations de L régions en P régions. La figure 5.15 donne un exemple très simple de ceci, pour $L = 3$.

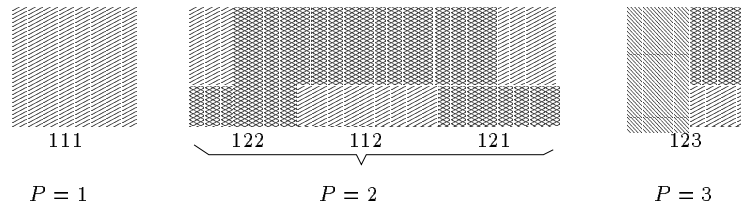


FIG. 5.15: Différentes configurations d'un objet décomposé en P régions effectives, pour un nombre maximum de régions $L = 3$ du modèle. Chaque configuration est définie par un nombre à L chiffres. Chaque chiffre est l'indice d'une région effective. La position de chacun de ces chiffres est l'indice de la région dans la description la plus complète ($P = 3$) de l'objet.

Nous ferons l'hypothèse que le terme de pénalisation ne dépend pas de la configuration testée, mais seulement du nombre de régions effectives. Nous proposons donc l'algorithme suivant pour la détection d'objets dont on ne connaît que le nombre maximal de régions. Cet algorithme, noté RS-GLRT (Region Selection GLRT) est décomposé en deux étapes :

Premièrement, nous effectuons la sélection de région (RS). Pour cela nous calculons le GLRT pour chaque configuration possible de l'objet, auquel nous retirons le terme de pénalisation approprié, ne dépendant que du nombre de régions effectives. Ceci permet de sélectionner la configuration la plus appropriée, ainsi que le nombre de régions effectives estimé par la sélection de région. Ce dernier est donné par :

$$P_{RS} = \arg \max_P [\mathcal{R}^{(P)} - \alpha P] \quad (5.14)$$

Cette opération est effectuée pour chaque point de l'image, afin que le nombre de régions puisse varier d'une position à l'autre.

Dans la seconde étape, la valeur du plan de sortie du RS-GLRT est donnée par le GLRT correspondant à la configuration sélectionnée plus haut.

$$\mathcal{R}_{RS} = \mathcal{R}^{(P_{RS})} \quad (5.15)$$

S'il s'agit de localiser la cible, on choisira le maximum de ce plan de sortie. Dans le cas d'une opération de détection, ce plan sera binarisé en fonction du seuil de détection choisi.

5.4.2 Choix du paramètre de pénalisation

Le paramètre α de l'équation 5.14 contrôle le terme de pénalisation. Une forte valeur de ce paramètre tendra à sélectionner un nombre de régions minimal, sauf en cas de fortes différences de contraste entre les régions, alors qu'une faible valeur aura tendance à utiliser un grand nombre de régions, même si le contraste entre les régions est faible ou nul. Notons que si ce paramètre est nul, l'algorithme RS-GLRT est exactement équivalent au $GLRT^{(L)}$ (l'algorithme utilisant le plus grand nombre de régions). Si ce paramètre est infini, le RS-GLRT est équivalent au $GLRT^{(1)}$ (l'algorithme supposant la cible homogène). Nous pouvons ainsi passer continuellement d'un algorithme 1 région à un algorithme L régions. Pour des raisons évidentes de simplicité d'utilisation, nous avons préféré ne pas laisser ce paramètre indéterminé.

En utilisant le critère d'information d'Akaike [49], nous avons vu que l'on obtient une valeur de $\alpha = 0.5$. L'autre approche évoquée plus haut, le MDL, nécessite l'estimation du nombre de bits utilisé pour coder chaque région. L'approximation de ce nombre de bits (cf. annexe A), permet d'obtenir pour les objets utilisés une valeur $\alpha_{MDL} \simeq 2.5$.

Ces deux approches, donnant des résultats assez différents, on pourra considérer le problème sous un autre angle. Le calcul du RS-GLRT implique l'estimation de la configuration de région vraie. Il est donc opportun d'effectuer le choix de ce paramètre à partir de la précision de cette estimation. L'estimation du nombre de régions ne dépend que des niveaux de gris des pixels à l'intérieur de la cible. On utilisera ainsi la notation C pour représenter le contraste interne dans la cible. Dans le cas où la cible est composée de plusieurs régions, nous avons choisi les moyennes m_i de chaque région i de telle sorte que $m_i/m_{i+1} = C$. La figure 5.16 montre le nombre moyen de régions estimées pour différentes valeurs du contraste C et du paramètre de pénalisation α . L'objet est composé de 2 régions de 200 pixels et la description fournie est composée au plus de 4 régions de 100 pixels chacune. Le nombre moyen de régions estimées a été évalué sur 100000 réalisations.

Le nombre de régions estimé décroît continuellement de 4 à 1 lorsque α augmente. Pour une valeur donnée de α , un contraste plus grand implique une estimation supérieure ou égale du nombre de région. Pour pouvoir conserver une estimation correcte du nombre de régions, quel que soit le contraste, choisir une valeur de l'ordre de comprise entre 2 et 3 semble indiqué.

Pour généraliser cela, nous avons tracé dans la figure 5.17, l'évolution du nombre de régions estimées en fonction de α pour un objet composé de 1, 2 ou 3 régions. Le contraste interne dans la cible est considéré infini (le rapport de moyenne entre 2 régions est de 100).

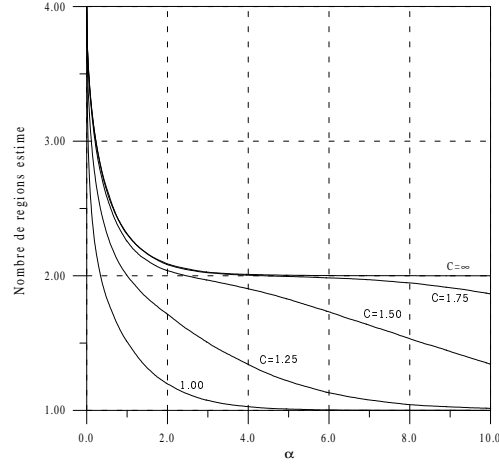


FIG. 5.16: Estimation du nombre moyen de régions estimé à partir de l'équation 5.14 en fonction du paramètre α , pour différents contrastes internes C . L'objet est formé de 2 régions de tailles égales, et la description maximale de l'objet suppose 4 régions.

Encore une fois, le choix $\alpha \simeq 3$ permet d'obtenir une estimation relativement correcte. Une valeur de $\alpha > 4$ donnerait de bons résultats pour un contraste infini, mais pas pour des contrastes plus faibles, comme on a pu le voir dans la figure 5.16. Nous pouvons remarquer que ces résultats sont assez proches de la valeur donnée par une approche de type MDL. La détermination plus précise d'une approche MDL rigoureuse constituerait une étude intéressante en soi pour pouvoir valider cette remarque. Nous avons préféré porté notre attention sur un autre point d'évaluation des algorithmes, leur vitesse de traitement des données.

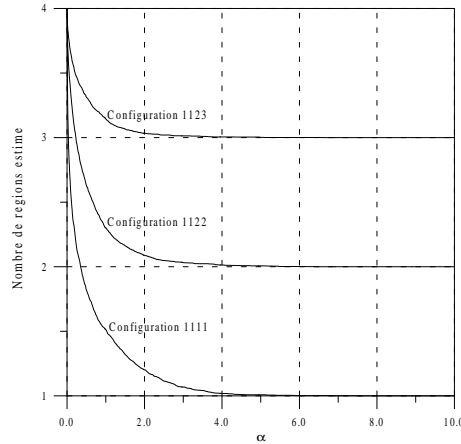


FIG. 5.17: Estimation du nombre de régions en fonction du paramètre α , pour différentes configurations. Le contraste est ici infini.

5.4.3 Complexité de cet algorithme

Au vu de la demande existant d'algorithmes capables de fonctionner en temps réel, il est important de mesurer la complexité de l'algorithme proposé.

Si l'on reprend la forme générale de l'algorithme RS-GLRT, nous pouvons remarquer que le calcul du rapport de vraisemblance pour une configuration possible de la cible (cf. figure 5.15) ne nécessite que la connaissance de la moyenne et du nombre de pixels de chacune des régions de cette configuration. Chacune des régions d'une configuration est une région de la description complète de l'objet, ou la fusion de plusieurs régions de cette description. Ces deux paramètres peuvent donc être obtenus directement à partir des moyennes et du nombre de pixels des régions de la description complète de l'objet. Or nous avons vu que ces moyennes peuvent être calculées par une simple corrélation entre le masque de chaque région et la scène à traiter.

Nous aurons donc à effectuer L corrélations, entre la scène et le masque de chacune des régions décrivant l'objet. Si la scène comporte N pixels, cette partie est de complexité $O(L N \log N)$, en utilisant une transformée de Fourier rapide. Il faut ensuite calculer le GLRT pour chaque configuration, ainsi que le GLRT auquel on adjoint la fonction de pénalisation, ce qui est fait pixel par pixel, donc linéaire en N . Soit \mathcal{N}_C le nombre de configurations possibles, le post-traitement des corrélations est donc de complexité $O(\mathcal{N}_C N)$.

Tant que ce nombre de configurations reste faible, nous pourrions considérer que seules les corrélations sont importantes. Effectuer L corrélations ne sera pas considéré comme une limite problématique.

Le nombre de configurations possibles pour P régions effectives nous est donné par l'analyse combinatoire [3]. Il s'agit du nombre de possibilités de partitionner un ensemble de L éléments en P sous ensembles non vides, donné par le nombre de Stirling de seconde espèce $\mathcal{S}_P^{(L)}$. Ce nombre est défini par la relation de récurrence suivante :

$$\mathcal{S}_P^{(L)} = P \mathcal{S}_P^{(L-1)} + \mathcal{S}_{P-1}^{(L)} \quad \text{avec } \mathcal{S}_1^{(1)} = 1 \text{ et } \mathcal{S}_1^{(2)} = 1 \quad (5.16)$$

Le tableau ci-dessous indique le nombre de total de configurations possibles $\mathcal{N}_C^{(L)} = \sum_{P=1}^L \mathcal{S}_P^{(L)}$ pour L variant de 1 à 10.

Nombre de régions	Nombre de configurations
maximum L	possibles
1	1
2	2
3	5
4	15
5	52
6	203
7	877
8	4140
9	21147
10	115975

Le nombre de configurations augmente de façon exponentielle avec le nombre de régions maximum utilisé pour décrire l'objet. Le tableau précédent montre qu'en pratique, utiliser plus de 5 ou 6 régions est difficile en utilisant cette méthode.

5.5 Estimation simplifiée du nombre de régions

Comme nous l'avons fait remarquer dans la section précédente, la sélection de la configuration adéquate pour un objet inconnu peut se révéler trop complexe pour pouvoir être utilisable en tant que telle. On pourrait alors envisager différentes manières d'accéder à la configuration de l'objet sans tester tous les cas possibles pour isoler les régions les plus différentes le plus rapidement possible, en utilisant par exemple une méthode de type descente de gradient.

Cependant, on peut obtenir une simplification beaucoup plus significative en reconsidérant notre problème du point de vue de la détection. Dans ce cas, nous ne voulons distinguer que deux classes :

- La cible, composée d'un nombre arbitraire de régions tirées de la description complète de l'objet.

- Le fond, supposé uniforme dans la fenêtre d’analyse.
- Il est alors tentant de considérer seulement deux configurations possibles :
- Une configuration 1 région, qui devra être choisie lorsque la fenêtre d’analyse est positionnée sur les zones fond de l’image.
- La configuration la plus détaillée (L régions), correspondant aux positions où une cible se trouve au centre de la fenêtre.

Le raisonnement mené plus haut pour la mise au point de l’algorithme RS-GLRT étant toujours valide, nous proposons donc l’algorithme SRS-GLRT (Simplified RS-GLRT), pour la détection d’objets composés de plusieurs régions : dans un premier temps, cet algorithme estime le nombre de régions de l’échantillon présenté :

$$L_{SRS} = \begin{cases} L & \text{si } \mathcal{R}_{inc}^{(L)} - \mathcal{R}_{inc}^{(1)} > \alpha_{SRS} \\ 1 & \text{sinon} \end{cases} \quad (5.17)$$

la valeur en sortie du filtre est ensuite donnée par :

$$\mathcal{R}_{inc}^{SRS} = \mathcal{R}_{inc}^{(L_{SRS})} \quad (5.18)$$

Le choix entre les deux configurations possibles est fait en chacun des pixels de l’image à traiter. Avec une valeur appropriée de α_{SRS} , le filtre devrait donc sélectionner la sortie de l’algorithme $GLRT^{(L)}$ sur la cible, et la sortie de l’algorithme $GLRT^{(1)}$ sur le fond. Si la cible est composée d’une seule région, aux fluctuations statistiques près, l’algorithme devrait se comporter comme le $GLRT^{(1)}$.

On peut noter que le choix de α_{SRS} est ici moins critique que dans le cas du RS-GLRT. La différence observée entre une configuration 1 région et une configuration L régions est en effet plus importante que la différence entre les configurations intermédiaires que devait discriminer l’algorithme RS-GLRT.

Le raisonnement mené précédemment montre que la différence entre les logvraisemblances d’1 région et de L régions est proportionnel à une variable aléatoire du χ^2 d’ordre $L - 1$. Nous prendrons donc comme valeur de α_{SRS} :

$$\alpha_{SRS} = (L - 1)\alpha \quad (5.19)$$

où α est la valeur du coefficient utilisé pour le RS-GLRT.

Notons bien que dans ce cas, la complexité de l’algorithme est très réduite par rapport au cas du RS-GLRT, puisque qu’ici, seulement deux configurations sont testées.

5.6 Performances de ces algorithmes

Nous pouvons maintenant comparer les performances des algorithmes que nous avons mis en oeuvre, dans les différentes situations pour lesquelles ils sont adaptés. Nous comparons ici les performances en détection des algorithmes $GLRT^{(1)}$, $GLRT^{(L)}$, $RS - GLRT$ et $SRS - GLRT$. Dans les expériences qui sont présentées ici, la cible est composée de 4 régions maximum, de 100 pixels chacune. La fenêtre d’analyse utilisée contient 800 pixels.

La figure 5.18 montre les performances obtenues pour un objet de 4 régions, dans lesquelles les moyennes sont respectivement 10, 12.5, 17.5, 20 sur la cible. Le fond présente une moyenne de 15. Dans ce cas, la cible complète ne présente pas de différence de moyenne avec le fond, ce qui explique l'échec du filtre $GLRT^{(1)}$. L'objet étant composé de 4 régions, le $GLRT^{(4)}$ est le filtre qui présente les meilleures performances. Le $RS - GLRT$ est quasiment équivalent au $GLRT^{(4)}$, ce qui indique que la sélection de région a été correctement effectuée. Le filtre $SRS - GLRT$ présente des performances un peu moins bonnes, mais largement supérieures au filtre $GLRT^{(1)}$.

Dans la figure 5.19, nous utilisons ces filtres sur une image contenant un objet homogène, présentant un faible contraste avec le fond. La moyenne sur la cible est de 10, contre 12.5 pour le fond. Bien entendu, le filtre $GLRT^{(1)}$ donne les meilleurs résultats. Le $SRS - GLRT$ suit de près ces performances. Le filtre $GLRT^{(4)}$ présente les moins bonnes performances, du fait des pertes discutées précédemment. Les résultats du $RS - GLRT$ sont plus mitigés. Sur la partie droite du graphe (faibles seuils de détection), il se comporte comme le $GLRT^{(1)}$. Dans la partie gauche du graphe (forts seuils de détection) il est plus proche du $GLRT^{(4)}$. Ceci est dû au mode de représentation des courbes COR. Si l'algorithme estime que la cible est composée de 4 régions, la valeur en sortie du filtre $RS - GLRT$ est plus grande que dans le cas où une seule région est utilisée. Les plus grandes valeurs en sorties correspondent dans les courbes COR aux forts seuils de détection, c'est à dire à la partie gauche des courbes.

Enfin, la figure 5.20 présente un cas intermédiaire, d'un objet composé de 2 régions de 200 pixels, dont les moyennes sont 10 et 15. Le fond (de moyenne 15.7) présente également un faible contraste avec la cible. Les algorithmes $GLRT^{(4)}$ et $RS - GLRT$ se révèlent les plus performants dans ce cas. Le filtre $GLRT^{(1)}$ est peu efficace et le $SRS - GLRT$ se situe entre ces deux bornes.

D'après les résultats précédents, le $RS - GLRT$ se comporte le mieux, sauf lorsque l'objet est homogène. Dans ce cas, ses performances sont très bonnes à fort taux de fausse alarme, mais médiocres à faible PFA. Le point délicat de cet algorithme est la détermination de la fonction pénalisant de trop nombreuses régions. Une étude plus précise de l'approche du MDL que nous n'avons fait qu'évoquer semble un prolongement extrêmement intéressant de ces travaux.

L'autre approche que nous avons proposée, le $SRS - GLRT$, a été développée en considérant l'accélération du $RS - GLRT$ comme le facteur le plus important. De fait, cela permet d'obtenir un filtre qui se comporte comme un compromis efficace entre 1 et L régions, quelle que soit la configuration présentée, tout en restant d'une complexité faible.

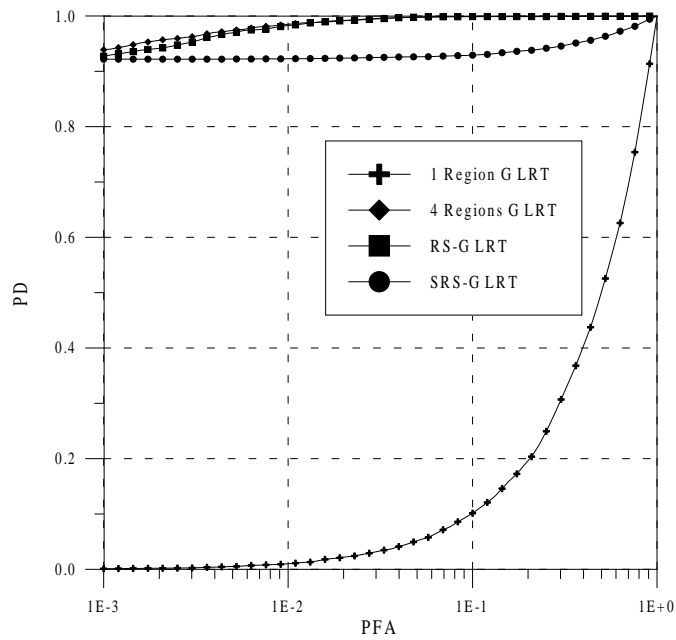


FIG. 5.18: courbes COR obtenues pour les algorithmes $GLRT^{(1)}$, $GLRT^{(4)}$, RS-GLRT et SRS-GLRT. L'objet est composé de 4 régions (de moyennes 10, 12.5, 17.5 et 20) et ne présente pas de contraste global avec le fond (de moyenne 15). Les régions sont composées de bruit Gamma d'ordre 1.

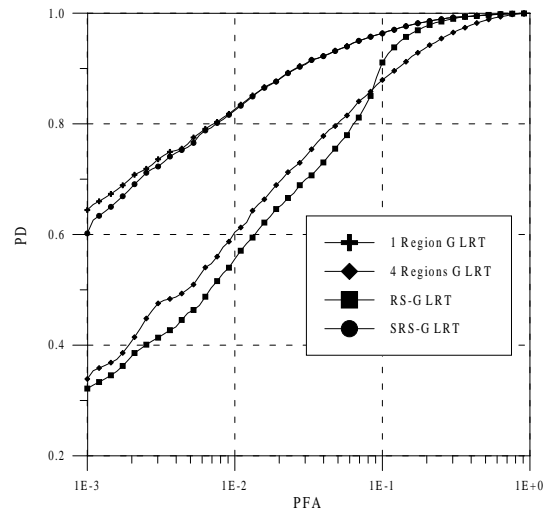


FIG. 5.19: courbes COR obtenues pour les algorithmes $GLRT^{(1)}$, $GLRT^{(4)}$, RS-GLRT et SRS-GLRT. L'objet est homogène (1 région) (de moyenne 10) et présente un faible contraste avec le fond (de moyenne 12.5). Les régions sont composées de bruit Gamma d'ordre 1.

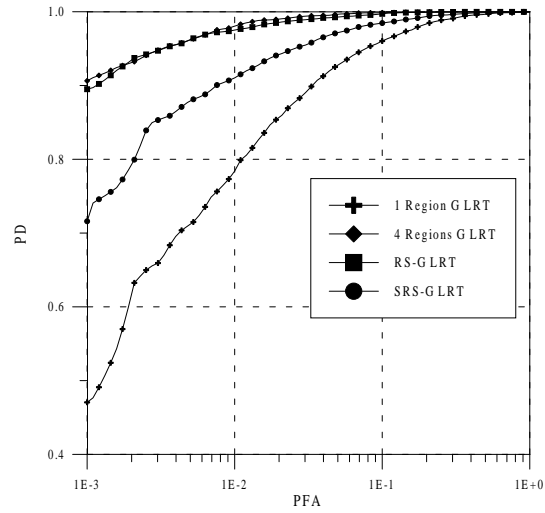


FIG. 5.20: courbes COR obtenues pour les algorithmes $GLRT^{(1)}$, $GLRT^{(4)}$, RS-GLRT et SRS-GLRT. L'objet est composé de 2 régions (de moyennes 10, 15) et présente un faible contraste avec le fond (de moyenne 16). Les régions sont composées de bruit Gamma d'ordre 1.

5.7 Conclusion

Nous avons dans ce chapitre présenté une approche permettant de traiter le cas d'objets structurés, dont on connaît la répartition spatiale des régions, mais dont l'intensité moyenne peut fluctuer fortement, rendant impossible l'utilisation de filtres utilisant la connaissance a priori des niveaux de gris de la cible. Ces algorithmes ont été développés en conservant une approche statistique, étendant le champ d'application des algorithmes GLRT, sans pour autant perdre le support théorique ayant servi de base à leur construction. Nous avons ensuite montré l'intérêt de ces algorithmes dans le cas de la visualisation d'objets 3D par exemple. Il serait d'ailleurs intéressant de les utiliser dans le cas d'images thermiques, pour lesquelles, les objets recherchés sont également modélisés en zones homogènes très simples.

La question rendue implicite par ces résultats était la limite du nombre de régions utilisables pour décrire les objets. En étudiant le comportement de ces algorithmes, nous avons déterminé théoriquement ces pertes, issues d'une sur-estimation de la complexité du modèle de cible utilisé. Un trop grand nombre de régions entraîne un accroissement des fluctuations en sortie du filtre sur les régions fond des images à traiter.

Nous avons ensuite envisagé de limiter ces pertes en estimant cette complexité en chaque point de l'image. Ceci nous a conduit à définir l'algorithme RS-GLRT, qui considère l'ensemble des configurations de régions possibles et procède à une sélection de la configuration la plus appropriée. Cette sélection se fait en considérant la logvraisemblance de chaque configuration, à laquelle on adjoint une fonction de pénalisation, limitant la complexité du modèle. La détermination de cette fonction de pénalisation n'est pas immédiate. Sous une forme linéaire, elle dépend d'un coefficient auquel la littérature existante sur des sujets proches donne des valeurs assez variables. Nous avons fixé cette valeur en considérant la qualité de l'estimation du nombre de régions.

Déterminer une méthode plus efficace pour déterminer ce nombre de régions pourrait faire l'objet de recherches intéressantes. Un autre problème a également retenu notre attention : pour des descriptions d'objets de plus de 5 à 6 régions, il peut être trop long de procéder à une recherche exhaustive de la configuration appropriée. Nous avons donc proposé une version simplifiée de cet algorithme, plus spécifiquement dédiée à la détection. Dans ce cas, le nombre de régions de l'objet n'est pas primordial, le seul point important étant de limiter les fluctuations dans les régions fond de l'image. En ne considérant que les deux descriptions extrêmes de la cible : une homogène et l'autre détaillée, on peut définir l'algorithme SRS-GLRT, qui permet de conserver des niveaux de performances assez bon, quelle que soit la configuration de la cible présentée, dans un temps de calcul de quelques corrélations.

Chapitre 6

Corrélateurs optiques

Sommaire

6.1	Structures de corrélateurs optiques	96
6.1.1	Différents types de codages de l'image d'entrée.	99
6.2	Comparaison entre le GLRT et le NLJTC	101
6.3	Choix d'un codage	103
6.3.1	Position du problème	103
6.3.2	Exemple	104
6.3.3	Etude de l'influence du codage en phase en présence de bruit non recouvrant	107
6.3.4	Illustration dans le cas d'images réalistes	113
6.4	Conclusion	114

Dans l'état actuel des technologies, le principal critère de choix d'un algorithme est bien souvent simplement sa vitesse. Dans des applications de poursuite de cibles, pour lesquelles on souhaite effectuer toutes les opérations de localisation ou de détection en temps réel (cadence vidéo), les filtres optimaux ne se révèlent pas toujours les plus adaptés. Même si les algorithmes GLRT présentés dans ce mémoire ont une complexité équivalente à celle de quelques corrélations, le calcul de celles-ci par FFT est parfois considéré comme trop lent.

Il est alors courant de recourir à des algorithmes sous optimaux, mais présentant des cadences de traitement plus élevées. C'est pourquoi, au cours des 40 dernières années, la corrélation optique est un domaine qui a fait l'objet de très nombreuses études. En effet, dans ce cas, la corrélation n'est plus un facteur limitatif du temps de traitement, celle-ci étant effectuée instantanément. De fait, le temps de traitement est alors limité par l'affichage des images dans le système optique. Le développement actuel des composants optroniques permet ainsi de penser que l'utilisation des corrélateurs optiques devrait s'accroître en conséquence. Depuis les travaux fondateurs de Van der Lugt [59], on a ainsi vu apparaître de nombreuses architectures de corrélateurs optiques, dont notamment celle

du JTC (*Joint Transform Correlator*) [14] et du NLJTC (*Non Linear Joint Transform Correlator*) [28].

Ce dernier bénéficie d'un fort support théorique, et ses propriétés d'adaptativité à la scène à traiter nous ont semblé en faire un candidat intéressant pour l'implémentation optique de filtres de corrélation.

Dans la première partie de ce chapitre, des simulations numériques du corrélateur optique nous permettront de comparer les performances en localisation du NLJTC à celles de l'algorithme du rapport de vraisemblance généralisé (GLRT) que nous avons présenté dans le chapitre 1. L'étude que nous avons menée dans le chapitre 2 a permis de montrer que pour une tâche de localisation, le GLRT est très proche de l'algorithme optimal du Maximum de Vraisemblance. On peut donc s'attendre à ce que les performances du GLRT soient supérieures à celles du NLJTC dans ce cas. Il s'agira donc ici surtout de quantifier cette différence.

Par ailleurs, toutes les architectures optiques ont en commun le problème de l'introduction de la scène à traiter et du filtre à appliquer dans le système optique. Cette introduction est effectuée par la modulation d'un faisceau de lumière incident à l'aide d'un modulateur spatial de lumière (SLM pour *Spatial Light Modulator*). Dès lors, le choix d'un type de modulation modifie les performances du filtrage, et définit un problème de traitement de signal auquel nous nous sommes intéressés dans la suite de ce chapitre.

Nous montrerons qu'en présence de bruit non recouvrant, le codage en phase permet d'obtenir des performances en localisation meilleures que celles obtenues pour un codage en amplitude. En particulier, nous montrerons que le filtre ainsi obtenu est moins sensible à de faibles contrastes entre la cible et le fond. Nous fournirons ensuite une interprétation phénoménologique de ces observations, dans le cas de scènes synthétiques, et dans le cas de fonds réalistes.

6.1 Structures de corrélateurs optiques

Corrélateurs de type VDL

Pour commencer, et pour définir un certain nombre de notations, nous présentons ici brièvement l'architecture d'un corrélateur optique de type Van Der Lugt (VDL) [59].

Ce type d'architecture repose sur une propriété bien connue de la diffraction : la répartition angulaire de l'amplitude lumineuse diffractée par une ouverture est égale à la transformée de Fourier de la "fonction de transparence" de l'ouverture diffractante. La corrélation étant une multiplication dans l'espace de Fourier, il est possible de réaliser une corrélation optique grâce à un montage tel que celui représenté sur la figure 6.1.

Le principe de fonctionnement est le suivant. Un faisceau collimaté de lumière cohérente (laser) traverse tout d'abord un modulateur spatial de lumière (SLM pour "Spatial Light Modulator") sur lequel est codée la scène s . Un SLM est un composant passif qui permet de moduler la polarisation, la phase ou l'intensité de la lumière qui le traverse. On peut le considérer comme une diapositive reprogrammable en temps réel, de manière optique ou électronique. Le faisceau diffracté par le SLM traverse ensuite une lentille, qui le focalise. Dans le plan focal de la lentille, la répartition spatiale du faisceau est égale à la

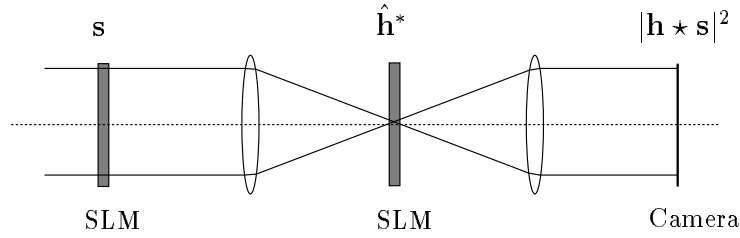


FIG. 6.1: Principe du corrélateur optique de type VDL

transformée de Fourier \hat{s} de la scène codée. On place dans ce même plan focal un second SLM sur lequel est inscrit le conjugué de la transformée du filtre \mathbf{h} . La répartition spatiale du faisceau en sortie du second SLM est donc le produit $\hat{h}_\nu^* \hat{s}_\nu$. La lumière traverse ensuite une seconde lentille qui projette sur le plan C la transformée de Fourier du produit $\hat{h}_\nu^* \hat{s}_\nu$, c'est à dire la fonction d'inter-corrélation entre \mathbf{s} et \mathbf{h} . Cette fonction est acquise par une caméra, qui est un composant quadratique, uniquement sensible à l'intensité lumineuse et non à l'amplitude. L'information visualisée par la caméra est donc le carré du plan de corrélation.

Corrélateur à transformée conjointe (JTC)

A peu près à la même époque que Van der Lugt, un autre moyen de calculer une corrélation de manière optique a été proposé [14] : le *Corrélateur à transformée conjointe* (JTC pour "Joint transform correlator"). Son principe consiste à présenter dans le même plan, et côte à côte, l'image et la réponse impulsionnelle du filtre séparées d'une distance $2d$ (cf. figure 6.2). La transformée de Fourier de ces deux objets est formée selon le même principe que dans le corrélateur de Van der Lugt, mais on place dans le plan focal un détecteur quadratique, par exemple une caméra CCD. Ce détecteur forme le module carré de la somme des transformées de Fourier des deux objets :

$$J_\nu = |\hat{h}_\nu e^{+2i\pi d\nu} + \hat{s}_\nu e^{-2i\pi d\nu}|^2 = |\hat{h}_\nu|^2 + |\hat{s}_\nu|^2 + \hat{h}_\nu^* \hat{s}_\nu e^{-4i\pi d\nu} + \hat{s}_\nu^* \hat{h}_\nu e^{4i\pi d\nu} \quad (6.1)$$

Cette grandeur, appelée *spectre conjoint*, est ensuite inscrite sur un second faisceau cohérent, qui, à l'aide d'une lentille, en forme la transformée de Fourier sur le plan focal d'une caméra. Le plan de sortie acquis par la caméra a donc l'expression suivante :

$$|C|_i^2 = [\mathbf{h} \otimes \mathbf{h}]_i^2 + [\mathbf{s} \otimes \mathbf{s}]_i^2 + [\mathbf{h} \otimes \mathbf{s}]_{i-2d}^2 + [\mathbf{h} \otimes \mathbf{s}]_{-i-2d}^2 \quad (6.2)$$

Le carré du produit de corrélation entre le filtre et la scène $[\mathbf{h} \otimes \mathbf{s}]_{i-2d}^2$ apparaît, décalé d'une distance $2d$ par rapport au centre du plan de sortie. Si la distance d est assez grande, il peut être séparé du pic parasite situé au centre de l'image et être utilisé pour localiser la cible.

Pour des raisons pratiques, essentiellement la réduction de la dynamique du spectre conjoint, une non-linéarité est parfois introduite au niveau du composant quadratique (la caméra dans le schéma précédent) donnant lieu à un corrélateur de type NLJTC (*Non Linear Joint Transform Correlator*).

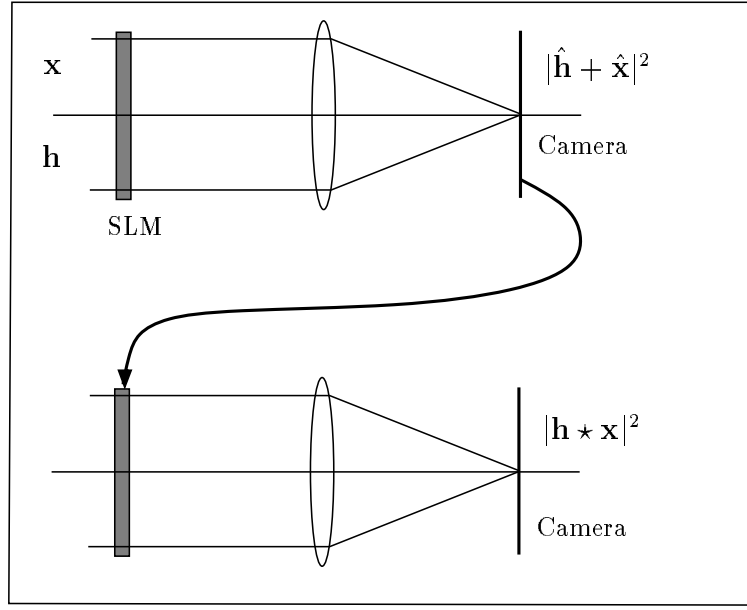


FIG. 6.2: Principe du corrélateur optique de type JTC.

Lorsque la non-linéarité est implantée en utilisant un cristal photoréfractif (cf. figure 6.3), il a été montré [45] qu'en première approximation, le spectre conjoint inscrit dans le cristal est de la forme suivante dans le plan de Fourier :

$$\hat{c}_\nu = \frac{\hat{h}_\nu^* \hat{s}_\nu}{\sigma^2 + |\hat{h}_\nu|^2 + |\hat{s}_\nu|^2} \quad (6.3)$$

où σ est une constante.

Pour localiser un objet dont on connaît une référence \mathbf{r} , ceci revient à définir un filtre h^σ , puis à effectuer la corrélation linéaire entre la scène \mathbf{s} et h^σ :

$$h_\nu^\sigma = \frac{\hat{r}_\nu}{\sigma^2 + |\hat{r}_\nu|^2 + |\hat{s}_\nu|^2} \quad (6.4)$$

Comme nous l'avons dit, cette non-linéarité a été introduite pour des raisons pratiques. Cependant, on a remarqué par la suite que les filtres non-linéaires exprimés par l'équation 6.4 permettaient souvent d'obtenir des performances supérieures aux filtres linéaires. On peut remarquer que le filtrage par le NLJTC est adaptatif, au sens où le filtre h^σ dépend de la scène \mathbf{s} à traiter.

Des justifications heuristiques ont été employées pour expliquer l'intérêt de disposer de filtres adaptatifs. En particulier, la recherche d'un filtre réalisant un compromis entre différents critères de localisation conduit à la construction du filtre suivant [45] :

$$\hat{h}_\nu = \frac{\hat{r}_\nu}{\lambda + \nu |\hat{r}_\nu|^2 + |\hat{s}_\nu|^2} \quad (6.5)$$

Ce filtre est similaire au NLJTC présenté dans l'équation 6.4.

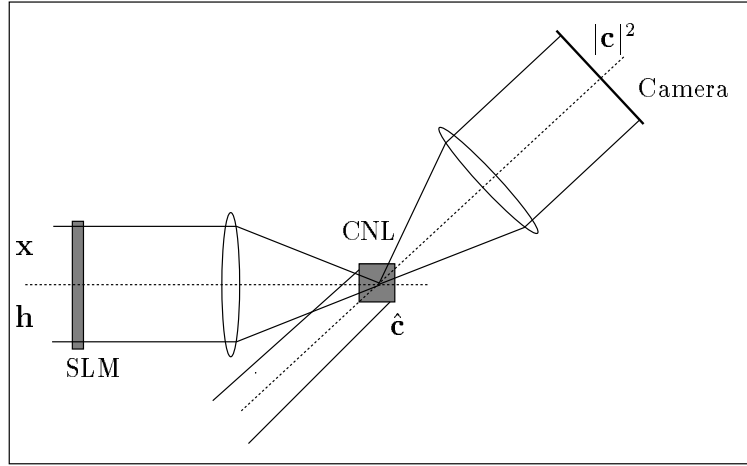


FIG. 6.3: Principe du corrélateur optique de type NLJTC. La non-linéarité est ici introduite par un cristal non-linéaire.

Une justification théorique a plus récemment été proposée [43, 42]. Les auteurs montrent qu'en présence d'un bruit additif gaussien de densité spectrale de puissance inconnue, le NLJTC est une bonne approximation d'un filtre de localisation optimal au sens de la théorie de la décision. Le filtrage réalisé par un NLJTC s'appuie donc sur des bases solides, tant théoriques que pratiques. Il faut cependant noter que ces études ont été menées dans le cas de bruit additif, qui n'est pas le type de bruits présent dans les scènes utilisées dans ce chapitre. Nous utilisons donc le NLJTC dans un contexte empirique. C'est pour cette raison qu'il est légitime de penser que ses performances peuvent être améliorées en modifiant les différents paramètres de ce système, notamment la façon dont la scène est codée sur les SLM.

6.1.1 Différents types de codages de l'image d'entrée.

La scène s et la référence r sont introduites dans le système optique sous la forme d'une fonction de transparence du SLM. Celle-ci s'exprime sous la forme d'un nombre complexe, fonction du niveau de gris du pixel à coder. Ce nombre complexe a un module inférieur ou égal à 1, les SLM ne permettant pas d'amplifier la lumière du faisceau incident.

On pourra regrouper les différents types de SLM disponibles sur le marché en deux grandes classes, correspondant à deux types de codages différents :

D'une part le codage en amplitude, pour lequel la transparence est réelle et directement proportionnelle au niveau de gris à coder. D'autre part le codage en phase, pour lequel la transparence a un module égal à 1 mais dont la phase est proportionnelle au niveau de gris à coder. Dans le cas du codage en phase, il faut de plus définir une profondeur de modulation k , représentant la modulation de phase maximale que l'on peut obtenir avec le SLM choisi.

Le domaine de codage d'un SLM est l'ensemble des valeurs complexes qui peuvent être atteintes pour l'ensemble des niveaux de gris possibles. Nous représentons sur la figure suivante les domaines de codage obtenus lorsque l'image est codée en amplitude ou en

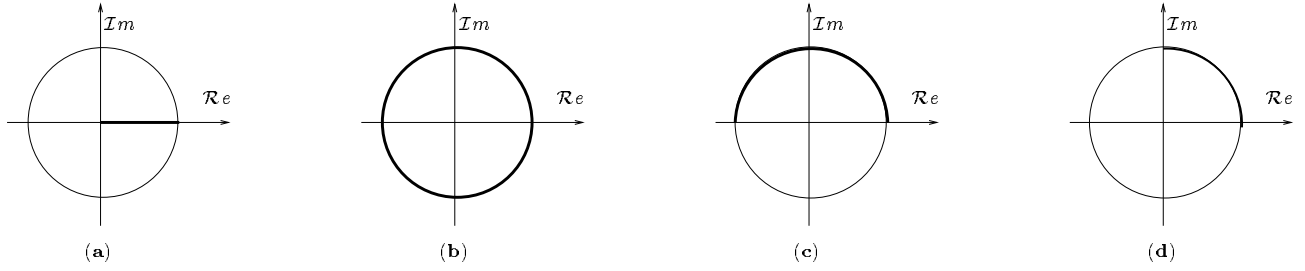


FIG. 6.4: Domaine de codage sur les SLM, **(a)** : Codage en amplitude pure, **(b)** : Codage en phase pure $k = 2\pi$. **(c)** : Codage en phase pure $k = \pi$. **(d)** : Codage en phase pure $k = \pi/2$.

phase.

L'expression de l'intensité de chaque pixel de la scène codée se déduit aisément de ces schémas : Pour un codage en amplitude pure, la scène codée prend alors la forme :

$$s_{amp}(x) = \frac{s(x)}{s_{max}} \quad (6.6)$$

avec s_{max} l'intensité maximale trouvée dans la scène \mathbf{s} .

Dans le cas du codage en phase, on obtient :

$$s_k(x) = \exp \left[i k \frac{s(x)}{s_{max}} \right] \quad (6.7)$$

avec $i = \sqrt{-1}$.

Le choix d'un type de codage particulier a été l'objet de nombreuses études ces dernières années. Le codage en phase offre l'avantage évident de fournir une meilleure efficacité optique que le codage d'amplitude [32], et cette propriété est la raison pour laquelle il est couramment utilisé dans les montages optiques.

Parallèlement, la communauté du traitement du signal a commencé à s'intéresser à ce problème en considérant la modification de phase résiduelle apparaissant souvent dans les modulateurs d'amplitude pure [24, 25, 23, 34, 35, 47]. En considérant un modèle de bruit additif, il a été démontré que la modulation de phase résiduelle tend à rendre les pics de corrélations linéaires plus étroits [24], mais à dégrader le rapport signal sur bruit en sortie du filtre [24, 35, 47].

Par ailleurs, quelques études récentes sur les corrélateurs linéaires ont montré que lorsque le bruit affectant l'image est non-recouvrant, le codage en phase peut conduire à une meilleure robustesse aux changements d'illumination de la scène [55], une meilleure capacité de discrimination [56], et rend le filtre moins sensible à un faible contraste entre la cible et le fond [56, 7].

Nous commencerons par vérifier ce dernier point dans le cas du NLJTC. De plus, pour définir une borne supérieure aux performances du NLJTC, nous comparerons ses performances avec celles du GLRT dans le cas de bruit non recouvrant, pour lequel le GLRT est optimal.

6.2 Comparaison entre le GLRT et le NLJTC

Il s'agit ici de mettre en évidence l'intérêt que peut avoir le NLJTC en tant qu'algorithme de localisation. Pour cela, nous utilisons ici des scènes synthétiques, dans lesquelles nous maîtrisons les paramètres de chacune des régions, ainsi que la position exacte de la cible. Le modèle d'images utilisé pour synthétiser du bruit non recouvrant est le modèle SIR de la section 1.1. Les cibles utilisées sont des carrés de taille variable. La région cible et la région fond sont des réalisations de champs aléatoires Gamma décorrélés, indépendants entre eux. Encore une fois, nous supposons que seule la forme de l'objet est connue, ce qui signifie que la référence employée pour localiser l'objet est le masque de la cible \mathbf{w} .

L'algorithme optimal pour ces scènes est l'algorithme GLRT que nous avons présenté à la section 1.3.3, dont nous redonnons ici l'expression :

$$\mathcal{R}^{inc}(x) = -N_a \log \left(\frac{1}{N_a} [\mathbf{s} \otimes \mathbf{w}](x) \right) - N_b \log \left(\frac{1}{N_b} [\mathbf{s} \otimes \bar{\mathbf{w}}](x) \right) + N_F \log \left(\frac{1}{N_F} [\mathbf{s} \otimes \mathbf{F}](x) \right) \quad (6.8)$$

où N_a est le nombre de pixels de la cible, N_F le nombre de pixels de la fenêtre d'analyse et $N_b = N_F - N_a$.

La forme générique du filtre réalisé par le NLJTC est celle de l'équation 6.9

$$h_\nu^\sigma = \frac{\hat{w}_\nu}{\sigma^2 + |\hat{w}_\nu|^2 + |\hat{s}_\nu|^2} \quad (6.9)$$

Cette équation fait apparaître un paramètre σ^2 . Pour un filtre numérique, ce paramètre peut être défini comme un paramètre de régularisation, permettant de limiter l'instabilité du filtre [43, 42]. Dans un montage optique, la valeur de ce paramètre dépend des caractéristiques du montage (cristal non-linéaire, puissance des lasers) [40]. Nous avons fixé cette valeur à 1.

Nous ne comparons ici le GLRT qu'avec le NLJTC utilisant un codage en amplitude et le NLJTC utilisant un codage en phase pour une profondeur de modulation $k = 2\pi$. Les autres profondeurs de modulation seront analysées plus loin. Les performances des algorithmes sont évaluées par leur probabilité de localisation correcte, qui correspond à la probabilité que le maximum du plan de sortie corresponde à la position réelle de la cible.

Nous nous sommes intéressés au comportement des différents filtres, en fonction du contraste existant entre la cible et le fond. Pour étudier ce comportement, nous avons fixé la moyenne du fond à 1, et la moyenne de la cible est supposée variable entre 0 et 1. Ceci permet de passer d'une localisation facile (pour une moyenne nulle, la cible est une zone homogène noire), à une localisation impossible (pour une moyenne identique à celle du fond, il est impossible de distinguer la cible du fond, les champs aléatoires présentant les mêmes statistiques). Enfin, pour chaque moyenne de la cible, nous avons estimé la probabilité de localisation associée à chaque filtre sur 100 réalisations des scènes. Nous obtenons ainsi les courbes de la figure 6.5.

Ces courbes confirment que le GLRT présente des performances meilleures que celles du NLJTC, quel que soit le contraste, et pour l'ensemble des cibles testées. On remarque également que cette différence est la plus importante pour des probabilités de localisations comprises entre 20 et 60 pour cent, qui correspondent à des niveaux de difficulté importants.

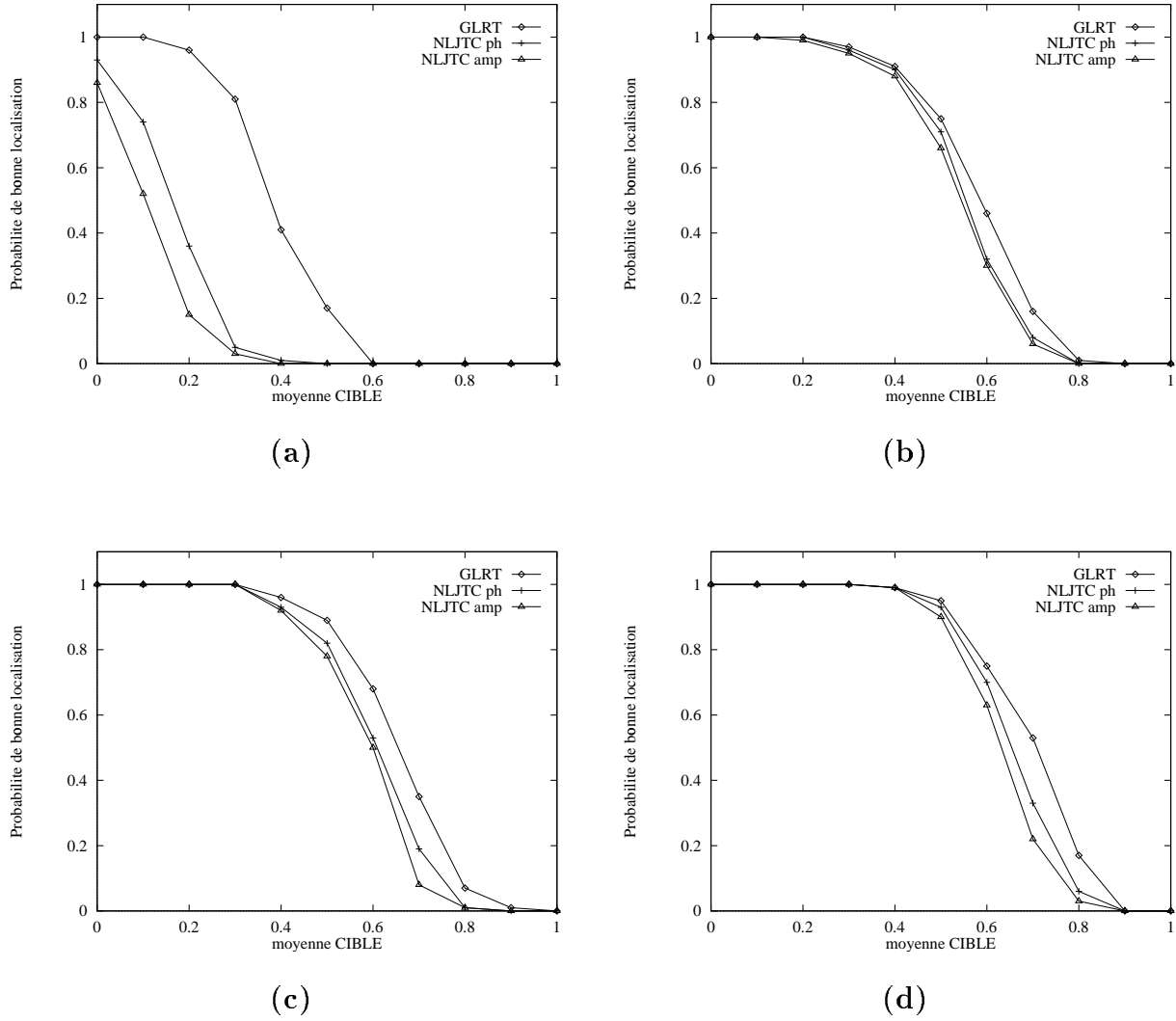


FIG. 6.5: Comparaison des performances en localisation du GLRT (\diamond) et du NLJTC. Le NLJTC peut utiliser une scène codée en phase (+) ou en amplitude (\triangle). Les scènes sont composées de champs aléatoires Gamma décorrelés. Le fond a une moyenne fixée à 1 et la moyenne de la cible est variable. Les probabilités de localisation sont estimées sur 100 réalisations des scènes. **(a)** : cible carrée de taille 5×5 pixels. **(b)** : cible carrée de taille 11×11 pixels. **(c)** : cible carrée de taille 15×15 pixels. **(d)** : cible carrée de taille 25×25 pixels.

La différence la plus importante est obtenue pour la cible la plus petite. On observe ainsi un écart maximal de quatre vingt pour cent en faveur du GLRT pour une moyenne de cible de 0.3. Ceci confirme que l'utilisation d'un algorithme optimal prend tout son sens lorsque le traitement à réaliser est le plus difficile. En revanche, cette différence s'amenuise rapidement pour des cibles plus grandes. La différence la plus importante pour les trois autres cibles n'est plus que de l'ordre de trente pour cent.

On peut également noter une différence légère entre les performances obtenues lorsque la scène à l'entrée du NLJTC est codée en phase ou en amplitude. Le codage en phase semble améliorer systématiquement les performances de l'algorithme.

Notons qu'ici, le GLRT était utilisé dans son cadre d'optimalité. De plus, dans le cas de bruits décorrélés tels que ceux que nous avons utilisés pour ces simulations, la propriété d'adaptativité du NLJTC n'est pas pleinement exploitée. Excepté dans le cas de cibles très petites, le NLJTC, dont la mise en oeuvre optique est déjà opérationnelle et qui présente des performances comparables à l'algorithme optimal apparaît donc une solution intéressante au problème de la localisation d'objets par un corrélateur optique.

Nous avons pu remarquer que le choix d'un codage en phase permettait d'augmenter la probabilité de bonne localisation lorsque la moyenne de la cible est proche de celle du fond, conformément à ce qui a été observé pour les filtres linéaires [7]. Dans la section suivante, nous fournirons une interprétation phénoménologique de ce phénomène et montrerons que sous certaines conditions, le codage en phase peut rendre les corrélateurs optiques moins sensibles à une faible différence de moyenne entre la cible et le fond.

6.3 Choix d'un codage

6.3.1 Position du problème

La scène $\mathbf{s} = \{s(x), x \in [0, N - 1]\}$ est supposée contenir un objet de niveaux de gris aléatoires $t(x)$ et est perturbée par un bruit non recouvrant :

$$s(x) = a w(x - x_0) t(x - x_0) + b(x) \bar{w}(x - x_0) \quad (6.10)$$

x_0 est la position vraie de la cible. w est le support de la cible et prend la forme d'un masque binaire de valeur 1 si le pixel considéré appartient à la cible, et 0 sinon. Les champs \mathbf{t} et \mathbf{b} sont supposés homogènes et décorrélés.

Dans ces scènes, nous aurons plus loin besoin d'une définition du contraste. Nous choisirons ici comme définition du contraste entre la cible et le fond la grandeur χ :

$$\chi = a \langle t(x) \rangle - \langle b(x) \rangle \quad (6.11)$$

où $\langle \rangle$ représente l'opérateur de moyenne d'ensemble.

Un estimateur de χ à partir de réalisations de \mathbf{t} et \mathbf{b} s'écrit sous la forme :

$$\chi_e = \frac{1}{N_{\mathbf{w}}} \sum_{x \in \mathbf{w}} a t(x) - \frac{1}{N_{\bar{\mathbf{w}}}} \sum_{x \in \bar{\mathbf{w}}} b(x). \quad (6.12)$$

où N_w est le nombre de pixels de la cible et $N_{\bar{w}}$ le nombre de pixels du fond. Une définition plus correcte de cette grandeur serait la "différence de niveaux de gris moyens". Nous conserverons l'appellation de contraste pour des raisons de simplicité.

Encore une fois, nous supposons que seule la forme de l'objet est connue, ce qui signifie que la référence employée pour la corrélation est le masque de la cible \mathbf{w} . Quel que soit le type de codage utilisé, cette référence sera codée en phase, avec une profondeur de modulation $l = \pi$. Ce choix n'est pas primordial. En effet, la référence étant binaire, sa version codée est également binaire, et n'est qu'une transformation affine de l'original, ce qui ne modifie pas le comportement général du filtrage.

Nous utilisons l'algorithme NLJTC en localisation, ce qui revient à choisir comme position de la cible le maximum du plan de sortie dont la transformée de Fourier a été donnée dans l'équation 6.9.

Critères de qualité

Caractériser les performances d'un filtre signifie choisir un critère d'évaluation. Dans le cas de la localisation, le critère le plus évident consiste simplement à observer si le maximum du plan de sortie se situe effectivement à la position de la cible.

Cependant, ce critère doit souvent être complété par d'autres mesures, notamment si le système doit être capable d'effectuer une discrimination entre plusieurs objets. En effet, si l'on prend l'exemple du filtre adapté, ses capacités en localisation sont bonnes, mais il présente de très mauvaises capacités de discrimination. Cela veut dire que s'il existe deux objets de forme voisine O_1 et O_2 dans la scène, le filtre adapté à O_1 produira un pic de corrélation non négligeable avec O_2 . Si la surface et la luminance de O_2 sont supérieures à celles de O_1 , ce pic pourra même atteindre une valeur supérieure à celle du pic produit par O_1 : le filtre adapté est incapable de discriminer les deux objets.

Pour mesurer la capacité de discrimination d'un filtre, le plus simple est de mesurer la finesse du pic de corrélation en sortie du filtre. En effet, on peut montrer que plus ce pic est fin, plus le filtre est discriminant envers les objets de formes différentes de la référence. Le critère classiquement utilisé est le rapport entre le carré de la valeur du maximum du pic de corrélation et l'énergie de la fonction de corrélation [61]. Ce critère est appelé PCE, pour *Peak to Correlation Energy*,

$$PCE = \frac{\max_x |C(x)|^2}{\sum_{x=0}^{N-1} |C(x)|^2} \quad (6.13)$$

Nous utiliserons plus loin ces deux critères de façon conjointe pour caractériser les performances des différents types de codage.

6.3.2 Exemple

Nous voulons étudier l'influence du type de codage sur le comportement du filtre réalisé par un NLJTC. A titre d'exemple, nous considérons une scène dans laquelle les niveaux de gris du fond sont une réalisation d'un bruit décorrélaté gaussien de moyenne 1 et d'écart type 0.5. Les niveaux de gris de la cible sont également distribués selon une loi gaussienne, de moyenne a , et d'écart type 0.1.

La figure 6.6 montre un exemple de ce type de scènes lorsque la moyenne de la cible est égale à celle du fond. Notons que dans cette scène, les moyennes de la cible et du fond sont égales. Cependant, ces deux régions se distinguent par leurs variances respectives et un bon algorithme devrait donc pouvoir localiser la cible.

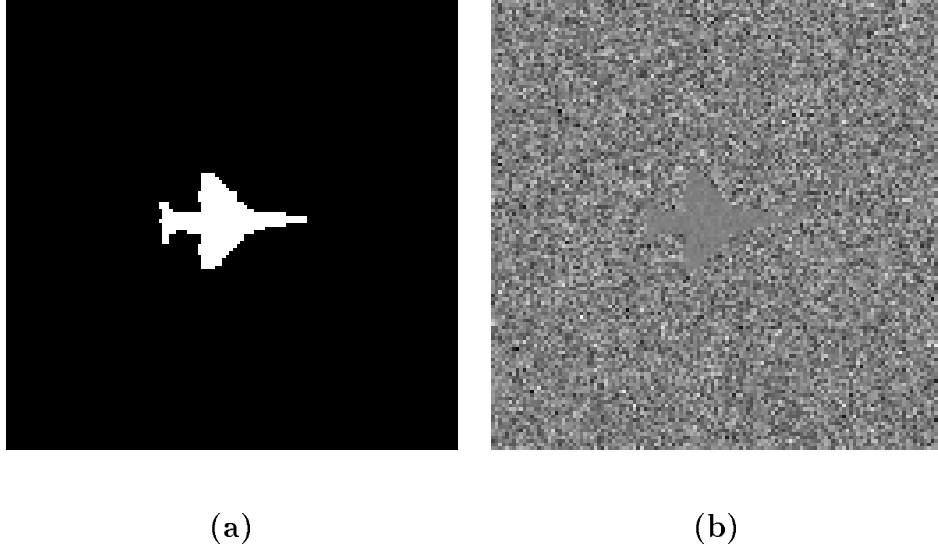


FIG. 6.6: **(a)** : Support de la cible \mathbf{w} . **(b)** : Exemple de scène \mathbf{s} . Les niveaux de gris de la cible sont distribués selon une loi gaussienne de moyenne 1 et de variance 0.1. Les niveaux de gris du fond sont distribués selon une loi gaussienne de moyenne 1 et de variance 0.5. L'image a une taille de 128×128 pixels.

Nous illustrons le comportement du NLJTC sur ce type de scène lorsque la moyenne de la cible varie entre 0 et 2. Les différents codages testés sont : le codage en amplitude, et les codages en phase pour des profondeurs de modulation différentes ($k = \pi/2, \pi$ et 2π). Pour ces différentes profondeurs de modulation, les niveaux de gris de la scène codée sont donnés par :

$$s_k(x) = \exp \left[i k \frac{s(x)}{s_{max}} \right] \quad (6.14)$$

avec s_{max} l'intensité maximale trouvée dans la scène \mathbf{s} et $i = \sqrt{-1}$.

Pour chaque moyenne de la cible et chaque profondeur de modulation k , nous calculons la valeur obtenue du critère PCE (cf. équation 6.13). Nous obtenons également une fonction binaire prenant la valeur 1 si la cible a été correctement localisée (si le maximum du plan de sortie correspond exactement au centre de la cible), et 0 sinon. Ces résultats sont présentés dans la figure 6.7.

On peut remarquer que le codage en amplitude ne permet pas de localiser la cible lorsque la moyenne de la cible est proche de celle du fond ($a = 1$). Ceci est dû à la nature du bruit non-recouvrant présent dans la scène. Dans le cas de bruit additif de moyenne m , la scène \mathbf{s} s'écrit :

$$s(x) = t(x) + m + n(x) \quad (6.15)$$

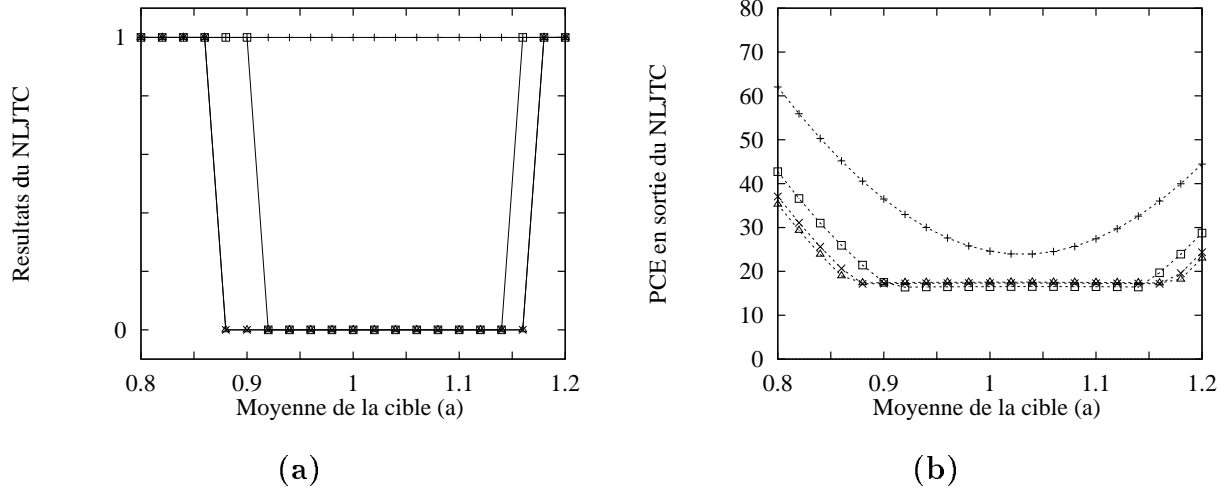


FIG. 6.7: Résultats en localisation obtenus par le NLJTC en fonction de la moyenne de la cible a , pour les différents types de codage. **(a)** : fonction binaire représentant le succès (1) ou l'échec (0) de la localisation. **(b)** : Valeur du critère PCE en sortie du filtre. (◇) : codage en amplitude; (×) : codage en phase, $k = \pi/2$; (□) : codage en phase, $k = \pi$; (+) : codage en phase, $k = 2\pi$.

En appliquant un filtre linéaire \mathbf{h} à cette scène, le plan de corrélation s'écrit :

$$C(x) = [\mathbf{h} \otimes \mathbf{t}](x) + m \sum_{i=0}^N h(i) + [\mathbf{h} \otimes \mathbf{t}](x) \quad (6.16)$$

Le seul terme de l'équation 6.16 dépendant de la moyenne du bruit est $m \sum_{i=0}^N h(i)$. Ne dépendant pas de la position testée x , ce terme ne fait qu'augmenter de façon uniforme le plan de corrélation. Les performances en localisation sont donc indépendantes de la moyenne du bruit m .

Dans le cas de bruit non-recouvrant, il existe toute une plage de valeurs de cette moyenne pour lesquelles les filtres linéaires sont mis en échec. L'existence de cette "zone interdite" pour les filtres linéaires a été analysée [16] et expliquée comme suit : une scène de bruit non recouvrant comme celle présentée dans l'équation 6.10 peut se mettre sous la forme suivante :

$$s(x) = a(x)w(x - x_0)t(x - x_0) + (m + n(x))\bar{w}(x - x_0) \quad (6.17)$$

Le plan de corrélation de cette scène par un filtre linéaire \mathbf{h} s'écrit :

$$\begin{aligned} C(x) &= [\mathbf{h} \otimes \mathbf{t}](x - x_0) + m [\mathbf{h} \otimes \bar{\mathbf{w}}](x - x_0) + [\mathbf{h} \otimes \mathbf{n}^{\bar{\mathbf{w}}}](x - x_0) \\ &= [\mathbf{h} \otimes (\mathbf{t} - m\bar{\mathbf{w}})](x - x_0) + [\mathbf{h} \otimes \mathbf{n}^{\bar{\mathbf{w}}}](x - x_0) + m \sum_{i=0}^{N-1} h(x) \end{aligned} \quad (6.18)$$

ou $\mathbf{n}^{\bar{\mathbf{w}}}$ est défini par $n^{\bar{\mathbf{w}}}(x) = n(x - x_0)\bar{\mathbf{w}}(x - x_0)$. Le bruit \mathbf{n} étant de moyenne nulle, le premier terme du membre de droite, pour $x = x_0$ représente la valeur du pic central

de corrélation, aux fluctuations statistiques prêt. Cette valeur peut s'annuler pour $m = [\mathbf{h} \otimes \mathbf{t}](0)/[\mathbf{h} \otimes \mathbf{w}](0)$, rendant la localisation impossible. De plus, pour avoir une bonne localisation, il faut que la valeur du pic de corrélation soit supérieure à la valeur maximale de la corrélation du filtre avec le fond, ce qui élargit l'ensemble des valeurs "interdites" de m .

Un codage en amplitude conduit à l'échec de la localisation lorsque la moyenne de la cible est proche de celle du fond. De ce point de vue, le NLJTC se comporte comme un filtre linéaire.

En revanche, les courbes de la figure 6.7 montrent une augmentation des performances en localisation avec l'augmentation de k , la profondeur de modulation, dans le cas du codage en phase. Pour $k = 2\pi$, le NLJTC permet de localiser la cible quel que soit le contraste entre la cible et le fond. Le succès de la localisation lorsque la cible et le fond ont la même moyenne implique que le NLJTC utilise la différence de variance existant entre la cible et le fond pour localiser la cible. C'est ce phénomène que nous désirons analyser plus en détail et qui fait l'objet de la section suivante.

6.3.3 Etude de l'influence du codage en phase en présence de bruit non recouvrant

Comme on l'a dit, la transformation opérée par le NLJTC peut être vue comme la corrélation linéaire entre la scène à traiter et un filtre adaptatif dont l'expression dépend de la scène à traiter (cf. équation 6.3). Les différences de performances observées ci-dessus (figure 6.7) peuvent donc avoir deux origines.

- Le codage en phase modifie l'expression du filtre et le rend plus efficace.
- Le codage en phase modifie la scène à traiter et rend celle-ci plus "facile".

Nous allons successivement examiner ces deux hypothèses, et montrerons ainsi que les différences de performances prennent leur source dans les modifications apportées à la scène.

Dans le cas du codage en amplitude, le plan de corrélation $C(x)$ peut être vu comme la corrélation de la scène \mathbf{s} avec un filtre h dont la transformée de Fourier $\hat{\mathbf{h}}$ est :

$$\hat{h}(\nu) = \frac{\hat{r}_\nu}{\sigma^2 + |\hat{r}_\nu|^2 + |\hat{s}_\nu|^2} \quad (6.19)$$

On peut démontrer facilement que la réponse impulsionnelle de ce filtre est réelle. Le plan de sortie est donc donné par :

$$|C(x)|^2 = [h(x) \otimes s(x)]^2 \quad (6.20)$$

\otimes représentant l'opérateur de corrélation.

Considérons maintenant l'action du NLJTC dans le cas de scènes codées en phase \mathbf{s}_k , pour les différentes valeurs de k . La référence est codée avec une même profondeur de modulation l , quelle que soit la valeur de k . Ici encore, le filtre agit comme la corrélation de la scène \mathbf{s}_k avec un filtre \mathbf{h}_{kl} . La transformée de Fourier $\hat{\mathbf{h}}_{kl}$ de ce filtre a l'expression suivante

$$\hat{h}_{kl}(\nu) = \frac{\hat{r}_l(\nu)}{\sigma^2 + |\hat{r}_l(\nu)|^2 + |\hat{s}_k(\nu)|^2} \quad (6.21)$$

Comme nous l'avons dit, la référence utilisée pour le filtrage est binaire et codée avec une profondeur de modulation $l = \pi$. On en déduit :

$$r_{\pi}^R(x) = \cos[\pi w(x)] = 1 - 2w(x) \quad (6.22)$$

$$r_{\pi}^I(x) = \sin[\pi w(x)] = 0 \quad (6.23)$$

L'expression du filtre $\mathbf{h}_{k,\pi}$ est donc explicitable par sa transformée de Fourier :

$$\hat{h}_{k,\pi}(\nu) = \frac{\delta(\nu) - 2\hat{w}(\nu)}{\sigma^2 + |\delta(\nu) - 2\hat{w}(\nu)|^2 + |\hat{s}_k(\nu)|^2} \quad (6.24)$$

L'expression de ce filtre est très similaire à celle du filtre obtenu pour un codage en amplitude. Les seules différences sont les transformations de $|\hat{w}(\nu)|^2$ en $|\delta(\nu) - 2\hat{w}(\nu)|^2$ et de $|\hat{s}(\nu)|^2$ en $|\hat{s}_k(\nu)|^2$.

Nous allons montrer que cette transformation n'affecte pas le comportement du filtre lorsque le contraste entre la cible et le fond est proche de zéro. Pour cela, nous reprenons la scène codée en amplitude de la figure 6.6, que nous corrélons avec les filtres obtenus pour un codage en amplitude et un codage en phase avec $k = 2\pi$. Le plan de sortie de ces filtres est présenté dans la figure 6.8. Dans les deux cas, la localisation conduit à un échec. En d'autres termes, les différents filtres ne se différencient pas par leur aptitude à utiliser l'information de variance pour localiser la cible.

La différence de performances observée dans la figure 6.7 n'est donc pas due au changement de filtre. Cette différence ne peut alors provenir que d'une transformation des scènes sur lesquelles ces filtres sont appliqués.

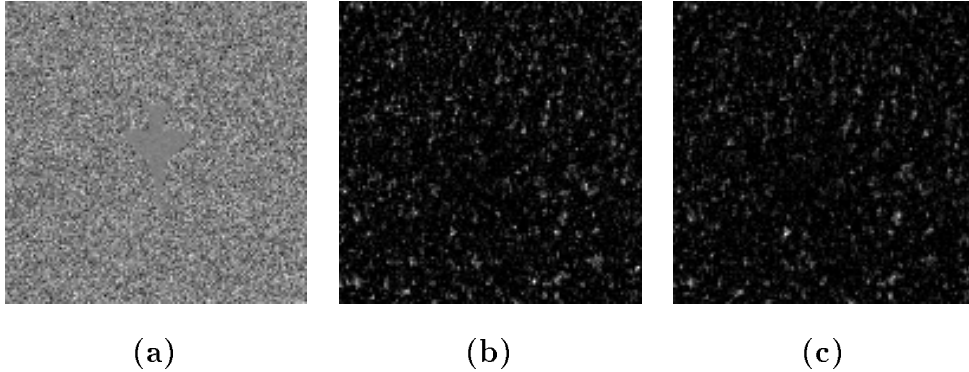


FIG. 6.8: La transformation du filtre \mathbf{h} en $\mathbf{h}_{k,\pi}$ (cf. équations 6.19 et 6.24) lors d'un changement de codage n'explique pas la différence de performances. **(a)** : Scène présentée s , codée en amplitude. **(b)** : Résultat de la corrélation par le filtre \mathbf{h} **(c)** : Résultat de la corrélation par le filtre $\mathbf{h}_{2\pi,\pi}$

Modifications du contraste dans les scènes codées

Si l'on reprend l'expression du filtre \mathbf{h}_{kl} de l'équation 6.24, on peut montrer que sa réponse impulsionnelle est réelle. Nous notons par ailleurs $\mathbf{s}_{\mathbf{k}}^R$ et $\mathbf{s}_{\mathbf{k}}^I$, les parties réelles et

imaginaires de la scène codée en phase. On peut exprimer le plan de corrélation $C(x)$ sous la forme :

$$|C(x)|^2 = [h_{kl}(x) \otimes s_k^R(x)]^2 + [h_{kl}(x) \otimes s_k^I(x)]^2 \quad (6.25)$$

Nous voyons ici apparaître deux corrélations différentes, combinées de façon non-linéaire. Dès lors, nous pouvons supposer que la différence de performances observée dans la figure 6.7 est due à la présence de ces deux corrélations distinctes. Nous présentons le résultat de ces différentes corrélations dans la figure 6.9.

Modifications d'histogrammes

La figure 6.9 permet de remarquer que la cible n'apparaît pas avec le même contraste dans les différentes versions codées de la scène. C'est sur cette observation que nous allons bâtir notre interprétation. Nous nous intéressons donc maintenant à l'expression de ce contraste dans les images \mathbf{s}_k^R et \mathbf{s}_k^I .

Par définition du codage en phase, ces images sont respectivement obtenues en prenant le cosinus et le sinus de l'intensité de chaque pixel de l'image originale.

$$\begin{aligned} s_k^R(x) &= \cos \left[k \frac{s(x)}{s_{max}} \right] \\ s_k^I(x) &= \sin \left[k \frac{s(x)}{s_{max}} \right] \end{aligned}$$

En tenant compte du fait que le bruit est non-recouvrant, on peut montrer :

$$s_k^R(x) = \cos \left[k \frac{a \, t(x)}{s_{max}} \right] w(x - x_0) + \cos \left[k \frac{b(x)}{s_{max}} \right] \bar{w}(x - x_0) \quad (6.26)$$

$$s_k^I(x) = \sin \left[k \frac{a \, t(x)}{s_{max}} \right] w(x - x_0) + \sin \left[k \frac{b(x)}{s_{max}} \right] \bar{w}(x - x_0) \quad (6.27)$$

En reprenant l'expression du contraste de l'équation 6.11, on peut exprimer le contraste dans les parties réelles et imaginaires de la scène codée en phase. Celui ci est donné par :

$$\chi^R = \left\langle \cos \left[k \frac{a \, t(x)}{s_{max}} \right] \right\rangle - \left\langle \cos \left[k \frac{b(x)}{s_{max}} \right] \right\rangle \quad (6.28)$$

$$\chi^I = \left\langle \sin \left[k \frac{a \, t(x)}{s_{max}} \right] \right\rangle - \left\langle \sin \left[k \frac{b(x)}{s_{max}} \right] \right\rangle \quad (6.29)$$

Les équations 6.26 et 6.27 montrent que les images codées en phase ont subi une transformation d'histogramme par rapport à l'image originale. Ces modifications sont non-linéaires, de la forme d'un cosinus et d'un sinus. La modification d'histogrammes entraîne un changement des valeurs moyennes à l'intérieur et à l'extérieur de la cible. Si les distributions statistiques sont différentes entre ces deux zones, le contraste χ^R et le contraste χ^I sont sensiblement modifiés par rapport au contraste original χ .

Nous illustrons ce phénomène dans la figure 6.10. Dans cette figure, nous avons tracé l'histogramme de la scène présentée dans la figure 6.6, pour les régions cibles et fond. Dans

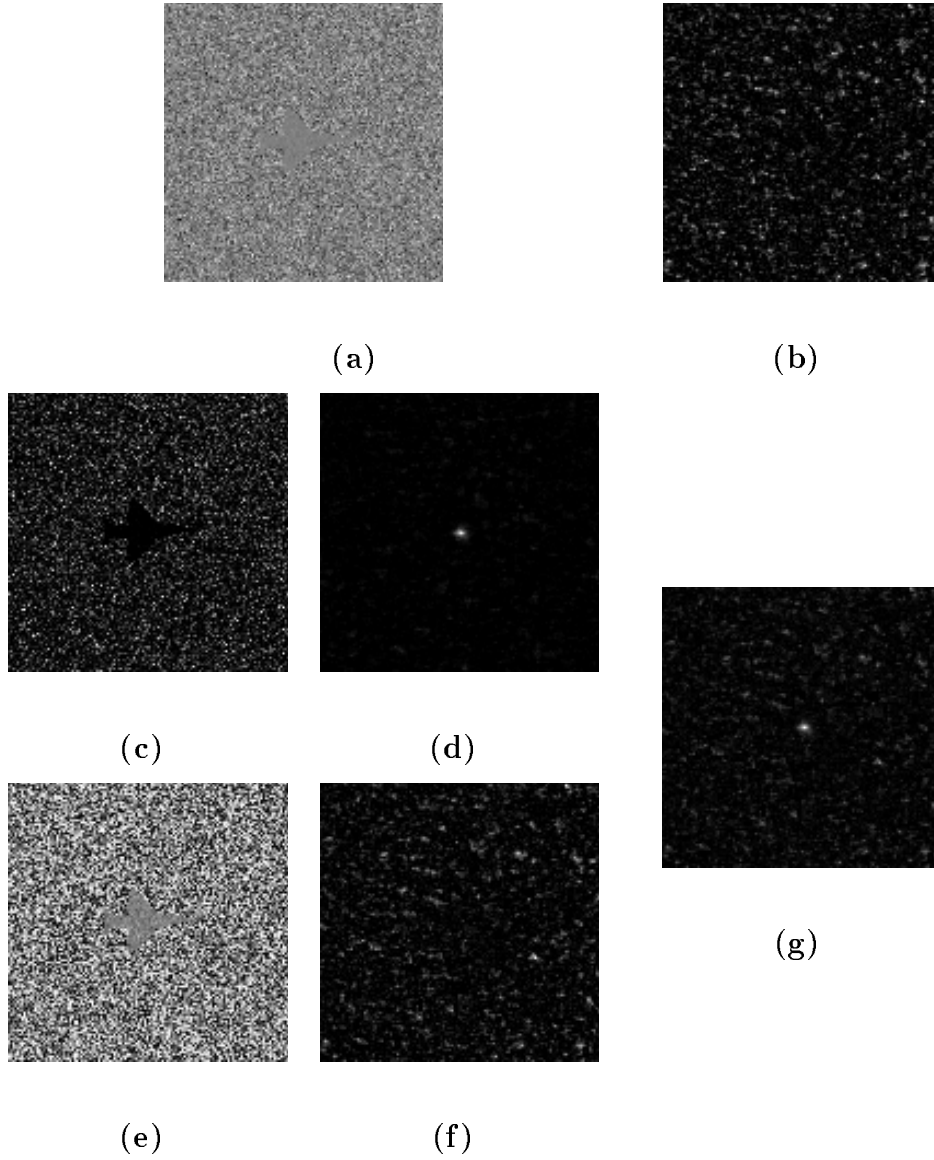


FIG. 6.9: Scènes codées et résultats du NLJTC.

(a) : Scène codée en amplitude.

(b) : Résultats du NLJTC sur (a).

(c) : Partie réelle de la scène codée en phase pour $k = 2\pi (s_k^R(x))$.

(d) : Résultats du NLJTC sur (c) $([h_{k\pi}(x) \otimes s_k^R(x)]^2)$.

(e) : Partie Imaginaire de la scène codée en phase pour $k = 2\pi (s_k^I(x))$.

(f) : Résultats du NLJTC sur (f) $([h_{k\pi}^R(x) \otimes s_k^I(x)]^2)$.

(g) : Résultat du NLJTC sur l'ensemble de la scène codée en phase $(|C(x)|^2, \text{ cf. Eq. 6.25})$.

cette scène, le contraste entre la cible et le fond est nul, les moyennes de ces deux régions étant identiques. Nous présentons également ces mêmes histogrammes, obtenus sur les parties réelles et imaginaires de la scène codée en phase, pour $k = 2\pi$. On peut remarquer que dans l'image $s_{2\pi}^R(x)$, les valeurs moyennes de la cible et du fond sont maintenant différentes, contrairement à ce qui se produit dans l'image codée en amplitude.

Le NLJTC, conçu avec une référence uniforme est principalement sensible à une différence de moyenne entre les régions. C'est donc bien la présence de la corrélation avec $s_{2\pi}^R(x)$ (cf. équation 6.25) qui explique l'efficacité du codage en phase que nous avons pu observer précédemment.

Nous nous intéressons ensuite au cas où le contraste dans la scène originale n'est pas nul ($a \neq 1$). Pour une scène codée en amplitude, le contraste χ_e (cf. équation 6.12) exprimé en fonction de a est une droite de pente 1 qui passe par 0 pour $a = 1$. Nous voulons observer ce que devient ce contraste dans les scènes codées en phase. La figure 6.11 représente l'évolution de ce contraste en fonction de a , dans les parties réelles et imaginaires de la scène codée en phase. Nous avons représenté cette évolution pour les trois différentes profondeurs de modulation possibles.

Pour $k = \pi/2$ et $k = \pi$, le contraste devient quasi nul simultanément dans $s_k^R(x)$ et $s_k^I(x)$, pour a proche de 1, de la même manière que pour un codage en amplitude. Ceci explique les résultats observés dans la figure 6.7, dans laquelle ces trois types de codages se comportaient de façon similaire.

En revanche, pour $k = 2\pi$, on peut remarquer que le contraste reste non nul, soit dans la partie réelle, soit dans la partie imaginaire de la scène codée en phase, quelle que soit la valeur de a . Ceci explique l'absence de zone d'échec observée pour ce codage dans la figure 6.7.

Nous pouvons conclure de cette étude les points suivants : le codage en phase peut permettre d'améliorer les performances du NLJTC en présence de bruit non-recouvrant. En effet, une différence de forme des histogrammes entre la cible et le fond (ici une différence de variance) est transformée par le codage en phase en une différence de moyenne. Cette différence de moyenne est la principale différence à laquelle le NLJTC est sensible lorsque les niveaux de gris de la cible sont inconnus. De plus, cette différence de moyenne est la plus importante pour un codage en phase de profondeur de modulation $k = 2\pi$. Ce codage est donc le plus performant.

Nous avons considéré des bruits gaussiens non corrélés pour modéliser les niveaux de gris de la cible et du fond. Bien entendu, les résultats peuvent différer sensiblement si l'on considère d'autres types de cibles et de fonds. En particulier, si la scène est composée d'une cible uniforme et bruitée par du bruit additif, les histogrammes de la cible et du fond ont la même forme. Dans ce cas, il a été démontré que le codage en phase dégrade les performances de la corrélation [24, 34, 47]. La différence avec les conclusions auxquelles nous sommes arrivés réside bien dans la nature du bruit qui est non-recouvrant.

Ce type de bruit correspond relativement bien à la modélisation d'images réelles. Il y a tout lieu de penser que dans ce cas également, le codage en phase permet de transformer une différence de forme d'histogrammes en une différence de moyenne, permettant d'obtenir des meilleures performances que celles d'un codage en amplitude. C'est ce que nous avons voulu vérifier dans la section suivante.

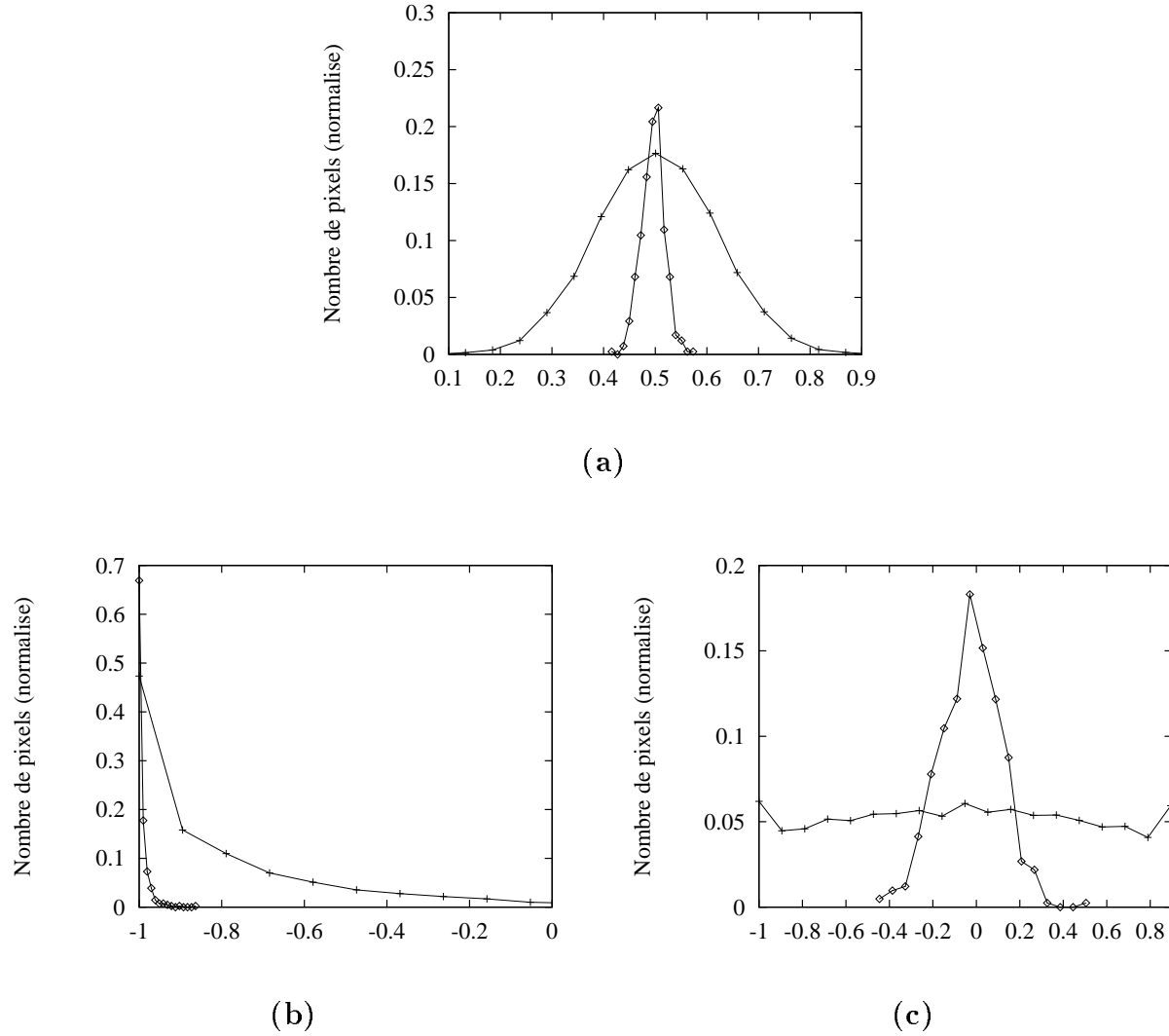


FIG. 6.10: Influence du codage sur le contraste : Histogrammes normalisés dans la cible (\diamond) et dans le fond (+), mesurés respectivement dans **(a)** : \mathbf{s} (scène codée en amplitude). **(b)** : $s_{2\pi}^R(x)$ (partie réelle de l'image codée en phase). **(c)** : $s_{2\pi}^I(x)$ (partie imaginaire de l'image codée en phase). La scène \mathbf{s} est celle présentée dans la figure 6.6.

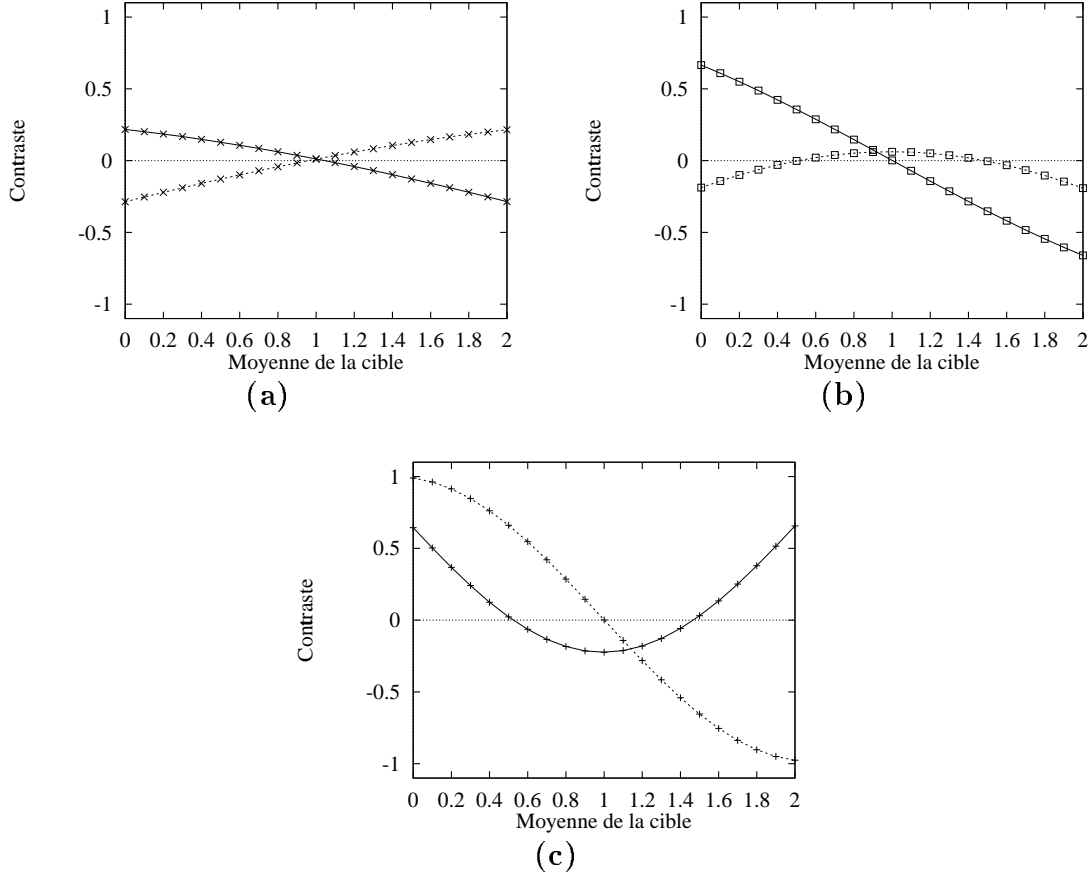


FIG. 6.11: Evolution du contraste χ_e dans la partie réelle (traits pleins) et dans la partie imaginaire (traits pointillés) de la scène codée en phase, en fonction de la moyenne de la cible a pour différentes profondeurs de modulations. (a) : $k = \pi/2$ (b) : $k = \pi$ (c) : $k = 2\pi$.

6.3.4 Illustration dans le cas d'images réalistes

Considérons le fond structuré **b** présenté dans la figure 6.12.a, ainsi que la cible **t** de la figure 6.12.b.

Les scènes à traiter sont obtenues par superposition de la cible sur le fond, au facteur d'illumination a prêt pour la cible, comme défini dans l'équation 6.10.

Nous pouvons refaire la même expérience que précédemment afin d'observer les performances en localisation du NLJTC pour les différents codages de la scène. Les résultats sont présentés dans la figure 6.13. Notons que le NLJTC utilise ici encore une référence uniforme, les niveaux de gris de la cible étant supposés inconnus. Pour ces types de cibles et de fonds, les modifications d'histogrammes sont plus complexes que dans le cas de bruits gaussiens. Néanmoins, le codage en phase pour $k = 2\pi$ permet de localiser la cible quelle que soit la valeur du paramètre a .

Nous pouvons en déduire que le codage en phase amène un accroissement de la robustesse du NLJTC face à de faibles contrastes, même dans le cas de fonds réalistes. Pour confirmer cette supposition, nous avons réalisé l'expérience suivante :

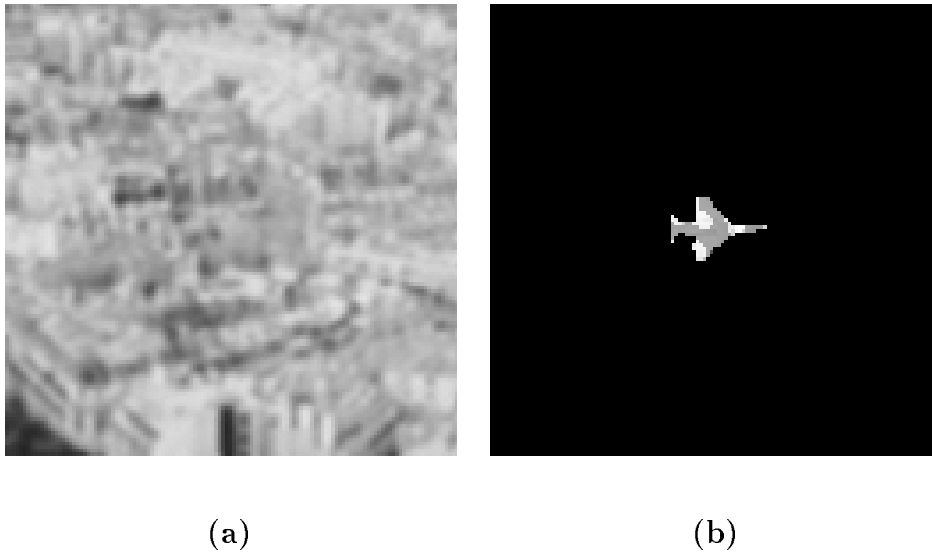


FIG. 6.12: Exemple de fond structuré et de cible contrastée.

Nous avons considéré une image de 761×569 pixels représentant un fond réaliste structuré. Dans cette image, nous avons extrait 100 vignettes de 128×128 pixels, dont la position dans l'image de départ a été choisie de façon aléatoire. Le tirage aléatoire de ces vignettes permet de supposer une grande variété d'histogrammes différents pour les niveaux de gris du fond. Au centre de chacune de ces vignettes, nous avons superposé la cible de la figure 6.12.b dont la moyenne, variable, est égale à a . Pour une valeur donnée de a , en appliquant un filtre sur toutes ces vignettes, on peut mesurer une probabilité de localisation correcte. La localisation est dite correcte si le maximum du plan de sortie correspond exactement à la position du centre de la cible dans la vignette. Encore une fois, nous avons comparé les probabilités de localisation correcte, obtenues par le NLJTC pour différents codages, en fonction de la moyenne a de la cible.

Les résultats de cette expérience sont présentés dans la figure 6.14. Comme précédemment, nous pouvons remarquer que le codage en phase conduit à des résultats sensiblement meilleurs que le codage en amplitude. Par exemple, la probabilité de localisation correcte minimale sur l'ensemble des valeurs de a est de 75% dans le cas du codage en phase ($k = 2\pi$), contre 50% dans le cas du codage en amplitude.

6.4 Conclusion

Dans ce chapitre, nous nous sommes intéressés à l'utilisation de filtres sous optimaux, dont l'intérêt peut être justifié par leur vitesse de traitement lorsqu'ils sont mis en œuvre de façon optique. Nous avons dans un premier temps illustré la perte de performances que l'on pouvait attendre en abandonnant les propriétés d'optimalité du traitement à appliquer. Nous avons ainsi montré que pour de très petites cibles, ces pertes sont très importantes. Pour des cibles de quelques dizaines de pixels de large, cette différence diminue, et le corrélateur optique présente des performances comparables à celles du filtre optimal.

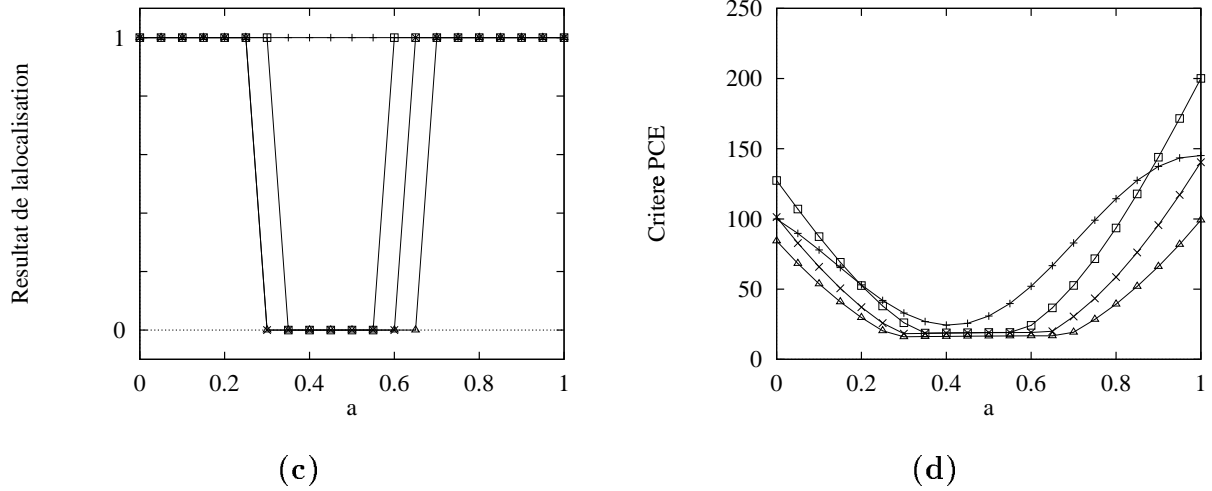


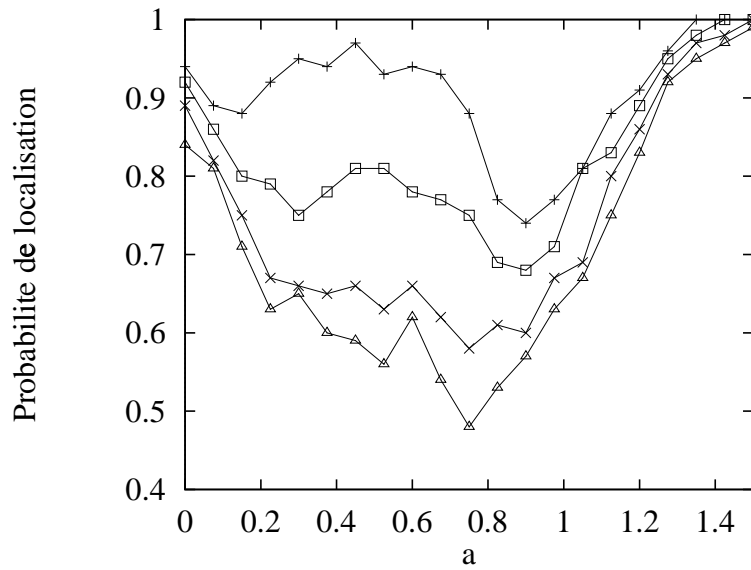
FIG. 6.13: Résultats en localisation à l'aide d'un NLJTC en fonction de la moyenne a de la cible, pour différents types de codage. **(a)** : Fond réaliste (128×128 pixels). **(b)** : Cible. **(c)** : fonction binaire représentant le succès (1) ou l'échec (0) de la localisation. **(d)** : Valeur du critère PCE en sortie du filtre. (\diamond) : Codage en amplitude; (\times) : Codage en phase, $k = \pi/2$; (\square) : Codage en phase, $k = \pi$; (+) : Codage en phase, $k = 2\pi$.

Par ailleurs, une mise en oeuvre optique nécessite de choisir une méthode de codage des scènes à présenter. De nombreuses études ont été faites sur le choix d'un codage. Dans le cas de bruit additif, il a été montré que le codage en amplitude est préférable. Dans le cas de bruit non recouvrant, des études de filtres linéaires ont abouties à des conclusions sensiblement différentes, des résultats expérimentaux ayant montré la supériorité du codage en phase. Cependant, aucune étude n'avait été réalisée concernant les filtres non-linéaires et aucune interprétation de la supériorité du codage en phase sur le codage en amplitude n'avait été fournie.

Nous avons donc effectué des simulations numériques d'un corrélateur optique non-linéaire, permettant de confirmer l'augmentation de performances obtenue à l'aide du codage en phase. Nous avons de plus fourni une interprétation simple de ce résultat, basée sur la modification d'histogrammes qu'induit le codage en phase. Cette modification entraîne un accroissement de la différence de moyenne entre la cible et le fond, qui est le paramètre auquel les filtres sont le plus sensibles. Cette démonstration a été effectuée sur des images composées de fonds et de cibles homogènes et décorrélés, puis généralisée au cas d'images réalistes, dans lesquelles les fonds sont structurés et inhomogènes et les cibles contrastées.



(a)



(b)

FIG. 6.14: Résultats du NLJTC sur une image réaliste. La cible est celle présentée Fig. 6.12.b. **(a)** : Image de fond (761×569 pixels). **(b)** : Probabilité de localisation correcte en fonction de a . La probabilité de localisation correcte est donnée pour les trois types de codage : (Δ) Codage en amplitude; (\times) Codage en phase, $k = \pi/2$; (\square) Codage en phase, $k = \pi$; ($+$) Codage en phase, $k = 2\pi$.

Conclusion

L'objectif de cette thèse était de d'améliorer et d'étendre les techniques de localisation et de détection d'objets de formes connues dans des images fortement bruitées, tout en respectant des contraintes de temps de calculs faibles. Le contexte d'images fortement bruitées nous a conduit à nous tourner vers des solutions basées sur une approche statistique du problème de la localisation et de la détection. Dans ce type d'images, il est en effet nécessaire de disposer d'algorithmes efficaces et dont le support théorique est bien connu. La théorie de la décision nous assure que les algorithmes obtenus par cette approche sont optimaux par rapport au critère choisi, si le modèle utilisé est bien représentatif des problèmes à traiter.

La contrainte du temps de calcul nous a conduit vers un modèle d'image simple. Pour obtenir des algorithmes d'une complexité équivalente à quelques corrélations, nous avons utilisé un modèle de bruit non-recouvrant, représentatif de scènes issues d'imageurs opto-électroniques. Ce modèle suppose l'image constituée de deux parties, la cible et le fond, dans chacun desquels l'intensité des pixels est une réalisation d'un champ aléatoire décorré de densité de probabilité paramétrique connue. En fonction de l'application concernée, la localisation ou la détection, ces modèles d'images conduisent à des algorithmes de type ML (*Maximum Likelihood*) ou GLRT (*Generalised Likelihood Ratio Test*).

Ce dernier est particulièrement intéressant comme en témoignent les nombreux champs d'applications dans lesquels il est utilisé. On le retrouve ainsi en imagerie SAR, et en imagerie hyperspectrale. Cet algorithme a notamment une propriété primordiale en détection : dans le cas d'images composées de bruit gamma ou gaussien, par exemple, il est à taux de fausse alarme constant (TFAC). C'est à dire que sa réponse sur une région homogène est indépendante des paramètres de la densité de probabilité de la région. Le seuil de détection ne dépend donc que de la probabilité de fausse alarme choisie par l'utilisateur. C'est donc sur ce type d'algorithme que nous nous sommes focalisés.

Son application principale consiste à détecter des cibles de forme connue dans un bruit de fond homogène. Il utilise pour cela une fenêtre d'analyse que l'on fait glisser en chaque point de l'image pour déterminer si une cible est présente ou pas au centre de la fenêtre. La taille de cette fenêtre constitue un paramètre indéterminé de l'algorithme. Caractériser son influence a constitué notre premier travail sur le GLRT.

Ensuite, le fil conducteur de notre étude a été d'étudier les différentes extensions possibles de ce filtre. Nous nous sommes intéressés aux trois extensions suivantes :

- cibles de tailles inconnues.
- cibles composées de plusieurs régions.
- mise en oeuvre optique des filtres.

1 Taille de la fenêtre d'analyse

Le calcul des filtres GLRT nécessite la gestion des paramètres des champs aléatoires représentant la cible et le fond dans l'image. Dans le cas de densité de probabilité appartenant à la famille exponentielle (par exemple des bruits gaussien ou gamma), ces paramètres sont peu nombreux (moyenne, variance, ordre). Ils sont nommés paramètres de nuisance, car ils ne nous intéressent pas directement, mais interviennent dans la confection du filtre. Lorsque ces paramètres sont inconnus *a priori*, et estimés au sens du maximum de vraisemblance, la qualité de leur estimation dépend de la taille des régions sur lesquelles on les mesure. La cible étant de taille fixée, la qualité de l'estimation des paramètres sur la zone cible est fixe. En revanche, sur la zone fond, des fenêtres d'analyse de plus en plus grandes permettent d'obtenir une meilleure estimation de ces paramètres et donc des performances meilleures.

Cependant, le gain ainsi obtenu stagne rapidement. La borne supérieure des performances du GLRT en localisation est donnée par les performances de l'algorithme du ML, optimal pour cette tâche. Nous avons montré qu'en présence d'un bruit homogène de la famille exponentielle, à partir d'une certaine taille de fenêtre, les performances du GLRT et du ML deviennent équivalentes. Les fenêtres possédant cette taille critique sont typiquement obtenues en effectuant de cinq à dix dilatations morphologiques de la cible cherchée dans les exemples que nous avons considérés. En bruit homogène, l'algorithme GLRT présente alors des performances très proches de l'optimalité, tout en utilisant des petites fenêtres d'analyse.

Dans le cas d'images où la statistique du fond varie d'un point à un autre, on parle de fond inhomogène. On se situe alors en dehors du modèle pour lequel le ML est optimal. L'algorithme ML, effectuant l'estimation des paramètres du fond sur l'ensemble de l'image, se révèle très sensible à la présence des inhomogénéités. Ceci peut conduire à une localisation erronée de la cible ou à l'apparition de fausses alarmes en détection. En revanche, modéliser le fond apparaissant dans une petite fenêtre d'analyse comme un champ aléatoire homogène est une meilleure approximation du contenu de ces images. En effet, on fait alors seulement l'hypothèse d'une homogénéité locale du fond. L'algorithme GLRT utilisant de petites fenêtres d'analyse est ainsi plus robuste à la présence d'inhomogénéités.

Le choix de la taille de la fenêtre apparaît donc comme un compromis entre l'optimalité du filtre en bruit homogène et la robustesse aux inhomogénéités du fond. Une fenêtre d'analyse obtenue par quelques dilatations morphologiques de la cible cherchée réalise ce compromis de façon satisfaisante.

2 Cibles de tailles inconnues

L'algorithme GLRT permet de localiser et de détecter une cible de forme connue. Par forme connue, on entend en particulier orientation et taille de la cible. Cet algorithme ne permet pas de résoudre les cas où la cible présente dans l'image n'est pas à la même échelle que la référence dont on dispose. Pour résoudre ce problème, nous avons proposé de considérer le paramètre de taille comme un paramètre de nuisance supplémentaire.

Pour cela, on peut utiliser un banc de filtres construits pour des cibles de tailles différentes. Pour résoudre le problème de fusion des sorties des différents filtres, la théorie bayésienne de l'estimation propose plusieurs solutions simples :

- L'estimation de la taille de la cible au sens du maximum de vraisemblance, qui consiste à choisir en chaque position la sortie du filtre qui a donné la réponse la plus forte.
- L'approche du marginal bayésien, qui consiste à sommer les vraisemblances obtenues par chaque filtre, pondérées par un coefficient correspondant à la probabilité d'observer la taille de cible considérée. Cette ensemble de coefficient constitue une information *a priori* que l'on introduit dans l'algorithme.

Il est en général admis que l'approche du marginal Bayésien conduit à des résultats meilleurs que le maximum de vraisemblance. Cependant, elle dépend de la concordance entre l'information *a priori* et la réalité des scènes à traiter. Nous avons voulu juger de la robustesse de chacune des méthodes à la qualité de cette information. En particulier, nous nous sommes intéressés au cas où la connaissance *a priori* est minimale. Dans ce cas, nous supposons que toutes les tailles de cibles sont équiprobables.

Nous avons alors montré que lorsque cette hypothèse est respectée, les deux approches sont équivalentes. En revanche, lorsque toutes les scènes présentées contiennent une cible de taille constante, l'estimation de taille au sens du maximum de vraisemblance permet d'obtenir des meilleures performances en détection que l'approche du marginal bayésien. En d'autres termes, l'estimation du maximum de vraisemblance est plus robuste à la distribution *a priori* des valeurs possibles des tailles de cibles.

De plus, nous avons montré que lorsque la fenêtre d'analyse est placée sur une zone homogène, la distribution statistique de la sortie des différents filtres est identiquement distribuée, et bien définie mathématiquement. Le GLRT pour chaque taille est donc à taux de fausse alarme constant également en fonction de la taille des régions utilisées. Au delà du simple intérêt académique, ceci est primordial d'un point de vue pratique. On peut en effet traiter les différentes sorties de la même façon. En particulier, en détection, on pourra utiliser le même seuil de détection quelle que soit la taille de cible estimée. Le seuil de détection choisi ne dépend ni de la taille des régions, ni de la moyenne des régions considérées, mais uniquement du taux de fausse alarme acceptable par l'utilisateur.

3 Détection de cibles dans des images issues de capteurs réels

Pour juger de l'intérêt de ces méthodes dans des situations réelles, nous avons ensuite considéré des images issues de différents capteurs. Dans les exemples considérés, l'intensité des pixels est correctement modélisée par la réalisation de variables aléatoires de densité de probabilité gamma. Ces images ne correspondent cependant pas au modèle d'image que nous avons choisi, le fond de ces images étant inhomogène. Les inhomogénéités prennent des formes très différentes dans les deux exemples que nous avons retenus et requièrent un traitement différent.

Les images SAR (Synthetic Aperture Radar) constituent un exemple d'application

idéal pour le GLRT. En effet, l'intensité des pixels dans ces images est en général modélisée par un champ aléatoire décorrélé et dont la densité de probabilité est une loi Gamma. Dans ce cas, nous avons montré que l'algorithme GLRT multi-tailles se révèle efficace.

Dans ces images, le fond est composé d'un patchwork de régions homogènes sur lesquelles apparaissent les cibles. Il n'est donc pas homogène dans son ensemble. Les régions constituant le fond sont néanmoins beaucoup plus grandes que les cibles cherchées. L'utilisation de petites fenêtres d'analyse prend donc tout son sens dans ce cas, puisque l'on ne considère le fond que localement. On peut néanmoins constater l'apparition de fausses alarmes lorsque la fenêtre d'analyse est placée à la frontière entre deux régions homogènes dont les paramètres diffèrent. L'algorithme GLRT multi tailles détecte alors la rupture de statistiques entre ces régions mais la discrimination par la forme de la cible n'est pas suffisante pour supprimer ces artefacts.

L'autre exemple d'application que nous avons retenu est celui d'un imageur polarimétrique actif, fournissant deux images représentatives des propriétés dépolarisantes de la scène observée. Le fond dans ces images est fortement inhomogène, essentiellement du fait de l'inhomogénéité du faisceau d'éclairement.

Dans ce cas, les inhomogénéités ne se présentent plus sous la forme de régions bien définies, mais sous forme de variations lentes de la moyenne du fond. Ces variations ont une longueur caractéristique du même ordre de grandeur que celle de la cible. L'utilisation de petites fenêtres ne permet alors plus d'obtenir des performances satisfaisantes, lorsque l'on travaille sur les deux images, les inhomogénéités entraînant l'apparition de fausses alarmes. Il est donc préférable d'utiliser une approche sous-optimale, consistant à combiner les deux images pour travailler sur une scène dans laquelle les inhomogénéités ont été atténuées. On applique alors à cette nouvelle image un algorithme GLRT qui lui est adapté, ce qui permet d'obtenir des performances correctes.

4 Cibles composées de plusieurs régions

Par ailleurs, on peut distinguer deux grands types de situations dans des applications de localisation/détection.

- On connaît les niveaux de gris de la cible.
- On ne connaît que la forme de la cible.

La connaissance des niveaux de gris de la cible est une information très forte pour des applications de localisation/détection. Si cette information est disponible, il est important de l'utiliser dans la conception des filtres. Cependant, un algorithme supposant les niveaux de gris de la cible connus est rapidement mis en échec lorsque les niveaux de gris de la cible fluctuent par rapport à ceux de la référence dont on dispose.

Les algorithmes que nous avons présentés dans les parties précédentes n'utilisent que l'information de forme de la cible cherchée. Ils sont donc peu sensibles aux fluctuations des niveaux de gris de la cible. En revanche, ils supposent que la cible constitue une zone homogène de l'image, ce qui peut être faux dans certains cas.

L'écart entre ces deux situations est donc énorme, et il nous a semblé intéressant de chercher des solutions intermédiaires. Nous avons donc proposé de prendre en compte l'information de la structure spatiale de la cible sans utiliser la connaissance de ses ni-

veaux de gris. Ceci peut être fait en modélisant la cible sous la forme d'un patchwork de régions homogènes, bien définies spatialement, et dont les paramètres sont inconnus. Cet algorithme suppose connu le nombre de régions de la cible.

Ceci trouve un grand intérêt dans le cas d'objets 3D présentant différentes facettes, et apparaissant dans les images dans des conditions d'éclairement différentes. Les différentes facettes apparaissent toujours, mais la distribution de leurs intensités est variable. Cette méthode est particulièrement intéressante dans le cas où les images sont fortement bruitées, et où les méthodes traditionnelles d'extraction de bords ne serait pas efficaces.

Il existe cependant une limite au nombre de régions utilisables pour définir la cible. En particulier, utiliser des régions qui n'apparaissent pas dans la cible rend plus difficile la discrimination entre la cible et le fond et induit donc des pertes de performances en détection. La connaissance des distributions en sortie des filtres GLRT conçus pour une cible homogène nous a permis de caractériser ces pertes.

Par ailleurs, pour limiter le nombre de régions utilisé au strict nécessaire, nous avons proposé une approche permettant d'estimer ce nombre de régions. Ceci limite les pertes évoquées ci-dessus, et permet de traiter efficacement le problème de la détection de cibles dont le nombre de régions est inconnu.

5 Implémentation optique des filtres

Les filtres GLRT que nous avons utilisés nécessitent un temps de calcul de l'ordre de quelques corrélations. Pour des applications où le calcul de celles-ci par voie numérique est trop long, on peut envisager de recourir à l'utilisation de corrélateurs optiques. La vitesse de traitement de ces corrélateurs est en effet uniquement limitée par la vitesse d'affichage des scènes sur les SLM, et non par l'opération de corrélation elle-même.

Parmi les corrélateurs optiques existants, le NLJTC est un candidat intéressant pour des opérations de localisation, bien que sous-optimal pour les modèles d'images que nous avons considéré dans ce mémoire. De ce fait, le GLRT présente de meilleures performances, comme l'a montré la comparaison entre ces deux algorithmes.

Par ailleurs, cette comparaison a permis de mettre en évidence un phénomène intéressant. La scène à traiter peut être codée sur les SLM, soit en phase, soit en amplitude. Il apparaît sur les expériences que nous avons menées que le codage en phase permet d'obtenir une meilleure probabilité de localisation que le codage en amplitude. Ceci est en contradiction apparente avec les résultats connus dans le cas de bruit additif, où le codage en amplitude est préférable. Dans le cas de bruit non recouvrant, la supériorité du codage en phase sur un codage en amplitude avait déjà été remarquée, mais sans que les raisons en soient éclaircies. Une étude phénoménologique de ce phénomène nous a permis de montrer qu'en présence de bruit non recouvrant, le codage en phase permet de transformer une différence de formes d'histogrammes entre la cible et le fond, en une différence de moyennes entre ces régions. Pour détecter des cibles dont on ne connaît que la forme, c'est à cette différence de moyenne que le NLJTC est le plus sensible. Le codage en phase permet de réhausser cette quantité, ce qui permet d'accroître la probabilité de localisation correcte du NLJTC.

6 Perspectives

Au terme de ce mémoire, nous pouvons tracer quelques perspectives quand au développement des résultats obtenus et à l'amélioration future des techniques que nous avons introduites.

Imagerie hyperspectrale

Un sujet d'étude très en vogue actuellement est celui des systèmes d'imagerie hyperspectrale, c'est à dire fournissant des images issues d'un grand nombre de capteurs de nature différente. La plupart de solutions proposées utilisent des techniques de type GLRT, qui sont suffisamment simples pour permettre de traiter des données très nombreuses rapidement. Etendre notre étude sur les filtres GLRT à ce type d'imagerie semble être un prolongement naturel de nos travaux. Il serait notamment intéressant d'observer les résultats de méthodes estimant le nombre de régions de la cible dans chaque canal.

Dans le cas de cibles de taille inconnue, nous avons noté que la capacité de discrimination de nos algorithmes pouvaient être insuffisante. En imagerie hyperspectrale, la multitude de canaux permet de définir une signature spectrale de l'objet à détecter. Cette signature devrait pouvoir être mise à profit pour augmenter la capacité de discrimination de nos filtres.

Bruit corrélé

Tous les algorithmes que nous avons présentés supposent que le bruit perturbant les images est décorrélié, ce qui permet d'obtenir des algorithmes simples et rapides. Il existe plusieurs types de solutions pour prendre en compte la présence de textures corrélées dans les images. La plus simple consiste à utiliser des prétraitements pour "blanchir" l'image [21]. Bien qu'efficaces et simples d'implémentation, ces méthodes ne sont pas caractérisées clairement. En particulier, l'étude de la variation du contraste entre la cible et le fond dans l'image prétraitée en fonction de celles de l'image originale semble nécessaire avant d'inclure un tel prétraitement dans une chaîne de traitement ou figurerait le GLRT.

Estimation du nombre de régions

Nous avons appliqués une méthode permettant d'estimer le nombre de régions composant la cible, dans une application de localisation-détection. Nous avons également montré que la méthode retenue pour cette estimation pouvait être justifiée par une approche de type Description de Longueur Minimale (MDL pour Minimum Description Length). Cependant, pour traiter l'ensemble des configurations de régions possibles, nous n'avons envisagé que deux possibilités : exploration systématique de toutes les solutions pour l'algorithme RS-GLRT, ou exploration restreinte à deux configurations possibles (cible homogène ou cible composée du nombre maximum de régions) pour l'algorithme SRS-GLRT. Il serait intéressant de travailler sur une méthode d'exploration rapide de l'ensemble des configurations possibles. On peut imaginer cet ensemble sous la forme d'un arbre dans lesquels certaines branches serait jugées rapidement inintéressantes et donc non explorées.

Annexe A

Principe de la description de longueur minimale (MDL)

Dans le chapitre 5, nous nous sommes intéressés au problème de la détection d'une cible composée de plusieurs régions. Après avoir présenté un algorithme (cf. section 5.2) supposant connus le nombre et la configuration spatiale des régions décrivant la cible, nous avons proposé un algorithme permettant d'estimer ces paramètres (cf. section 5.4).

L'estimation du nombre de régions nécessaire pour décrire la cible a été faite de façon empirique. L'objet de cette annexe est de montrer qu'il est possible de retrouver un résultat identique en s'appuyant sur une approche plus théorique, celle de la description de longueur minimale, ou MDL pour *Minimum Description Length* [48].

Pour cela, nous supposons connu l'ensemble des configurations dans lesquelles la cible cible peut apparaître, comme présenté dans la figure A.1.

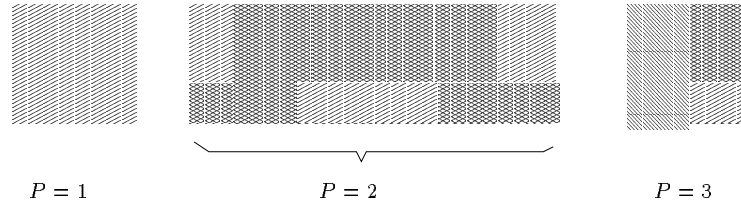


FIG. A.1: Cible multi région et configurations possibles. P est le nombre de région de la configuration considérée. L'ensemble des configurations possible est noté Ω , ici de cardinal 5. La cible la plus complexe est à droite, ou chacune des trois régions est supposée être une réalisation d'un champ aléatoire indépendant des autres régions. Cette description multi région est parfois inutile, et un modèle plus simple (une ou deux région) est alors préférable.

Comme nous l'avons vu, le nombre de régions utilisé constitue l'ordre de notre modèle de cible, c'est à dire la complexité de la description de la cible que l'on cherche. Pour estimer cet ordre, nous avons dans la section 5.4 procédé de façon empirique, en ajoutant au rapport de vraisemblance un terme pénalisant les modèles d'ordres trop grands. Le principe MDL a été proposé par Rissanen pour l'estimation de l'ordre des modèles statistiques paramétriques et nous allons voir ici comment il peut s'appliquer à notre problème.

Le principe du MDL consiste à minimiser la quantité d'information nécessaire pour

décrire l'échantillon dont on dispose ³. Cette quantité est mesurée par le nombre de bits qu'il faudrait utiliser pour coder l'échantillon. Dans notre cas, l'échantillon est le contenu de la fenêtre d'analyse à une position donnée. Cette fenêtre, dans l'hypothèse où la cible est présente, contient la cible au centre de la fenêtre, dans l'une ou l'autre des configurations possibles, et une zone fond, indépendante de la configuration de la cible. Pour choisir quelle configuration de cible doit être retenue, nous ne nous intéressons donc qu'à la partie de la fenêtre correspondant au support de la cible. Ainsi la quantité d'information à minimiser est le nombre de bits qu'il faudrait utiliser pour coder toute la partie de l'échantillon qui constitue la cible éventuellement présente au centre de la fenêtre d'analyse.

La théorie de l'information nous permet de déterminer une approximation de ce nombre de bits [52, 8]. En effet, il a été 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 [52].

Dans notre problème, l'intensité de chaque pixel est vue comme un événement aléatoire. Par ailleurs, pour une configuration \mathcal{C}_k donnée, les pixels sont groupés en P_k régions homogènes, il est donc pratique de considérer le nombre de bits nécessaires pour coder les niveaux de gris des pixels de chaque région possible.

Soit une région R_i^k de la configuration \mathcal{C}_k , de support \mathbf{w}_i^k , composée de N_i^k pixels dont les intensités $x(j)$, $j \in \mathbf{w}_i^k$ sont identiquement distribuées selon une loi de densité de probabilité $\mathcal{P}_{\theta_i^k}$. Les pixels étant supposés indépendants entre eux, le nombre minimal moyen de bits nécessaire pour coder les N_i^k variables aléatoires de la région R_i^k est égal au produit de N_i^k par l'entropie S_i^k de la loi de probabilité. Le nombre de bits utilisé pour coder l'intensité des pixels de R_i^k est donc $\Delta_i^k \simeq N_i^k S_i^k$.

L'entropie S_i^k est donnée par la relation suivante : $S_i^k = - \int \mathcal{P}_{\theta_i^k}(x) \log_2 [\mathcal{P}_{\theta_i^k}(x)] dx$. Ce terme est la moyenne de $\log_2 [\mathcal{P}_{\theta_i^k}(x)]$. Lorsqu'on le remplace par la moyenne empirique calculée sur la région toute entière, on obtient :

$$S_i^k \simeq -1/N_i^k \sum_{j \in \mathbf{w}_i^k} \log_2 [\mathcal{P}_{\theta_i^k}(x_j)] \quad (\text{A.1})$$

On a donc :

$$\Delta_i^k = N_i^k S_i^k \simeq - \sum_{j \in \mathbf{w}_i^k} \log_2 [\mathcal{P}_{\theta_i^k}(x_j)] \quad (\text{A.2})$$

On peut noter que cette expression est celle de la logvraisemblance (en base 2) de l'hypothèse que \mathbf{w}_i^k soit le support d'une région homogène de paramètre θ_i^k . Si l'on utilise le logarithme népérien à la place du logarithme en base 2 pour calculer toutes les grandeurs concernées, on peut noter que les vraisemblances que nous avons utilisées dans ce mémoire peuvent s'interpréter comme l'expression d'une longueur de codage ⁴.

Ici, nous avons défini une région par le codage de son contenu. Il n'est pas nécessaire de coder la forme de chaque région puisqu'elles sont connues par avance. Tout au plus, il

³Rissanen justifie ceci "philosophiquement" par une citation d'Einstein : 'All theories should be as simple as possible - but not simpler'

⁴Dans un cas on maximise la vraisemblance et dans le cas du MDL, on minimise son opposé.

faut définir un code pour déterminer quelle configuration est utilisée. Ceci peut être fait avec le même nombre de bits, quelle que soit la configuration (cf. figure 5.15). Ce nombre de bits étant constant, nous ne le considérons pas dans la suite.

En revanche, la quantité Δ_i^k ainsi obtenue suppose que l'on connaît l'ensemble des paramètres θ_i^k de la région. Si ces paramètres sont inconnus, il faut ajouter à ce nombre le nombre de bits utile pour coder ces paramètres. Ce nombre de bits dépend de la précision qu'il faut utiliser pour le codage. Pour développer cela, prenons pour exemple celui d'une statistique Gamma d'ordre 1. Celle-ci est complètement définie par sa moyenne m_i^k . Si les moyennes que l'on peut trouver dans l'image peuvent varier entre 0 et M, la précision à utiliser pour coder la moyenne d'une région dépend des fluctuations statistiques de l'estimateur de moyenne.

Ces fluctuations sont maximales pour $m_i^k = M$, et dans ce cas, l'écart-type de cet estimateur est

$$\sigma_{\hat{m}_i^k} = M / \sqrt{(N_i^k)} \quad (\text{A.3})$$

Le nombre de valeurs distinctes utilisé pour coder une moyenne comprise entre 0 et M lorsque la précision de codage est liée aux fluctuations statistiques peut être approximé par $M / \sigma_{\hat{m}_i^k} = \sqrt{(N_i^k)}$. Le nombre de bits correspondant au codage de ces valeurs est $\log(\sqrt{N_i^k}) = 0.5 \log(N_i^k)$, où, comme pour la logvraisemblance, nous avons utilisé le logarithme népérien à la place du logarithme en base 2.

Le nombre de bits minimum utilisé pour coder l'intensité des pixels d'une région composée de bruit décorrélié Gamma d'ordre 1 et est donc :

$$\Delta_i^k = - \sum_{j \in \mathbf{w}_i^k} \log [\mathcal{P}_{\theta_i^k}(x_j)] + 0.5 \log(N_i^k) \quad (\text{A.4})$$

Lorsque les moyennes de chaque région sont inconnues et sont estimées au sens du maximum de vraisemblance, l'expression de la pseudo logvraisemblance d'observer la scène a déjà été calculé (cf. équation 5.4). On obtient alors comme expression du critère à minimiser :

$$\Delta_i^k = N_i^k \log [\hat{m}_i^k] + 0.5 \log(N_i^k) \quad (\text{A.5})$$

avec \hat{m}_i^k la moyenne estimée sur la région i de la configuration testée \mathcal{C}_k .

Le raisonnement est identique pour l'ensemble des régions de la cible dans la configuration \mathcal{C}_k . On obtient donc comme valeur du critère MDL pour cette configuration :

$$\Delta(\mathcal{C}_k) = \sum_{i=1}^{P_k} N_i^k \log \left[\frac{1}{N_i^k} \sum_{j \in \mathbf{w}_i^k} s_j \right] + 0.5 \sum_{i=1}^{P_k} \log(N_i^k) \quad (\text{A.6})$$

Le dernier terme correspond à un terme pénalisant un trop grand nombre de régions. Le calcul de ce terme doit être effectué dans chacune des configurations de la cible. En toute rigueur, il faudrait donc recompter le nombre de pixels de chaque région dans chaque configuration. On peut cependant le borner simplement, en remarquant que

$$0.5 \log(\min_{i \in [1..k]} N_i^k) < \alpha = 0.5 \frac{1}{P_k} \sum_{i=1}^{P_k} \log(N_i^k) < 0.5 \log(\max_{i \in [1..k]} N_i^k) \quad (\text{A.7})$$

Si la taille des régions varie assez peu, cet intervalle est suffisamment petit pour qu'une définition exacte ne soit pas nécessaire. Par exemple, si la taille de régions est comprise entre 100 et 200 pixels, on obtient $2.30 < \alpha < 2.65$. Considérons que ces valeurs sont "suffisamment proches" pour que l'on puisse ne considérer qu'un **nombre de pixels caractéristique** N de la taille des régions. On peut alors écrire la longueur du codage d'une configuration sous la forme :

$$\Delta(\mathcal{C}_k) = \sum_{i=1}^{P_k} N_i^k \log \left[\frac{1}{N_i^k} \sum_{j \in \mathbf{w}_i^k} s_j \right] + 0.5 P_k \log(N) \quad (\text{A.8})$$

On choisira alors la configuration \mathcal{C}_{mdl} telle que :

$$\Delta(\mathcal{C}_{mdl}) = \arg \min_{k \in \Omega} \left[\sum_{i=1}^{P_k} N_i^k \log \left[\frac{1}{N_i^k} \sum_{j \in \mathbf{w}_i^k} s_j \right] + 0.5 P_k \log(N) \right] \quad (\text{A.9})$$

Lorsque l'on minimise ce critère en fonction de k , on ne considère que les termes correspondant à l'intérieur de la cible. Tout ce qui est extérieur au support total de la cible n'intervient pas dans la sélection d'une configuration. Ceci est donc équivalent à minimiser le critère suivant :

$$J(\mathcal{C}_k) = \sum_{i=1}^{P_k} N_i^k \log \left[\frac{1}{N_i^k} \sum_{j \in \mathbf{w}_i^k} s_j \right] + N_O \log \left[\frac{1}{N_O} \sum_{j \in \mathbf{w}_O} s_j \right] - N_A \log \left[\frac{1}{N_A} \sum_{j \in \mathbf{w}_A} s_j \right] + 0.5 P_k \log(N) \quad (\text{A.10})$$

où N_A est le nombre de pixels de la fenêtre d'analyse, \mathbf{w}_A le support de la fenêtre d'analyse, N_O le nombre de pixels du complémentaire de la cible dans la fenêtre et \mathbf{w}_O son support.

L'ensemble des trois premiers termes correspond à l'opposé du logarithme du rapport de vraisemblance $\mathcal{R}(\mathcal{C}_k)$ entre l'hypothèse \mathcal{H}_{1k} (une cible dans la configuration \mathcal{C}_k est présente au centre de la cible) et \mathcal{H}_0 (la fenêtre ne contient que du bruit homogène) (cf. eq 5.4).

On peut donc effectuer la sélection de la configuration la plus appropriée de telle sorte que :

$$J(\mathcal{C}_{mdl}) = \arg \max_{k \in \Omega} [\mathcal{R}(\mathcal{C}_k) - 0.5 P_k \log(N)] \quad (\text{A.11})$$

On reconnaît alors une fonction de pénalisation du type $g(P_k) = \alpha P_k$, telle que celle que nous avons utilisé dans la présentation de l'algorithme RS-GLRT, dans la section 5.4. Comme on l'a vu précédemment, dans le cas d'une cible composée de régions de 200 pixels, on obtient $\alpha_{MDL} = 0.5 \log(200) = 2.65$. Pour $N = 100$, on obtient une valeur $\alpha_{MDL} = 2.30$.

Ces valeurs sont très proches de la valeur $\alpha = 3$ que nous avons choisie de façon plus empirique en imposant une fonction de pénalisation linéaire, puis en choisissant la valeur de α qui expérimentalement donne la meilleure estimation du nombre de région (cf. section 5.4.2). L'étude menée dans cette section montrait que toute valeur de α comprise entre 2 et 3 permet d'obtenir des résultats corrects. Dans les exemples que nous avons considérés, les tailles des régions ne sont pas très différentes. Nous pouvons donc considérer

que seul importe l'ordre de grandeur de la taille des régions. Cependant, une définition plus rigoureuse pourrait constituer une suite intéressante de cette étude, afin de pouvoir l'étendre à des cibles dont les régions peuvent être de tailles plus inhomogènes.

Annexe B

Calculs

B.1 Détection a Taux de Fausse Alarme Constant en bruit gaussien

Nous reprenons ici une propriété énoncée dans la section 1.4.3. Nous avons alors affirmé que les détecteurs de type GLRT offraient la propriété d'être à Taux de Fausse Alarme Constant. Ceci signifie qu'en l'absence de cible, leur réponse est indépendante des paramètres du fond. Nous avons démontré cette propriété dans le cas de statistiques Gamma. Ici, nous allons développer les calculs correspondants au cas de lois gaussiennes (cf. tableau 1.1).

La densité de probabilité d'une Variable aléatoire Gaussienne est la suivante :

$$\mathcal{P}_{m,\sigma}(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{(x-m)^2}{2\sigma^2}\right] \quad (\text{B.1})$$

L'expression du logarithme du rapport de vraisemblance dans le cas ou chaque région suit une statistique gaussienne est la suivante [19] :

$$\mathcal{R} = -N_w \log(\hat{\sigma}_w^2(\mathbf{s})) - N_{\bar{w}} \log(\hat{\sigma}_{\bar{w}}^2(\mathbf{s})) + N_F \log(\hat{\sigma}_F^2(\mathbf{s})) \quad (\text{B.2})$$

où

$$[\hat{\sigma}_w(\mathbf{s})]^2 = \frac{1}{N_w} \sum_{(i) \in \mathbf{w}} [s(i) - \hat{\mu}_w(\mathbf{s})]^2 \quad (\text{B.3})$$

$$[\hat{\sigma}_{\bar{w}}(\mathbf{s})]^2 = \frac{1}{N_{\bar{w}}} \sum_{(i) \in \bar{\mathbf{w}}} [s(i) - \hat{\mu}_{\bar{w}}(\mathbf{s})]^2 \quad (\text{B.4})$$

$$[\hat{\sigma}_F(\mathbf{s})]^2 = \frac{1}{N_F} \sum_{(i) \in \mathbf{F}} [s(i) - \hat{\mu}_F(\mathbf{s})]^2 \quad (\text{B.5})$$

Dans le cas ou l'hypothèse \mathcal{H}_0 est vraie, les échantillons $s(i)$ sont des réalisations de variables aléatoires gaussiennes de moyenne m_b et de variance σ_b^2 .

On peut écrire

$$s(i) = m_b + \sigma_b n(i) \quad (\text{B.6})$$

où les $n(i)$ sont des variables aléatoires gaussiennes, centrées réduites. On a alors :

$$\begin{aligned} \hat{\sigma}_w^2(\mathbf{s}) &= \sigma_b^2 \hat{\sigma}_w^2(\mathbf{n}) \\ \hat{\sigma}_{\bar{w}}^2(\mathbf{s}) &= \sigma_b^2 \hat{\sigma}_{\bar{w}}^2(\mathbf{n}) \\ \hat{\sigma}_F^2(\mathbf{s}) &= \sigma_b^2 \hat{\sigma}_F^2(\mathbf{n}) \end{aligned} \quad (\text{B.7})$$

En réintroduisant ceci dans l'équation B.2, on obtient

$$\begin{aligned} \mathcal{R} &= -N_w \log(\hat{\sigma}_w^2(\mathbf{n})) - N_w \log(\sigma_b^2) - N_{\bar{w}} \log(\hat{\sigma}_{\bar{w}}^2(\mathbf{n})) - N_{\bar{w}} \log(\sigma_b^2) \\ &\quad + N_F \log(\hat{\sigma}_F^2(\mathbf{n})) + N_F \log(\sigma_b^2) \\ &= -N_w \log(\hat{\sigma}_w^2(\mathbf{n})) - N_{\bar{w}} \log(\hat{\sigma}_{\bar{w}}^2(\mathbf{n})) + N_F \log(\hat{\sigma}_F^2(\mathbf{n})) \end{aligned} \quad (\text{B.8})$$

Cette expression étant indépendante des paramètres m_b et σ_b^2 , on démontre bien ainsi la propriété énoncée.

B.2 Algorithme du Maximum de Vraisemblance pour des bruits de Poisson

Dans la section 5.2.3, nous avons donné la formule de l'algorithme GLRT multi-région adapté à du bruit de Poisson décorrélié. Nous allons ici développer les calculs permettant d'obtenir cet algorithme. Nous ne ferons cette démonstration que pour une cible homogène, la généralisation au cas de cibles multirégions étant immédiate.

Nous supposons que la scène à traiter correspond au modèle SIR 1.1, dans lesquels les niveaux de gris de la cible et du fond sont des réalisations de champs aléatoires décorréliés. Nous supposons de plus que dans chacune des régions, l'intensité de chaque pixel est distribuée suivant une loi de Poisson de moyenne m_a sur la cible et m_b sur le fond.

La cible a pour support le masque \mathbf{w} , et la fenêtre d'analyse est notée \mathbf{F} . $\bar{\mathbf{w}}$ est le complémentaire de la cible dans la fenêtre d'analyse. Chacune de ces grandeurs indicées par j représente le support de la région correspondante, centrée sur le j^{eme} pixel de l'image.

La densité de probabilité d'une variable aléatoire distribuée suivant une loi de Poisson de moyenne m est donnée par :

$$P_m(x) = \exp[-m] \frac{m^x}{x!} \text{ où } x \in \mathbb{N} \quad (\text{B.9})$$

Nous pouvons reprendre l'expression de la logvraisemblance d'observer la scène \mathbf{s} que nous avons établi dans l'équation 1.4 lorsqu'une cible est présente.

$$\mathcal{L}(\mathbf{s}|j, m_a, m_b) = \mathcal{L}(\mathbf{s} \cap \mathbf{w}_j | m_a) + \mathcal{L}(\mathbf{s} \cap \mathbf{1} - \mathbf{w}_j | m_b) \quad (\text{B.10})$$

Lorsque la cible est absente :

$$\mathcal{L}0(\mathbf{s}|j, m_b) = \mathcal{L}(\mathbf{s} \cap \mathbf{F}_j | m_b) \quad (\text{B.11})$$

Le logarithme du rapport de vraisemblance est alors donné par :

$$\mathcal{R}(\mathbf{s}|j, m_a, m_b) = \mathcal{L}1(\mathbf{s} \cap \mathbf{F}_j | m_a, m_b) - \mathcal{L}0(\mathbf{s} | \mathbf{j}, m_b) \quad (\text{B.12})$$

Ceci, compte tenu de l'hypothèse de champs décorrélés se décompose aisément de la façon suivante.

$$\mathcal{R}(\mathbf{s} | \mathbf{j}, m_a, m_b) = \sum_{i \in \mathbf{w}_j} [-m_a + s(i) \log(m_a) + s(i)(s(i) + 1)/2] \quad (\text{B.13})$$

$$+ \sum_{i \in \bar{\mathbf{w}}_j} [-m_b + s(i) \log(m_b) + s(i)(s(i) + 1)/2] \quad (\text{B.14})$$

$$- \sum_{i \in \mathbf{F}_j} [-m_b + s(i) \log(m_b) + s(i)(s(i) + 1)/2] \quad (\text{B.15})$$

$$= -N_w m_a + \sum_{i \in w_j} [s(i) \log(m_a)] \quad (\text{B.16})$$

$$-N_{\bar{w}} m_b + \sum_{i \in \bar{w}_j} [s(i) \log(m_b)] \quad (\text{B.17})$$

$$-N_F m_b - \sum_{i \in \mathbf{F}_j} [s(i) \log(m_b)] \quad (\text{B.18})$$

Lorsque les paramètres sont inconnus nous les estimons au sens du maximum de vraisemblance. Encore une fois, il faut distinguer la moyenne du fond estimée dans la fenêtre d'analyse \mathbf{F}_j , que nous noterons \hat{m}_c et la moyenne du fond estimée dans la région $\bar{\mathbf{w}}_j$, que nous noterons \hat{m}_b

$$\begin{aligned} \frac{\partial \mathcal{R}}{\partial m_a} = 0 & \iff \hat{m}_a = \frac{1}{N_a} \sum_{i \in \mathbf{w}_j} s(i) \\ \frac{\partial \mathcal{L}}{\partial m_b} = 0 & \iff \hat{m}_b = \frac{1}{N_b} \sum_{i \in \bar{\mathbf{w}}_j} s(i) \\ \frac{\partial \mathcal{L}}{\partial m_c} = 0 & \iff \hat{m}_c = \frac{1}{N_F} \sum_{i \in \mathbf{F}_j} s(i) \end{aligned} \quad (\text{B.19})$$

En ré-introduisant ces valeurs dans l'expression de l'équation B.18, on obtient

$$\mathcal{R}(\mathbf{s} | \mathbf{j}, m_a, m_b) = N_a \hat{m}_a \log(\hat{m}_a) + N_b \hat{m}_b \log(\hat{m}_b) - N_F \hat{m}_c \log(\hat{m}_c) \quad (\text{B.20})$$

Bibliographie

- [1] F. Murtagh A. Bijaoui, J.L. Starck. Restauration des images multi-echelles par l'algorithme à trous. *Traitement du Signal et des Images*, 40:229–243, 1994.
- [2] A. Gupta A. Leonardis and R. Bajcsy. Segmentation of range images as the search of geometric parametric models. *Int'l J. Computer Vision*, 14(3):253–277, 1995.
- [3] M. Abramowitz and I. Stegun. *Handbook of mathematical functions*. Dover Publications, 1972.
- [4] S. Breugnot and Ph. Clémenceau. Modeling and performances of a polarization active imager at $\lambda=806$ nm. In Gary W. Kamerman; Christian Werner, editor, *Laser Radar technology and applications IV*, volume 3707, pages 449–460. Proc. SPIE, 1999.
- [5] D. Casasent and D. Psaltis. Position, rotation and scale invariant optical correlation. *Applied Optics*, 15(7):1795–1799, 1976.
- [6] D. Casasent and D. Psaltis. Scale invariant optical correlation using mellin transforms. *Optical Communications*, 17:59–63, 1976.
- [7] K. Chalasinska-Macukow and K. Styczinsky. Optoelectronic correlators with phase-encoded input scenes. In *International Conference on Correlation Optics*, volume 3317, pages 142–151. SPIE, 1997.
- [8] T. M. Cover and J. A. Thomas. *Elements of Information Theory*. Wiley-interscience, New York, 1991.
- [9] E. Marom D. Mendlovic and N. Konforti. Shift and scale invariant pattern recognition with logarithmic radial harmonic filters. *Opt. Commun.*, 67:172–176, 1998.
- [10] R. O. Duda and P. E. Hart. *Pattern classification and scene analysis*. John Wiley and sons, Inc., New York, 1973.
- [11] A. H. Fazlollahi, B. Javidi, and P. Willet. Minimum-error-probability receiver for detecting a noisy target in colored background noise. *J. Opt. Soc. Am. A*, 14(4):845–852, 1997.
- [12] K. H. Fielding and J. L. Horner. Clutter effects on optical correlators. In-house report, Rome Air Development Center, 1990.
- [13] V. S. Frost, Shanmugan K.S., and Holtzman J. C. Edge detection for SAR and other noisy images. In *Proc. International Geoscience and Remote Sensing Symposium, Munich, Germany*, volume FA2, pages 4.1–4.9, 1982.
- [14] J. W. Goodman. *Fourier Optics*. McGraw-Hill Physical and Quantum Electronics Series, San Francisco, 1968.

- [15] F. Goudail. *Localisation d'objets dans des images fortement bruitées : une approche probabiliste paramétrique*. Thèse de doctorat, Université Aix-Marseille III, Marseille, France, 1997.
- [16] F. Goudail, V. Laude, and Ph. Réfrégier. Influence of non-overlapping noise on regularized linear filters for pattern recognition. *Opt. Lett.*, 20(21):2237–2239, 1995.
- [17] F. Goudail and Ph. réfégier. Statistical techniques for target detection in polarization diversity images. *Opt. Lett.*, 26(9):644–646, 2001.
- [18] F. Goudail and Ph. Réfrégier. Optimal detection of a target with random gray levels on a spatially disjoint noise. *Opt. Lett.*, 21(7):495–497, 1996.
- [19] F. Guérault. *Techniques statistiques pour l'estimation de la position d'un objet dans des images bru itées*. Thèse de doctorat, Université Aix-Marseille III, Marseille, France, 1999.
- [20] F. Guérault and Ph. Réfrégier. Optimal χ^2 filtering method and application to targets and backgrounds with random correlated gray levels. *Opt. Lett.*, 22(9):630–632, 1997.
- [21] F. Guérault and Ph. Réfrégier. Statistically independent region processor for target and background with random textures: whitening preprocessing approach. *Optics Commun.*, 142(4,5,6):197–202, 1997.
- [22] F. Guérault and Ph. Réfrégier. Unified statistically independent region processor for deterministic and fluctuating target in non-overlapping background. *Opt. Lett.*, 23(6):412–414, 1998.
- [23] J. Horner and H. Bartlet. 2-bit correlation. *Applied Optics*, 24:2889, 1989.
- [24] J. L. Horner and P. D. Gianino. Signal-dependent phase distortion in optical correlators. *Appl. Opt.*, 26(12):2484–2490, 1987.
- [25] J. L. Horner and R. A. Soref. Phase-dominant spatial light modulators. *Elec. Lett.*, 24(10):626–627, 1988.
- [26] A.O.Hero H.S. Kim. Comparison of glr and invariance methods applied to adaptive target detection in structured clutter. Technical report 323, University of Michigan, 2000.
- [27] Y. Irving.S.Reed, Xiaoli. Adaptative multiple band cfar detection of an optical pattern with unknown spectral distribution. *IEEE Trans. on Acoustics, Speech and Signal Processing*, 38(10):2464–2482, 1990.
- [28] B. Javidi. Nonlinear joint power spectrum based optical correlation. *Appl. Opt.*, 28(12):2358–2367, 1989.
- [29] B. Javidi, Ph. Réfrégier, and P. Willet. Optimum receiver design for pattern recognition with nonoverlapping target and scene noise. *Opt. Lett.*, 18(19):1660–1662, 1993.
- [30] B. Javidi and J. Wang. Limitation of the classic definition of the correlation signal-to-noise ratio in optical pattern recogition with disjoint signal and scene noise. *Appl. Opt.*, 31(32):6826–6829, 1992.
- [31] B. Javidi and J. Wang. Design of filters to detect a noisy target in nonoverlapping background noise. *JOSA A.*, 11(10):2604–2612, 1994.

-
- [32] R. R. Kallman and D. H. Goldstein. Phase encoding input images for optical pattern recognition. *Opt. Eng.*, 33(6):1806–1812, 1994.
 - [33] S. Mallat. A theory for multiresolution signal decomposition : the wavelet representation. *IEEE PAMI*, 11(6):674–693, 1989.
 - [34] S. Mazé and Ph. Réfrégier. Noise robustness of optical correlation for amplitude or phase modulation of the input image. *Opt. Lett.*, 17(6):426–428, 1992.
 - [35] S. Mazé and Ph. Réfrégier. Optical correlation: influence of the coding of the input image. *Appl. Opt.*, 19(6):405–408, 1994.
 - [36] Ch. Minetti and F. Dubois. Reduction in correlation-filter sensitivity to background clutter by the automatic spatial frequency selection algorithm. *Appl. Opt.*, 35(11):1900–1903, 1996.
 - [37] C. J. Oliver, I. Mc Connell, D. Blacknell, and R. G. White. Optimum edge detection in SAR. In *Conf. on Satellite Remote Sensing, Paris, France*, volume 2584, pages 152–163. SPIE, 1995.
 - [38] C. J. Oliver and S. Quegan. *Understanding SAR images*, chapter Principles of SAR image formation, pages 11–42. Artech House, Norwood, 1998.
 - [39] Athanasios Papoulis. *Probability, random variables, and stochastic processes*. WBC/McGraw-Hill, Boston, 1991.
 - [40] H. Rajbenbach, S. Bann, Ph. Réfrégier, P. Joffre, J.-P. Huignard, H. St. Buchkremer, A. S. Jensen, E. Rasmussen, K. H. Brenner, and G. Lohman. Compact photorefractive correlator for robotic applications. *Appl. Opt.*, 31(26):5666–5674, 1992.
 - [41] Ph. Réfrégier. Application of the stabilizing functional approach to pattern recognition filters. *J. Opt. Soc. Am. A*, 11(4):1243–1251, 1994.
 - [42] Ph. Réfrégier. Bayesian theory for target location in noise with unknown spectral density. *JOSA A*, 16(2):276–283, 1999.
 - [43] Ph. Réfrégier and F. Goudail. A decision theoretical approach to nonlinear joint-transform correlation. *JOSA A*, 15:61–67, 1998.
 - [44] Ph. Réfrégier, F. Goudail, and Th. Gaidon. Optimal location of random targets in random background: random markov fields modelization. *Opt. Com.*, 128(4,5,6):211–215, 1996.
 - [45] Ph. Réfrégier, B. Javidi, and V. Laude. Nonlinear joint-transform correlation: an optimal solution for adaptive image discrimination and input noise robustness. *Opt. Lett.*, 19(6):405–407, 1994.
 - [46] Ph. Réfrégier, B. Javidi, and G. Zhang. Minimum mean-square-error filter for pattern recognition with spatially disjoint signal and scene noise. *Opt. Lett.*, 18(17):1453–1456, 1993.
 - [47] Ph. Réfrégier and V. Laude. Spatial fluctuations of optical field modulated with spatial light modulators and noisy input signals. *JOSA A*, 12(6):1338–1345, 1995.
 - [48] J. Rissanen. *Stochastic complexity in statistical inquiry*. World Scientific, Singapore, 1989.

- [49] Kitagawa G. Sakamoto Y., Ishiguro M. *Akaike information Criterion*. KTK Scientific Publishers, Tokio, 1986.
- [50] G. Saporta. *Probabilités, analyse des données et statistique*, page 169. Editions Technip, Paris, 1990.
- [51] J. Serre. *Images analysis and mathematical morphology*. Academic Press, New york, 1982.
- [52] C. E. Shannon. A mathematical theory of communication. *Bell Syst. Tech. J.*, 27:379– 423; 623 – 656, 1948.
- [53] M.J. Shensa. Discrete wavelet transforms: Wedding the 'a trous' and mallat algorithm. *Proc IEEE Trans. Signal Processing*, 40:2464–2482, 1992.
- [54] H. Sjöberg, F. Goudail, and Ph. Réfrégier. Optimal algorithms for target location in non-homogeneous binary images. *Journal of the Optical Society of America A*, 15(12):2976–2985, December 1998.
- [55] K. Styczinsky, J. Campos, M. J. Yzuel, and K. Chalasinska-Macukow. New arrangement for limited intensity invariance pattern recognition with high diffraction efficiency. *Opt. Commun.*, 118:193–198, 1995.
- [56] K. Styczinsky, K. Chalasinska-Macukow, J. Campos, and M. J. Yzuel. Minimum-mean-square-error filters for spatially non-overlapping phase target and input scene noise. *Opt. Commun.*, 127:325–333, 1996.
- [57] A. Rosenfeld and T.-Y. Phillips and A.C. Sher. $O(\log(n))$ bimodality analysis. *Pattern Recognition*, 22:741–746, 1989.
- [58] G. L. Turin. An introduction to matched filters. *IRE Trans. Inf. Th.*, 6(3):311–329, 1960.
- [59] A. Vander Lugt. Signal detection by complex filtering. *IEEE Trans. Inform. Theory*, IT-10:139–145, 1964.
- [60] H. Ventsel. *Théorie des Probabilités*. Edition MIR, Moscou.
- [61] B. V. K. Vijaya Kumar and L. Hassebrook. Performance measures for correlation filters. *Appl. Opt.*, 29(20):2997–3006, 1990.
- [62] S. C. Zhu and A. Yuille. region competition: Unifying snakes, region growing, and bayes/mdl for multiband image segmentation. *IEEE transactions on Pattern Analysis and Machine Intelligence*, 18(9):884–900, 1996.