

N° attribué par la bibliothèque

--	--	--	--	--	--	--	--	--	--

THÈSE

pour obtenir le grade de

DOCTEUR de l'UNIVERSITÉ PARIS 13Spécialité : **Signaux et Images**

présentée et soutenue publiquement par

Laure GENIN

le 22 mars 2013

Titre de la thèse :

**Détection d'objets de petite taille sur
des séquences aériennes ou satellitaires**

Composition du jury :

M.	François GOUDAIL,	Rapporteur
M.	Charles KERVRANN,	Rapporteur
M.	Guy LE BESNERAIS,	Directeur de thèse
Mme.	Françoise DIBOS,	Examineur
M.	Frédéric CHAMPAGNAT,	Examineur
M.	Pierre-Luc GEORGY,	Examineur

Table des matières

Notations	v
Acronymes	ix
Introduction	1
1 Problématique de détection	7
1.1 Modèle additif	7
1.2 Rappels sur la théorie de la détection	8
1.2.1 Formalisation du problème	8
1.2.2 Erreurs moyennes de première et seconde espèce	9
1.2.3 Théorie de Neyman Pearson	10
1.2.4 Test du rapport de vraisemblance en présence de paramètres inconnus	11
1.3 Différentes manières d'appliquer la théorie de la détection	12
1.3.1 Approches globales	13
1.3.2 Approches locales	16
1.4 Méthodes de détection statistiques dans le cadre de la localisation de cibles	16
1.4.1 Hypothèses de départ	17
1.4.2 Filtrage Adapté	19
1.4.3 Détection/Localisation en présence de paramètres inconnus	20
1.4.4 Synthèse	22
1.5 Processus de détection en deux étapes basé sur la modélisation du fond au premier et second ordre	23
1.6 Présentation des principales méthodes de détection	25
1.6.1 Méthodes par suppression de fond : modélisation du fond au premier ordre	25
1.6.2 Autres méthodes de détection	28
1.7 Approche considérée	29
2 Méthode d'évaluation de performances des algorithmes de détection d'objets ponctuels	31
2.1 Choix des critères de performances étudiés	31
2.1.1 Critères basés sur des mesures de statistiques sur le résidu	31
2.1.2 Critères évalués à partir des décisions binaires obtenues après seuillage	33
2.2 Méthode d'évaluation des critères de performances	36
2.3 Evaluation par zones	38
2.4 Images de fond	39

2.4.1	Base de fonds terrestre	39
2.4.2	Base de fonds satellitaire	41
2.5	Modèle de cible	43
3	Modélisation du fond au premier ordre : Développement d’algorithmes de détection de cibles ponctuelles basés sur des méthodes de débruitage à patchs	47
3.1	Introduction de l’approche proposée	48
3.1.1	Approche par suppression de fond	48
3.1.2	Lien avec le débruitage	48
3.1.3	Principe des algorithmes de débruitage à patchs	49
3.1.4	Références sur la détection par méthodes à patchs	51
3.2	Présentation de deux algorithmes de débruitage à patchs	51
3.2.1	Principe de l’algorithme NL-means	52
3.2.2	Principe de l’algorithme BM3D	53
3.3	Adaptation du NL-means et BM3D à la détection	56
3.3.1	Algorithme de détection DNLM	57
3.3.2	Algorithme de détection DBM3D	68
4	Modélisation du fond au second ordre : Développement d’une méthode de détection basée sur une classification des pixels par mélange de gaussiennes	75
4.1	Introduction de la méthode proposée	76
4.1.1	Approche par filtrage adapté adaptatif	76
4.1.2	Limitations des méthodes de détection par filtrage adapté existantes	76
4.1.3	Principe général de l’approche proposée	78
4.2	Algorithme de classification	79
4.3	Classification robuste	80
4.4	Filtrage GMMF	83
4.5	Discussion : variantes de l’algorithme GMMF	84
4.5.1	Variantes possibles de la classification des pixels par CEM	84
4.5.2	Possibilité d’utiliser d’autres méthodes de classification robustes	85
4.5.3	Une approche de filtrage basée sur un mélange de gaussiennes	86
4.6	Paramètres et détails d’implémentation du processus de classification	88
4.6.1	Paramètre N : dimension des patchs	88
4.6.2	Initialisation de l’étape de classification	89
4.6.3	Nombre d’itérations	91
4.6.4	Choix du nombre de classes	93
4.7	Modélisation de la signature des cibles	97
4.7.1	Prise en compte de la déformation de la cible liée au prétraitement des images	97
4.7.2	Prise en compte de la position subpixelique des cibles	99
4.8	Extension au contexte spatio-temporel	107
4.8.1	Classification et filtrage adapté spatio-temporels	108
4.8.2	Classification spatiale et filtrage adapté spatio-temporel	110

5	Évaluation des performances et analyse critique des méthodes de détection proposées	113
5.1	Algorithmes de détection par suppression de fond	113
5.1.1	Rappel des paramètres des filtres étudiés	114
5.1.2	Comparaison avec l'existant	114
5.1.3	Étude de l'influence des paramètres associés aux cibles	115
5.1.4	Caractérisation complémentaire des algorithmes : analyse de la répartition des fausses alarmes et non-détections	119
5.2	Filtres de détection basés sur une modélisation au second ordre	122
5.2.1	Paramètres des filtres étudiés	122
5.2.2	Comparaison des méthodes de détection au second ordre	124
5.2.3	Études des performances en fonction de l'algorithme de BS	128
5.2.4	Recherche d'un meilleur compromis coût de calcul / performances	128
5.2.5	Caractérisation complémentaire des algorithmes : analyse de la répartition des fausses alarmes et non-détections	129
5.3	Évaluation des algorithmes spatio-temporels	133
5.3.1	Algorithmes de suppression de fond	133
5.3.2	Algorithme basé sur une modélisation du fond au second ordre : GMMF	135
	Conclusion générale et perspectives	139
A	Calcul du GLRT dans le cas où le signe de la cible est connu	145
B	Étude paramétrique du DNLM	147
C	Filtre GLRT-GM	153
C.1	Principe du GLRT-GM	153
C.1.1	Expression théorique	153
C.1.2	Approximation	154
C.2	Résultats expérimentaux	154
	Bibliographie	157
	Publications	163

Notations

Notations générales

a	Une lettre italique note un scalaire
\mathbf{a}	Une lettre en gras note un vecteur
\mathbf{M}	Une lettre capitale en gras note une matrice
\mathbf{M}^T	Transposé de la matrice \mathbf{M}
\mathbf{M}^{-1}	Inverse de la matrice \mathbf{M}
$ \mathbf{M} $	Déterminant de la matrice \mathbf{M}
$ a $	Valeur absolue du scalaire a
$\ \cdot\ _2$	Norme euclidienne
$\ \cdot\ _{2,a}$	Norme euclidienne pondérée par une gaussienne
$\text{sign}(a)$	Signe de a
$\hat{\cdot}$	Estimée
$\text{Card}(\Omega)$	Cardinal de l'ensemble Ω
$\{a, \dots, b\}$	Intervalle discret
f^{-1}	Inverse de la fonction f

Notations liées au modèle associé aux images observées

\mathbf{x}	Position du pixel ($\mathbf{x} = (i, j)$ ou $\mathbf{x} = (i, j, t)$, $\mathbf{x} \in \mathcal{X}$)
\mathcal{X}	Domaine de définition de l'image ou de la séquence d'images observée
\mathbf{u}	Données observées (une image ou séquence d'images)
\mathbf{b}	Composante de fond des données observées
s_q	Composante des données observées associée à la cible numéro q
\mathbf{S}	Composante des données observées associée à l'ensemble des cibles présentes ($\mathbf{S} = \sum_q s_q(\mathbf{x})$)
$u(\mathbf{x}), b(\mathbf{x}), s_q(\mathbf{x}), S(\mathbf{x})$	Intensité respectivement de \mathbf{u} , \mathbf{b} , s_q et \mathbf{S} au pixel \mathbf{x}
\mathbf{u}	Versión vectorielle de \mathbf{u} : $\mathbf{u} = [u(\mathbf{x}), \mathbf{x} \in \mathcal{X}]$
\mathbf{b}	Versión vectorielle de \mathbf{b} : $\mathbf{b} = [b(\mathbf{x}), \mathbf{x} \in \mathcal{X}]$
\mathbf{S}	Versión vectorielle de la somme de l'ensemble des composantes associées aux cibles : $\mathbf{S} = [\sum_q s_q(\mathbf{x}), \mathbf{x} \in \mathcal{X}]$
\mathcal{I}	Domaine de définition de l'image en cours de traitement, $\mathcal{I} \subseteq \mathcal{X}$

Notations liées au modèle de cible

r_c	Fréquence de coupure normalisée
f_c	Fréquence de coupure optique
f_e	Fréquence d'échantillonnage du capteur
$J_1(\cdot)$	Fonction de Bessel de première espèce
$\mathbf{x}_c = (i_c, j_c)$	Position pixellique de la cible
$\boldsymbol{\epsilon} = (\epsilon_i, \epsilon_j)$	Décalage subpixellique de la cible par rapport au centre du pixel \mathbf{x}_c
s_o	Intensité globale de la cible
$s(\mathbf{x})$	Intensité de la cible au pixel \mathbf{x}
$h_o(\cdot)$	Réponse impulsionnelle de l'optique

Notations liées aux tests statistiques

H_1	Hypothèse : cible présente
H_0	Hypothèse : cible absente
$a \underset{H_0}{\overset{H_1}{\geq}} b$	H_1 est validé si $a \geq b$, H_0 est validé si $a < b$
\mathbf{u}	Vecteur d'observation
U	Variable aléatoire dont est issue l'observation \mathbf{u}
M	Dimension du vecteur \mathbf{u}
$\Lambda(\mathbf{u})$	Observation filtrée à seuiller
μ	Seuil appliqué pour la détection
Ω	Ensemble des vecteurs observables
Ω_0 et Ω_1	Sous-ensembles complémentaires de Ω correspondant à choisir respectivement H_1 et H_0
$p(\cdot)$	Densité de probabilité
$p(\cdot \cdot)$	Densité de probabilité conditionnelle
$\mathbb{P}(\cdot)$	Mesure de probabilité
$\mathbb{P}(\cdot \cdot)$	Mesure de probabilité conditionnelle
θ_0, θ_1	Paramètres inconnus des lois de probabilités sous respectivement l'hypothèse H_0 et H_1
$\phi_{H_0}(\cdot), \phi_{H_1}(\cdot)$	Fonction de répartition de la variable aléatoire $\Lambda(\mathbf{u})$ sachant H_0 respectivement H_1

Notations liées à la détection / localisation de cibles

\mathbf{u}_x	Observation associée au pixel \mathbf{x}
\mathbf{b}_x	Composante de fond de l'observation associée au pixel \mathbf{x}
\mathbf{s}	Signature de la cible recherchée
α	Intensité de la cible
U_x	Variable aléatoire dont est issue l'observation \mathbf{u}_x
N	Dimension du côté du voisinage centré sur \mathbf{x} utilisé pour définir \mathbf{u}_x
$\mathcal{N}(\mathbf{m}_x, \Sigma_x)$	Loi gaussienne multivariée de paramètre \mathbf{m}_x et Σ_x associée à U_x
\mathbf{m}_x	Moyenne : statistique du 1 ^{er} ordre associée à U_x
Σ_x	Matrice de covariance : statistique du 2 nd ordre associée à U_x

Notations liées au débruitage

u	Données bruitées, données observées (une image ou séquence d'images)
v	Données « vraies » (exemptes de bruit), données non observées
n	Bruit présent dans les données
\widehat{v}	Données débruitées estimées

Notations liées aux méthodes « à patches »

P_x	Patch d'indice x
\widehat{P}_x	Estimée du patch P_x
C_{xR}	Ensemble des indices des patches considérés similaires au patch de référence d'indice xR
$\mathbf{P}_{C_{xR}}$	Groupe (tableau 3D) constitué des patches $\{P_x\}_{x \in C_{xR}}$
$\widehat{\mathbf{P}}_{C_{xR}}$	Estimée du groupe (tableau 3D) $\mathbf{P}_{C_{xR}}$
\widehat{P}_x^{xR}	Estimée du patch P_x extrait du groupe $\widehat{\mathbf{P}}_{C_{xR}}$
N_1	Dimension du côté des patches carrés
δ_{N_1}	Rayon des patches de dimension impaire : $\delta_{N_1} = (\frac{N_1-1}{2}, \frac{N_1-1}{2})$
$P_{x-\delta_{N_1}}, \widehat{P}_{x-\delta_{N_1}}$	Patch bruité, estimé, de dimension impaire N_1 et centré sur le pixel x
S_x	Support de recherche, voisinage du pixel x dans lequel on recherche des patches similaires
N_S	Dimension spatiale de S_x (carré de dimension $N_S \times N_S$ pixels)
L	Dimension temporelle de S_x
$d(P_x, P_y)$	Mesure de dissimilarité entre P_x et P_y
$w(P_{x-\delta_{N_1}}, P_{y-\delta_{N_1}})$	Poids associé au pixel y dans l'estimation du pixel x par NL-means
h	Paramètre du NL-means ajustant l'exigence de similarité entre les patches
σ	Ecart-type <i>a priori</i> du bruit intervenant dans BM3D
τ_{match}	Seuil sur $d(P_x, P_y)$ utilisé pour déterminer si P_x et P_y sont similaires ou non
N_2	Nombre de patches maximum par groupe
N_{step}	Pas entre les pixels xR qui définissent les patches de références
N_{har}^{xR}	Nombre de coefficients en ondelette retenus (non-nuls)
w_{xR}	Pondération associée au groupe $\widehat{\mathbf{P}}_{C_{xR}}$

Notations liées aux méthodes basées sur la modélisation du fond au second ordre

r	Image du résidu en sortie d'un algorithme de suppression de fond
\mathbf{r}_x	Observation (complète) associée à \mathbf{x} : version vectorielle du patchs $P_{x-\delta}$ extrait de r
N	Dimension du côté du voisinage centré sur \mathbf{x} utilisé pour définir \mathbf{r}_x
k_x, \widehat{k}_x	Indice, resp. indice estimé, de la distribution associée au vecteur \mathbf{r}_x
K	Nombre de distributions gaussiennes constituant le modèle de mélange de gaussiennes
$\mathbf{r}_x^{\text{trunc}}$	Observation restreinte associée à \mathbf{x}
n	Dimension du côté du trou central de la couronne utilisé pour définir $\mathbf{r}_x^{\text{trunc}}$
$\Sigma_k, \widehat{\Sigma}_k$	Matrice de covariance, matrice de covariance estimée, de la $k^{\text{ème}}$ distribution du mélange associé aux observations complètes
$\Sigma_k^{\text{trunc}}, \widehat{\Sigma}_k^{\text{trunc}}$	Matrice de covariance, matrice de covariance estimée, de la $k^{\text{ème}}$ distribution du mélange associé aux observations restreintes
C_k	Ensemble des pixels \mathbf{x} associés dont le label est k_x
$\pi_k, \widehat{\pi}_k$	Poids, poids estimé, de la $k^{\text{ème}}$ distribution gaussienne du mélange
ν_K	Degré de liberté du modèle de mélange de gaussiennes
d	Dimension du vecteur d'observation utilisé dans l'étape de classification
NO	Nombre total d'observations classifiées
\mathbf{s}_ϵ	Signature des cibles ayant un décalage subpixelique ϵ

Acronymes

Pour faciliter la correspondance entre le présent manuscrit et la bibliographie, la majorité des acronymes est conservée en langue anglaise.

ALRT	<i>Approximate Likelihood Ratio Test</i>
AMF	<i>Adaptive Matched Filter (Filtrage Adapté Adaptatif)</i>
ANF	<i>Adaptive Normalized Filter</i>
BM3D	<i>Block Matching and 3D Filtering</i>
BS	<i>Background Suppression (Suppression de fond)</i>
CEM	<i>Classification Expectation-Maximisation</i>
CFAR	<i>Constant False Alarm Rate (TFAC : Taux de Fausses Alarmes Constant)</i>
DBM3D	<i>Detection by BM3D (Détection par BM3D)</i>
DBT	<i>Detect Before Track (Détection avant pistage)</i>
DNFA	<i>Distance to the Nearest False Alarm</i>
DNLM	<i>Detection by NLM (détection par NL-Means)</i>
ELRT	<i>Exact Likelihood Ratio Test</i>
EM	<i>Expectation-Maximisation</i>
FA	<i>False Alarm (Fausses Alarmes)</i>
GLRT	<i>Generalized Likelihood Ratio Test (TRVG : Test du Rapport de Vraisemblance Généralisé)</i>
GMF	<i>Global Matched Filter (Filtrage Adapté Global)</i>
GMMF	<i>Gaussian Mixture Matched Filter (Filtrage Adapté par Mélange de Gaussiennes)</i>
ICL	<i>Integrated Classification Likelihood</i>
LRT	<i>Likelihood Ratio Test (TRV : Test du rapport de vraisemblance)</i>
MF	<i>Matched Filter (Filtrage Adapté)</i>
ML	<i>Maximum Likelihood (MV : maximum de vraisemblance)</i>
NL-means	<i>Non-Local Means</i>
Pd	<i>Probability of Detection (Probabilité de Détections)</i>
Pfa	<i>Probability of False Alarm (Probabilité de Fausses Alarmes)</i>
Pnd	<i>Probability of Non-Detection (Probabilité de Non-Détection)</i>
PSF	<i>Point Spread Function (Réponse impulsionnelle de l'optique)</i>
ROC	<i>Receiver Operating Characteristic (COR : Caractéristique Opérationnelle du Récepteur)</i>
TBD	<i>Track Before Detect (Pistage avant détection)</i>

Introduction

Contexte

Les travaux de recherche présentés dans ce manuscrit s'inscrivent dans le cadre d'un partenariat entre l'ONERA (Office National d'Études et de Recherches Aérospatiales) et la société Astrium, filiale du groupe EADS, ayant pour but de développer une nouvelle génération d'algorithmes de détection d'objets ponctuels dans des séquences d'images.

Applications et enjeux

La détection automatique d'objets de petite taille est un sujet d'intérêt fortement croissant dans de nombreux domaines tels que celui de la défense et de la sécurité [Vasquez *et al.*, 2010], du biomédical [Olivo-Marín, 2002, Boulanger *et al.*, 2006, Grosjean et Moisan, 2009] ou de l'astronomie [Herranz *et al.*, 2011]. L'application considérée au cours de ces travaux est la veille passive. Elle consiste à détecter de façon automatique des petits objets sur un champ large en vue de piloter d'autres systèmes d'alertes de champ plus étroit permettant d'apporter plus d'informations sur les menaces détectées qui peuvent correspondre à des avions, hélicoptères, missiles ou tout autre objet volant. Les applications visées concernent aussi bien le cas de caméras au sol ou embarquées à bord d'avions pointant en incidence rasante vers l'horizon que de capteurs embarqués à bord de satellites et observant le sol. Dans ce contexte le but est de détecter les cibles au plus tôt, c'est à dire de les détecter lorsqu'elles sont très éloignées du capteur. Les cibles se présentent alors comme des spots lumineux de dimension inférieure ou de l'ordre de celle du pixel et sont dites ponctuelles ou non-résolues.

Vision Detect Before Track versus Track Before Detect

Parmi les diverses et nombreuses méthodes de détection d'objets ponctuels existantes, on peut distinguer deux types d'approches [Blostein et Richardson, 1994, Samson, 2002, Askar *et al.*, 2002].

La première, l'approche historique, divise le problème de détection en deux étapes (cf. figure 1a). Une étape de pré-détection consiste tout d'abord à réduire le flot d'informations arrivant sur un délai relativement court (une seule image ou un nombre limité d'images). En tirant parti le mieux possible de la corrélation spatio-temporelle du fond, elle déclare pour chaque image si des cibles sont présentes ou non et fournit leurs positions (ces pré-détections sont appelées plots). Puis, dans un second temps, une étape de pistage des plots obtenus permet d'éliminer un certain nombre de fausses alarmes. Cette approche est souvent désignée par l'expression *Detect Before Track (DBT)* ou *Track After Detect*. Elle a l'avantage de permettre de réduire considérablement la quantité de données à traiter lors

du pistage. Cependant un des risques avec cette approche est la non-détection de cibles lors de l'étape de pré-détection. C'est pourquoi il est important d'être prudent quant au choix du paramètre de seuillage de l'algorithme de pré-détection qui gère un compromis entre les non-détections et les fausses alarmes des cibles avant l'étape de pistage.

La seconde approche, nommée *Track Before Detect* (TBD) (cf. figure 1b), est basée sur une certaine connaissance *a priori* des dynamiques des objets d'intérêt [Davey *et al.*, 2008, Zhang *et al.*, 2005, Johnston et Krishnamurthy, 2002, Bruno et Moura, 2000, Tonissen et Evans, 1996, Blostein et Huang, 1991, Reed *et al.*, 1983]. La détection et le pistage sont rassemblés dans une même et unique étape dans le but d'essayer d'améliorer la détection des cibles de faible intensité [Davey *et al.*, 2008]. Cette approche consiste à chercher dans la séquence d'images les trajectoires qui correspondent *a priori* à celles d'une cible en travaillant directement sur l'ensemble du flot d'informations. Pour améliorer les performances de détection, notamment en présence de fonds texturés tels que l'on peut observer en imagerie optique, il est souvent nécessaire d'appliquer des prétraitements permettant de supprimer le fond et de faire ressortir les cibles par rapport à celui-ci (cf. figure 1b). Carnie *et al.* proposent par exemple d'appliquer des filtrages morphologiques afin de faire ressortir les cibles par rapport au fond préalablement à l'application d'un algorithme d'accumulation de l'intensité des cibles [Carnie *et al.*, 2006]. Cette approche TBD ne faisant intervenir aucune étape de pré-détection, nécessite la résolution de problèmes d'optimisation dans un domaine de recherche très grand et donc des calculs intensifs.

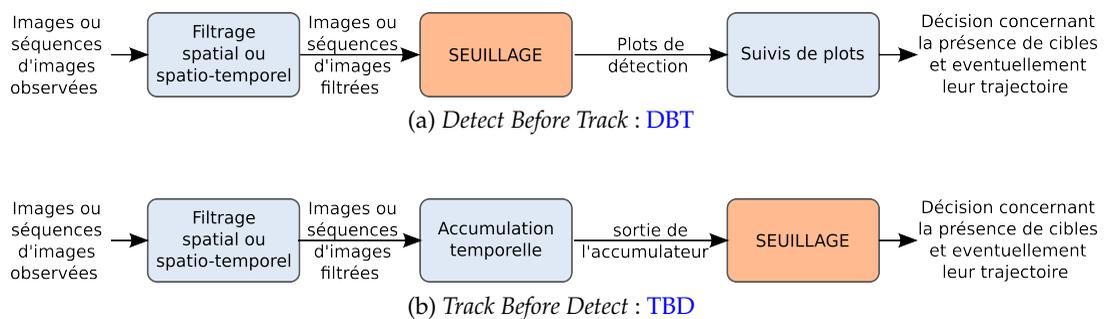


FIGURE 1 – Schéma de principe des approches de détection DBT et TBD.

Dans cette thèse, notre attention s'est focalisée sur l'amélioration de la prise en compte de la corrélation spatio-temporelle du fond qui intervient dans l'étape de pré-détection des approches DBT mais qui peut aussi tout à fait intervenir dans des étapes de prétraitement d'approches TBD. En effet le rehaussement des cibles par rapport au fond en tirant parti de la corrélation spatio-temporelle des fonds est une étape clé dans les processus de détection. L'amélioration de cette étape de traitement peut ainsi permettre d'améliorer les performances de nombreuses chaînes de détection plus complexes.

Limitations des approches actuelles basées sur la prise en compte de la corrélation spatio-temporelle du fond

Les méthodes de détection basées sur la prise en compte de la corrélation temporelle des fonds échouent dans les contextes où les fonds ne sont que très faiblement corrélés temporellement. On se retrouve très souvent face à des fonds non-corrélés quand la ca-

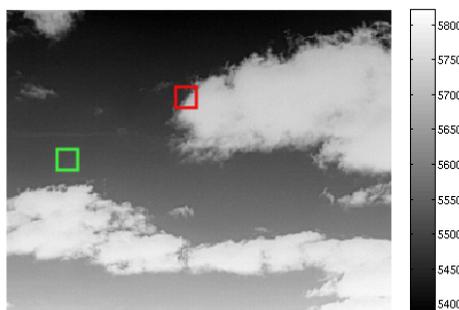


FIGURE 2 – Image typique de fond présentant des zones uniformes (carré vert) et des zones avec de forts gradients de niveaux de gris (carré rouge).

dence d'acquisition diminue. Cela se produit quand on a par exemple à faire à des capteurs bon marché qui ont une fréquence d'acquisition faible ou à des capteurs tournant à 360° comme ceux utilisés pour la veille panoramique [Vasquez *et al.*, 2010], ou encore lorsque l'on travaille avec des systèmes embarqués fournissant des images très volumineuses qui ne peuvent donc pas être transmises à une cadence élevée, comme c'est le cas avec les satellites météorologiques. Un autre cas où l'on doit faire face à des fonds non-corrélés est par exemple le cas où la scène est imagée en présence de très fortes turbulences. De plus ces méthodes échouent aussi dans le contexte où les cibles sont quasiment statiques vis à vis du fond. Dans ce contexte, soit aucune précaution n'est prise et ces cibles sont considérées comme faisant partie du fond et ne sont donc pas détectées ; soit on diminue la cadence de traitement pour que le mouvement des cibles permette de les distinguer du fond et l'on se retrouve dans un contexte où la corrélation temporelle des fonds est faible.

Dans les contextes de détection qui viennent d'être décrits il est le plus souvent favorable de se restreindre à des approches de détection dites spatiales utilisant uniquement l'information contenue dans l'image traitée. Cependant les méthodes de détection basées sur la corrélation spatiale des fonds ont des difficultés à faire face à des fonds présentant des zones de fortes variations de niveaux de gris telles que les bords de nuages (cf. Fig.2). En effet, des méthodes simples permettent de détecter des objets ponctuels dans des fonds homogènes mais leur taux de fausses alarmes augmente rapidement quand de forts gradients d'intensité sont présents.

Axe de recherche considéré

Le travail de cette thèse a consisté à développer de nouveaux algorithmes de détection d'objets ponctuels plus performants dans le contexte de fonds inhomogènes présentant de fortes variations d'intensité et pour des dynamiques de cible qui mettent en échec les approches usuelles de détection temporelles. Il s'est principalement focalisé sur l'amélioration de la prise en compte de la corrélation spatiale des fonds. Cependant, comme la corrélation temporelle des fonds est souvent importante, des extensions temporelles des méthodes proposées ont été réalisées.

Les filtres que nous cherchons à développer reposent sur une forme de modélisation des inhomogénéités du fond mais il est important que cette modélisation soit robuste à la présence de cibles. En effet, comme une cible ponctuelle apparaît elle-même comme une inhomogénéité, les filtres performants risquent souvent de confondre les cibles avec des instationnarités du fond et de conduire à un taux de cibles non détectées très élevé.

L'objectif de ce travail de recherche est donc de proposer des filtres qui réalisent un bon compromis entre adaptabilité aux inhomogénéités et robustesse à la présence de cibles.

Contributions

Fil conducteur de notre approche

En nous basant sur un formalisme du test du rapport de vraisemblance généralisé (**GLRT** : *Generalized Likelihood Ratio Test*), nous proposons de ramener le problème de détection à deux étapes qu'on peut associer à l'estimation des statistiques locales du fond au premier et second ordre. La première étape de ce schéma consiste à appliquer une méthode de suppression de fond nécessitant une modélisation du fond au premier ordre autrement dit une estimation de la moyenne du fond. La seconde consiste à appliquer un filtrage adapté (**MF** : *Matched Filter*) qui nécessite une modélisation du fond au second ordre autrement dit une estimation de la matrice de covariance en chaque pixel du fond.

Modélisation au premier ordre par méthode à patchs

Nous proposons de nouvelles méthodes de suppression de fond très performantes fondées sur l'utilisation d'approches "à patchs" issues du domaine du débruitage. Ces filtres de débruitage permettent d'obtenir une modélisation au premier ordre efficace de fonds très inhomogènes. Néanmoins il convient de les adapter pour les rendre robustes aux cibles. Nous proposons tout d'abord une nouvelle adaptation du filtre de débruitage **NL-means** (*Non-Local Means*) introduit par Buades *et al.* en 2005 [[Buades *et al.*, 2005b](#)] qui a marqué un tournant dans l'histoire du débruitage. Nous montrons ensuite que le filtre **BM3D** (*Block Matching and 3D Filtering*) [[Dabov *et al.*, 2007b](#)] considéré actuellement comme un des filtres de débruitage les plus performants, peut être adapté aisément pour conduire à une méthode d'estimation de la moyenne du fond efficace permettant de bien préserver les zones de fortes transitions [[Genin *et al.*, 2012a](#)]. On propose des extensions spatio-temporelles de ces approches qui gardent l'avantage des méthodes de débruitage à patchs de ne pas nécessiter d'estimation précise préalable du mouvement du fond.

Une approche par classification non supervisée de la modélisation au second ordre

Quelque soit la performance de l'étape de suppression de fond, nous montrons que les résidus présentent des textures inhomogènes qu'il est important de prendre en compte pour la détection. Pour ce faire, nous proposons une nouvelle méthode d'estimation des statistiques au second ordre du résidu fondée sur une étape de classification non supervisée. Notre approche consiste à classifier les pixels de l'image par un algorithme **CEM** (*Classification Expectation-Maximisation*) basé sur un modèle de mélange de gaussiennes et à estimer une matrice de covariance pour chaque classe. Cette matrice est ensuite utilisée pour appliquer un **MF** local à chaque pixel. Une question clé est la robustesse de la classification face à la présence de cibles. Ce problème, qui a donné lieu à des développements parfois très complexes dans la littérature, est résolu ici de façon très simple par le choix d'un vecteur d'observation spécifique dans l'algorithme **CEM**. Ce choix conduit à de très bonnes propriétés de robustesse sans modifier la structure du filtre. Nous proposons ensuite une version temporelle de cette étape de modélisation au second ordre capable

de tirer avantage de la corrélation temporelle des fonds s'il y en a, tout en restant très performante en présence de fonds temporellement non-corrélés.

L'association des méthodes de détection basées sur la modélisation du fond au premier et second ordre conduit à de nouveaux filtres de détection très performants car capables de bien gérer la présence de textures et de zones de fortes transitions dans les fonds.

Une étude de performance sur fonds réels

Pour évaluer l'apport des méthodes proposées nous avons effectué une importante étude de performance utilisant deux bases d'images de fonds réels et des cibles simulées. Le choix des fonds et les caractéristiques de l'imageur utilisées pour la simulation des cibles sont représentatifs des contextes applicatifs qui nous intéressent. Cette étude comparative, intégrant des techniques classiques de la littérature, nous permet de mettre en évidence le rôle et l'importance de chacune des deux étapes de modélisation qui structurent notre travail.

Organisation du manuscrit

Ce mémoire s'articule en cinq chapitres de la façon suivante :

Le premier chapitre fait tout d'abord un rappel de la théorie de la détection dans un cadre statistique. De cette théorie, et plus précisément du test du rapport de vraisemblance généralisé ([GLRT](#) : *Generalized Likelihood Ratio Test*), on fait ressortir le schéma de détection en deux étapes constituant le fil conducteur de notre démarche dans toute cette thèse. Un état de l'art des différentes méthodes de détection est ensuite réalisé dans lequel on montre que le schéma de détection proposé fournit un cadre interprétatif adapté à un grand nombre de méthodes existantes.

Le deuxième chapitre traite de l'évaluation des performances des algorithmes et présente les deux bases de données de fonds infrarouges (IR) étudiées. Il s'agit d'images de fond nuageux, acquises par une caméra au sol pour la première base et par une caméra satellitaire pour la seconde. Ce chapitre discute des avantages et inconvénients des solutions actuelles utilisées pour évaluer les performances et présente la méthode sélectionnée. Une solution basée sur des fonds réels dans lesquels des cibles simulées sont incrustées est retenue, permettant ainsi d'avoir une évaluation des performances de détection en tout pixel.

Le troisième chapitre introduit les nouveaux algorithmes de détection par modélisation du fond au premier ordre que nous proposons. Un rappel du principe de l'intérêt des méthodes de débruitage à patchs est tout d'abord réalisé. Ce chapitre présente ensuite, dans un premier temps, l'adaptation du filtre de débruitage [NL-means](#) que nous proposons conduisant au filtre de détection que nous nommons [DNLM](#) (*Detection by NLM*). Nous présentons finalement le filtre de détection [DBM3D](#) (*Detection by BM3D*), basé sur le filtre de débruitage [BM3D](#).

Le quatrième chapitre introduit la nouvelle méthode de détection basée sur une modélisation du fond au second ordre que nous proposons. Cette méthode est nommée [GMMF](#) (*Gaussian Mixture Matched Filter*).

Dans le cinquième chapitre une évaluation de performances détaillée des différents algorithmes de détection est réalisée. Elle permet de comparer les propriétés de chacun des algorithmes et ainsi définir celui qui sera le plus approprié en fonction du contexte.

Enfin nous concluons ce travail en dégageant d'une part les points importants de cette recherche et d'autre part les perspectives qu'elle offre.

Chapitre 1

Problématique de détection

Introduction

Ce chapitre présente la problématique de détection d'objets ponctuels dans un cadre statistique et justifie la démarche de détection en deux étapes qui a été adoptée au cours de ces travaux de recherche.

Organisation du chapitre — Le modèle considéré pour les données dans le cadre de notre problématique de détection est tout d'abord présenté dans la section 1.1. La section 1.2 fait un rappel de la théorie statistique de la détection, puis la section 1.3 discute des différentes manières possibles d'utiliser cette théorie en distinguant une approche dite globale d'une approche dite locale. Après une présentation en section 1.4 des différents tests qui peuvent découler de la théorie de la détection en considérant l'hypothèse d'un fond gaussien, la section 1.5 introduit le schéma de détection en deux étapes qui sera le fil conducteur de notre démarche dans toute la suite du manuscrit. La section 1.6 réalise un état de l'art des méthodes de détection existantes en montrant qu'une grande majorité peuvent être interprétées par le schéma proposé de détection en deux étapes.

1.1 Modèle additif

L'image ou la séquence d'images observée peut être modélisée par l'équation suivante sous l'hypothèse que Q cibles sont présentes :

$$u(\mathbf{x}) = \sum_{q=1}^Q s_q(\mathbf{x}) + b(\mathbf{x}) = S(\mathbf{x}) + b(\mathbf{x}) \quad (1.1)$$

où $u(\mathbf{x})$ est l'intensité observée au pixel de position $\mathbf{x} = (i, j, t)$, $\mathbf{x} \in \mathcal{X}$ avec \mathcal{X} le domaine de définition de l'image ou la séquence d'images observée, $b(\mathbf{x})$ est l'intensité du fond au pixel \mathbf{x} , $s_q(\mathbf{x})$ est la contribution d'intensité au pixel \mathbf{x} de la cible numéro q et $S(\mathbf{x})$ est la contribution d'intensité au pixel \mathbf{x} de l'ensemble des cibles ($S(\mathbf{x}) = \sum_{q=1}^Q s_q(\mathbf{x})$). Comme nous cherchons à détecter des cibles, celles-ci se trouvent à une grande distance du capteur et l'approximation consistant à les considérer comme un signal additif est tout à fait justifiée. En effet chaque cible ne masque qu'une surface négligeable du fond par rapport à la surface globale de fond visible par un pixel du capteur.

De façon générale, on peut différencier la problématique de détection de celle de la localisation. La première, la détection, consiste à discriminer deux hypothèses correspondant respectivement à l'absence de cible et à la présence d'une ou plusieurs cibles. La seconde, la localisation, consiste quant à elle à déterminer la position des différentes cibles éventuellement présentes. Dans la section suivante on se limite à la problématique de détection, elle a pour but de reprendre les bases de la théorie de la détection dans un cadre statistique.

1.2 Rappels sur la théorie de la détection

1.2.1 Formalisation du problème

Un problème de détection consiste à faire un choix parmi deux hypothèses. Dans le cadre de notre contexte de détection d'objets ponctuels, les deux alternatives à discriminer sont les suivantes :

Hypothèse H_1 : Un ou plusieurs objets ponctuels sont présents.

Hypothèse H_0 : Il n'y a aucun objet ponctuel présent.

Le but de tout algorithme de détection est de chercher à tirer parti au mieux de l'information disponible afin de choisir parmi ces deux hypothèses celle qui est la plus vraisemblable. Dans ce qui suit nous noterons $\mathbf{u} = [u(\mathbf{x}), \mathbf{x} \in \mathcal{X}]$, le vecteur contenant l'intensité de tous les pixels de l'image ou de la séquence d'images observée.

Ce vecteur \mathbf{u} correspondant aux données observées est un vecteur appartenant à l'ensemble Ω rassemblant tous les vecteurs observables. Cet ensemble est un sous espace de \mathbb{R}^M où M est le nombre de pixels présents dans l'image ou de la séquence d'images observée. L'espace Ω est donc de très grande dimension.

Un algorithme de détection doit être capable pour tout vecteur d'observation \mathbf{u} de l'espace Ω de décider si celui-ci correspond à l'absence ou à la présence de cibles, autrement dit si c'est l'hypothèse H_1 ou H_0 qui est vraie. Le modèle mathématique correspondant à cette décision est une fonction de $\Omega \rightarrow \{H_0, H_1\}$. Elle est équivalente à diviser l'ensemble Ω en deux sous-ensembles Ω_1 et Ω_0 associés respectivement aux hypothèses H_1 et H_0 [Goudail, 2011].

On peut toujours décomposer une telle fonction de décision en deux étapes :

$$\Omega \longrightarrow \mathbb{R} \longrightarrow \{H_0, H_1\}. \quad (1.2)$$

La première étape correspond à l'application d'une fonction Λ qui projette les données mesurées dans \mathbb{R} et la seconde à un test à seuil :

- **Si $\Lambda(\mathbf{u}) \geq \mu$** : Choix de l'hypothèse H_1 qui correspond à la présence d'un ou plusieurs objets ponctuels.
- **Si $\Lambda(\mathbf{u}) < \mu$** : Choix de l'hypothèse H_0 qui correspond à l'absence d'objet ponctuel.

La fonction Λ et le seuil μ appliqués, définissent la frontière entre les deux sous-ensembles Ω_1 et Ω_0 .

Cette décomposition n'est pas simplement formelle elle fait apparaître un paramètre μ qui en pratique gère un compromis entre deux types d'erreurs :

- L'erreur de première espèce correspond à décider que l'objet est présent alors que celui-ci ne l'est pas, c.a.d à choisir l'hypothèse H_1 alors que c'est l'hypothèse H_0 qui est vraie. Dans ce cas on dit que l'on se trouve en présence d'une fausse alarme (**FA**).
- L'erreur de seconde espèce correspond à décider que l'objet est absent alors que celui-ci ne l'est pas, c.a.d à choisir l'hypothèse H_0 alors que c'est l'hypothèse H_1 qui est vraie. Dans ce cas on dit que l'on se trouve en présence d'une non-détection.

L'enjeu d'un problème de détection est de trouver la fonction Λ optimale, qui permette de discriminer au mieux les deux hypothèses H_1 et H_0 .

1.2.2 Erreurs moyennes de première et seconde espèce

Comme une image peut être considérée comme étant la réalisation de phénomènes aléatoires, on considère que l'observation \mathbf{u} est la réalisation d'un vecteur aléatoire, notée U . Dans ce cas, l'observation filtrée $\Lambda(\mathbf{u})$ est elle aussi la réalisation d'une variable aléatoire et l'occurrence d'une erreur de première ou deuxième espèce dépend de cette réalisation. Concernant ces deux erreurs, il est donc important de s'intéresser à leur probabilité d'occurrence aussi appelée erreur moyenne.

- L'erreur moyenne de première espèce est plus couramment nommée probabilité de fausses alarmes (**Pfa**) et est égale à :

$$P_{FA}(\mu) = \mathbb{P}(\Lambda(\mathbf{u}) \geq \mu | H_0) = \int_{\mu}^{\infty} p(t|H_0) dt. \quad (1.3)$$

où $p(t|H_0)$ est la densité de probabilité de la variable aléatoire $\Lambda(\mathbf{u})$ sachant que l'hypothèse H_0 est vraie.

- L'erreur moyenne de seconde espèce est plus couramment nommée probabilité de non-détection (**Pnd**) et est égale à :

$$P_{ND}(\mu) = \mathbb{P}(\Lambda(\mathbf{u}) < \mu | H_1) = \int_{-\infty}^{\mu} p(t|H_1) dt. \quad (1.4)$$

où $p(t|H_1)$ est la densité de probabilité de la variable aléatoire $\Lambda(\mathbf{u})$ sachant que l'hypothèse H_1 est vraie.

En détection, on s'intéresse le plus souvent à la probabilité de détection (**Pd**) qui correspond à la probabilité d'occurrence de choisir l'hypothèse H_1 quand celle-ci est vraie :

$$P_D(\mu) = 1 - P_{ND}(\mu) = \mathbb{P}(\Lambda(\mathbf{u}) \geq \mu | H_1) = \int_{\mu}^{\infty} p(t|H_1) dt. \quad (1.5)$$

On voit d'après (1.3) et (1.4) que ces erreurs sont antagonistes. De plus, elles ont en général un impact différent selon l'application visée. Dans le domaine médical on peut désirer ne pas manquer une pathologie quitte à alerter le patient à tort. En télécommunications les deux types d'erreur ont le même impact. Dans les systèmes de veille, la puissance de calcul disponible limite le débit de fausses alarmes qu'il est possible de gérer, ceci aboutit en général à des exigences plus importantes sur les fausses alarmes que sur les non détections.

A ce niveau on a trois quantités qui sont liées, qui sont le seuil μ , la **Pd** et la **Pfa** et c'est le réglage du seuil de détection μ qui permet de gérer le compromis entre les deux types d'erreurs. Il est très classique de représenter la **Pd** en fonction de la **Pfa** pour

différentes valeurs de seuil, la courbe ainsi obtenue est appelée courbe **ROC** (*Receiver Operating Characteristic*). La relation théorique entre la **Pd** et la **Pfa** est donnée par la formule suivante :

$$P_D(P_{FA}) = 1 - \phi_{H_1}(\phi_{H_0}^{-1}(1 - P_{FA})) \quad (1.6)$$

où ϕ_{H_0} (respectivement ϕ_{H_1}) est la fonction de répartition de la variable aléatoire $\Lambda(\mathbf{u})$ sachant H_0 (respectivement H_1) vrai. La figure 1.1 présente la forme classique des courbes **ROC**. Pour obtenir un taux de fausses alarmes faible, on applique sur l'image en sortie de filtrage un seuil de détection μ élevé et le taux de détection résultant est donc faible. Quand le taux de fausses alarmes augmente, le seuil de détection est abaissé et le taux de détection est donc augmenté.

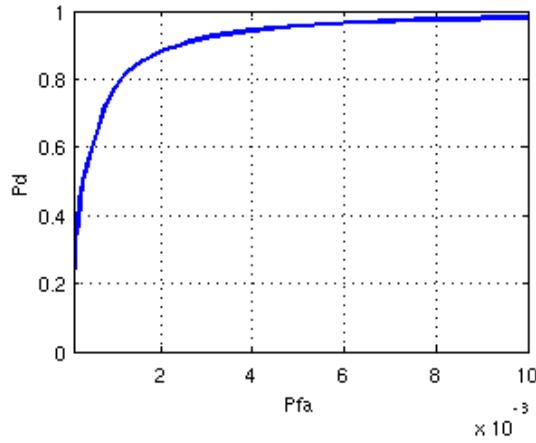


FIGURE 1.1 – Exemple de courbe **ROC**.

Dans les applications de détection de cible, le paramètre μ est généralement fixé en fonction du taux de fausses alarmes que l'on s'autorise. Pour un point de fonctionnement donné, autrement dit une **Pfa** fixée, un filtre de détection est meilleur qu'un autre si la probabilité de détection obtenue avec ce filtre est supérieure à celle obtenue avec l'autre filtre.

1.2.3 Théorie de Neyman Pearson

Dans l'idéal on souhaiterait être capable de réduire simultanément les erreurs moyennes de première et seconde espèce autrement dit la **Pfa** et la **Pnd**. Cependant cela n'est pas possible car les deux types d'erreurs ont des fonctions de coûts contradictoires. Un compromis doit donc forcément être fait entre **Pnd** et **Pfa**. Un bon compromis est donné par la stratégie de Neyman-Pearson qui consiste à maximiser la **Pd** (ce qui est équivalent à minimiser la **Pnd**) pour une **Pfa** constante fixée [Van Trees, 1968]. Cette stratégie est équivalente à seuiller le rapport de vraisemblance défini comme le rapport des densités des lois conditionnelles. Ce test est nommé test du rapport de vraisemblance (**LRT** : *Likelihood Ratio Test*) et s'exprime de la façon suivante [Van Trees, 1968] :

$$\Lambda_{LRT}(\mathbf{u}) = \frac{p(\mathbf{u} | H_1)}{p(\mathbf{u} | H_0)} \underset{H_0}{\overset{H_1}{\geq}} \mu \quad (1.7)$$

où $p(. | H_0)$ (respectivement $p(. | H_1)$) correspond à la densité de probabilité de la variable aléatoire U sous l'hypothèse H_0 (respectivement H_1) et μ est le seuil de détection.

Dans le cadre où l'on connaît parfaitement les lois de distribution de la variable aléatoire U selon chacune des hypothèses, ce test correspond au test optimal : aucun algorithme ne peut fournir de courbe ROC au-dessus de celle du rapport de vraisemblance. Remarquons que cette stratégie de seuillage du rapport de vraisemblance est aussi la stratégie optimale dans le cadre bayésien : c'est celle qui minimise le risque bayésien [Barret, 2009, Van Trees, 1968].

La stratégie de Neyman Pearson permet de présenter la problématique de détection dans un cadre théorique qui n'est quasiment jamais applicable tel quel en pratique car on ne dispose pas des densités de probabilité de la variable aléatoire U sous chacune des hypothèses. Afin de pouvoir appliquer le test du rapport de vraisemblance, ces probabilités doivent être modélisées par des lois de distribution paramétriques.

1.2.4 Test du rapport de vraisemblance en présence de paramètres inconnus

Un certain nombre de paramètres des modèles de distribution sous chacune des hypothèses H_0 et H_1 sont cependant parfois inconnus. Pour appliquer la stratégie de Neyman Pearson décrite ci-dessus, différentes approches sont envisageables pour traiter ces paramètres inconnus. On note θ_0 et θ_1 les vecteurs de paramètres inconnus associés respectivement à l'hypothèse H_0 et H_1 .

1.2.4.1 Test du rapport de vraisemblance exact

Une première approche consiste à se placer dans un contexte bayésien, à considérer que les vecteurs de paramètres inconnus θ_0 et θ_1 sont des réalisations de vecteurs aléatoires de distribution de probabilité respectivement $p(\theta_0)$ et $p(\theta_1)$ et à appliquer l'approche de Neyman Pearson. Ce test est nommé rapport de vraisemblance exact (ELRT : *Exact Likelihood Ratio Test*) dans [Samson et al., 2004] et s'exprime de la façon suivante :

$$\Lambda_{ELRT}(\mathbf{u}) = \frac{p(\mathbf{u} | H_1)}{p(\mathbf{u} | H_0)} = \frac{\int_{\mathcal{D}_1} p(\mathbf{u} | H_1, \theta_1) p(\theta_1) d\theta_1}{\int_{\mathcal{D}_0} p(\mathbf{u} | H_0, \theta_0) p(\theta_0) d\theta_0} \underset{H_0}{\overset{H_1}{\gtrless}} \mu, \quad (1.8)$$

où \mathcal{D}_0 , respectivement \mathcal{D}_1 , est le domaine de variation de θ_0 , respectivement θ_1 . Cette approche nécessite de spécifier des probabilités *a priori* $p(\theta_0)$ et $p(\theta_1)$ et donc le plus souvent nécessite de faire des hypothèses concernant ces deux distributions.

1.2.4.2 Test du rapport de vraisemblance généralisé

Pour certains paramètres, imposer des probabilités *a priori* n'est pas toujours aisé et les *a priori* considérés peuvent s'avérer non pertinents. C'est pourquoi il est parfois intéressant de considérer que l'on ne dispose d'aucun *a priori* concernant les paramètres inconnus. Pour faire face à des paramètres totalement inconnus, l'approche sous-optimale la plus commune est celle qui consiste à estimer les paramètres inconnus par maximum de vraisemblance (ML : *Maximum Likelihood*). Dans ce cas le test à appliquer est le test du rapport de vraisemblance généralisé GLRT (*Generalized Likelihood Ratio Test*) [Van Trees, 1968] :

$$\Lambda_{GLRT}(\mathbf{u}) = \frac{p(\mathbf{u} | H_1, \widehat{\theta}_1)}{p(\mathbf{u} | H_0, \widehat{\theta}_0)} \underset{H_0}{\overset{H_1}{\gtrless}} \mu \quad (1.9)$$

avec

$$\begin{aligned}\widehat{\theta}_1 &= \arg \max_{\theta_1} p(\mathbf{u} | H_1, \theta_1) \\ \widehat{\theta}_0 &= \arg \max_{\theta_0} p(\mathbf{u} | H_0, \theta_0).\end{aligned}\tag{1.10}$$

Test à taux de fausses alarmes constant

Comme discuté dans la partie 1.2.2, la relation entre le seuil de détection et le taux de fausses alarmes est entièrement caractérisée par la fonction de répartition de la variable aléatoire $\Lambda(\mathbf{u})$ sachant H_0 . Les tests statistiques sont dits à taux de fausses alarmes constant (CFAR : *Constant False Alarm Rate*) quand cette fonction de répartition ϕ_{H_0} ne dépend d'aucun paramètre inconnu [Matteoli et al., 2010]. En effet dans ce cas il est possible de calculer le seuil qui va maintenir le taux de fausses alarmes constant et égal à une valeur choisie. Cette propriété est très intéressante pour un bon nombre d'applications. Le GLRT qui estime les paramètres inconnus par ML a la propriété intéressante de toujours conduire à des tests CFAR.

1.3 Différentes manières d'appliquer la théorie de la détection

La théorie de la détection fournit des outils méthodologiques généraux, dans un cadre théorique, pour la construction de tests permettant de décider si oui ou non des cibles sont présentes dans l'image ou la séquence d'images analysée. Cependant la construction de ces tests nécessite la connaissance ou l'estimation de la densité de probabilité des données observées sous l'hypothèse qu'aucune cible n'est présente ($p(\mathbf{u} | H_0)$) et sous l'hypothèse qu'au moins une cible est présente ($p(\mathbf{u} | H_1)$). Elle nécessite donc de faire des hypothèses *a priori* concernant ces densités de probabilité afin de les modéliser le plus précisément possible. Certains algorithmes font des *a priori* concernant la dynamique de la cible. D'autres font l'hypothèse par exemple que la variation d'intensité du fond peut être modélisée par des fonctions paramétriques [Vasquez et al., 2010]. La nécessité de faire des choix et des hypothèses pour appliquer la théorie de la détection a conduit au développement de très nombreuses méthodes de détection d'objets ponctuels.

Dans l'introduction, nous avons introduit une classification qui consiste à distinguer les approches TBD (*Track Before Detect*) et DBT (*Detect Before Track*). L'approche TBD consiste à prendre une décision basée sur l'observation \mathbf{u} rassemblant l'intensité de tous les pixels disponibles, et l'étape de pré-détection intervenant dans les approches DBT correspond à prendre des décisions sur des sous-ensembles de dimension beaucoup plus réduite (typiquement des fenêtres de dimension 7×7 pixels $\times 3$ images). Ceci nous amène à distinguer dans la suite les approches globales et les approches locales.

Notons par ailleurs qu'en surveillance passive comme pour un grand nombre d'autres applications, les algorithmes de détection ne doivent pas uniquement être capables de décider de la présence de cibles mais doivent aussi les localiser dans l'image ou la séquence d'images en cours d'analyse. Par nature les approches locales localisent les détections, concernant les approches globales ce n'est pas toujours le cas et l'application d'un posttraitement permettant de localiser les cibles sera parfois nécessaire.

1.3.1 Approches globales

Ces méthodes de détection cherchent à tirer parti de toute l'information disponible pour décider si oui ou non au moins une cible est présente dans les données. Elles nécessitent d'estimer la densité de probabilité associée aux données observées sous chacune des hypothèses H_0 et H_1 . La première hypothèse qui est très souvent faite dans la littérature est que les cibles sont additives par rapport au fond. Nous considérons ici que nous sommes bien dans ce cas et donc que les données observées suivent le modèle additif présenté en section 1.1.

Dans ce contexte le test d'hypothèse se formalise de la façon suivante :

$$\begin{cases} H_0 & : \mathbf{u} = \mathbf{b} & \text{(Fond uniquement)} \\ H_1 & : \mathbf{u} = \mathbf{b} + \mathbf{S} & \text{(Présence d'une cible)} \end{cases} \quad (1.11)$$

où $\mathbf{b} = [b(\mathbf{x}), \mathbf{x} \in \mathcal{X}]$ et $\mathbf{S} = [\sum_{q=1}^Q s_q(\mathbf{x}), \mathbf{x} \in \mathcal{X}]$. Le LRT défini en équation (1.7) est alors égal à :

$$\Lambda_{LRT}(\mathbf{u}) = \frac{p(\mathbf{u} | H_1)}{p(\mathbf{u} | H_0)} = \frac{p_{\mathbf{b}}(\mathbf{u} - \mathbf{S})}{p_{\mathbf{b}}(\mathbf{u})} \underset{H_0}{\overset{H_1}{\gtrless}} \mu \quad (1.12)$$

où $p_{\mathbf{b}}$ est la densité de probabilité associée à la composante de fond \mathbf{b} .

La méthode de détection par filtrage adapté tridimensionnel proposée par Reed *et al.* [Reed *et al.*, 1983, Reed *et al.*, 1988] est équivalente à appliquer le LRT sous l'hypothèse que la composante de fond \mathbf{b} peut être modélisée par une loi gaussienne multivariée de matrice de covariance connue et que la composante \mathbf{S} est parfaitement définie. L'application de cette méthode suppose que la forme, l'intensité, la position initiale dans la séquence et la trajectoire de la cible sont connues. Les performances de cette méthode sont très fortement dégradées quand la dynamique réelle des cibles diverge de celle considérée pour construire le filtrage adapté.

En pratique, dans les contextes considérés ici, les informations concernant les cibles sont très limitées : non seulement leur nombre et leur position, mais leur vitesse ou la direction de leur déplacement sont généralement inconnues. Le vecteur \mathbf{S} contenant la contribution de la cible ou des cibles présentes dans les données observées dépend donc de paramètres inconnus qu'il est nécessaire de prendre en compte. On a alors deux grandes familles de solutions : celles basées sur le GLRT et celles basées sur le ELRT (cf. section 1.2.4).

1.3.1.1 Méthodes basées sur le GLRT

Ce problème de paramètres inconnus peut être surmonté en partie en utilisant un banc de filtres construit pour décrire l'ensemble des vecteurs \mathbf{S} qu'il est possible d'observer. Pour réduire la dimension de cet « espace d'état », noté $\Omega_{\mathbf{S}}$, il est nécessaire d'introduire des *a priori* sur un certain nombre de paramètres inconnus. Concernant par exemple la dynamique de la cible, une hypothèse très souvent considérée est que la cible a une vitesse constante [Blostein et Huang, 1991, Tonissen et Evans, 1996, Tonissen et Bar-Shalom, 1998, Davey *et al.*, 2008]. En pratique, l'espace d'état doit être discrétisé. Le problème est donc de trouver un compromis entre la précision du résultat et le coût de calcul lié à la dimension de l'ensemble des vecteurs \mathbf{S} à tester ($\Omega_{\mathbf{S}}$). L'idée est ensuite de calculer pour chacun des vecteurs $\mathbf{S} \in \Omega_{\mathbf{S}}$, le rapport de vraisemblance défini en équation (1.12). L'algorithme considère ensuite qu'au moins une cible est présente, si

au moins un des rapports dépasse le seuil de détection. Cela revient à appliquer un **GLRT** (cf. section 1.2.4.2), pour décider si au moins une cible est présente dans les données :

$$\Lambda_{GLRT}(\mathbf{u}) = \max_{\mathbf{S} \in \Omega_S} \frac{p_{\mathbf{b}}(\mathbf{u} - \mathbf{S})}{p_{\mathbf{b}}(\mathbf{u})} \underset{H_0}{\overset{H_1}{\gtrsim}} \mu. \quad (1.13)$$

En considérant ensuite que chaque test de vraisemblance 1.12 dépassant le seuil fixé correspond à la présence d'une cible, ces algorithmes en plus de faire de la détection permettent de localiser les cibles et de déterminer leurs trajectoires.

Cette approche de détection est très coûteuse en temps de calcul quand l'ensemble Ω_S est grand. Elle nécessite donc dans ce cas de considérer que le fond est distribué selon une distribution de probabilité $p_{\mathbf{b}}$ simple. La majorité des algorithmes font l'hypothèse d'un fond blanc gaussien de moyenne nulle avec une matrice de covariance qui est égale à $\Sigma = \sigma^2 \mathbf{I}$ [Blostein et Huang, 1991, Tonissen et Evans, 1996, Tonissen et Bar-Shalom, 1998]. Sous cette hypothèse, le **GLRT** (éq. (1.13)) est alors équivalent au test suivant :

$$\log(\Lambda_{GLRT}(\mathbf{u})) = \max_{\mathbf{S} \in \Omega_S} \mathbf{u}^T \Sigma^{-1} \mathbf{S} - \frac{1}{2} \mathbf{S}^T \Sigma^{-1} \mathbf{S} = \max_{\mathbf{S} \in \Omega_S} \frac{1}{\sigma^2} \left[\mathbf{u}^T \mathbf{S} - \frac{1}{2} \mathbf{S}^T \mathbf{S} \right] \underset{H_0}{\overset{H_1}{\gtrsim}} \mu. \quad (1.14)$$

De plus, un grand nombre d'algorithmes font aussi l'hypothèse que l'intensité de la cible est constante dans la séquence et que la cible est purement ponctuelle. Autrement dit, ils font l'hypothèse que la cible contribue à augmenter l'intensité d'un seul pixel : celui dans lequel la cible se trouve présente. Le **GLRT** se réécrit alors :

$$\max_{\mathbf{S} \in \Omega_S} \frac{1}{\sigma^2} \sum_{l=1}^L \left(I_0 \mathbf{u}(\mathbf{x}_s^l) - \frac{1}{2} I_0^2 \right) \underset{H_0}{\overset{H_1}{\gtrsim}} \mu. \quad (1.15)$$

où \mathbf{x}_s^l correspond à la position de la cible dans la $l^{\text{ème}}$ image, et I_0 l'intensité de la cible. Ce qui au final revient à appliquer le test suivant :

$$\max_{\mathbf{S} \in \Omega_S} \sum_{l=1}^L \mathbf{u}(\mathbf{x}_s^l) \underset{H_0}{\overset{H_1}{\gtrsim}} \mu. \quad (1.16)$$

Toutes les méthodes de détection consistant à sommer au cours du temps l'intensité de pixels vis à vis de différentes hypothèses de dynamique de cibles, dans le but d'accumuler l'intensité des différents pixels correspondant à une même cible, peuvent donc s'interpréter comme étant une application du **GLRT** (éq. (1.13)). Toutes les trajectoires pour lesquelles le résultat de l'intégration temporelle est supérieur au seuil que l'on s'est fixé correspondent à la présence d'une cible. Pour éviter de faire une recherche exhaustive de ces trajectoires, des approches par programmation dynamique [Barniv, 1985, Barniv et Kella, 1987, Tonissen et Evans, 1996] ont été proposées. Une autre approche consiste à utiliser un test séquentiel du rapport des probabilités (SPRT : *sequential probability ratio test*) [Blostein et Huang, 1991]. D'autres méthodes de détection basées sur des intégrations temporelles d'intensité dérivant de méthodes de détection de segments existent aussi. Elles incluent notamment des approches faisant intervenir la transformée de Hough [Carlson et al., 1994], et des méthodes qui consistent à maximiser l'intégrale temporelle en ajoutant une contrainte favorisant la régularité des trajectoires basée sur le concept de courbure totale [Aubert et al., 2010].

1.3.1.2 Méthodes basées sur le ELRT

Plutôt que de chercher à estimer la meilleure trajectoire de cible comme le fait l'approche par GLRT, d'autres méthodes cherchent à accumuler les probabilités de présence de cibles associées aux différentes trajectoires possibles [Bruno et Moura, 2000]. Elles correspondent à une approche ELRT (cf. section 1.2.4.1) qui se reformule de la façon suivante dans notre contexte de cibles additives :

$$\Lambda_{ELRT}(\mathbf{u}) = \int_{\mathbf{S} \in \Omega_{\mathbf{S}}} \frac{p_{\mathbf{b}}(\mathbf{u} - \mathbf{S})}{p_{\mathbf{b}}(\mathbf{u})} p(\mathbf{S}) d\mathbf{S} \underset{H_0}{\overset{H_1}{\gtrless}} \mu \quad (1.17)$$

Pour résoudre le problème de décision de présence ou non de cible par ELRT, une première méthode consiste en utilisant la règle de Bayes à se ramener à un test sur le rapport des probabilités *a posteriori* [Bruno et Moura, 2000]. Elle se base sur une discrétisation de l'espace d'état et sur la prise en compte *a priori* concernant la dynamique de la cible via des lois de transitions entre les états de deux images consécutives. Une fois qu'il a été décidé qu'une cible était présente, la position de la cible est estimée par l'algorithme en maximisant la probabilité *a posteriori* qu'une cible se trouve à une position donnée. Cette méthode a le défaut de considérer qu'une seule cible est présente et a de plus le désavantage de fournir des positions de cibles qui ne sont pas forcément cohérentes avec la dynamique du modèle contrairement aux approches par GLRT. D'autres méthodes de détection par approches bayésiennes ont été proposées depuis. Elles mettent en oeuvre du filtrage particulaire pour estimer $\int_{\mathbf{S} \in \Omega_{\mathbf{S}}} p_{\mathbf{b}}(\mathbf{u} - \mathbf{S}) p(\mathbf{S}) d\mathbf{S}$ [Boers et Driessen, 2003, Rabaste et al., 2012]. L'utilisation du filtrage particulaire a l'avantage de permettre de s'affranchir d'une discrétisation de l'état et de plus de pouvoir intégrer des modèles non-linéaires d'évolution temporelle des cibles. De nombreuses recherches sont actuellement en cours afin d'améliorer ces méthodes, notamment en ce qui concerne l'amélioration de l'initialisation de l'algorithme, de l'étape de rééchantillonnage des particules et de la gestion du contexte multi-cibles.

Toutes ces méthodes qui travaillent directement sur l'ensemble du flot d'information, nécessitent la résolution de problèmes d'optimisation dans un domaine de recherche très grand et sont donc très gourmandes au niveau coût de calcul. Des méthodes cherchent à faire intervenir des modèles de signatures de cibles plus évolués qui tiennent compte par exemple de la réponse impulsionnelle (PSF : *Point Spread Function*) du système imageur et du fait que l'intensité de la cible peut varier au cours du temps [Tonissen et Bar-Shalom, 1998].

Cependant, la complexification du modèle de signature de la cible est quasiment toujours nécessairement associée à une modélisation simple du fond par bruit blanc. Pour le moment aucune méthode de détection ne permet de faire intervenir des modèles fins, à cette échelle globale, à la fois pour la distribution de probabilité du fond et pour la signature et la dynamique des cibles. Des méthodes permettent, par exemple de prendre en compte une corrélation spatiale des fonds, mais cela se fait sous une hypothèse d'homogénéité spatiale des fonds [Bruno et Moura, 2000].

Une modélisation des fonds non-stationnaires présentant des zones de fortes transitions n'est pas possible en pratique dans le cadre des approches globales. De nombreuses méthodes utilisent ce type de modélisation dans une approche locale, spatialement et temporellement, de la détection comme nous allons le voir dans la suite.

1.3.2 Approches locales

Les méthodes de détection locales travaillent avec des observations de dimensions restreintes. Plusieurs observations vectorielles sont donc définies à partir de l'ensemble des données disponibles : typiquement, une observation vectorielle, notée \mathbf{u}_x , est associée à chaque pixel x de l'image [Margalit *et al.*, 1985].

Considérant qu'une observation \mathbf{u}_x est la réalisation d'un vecteur aléatoire U_x , on applique un test de détection spécifique en chaque pixel afin de discriminer les deux hypothèses suivantes : (H_1) une cible est présente au pixel x et (H_0) aucune cible n'est présente au pixel x . Pour $\forall x \in \mathcal{X}$:

$$\Lambda_{LRT}(\mathbf{u}_x) = \frac{p(\mathbf{u}_x | H_1)}{p(\mathbf{u}_x | H_0)} \underset{H_0}{\overset{H_1}{\gtrless}} \mu \quad (1.18)$$

Une telle approche locale a ainsi l'avantage d'effectuer de façon conjointe la tâche de détection et celle de localisation.

Pour pouvoir calculer le test ci-dessus, de nombreux choix et hypothèses sont à faire. Il faut tout d'abord déterminer, pour chaque pixel x , quelles données on sélectionne pour construire l'observation \mathbf{u}_x , puis considérer des *a priori* concernant la cible et le fond pour pouvoir être capable d'estimer en chaque pixel les probabilités $p(\mathbf{u}_x | H_1)$ et $p(\mathbf{u}_x | H_0)$. Dans notre domaine d'application, l'observation \mathbf{u}_x est un vecteur qui est constitué de l'intensité des pixels du voisinage spatial ou spatio-temporel du pixel x (cf. figure 1.2). Cette observation \mathbf{u}_x peut parfois être restreinte à un scalaire et correspond alors tout

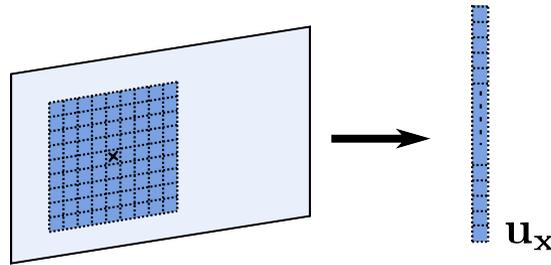


FIGURE 1.2 – Construction dans la cas spatial de l'observation vectorielle \mathbf{u}_x associée au pixel x marqué par une croix.

simplement à l'intensité du pixel x . L'avantage de travailler à une échelle locale et donc d'appliquer un test de détection en chaque pixel, est que l'on peut définir une densité de probabilité pour la composante de fond de \mathbf{u}_x qui est différente pour chacun des pixels x . Cette variation possible de la distribution du fond en fonction de sa position dans l'image, peut permettre d'obtenir des algorithmes de détection performants en présence de fonds non-stationnaires. C'est pourquoi, dans cette thèse nous avons choisi de nous orienter vers les méthodes de détection basées sur cette approche locale.

1.4 Méthodes de détection statistiques dans le cadre de la localisation de cibles

Dans cette section nous allons présenter les principaux filtres de détection issus de la théorie de la décision que l'on pourrait envisager d'utiliser pour répondre à notre problématique de détection dans le cadre d'une approche locale.

Sous hypothèse de cible additive (cf. section 1.1), les deux hypothèses locales s'écrivent :

$$\begin{cases} H_0 & : \mathbf{u}_x = \mathbf{b}_x & \text{(Fond uniquement)} \\ H_1 & : \mathbf{u}_x = \mathbf{b}_x + \alpha \mathbf{s} & \text{(Présence d'une cible)} \end{cases} \quad (1.19)$$

où \mathbf{b}_x correspond à l'intensité des pixels correspondant au fond, \mathbf{s} est la signature normalisée de la cible et α est l'intensité de la cible. On a $\alpha \geq 0$ pour les cibles dites positives qui ont une intensité supérieure à celle du fond (contraste positif) et $\alpha \leq 0$ pour les cibles dites négatives qui ont une intensité inférieure à celle du fond (contraste négatif). Dans les cas où l'on cherche à détecter uniquement des cibles positives ou alors uniquement des cibles négatives (cas nommé dans la suite : signe de α connu), on définit le paramètre scalaire sign_α suivant :

$$\text{sign}_\alpha = \begin{cases} 1 & \text{pour les cibles positives } (\alpha \geq 0) \\ -1 & \text{pour les cibles négatives } (\alpha \leq 0) \end{cases} \quad (1.20)$$

Dans les applications de détection infrarouges, il est très courant de considérer le contraste des cibles positif. Cela est pertinent car les objets recherchés sont le plus souvent motorisés et ont donc une température supérieure à celle du fond.

En section 1.4.1 les *a priori* que nous considérons concernant la modélisation statistique du fond et la signature de la cible sont présentés, puis en section 1.4.2 nous présentons les principaux filtres de détection de la littérature qui considèrent ces *a priori*.

1.4.1 Hypothèses de départ

1.4.1.1 Modèle de fond

Dans la littérature un grand nombre de méthodes de détection font l'hypothèse que l'ensemble du fond contenu dans une image peut être modélisé par un modèle gaussien stationnaire. Pour les fonds non-stationnaires ces méthodes nécessitent d'appliquer des prétraitements afin de se ramener à des fonds gaussiens ou autrement dit de blanchir les images. Une première approche consiste à soustraire la moyenne locale du fond [Margalit *et al.*, 1985, Chen et Reed, 1987], une autre approche consiste à réaliser une transformation point à point des données pour se rapprocher d'un champ gaussien [Chapple *et al.*, 2001].

Dans le cadre d'une approche locale, une manière naturelle de prendre en compte la non-stationnarité des fonds est d'utiliser une distribution dont les paramètres dépendent du pixel x pour modéliser le vecteur \mathbf{b}_x .

Pour gérer la non-stationnarité on peut aussi utiliser des modèles conditionnellement gaussiens à variance inhomogène comme ceux fournis par les vecteurs aléatoires sphériquement invariants (SIRV : *Spherically Invariant Random Vectors*) [Pascal *et al.*, 2005, Ovarlez *et al.*, 2011]. Un SIRV est le produit d'un vecteur aléatoire complexe gaussien (appelé speckle) et de la racine carré d'une variable aléatoire scalaire positive (appelée texture). Il est intéressant d'utiliser les SIRV dans le cadre où la loi de distribution de la texture est supposée inconnue afin de conserver une modélisation très générale des fonds [Pascal, 2006]. Les modèles SIRV ont été utilisés dans des applications de détection dans les domaines du radar [Pascal *et al.*, 2005] et de l'imagerie hyperspectrale [Ovarlez *et al.*, 2011], ils n'ont jamais été utilisés à notre connaissance pour faire de la détection en imagerie optique mono-bande.

Une modélisation du fond par SIRV est plus complexe à mettre en oeuvre qu'un simple modèle de fond gaussien. C'est pourquoi nous avons choisi dans cette thèse de nous restreindre à une modélisation gaussienne locale du fond qui permet de répondre à la problématique à moindre coût. Dans toute la suite de cette section, on considère donc que le fond se modélise localement par une gaussienne $\mathcal{N}(\mathbf{m}_x, \Sigma_x)$.

1.4.1.2 Signature de la cible

Concernant la signature de la cible, celle-ci est parfois considérée inconnue et dans ce cas le test qui est le plus couramment utilisé est le test de détection d'anomalie qui consiste simplement à seuiller la log-vraisemblance du fond sous l'hypothèse H_0 ($\log p(\mathbf{u}_x|H_0)$) [Matteoli *et al.*, 2010]. Un des filtres de détection le plus souvent considéré est le filtrage dit RX (Reed-Xiaoli) [Matteoli *et al.*, 2010, Cafer *et al.*, 2008], qui correspond au filtre proposé par Reed et Yu [Reed et Yu, 1990] quand la répartition spatiale de la cible est négligée. Ce filtre est très utilisé et donne de bonnes performances de détection dans le domaine de l'imagerie hyperspectrale [Matteoli *et al.*, 2010, Cafer *et al.*, 2008, Manolakis *et al.*, 2003] où le vecteur d'observation \mathbf{u}_x est constitué de l'ensemble des intensités dans les différentes bandes de fréquences mesurées en \mathbf{x} .

Dans notre contexte de détection en imagerie monobande, considérer que la signature de la cible est inconnue conduit à de mauvaises performances de détection. Il est souvent nettement préférable de considérer la signature de la cible connue *a priori*, même si le motif de cible considéré reste assez grossier. Dans le cadre de la détection mono-image le vecteur d'observation associé au pixel \mathbf{x} est un vecteur de dimension N^2 constitué de l'intensité des pixels contenus dans un voisinage de dimension $N \times N$ centré sur le pixel \mathbf{x} (cf. figure 1.2). Les cibles étant très éloignées du capteur leur signature est définie par la réponse impulsionnelle (PSF : *Point Spread Function*) du système imageur. Le motif de cible le plus simple, qui sera le plus souvent considéré dans la suite de ce travail, est une simple impulsion idéale. Sa version discrète est présentée en figure 1.3 pour $N = 5$. La signature \mathbf{s} de la cible est alors un vecteur de dimension N^2 dont toutes les composantes sont nulles sauf celle correspondant au pixel central. Les méthodes de détection proposées peuvent se généraliser sans difficulté à tout autre motif connu. Un autre motif sera par exemple considéré dans la section 4.7.1 du chapitre 4.

0	0	0	0	0
0	0	0	0	0
0	0	1	0	0
0	0	0	0	0
0	0	0	0	0

FIGURE 1.3 – Motif de cible considéré pour définir la signature \mathbf{s} de la cible.

En considérant une telle signature de cible contenant une unique composante non nulle, on fait l'hypothèse que les cibles ponctuelles recherchées sont centrées sur un pixel. Or en pratique on sait qu'une cible peut tout à fait se trouver à l'intersection de deux ou même quatre pixels et que leur apparence dépend fortement de leur position subpixelique. Ce phénomène de phasage des cibles n'est que très rarement pris en compte dans la littérature. Des méthodes ayant pour but de prendre en compte ce phénomène ont été

proposées par Samson et *et al.* [Samson *et al.*, 2004]. Dans le chapitre 4 en section 4.7.2 nous utiliserons ces méthodes pour évaluer le gain qu’elles peuvent apporter à la performance de détection dans notre contexte. Cependant, un inconvénient de ces modèles est qu’ils dépendent d’un paramètre inconnu en général : le décalage subpixellique de la cible. Dans les premiers chapitres, on se limite donc au cas où la signature \mathbf{s} associée aux cibles que l’on cherche à détecter est unique et connue.

1.4.2 Filtrage Adapté

Comme nous venons d’en discuter dans la section 1.4.1.1, on considère que le fond suit une loi gaussienne multivariée $\mathcal{N}(\mathbf{m}_x, \Sigma_x)$. Sous cette hypothèse de fond gaussien, le vecteur aléatoire U_x suit la même loi de probabilité que le fond $\mathcal{N}(\mathbf{m}_x, \Sigma_x)$ dans le cas où l’hypothèse H_1 est vérifiée et la loi de probabilité $\mathcal{N}(\mathbf{m}_x + \alpha\mathbf{s}, \Sigma_x)$ dans le cas où c’est l’hypothèse H_0 qui est vérifiée. Supposant que les images peuvent être pré-traitées afin de supprimer la moyenne du fond, les tests proposés dans la littérature sont parfois présentés dans le cas où le fond suit une loi normale centrée soit $\mathbf{m}_x = 0$. Cependant ces tests de détection sont généralisables au cas de fonds gaussiens non-centrés et dans toute la suite de ce chapitre nous choisissons de présenter les tests dans le cas général où $\mathbf{m}_x \neq 0$.

Dans le cas où tous les paramètres sont connus ($\mathbf{m}_x, \Sigma_x, \alpha$ et \mathbf{s}), la stratégie de Neyman Pearson présentée dans la partie 1.2.3 revient ici à seuiller le rapport de vraisemblance suivant :

$$\Lambda_{LRT}(\mathbf{u}_x) = \frac{p(\mathbf{u}_x | H_1)}{p(\mathbf{u}_x | H_0)} = e^{-\frac{1}{2}(\alpha^2 \mathbf{s}^T \Sigma_x^{-1} \mathbf{s} - 2\alpha \mathbf{s}^T \Sigma_x^{-1} (\mathbf{u}_x - \mathbf{m}_x))} \underset{H_0}{\overset{H_1}{\geq}} \mu \quad (1.21)$$

Pour les cibles d’intensité non nulle ($\alpha \neq 0$), ce test du rapport de vraisemblance peut se réécrire sous la forme suivante :

$$\text{sign}_\alpha \mathbf{s}^T \Sigma_x^{-1} (\mathbf{u}_x - \mathbf{m}_x) \underset{H_0}{\overset{H_1}{\geq}} \mu_{MF} \quad (1.22)$$

avec $\mu_{MF} = \frac{\log \mu}{|\alpha|} + \frac{|\alpha|}{2} \mathbf{s}^T \Sigma_x^{-1} \mathbf{s}$ et $\text{sign}_\alpha = 1$ si la cible est positive et $\text{sign}_\alpha = -1$ si elle est négative (cf. équation (1.20)). Le test du rapport de vraisemblance appliqué à la problématique de détection d’un signal dans du bruit est donc équivalent sous hypothèses gaussiennes à seuiller la sortie du filtre adapté (MF : *Matched Filter*) : $\text{sign}_\alpha \mathbf{s}^T \Sigma_x^{-1} (\mathbf{u}_x - \mathbf{m}_x)$. Le filtre adapté est par définition le filtre linéaire qui maximise le rapport signal à bruit [Reed *et al.*, 1983]. Ce filtre s’interprète aussi comme le filtre blanchisseur qui minimise l’erreur quadratique moyenne [Samson, 2002].

La loi conditionnelle du test sous l’hypothèse H_0 étant $\mathcal{N}(0, \mathbf{s}^T \Sigma_x^{-1} \mathbf{s})$, le seuil dépend non seulement de la Pfa désirée mais aussi des paramètres connus \mathbf{s} et Σ_x . Dès qu’un de ces deux derniers paramètres change, le seuil correspondant à une Pfa fixée doit donc être ré-estimé.

Test normalisé

Afin d’obtenir un test pour lequel le seuil correspondant à une Pfa fixée ne dépend pas des paramètres du test d’hypothèse, il suffit de normaliser le MF (éq. (1.22)) de la façon suivante :

$$\text{sign}_\alpha \frac{\mathbf{s}^T \Sigma_x^{-1} (\mathbf{u}_x - \mathbf{m}_x)}{[\mathbf{s}^T \Sigma_x^{-1} \mathbf{s}]^{1/2}} \underset{H_0}{\overset{H_1}{\geq}} \mu_{MF2}, \quad (1.23)$$

avec $\mu_{MF_2} = \frac{\log \mu}{|\alpha|[\Sigma_x^{-1} \mathbf{s}]^{1/2}} + \frac{|\alpha|}{2} \mathbf{s}^T [\Sigma_x^{-1} \mathbf{s}]^{1/2}$. On obtient ainsi un test dont la loi conditionnelle sous H_0 est $\mathcal{N}(0, 1)$.

1.4.3 Détection/Localisation en présence de paramètres inconnus

Dans la majorité des cas en théorie de la détection un certain nombre de paramètres intervenant dans le test d'hypothèse sont inconnus, c'est alors le plus souvent le test du **GLRT** donné par l'équation (1.29) qui est considéré (cf. section 1.2.4.2). Les principaux tests de détection issus du **GLRT** présents dans la littérature sont détaillés ci-dessous en considérant des hypothèses de plus en plus faibles.

1.4.3.1 Hypothèse : intensité de la cible α inconnue et statistiques du fond connues

Le paramètre qui est le plus souvent considéré inconnu est l'intensité de la cible. Les deux cas suivants peuvent se distinguer.

a/ Signe de α connu : Le premier cas est celui où l'on connaît le signe de α , c'est à dire que l'on sait si les cibles ont un contraste positif ($\alpha \geq 0$) ou négatif ($\alpha \leq 0$) par rapport au fond. Dans ce cas, l'estimée du paramètre α par **ML** est :

$$\hat{\alpha} = \max\left(0, \frac{\mathbf{s}^T \Sigma_x^{-1} (\mathbf{u}_x - \mathbf{m}_x)}{\mathbf{s}^T \Sigma_x^{-1} \mathbf{s}}\right) \text{ si } \alpha \geq 0 \text{ et } \hat{\alpha} = \min\left(0, \frac{\mathbf{s}^T \Sigma_x^{-1} (\mathbf{u}_x - \mathbf{m}_x)}{\mathbf{s}^T \Sigma_x^{-1} \mathbf{s}}\right) \text{ si } \alpha \leq 0. \quad (1.24)$$

L'approche **GLRT** conduit alors au test de détection suivant (démonstration en annexe A) :

$$\text{sign}_\alpha \frac{\mathbf{s}^T \Sigma_x^{-1} (\mathbf{u}_x - \mathbf{m}_x)}{[\mathbf{s}^T \Sigma_x^{-1} \mathbf{s}]^{1/2}} \underset{H_0}{\overset{H_1}{\geq}} \sqrt{2 \log(\mu)}. \quad (1.25)$$

On peut noter que ce test correspond exactement au filtrage adapté normalisé (éq. (1.23)) obtenu quand l'intensité de la cible, α , est connue. Connaître l'intensité de la cible dans le cadre du test d'hypothèse formalisé en équation (1.19), ne nous permet donc pas d'obtenir de meilleures performances de détection par rapport au cas où seul le signe de la signature de la cible est connu.

b/ Signe de α inconnu : Le deuxième cas est celui où l'on considère que le signe de α est inconnu, c'est à dire que l'on ne connaît pas le contraste des cibles, on cherche alors à détecter aussi bien les cibles de contraste positif que les cibles de contraste négatif. Une application basée sur une telle hypothèse, est par exemple le cas de la détection de voitures sur des images acquises dans le spectre visible : on cherche aussi bien à détecter les voitures de couleurs claires que celles de couleurs foncées. L'estimée du paramètre α par **ML** est alors :

$$\hat{\alpha} = \frac{\mathbf{s}^T \Sigma_x^{-1} (\mathbf{u}_x - \mathbf{m}_x)}{\mathbf{s}^T \Sigma_x^{-1} \mathbf{s}}, \quad (1.26)$$

et l'approche **GLRT** conduit alors au test de détection suivant :

$$\frac{[\mathbf{s}^T \Sigma_x^{-1} (\mathbf{u}_x - \mathbf{m}_x)]^2}{\mathbf{s}^T \Sigma_x^{-1} \mathbf{s}} \underset{H_0}{\overset{H_1}{\geq}} 2 \log(\mu). \quad (1.27)$$

Les performances théoriques de ce test de détection (P_d à P_{fa} fixée) qui se calculent à partir des équations (1.3) et (1.5) sont inférieures à celles du test de l'équation (1.25) correspondant au cas d'une cible de signature connue. Contrairement à la non-connaissance de l'intensité de la cible dans le cas d'une cible de signe connue, la non-connaissance du signe de la cible conduit à un test de détection moins performant. De façon générale les performances de détection ont tendance à baisser quand la connaissance *a priori* sur le signal diminue.

1.4.3.2 Hypothèse : intensité de la cible et statistiques du fond inconnues

a/ GLRT où les statistiques du fond sont considérées connues

En présence de fond de statistiques inconnues, un filtre classique de détection est le filtre adapté adaptatif (AMF : *Adaptive Matched Filter*) [Robey *et al.*, 1992]. Ce filtre estime dans un premier temps les statistiques du fond par maximum de vraisemblance en considérant que l'on peut accéder à un certain nombre d'observations du fond régies par les mêmes statistiques, puis considère ces statistiques comme connues et applique le test donné par l'équation (1.27). Le filtre AMF pour lequel les statistiques du fond sont à l'origine non-connues, est bien un test CFAR [Robey *et al.*, 1992].

Dans la suite nous nommerons, filtrage adapté adaptatif (AMF), les méthodes de filtrage qui, suite à une étape d'estimation des statistiques du fond, considèrent celles-ci connues et appliquent le test donné par l'équation (1.25) ou (1.27) selon si le signe de l'intensité de la cible est supposé connu ou non.

b/ GLRT de Kelly :

En présence de fond de statistiques inconnues, une autre approche de détection consiste à considérer les paramètres \mathbf{m}_x et Σ_x comme étant des paramètres inconnus au sein même du GLRT. Les tests d'hypothèses qui estiment ces paramètres par maximum de vraisemblance dans le cadre du GLRT, font intervenir deux hypothèses H_0 et H_1 légèrement différentes de celles définies à l'équation (1.19). Ces hypothèses sont :

$$\begin{cases} H_0 : \mathbf{u}_x = \mathbf{b}_x & \text{et } \mathbf{u}_{x_i} = \mathbf{b}_{x_i} \text{ pour } i = 1 \dots M \text{ (Fond uniquement)} \\ H_1 : \mathbf{u}_x = \mathbf{b}_x + \alpha \mathbf{s} & \text{et } \mathbf{u}_{x_i} = \mathbf{b}_{x_i} \text{ pour } i = 1 \dots M \text{ (Présence d'une cible)} \end{cases} \quad (1.28)$$

Les données $\mathbf{u}_{x_1}, \mathbf{u}_{x_2}, \dots, \mathbf{u}_{x_M}$ sont appelées données secondaires ou données d'entraînement, elles sont supposées ne contenir aucune composante de cibles, soit dit autrement ne contenir que des composantes de fond. Les données \mathbf{u}_x sont quant à elles appelées données primaires ou données de test, elles sont supposées contenir une cible sous l'hypothèse H_1 et uniquement du fond sous H_0 . L'ensemble des $M + 1$ vecteurs aléatoires, $\mathbf{b}_x, \mathbf{b}_{x_1}, \mathbf{b}_{x_2}, \dots, \mathbf{b}_{x_M}$ sont supposés être indépendants et identiquement distribués (i.i.d.). Le test du GLRT associé s'exprime alors de la façon suivante :

$$\Lambda_{GLRT}(\mathbf{u}_x, \mathbf{u}_{x_1}, \mathbf{u}_{x_2}, \dots, \mathbf{u}_{x_M}) = \frac{\max_{\theta_1} p(\mathbf{u}_x, \mathbf{u}_{x_1}, \mathbf{u}_{x_2}, \dots, \mathbf{u}_{x_M} | H_1, \theta_1)}{\max_{\theta_0} p(\mathbf{u}_x, \mathbf{u}_{x_1}, \mathbf{u}_{x_2}, \dots, \mathbf{u}_{x_M} | H_0, \theta_0)} \underset{H_0}{\overset{H_1}{\geq}} \mu, \quad (1.29)$$

où θ_0 et θ_1 sont les vecteurs contenant les paramètres qui sont inconnus associés respectivement à l'hypothèse H_0 et H_1 .

On peut noter que quand les statistiques du fond sont connues, considérer le test d'hypothèse qui fait intervenir des données secondaires (éq. (1.28)) ou celui qui ne considère

que les données primaires (éq. (1.19)) est totalement équivalent. Les mêmes tests de détection sont obtenus car les données secondaires n'interviennent pas quand les statistiques du fond sont connues.

Le test proposé par Kelly [Kelly, 1986] se base sur l'hypothèse que les vecteurs aléatoires secondaires et la composante de fond du vecteur aléatoire primaire sont régis par une même loi gaussienne $\mathcal{N}(\mathbf{m}_x, \Sigma_x)$. Les paramètres \mathbf{m}_x et Σ_x sont inconnus tout comme le signe et l'intensité de la cible. Kelly dans [Kelly, 1986] traite le cas de fonds à valeurs complexes mais le test obtenu se généralise très facilement au cas de fonds à valeurs réelles et on obtient que l'approche GLRT conduit au test de détection suivant [Kelly, 1986] :

$$\frac{[\mathbf{s}^T \widehat{\Sigma}_x^{-1} (\mathbf{u}_x - \widehat{\mathbf{m}}_x)]^2}{(\mathbf{s}^T \widehat{\Sigma}_x^{-1} \mathbf{s}) \left[1 + \frac{1}{M} (\mathbf{u}_x - \widehat{\mathbf{m}}_x)^T \widehat{\Sigma}_x^{-1} (\mathbf{u}_x - \widehat{\mathbf{m}}_x) \right]} \underset{H_0}{\overset{H_1}{\geq}} M \times \frac{\mu^{\frac{2}{M+1}} - 1}{\mu^{\frac{2}{M+1}}} \quad (1.30)$$

avec

$$\widehat{\mathbf{m}}_x = \frac{1}{M} \sum_{i=1}^M \mathbf{u}_{x_i} \quad \text{et} \quad \widehat{\Sigma}_x = \frac{1}{M} \sum_{i=1}^M (\mathbf{u}_{x_i} - \widehat{\mathbf{m}}_x)(\mathbf{u}_{x_i} - \widehat{\mathbf{m}}_x)^T. \quad (1.31)$$

On peut noter que quand M est très grand, le test est équivalent à seuiller le filtre AMF qui consiste simplement à appliquer le test donné par l'équation (1.27) en substituant aux statistiques de fond supposées connues celles estimées par ML.

1.4.4 Synthèse

Sous l'hypothèse que, en chaque pixel, la composante de fond du vecteur d'observation peut être assimilée à un vecteur aléatoire gaussien, les principaux tests issus de la théorie de la détection ont été présentés dans la partie 1.4.3. Le tableau 1.1 fait une synthèse de ces tests.

Le test du GLRT de Kelly se base sur l'hypothèse que l'on dispose de données d'entraînement qui sont des vecteurs aléatoires indépendants et identiquement distribués (i.i.d.). Cependant ces données secondaires sont souvent constituées des vecteurs d'observation associés aux pixels voisins du pixel \mathbf{x} en cours de test. Elles contiennent donc souvent des vecteurs qui ne sont pas indépendants. Notamment en présence de fonds fortement corrélés cette hypothèse que les vecteurs constituant les données secondaires sont i.i.d. n'est le plus souvent pas valide.

Comme il est difficile de disposer d'observations i.i.d. de la composante de fond, on se tourne dans la suite vers les approches du type « AMF », qui dans un premier temps estiment les statistiques du fond puis dans un second temps construisent un test à partir du LRT ou du GLRT en considérant les statistiques du fond connues. Robey *et al.* [Robey *et al.*, 1992] estiment les statistiques du fond par maximum de vraisemblance mais dans une telle approche on peut tout à fait envisager d'autres méthodes pour estimer les statistiques du fond. Les travaux de cette thèse ont consisté à proposer de nouvelles méthodes d'estimation des statistiques du fond efficaces en présence de fonds non-stationnaires.

Enfin, concernant le signe de l'intensité de la cible on peut distinguer deux cas : celui où il est supposé connu et celui où il est considéré inconnu. Comme on travaille en imagerie infrarouge, domaine dans lequel les menaces à identifier sont le plus souvent motorisées

Hypothèses à discriminer :				
$\begin{cases} H_0 & : \mathbf{u}_x = \mathbf{b}_x & \text{et } \mathbf{u}_{x_i} = \mathbf{b}_{x_i} \text{ pour } i = 1 \dots M & \text{(Fond uniquement)} \\ H_1 & : \mathbf{u}_x = \mathbf{b}_x + \alpha \mathbf{s} & \text{et } \mathbf{u}_{x_i} = \mathbf{b}_{x_i} \text{ pour } i = 1 \dots M & \text{(Présence d'une cible)} \end{cases}$				
Hypothèse de fond gaussien :				
$\begin{aligned} \text{sous } H_0 & : \mathbf{u}_x \sim \mathcal{N}(\mathbf{m}_x, \Sigma_x) & \text{et } \mathbf{u}_{x_i} \sim \mathcal{N}(\mathbf{m}_x, \Sigma_x) & \text{pour } i = 1 \dots M \\ \text{sous } H_1 & : \mathbf{u}_x \sim \mathcal{N}(\mathbf{m}_x + \alpha \mathbf{s}, \Sigma_x) & \text{et } \mathbf{u}_{x_i} \sim \mathcal{N}(\mathbf{m}_x, \Sigma_x) & \text{pour } i = 1 \dots M \end{aligned}$				
Nom du filtre	Expression analytique du filtre : $\Lambda(\mathbf{u}_x)$	Hypothèses considérées dans GLRT		Commentaires
		pour le fond	pour la cible	
MF (Matched Filter)	$\text{sign}_\alpha \mathbf{s}^T \Sigma_x^{-1} (\mathbf{u}_x - \mathbf{m}_x)$	Σ_x : connu \mathbf{m}_x : connu	\mathbf{s} : connu α : connu	Attention le seuil dépend des statistiques du fond
MF normalisé	$\frac{\text{sign}_\alpha \mathbf{s}^T \Sigma_x^{-1} (\mathbf{u}_x - \mathbf{m}_x)}{[\mathbf{s}^T \Sigma_x^{-1} \mathbf{s}]^{1/2}}$	Σ_x : connu \mathbf{m}_x : connu	\mathbf{s} : connu α : connu	
AMF (Adaptive Matched Filter)	$\frac{\text{sign}_\alpha \mathbf{s}^T \Sigma_x^{-1} (\mathbf{u}_x - \mathbf{m}_x)}{[\mathbf{s}^T \Sigma_x^{-1} \mathbf{s}]^{1/2}}$ $\frac{[\mathbf{s}^T \Sigma_x^{-1} (\mathbf{u}_x - \mathbf{m}_x)]^2}{\mathbf{s}^T \Sigma_x^{-1} \mathbf{s}}$	Σ_x : connu \mathbf{m}_x : connu	\mathbf{s} : connu α : inconnu sign_α : connu \mathbf{s} : connu α : inconnu sign_α : inconnu	Tests les plus couramment utilisés. Quand Σ_x et \mathbf{m}_x sont inconnus, ils sont estimés par ML en dehors du GLRT et sont ensuite considérés connus dans le GLRT
GLRT de Kelly	$\frac{[\mathbf{s}^T \widehat{\Sigma}_x^{-1} (\mathbf{u}_x - \widehat{\mathbf{m}}_x)]^2}{(\mathbf{s}^T \widehat{\Sigma}_x^{-1} \mathbf{s}) \left[1 + \frac{1}{M} (\mathbf{u}_x - \widehat{\mathbf{m}}_x)^T \widehat{\Sigma}_x^{-1} (\mathbf{u}_x - \widehat{\mathbf{m}}_x) \right]}$	Σ_x : inconnu \mathbf{m}_x : inconnu	\mathbf{s} : connu α : inconnu sign_α : inconnu	

TABLE 1.1 – Récapitulatif des tests statistiques couramment utilisés en théorie de la détection sous hypothèse de fonds gaussiens.

et donc plus chaudes que le fond, on suppose que le signe de α est connu et positif ($\alpha \geq 0$, soit $\text{sign}_\alpha = +1$). Dans la suite, on le précisera si d'autres hypothèses concernant le signe de α sont considérées. Supposer que la signature de la cible est connue et que son intensité est positive conduit à considérer le test de détection donné par l'équation (1.25).

1.5 Processus de détection en deux étapes basé sur la modélisation du fond au premier et second ordre

Comme nous venons de le voir l'approche par **GLRT** nous conduit sous différentes hypothèses à considérer le filtrage **AMF** défini par l'équation (1.25). Nous en déduisons une procédure de détection en deux étapes en dissociant l'estimation du modèle de fond au premier et second ordre [Genin et al., 2012a]. Le schéma de principe de cette procédure est présenté en figure 1.4. Cette procédure consiste :

1. à centrer les observations en appliquant une approche de suppression de fond, c.a.d. à estimer la moyenne du fond et à la retirer des données observées,
2. à blanchir le résidu sous l'hypothèse que celui-ci est bien centré en appliquant un filtrage adapté adaptatif (**AMF**) après avoir préalablement estimé les matrices de variance-covariance du fond,

et bien sûr à seuiller la sortie.

Cette procédure fait intervenir la moyenne et la matrice de covariance associées au fond, ou, autrement dit, les paramètres d'un modèle statistique du fond au premier et second ordre.

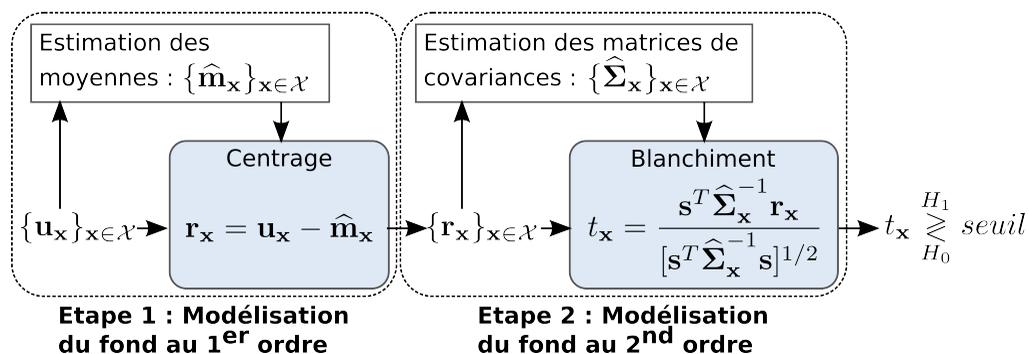


FIGURE 1.4 – Schéma général du principe de détection en deux étapes basé sur le principe du GLRT.

Si l'on reste sous une hypothèse de fond gaussien, les composantes du modèle peuvent par exemple être issues d'une simple méthode d'estimation linéaire des statistiques. La figure 1.5 illustre les résultats qui sont obtenus avec une simple méthode d'estimation linéaire des statistiques. Elle présente pour une image jouet ne contenant aucune cible, le résultat obtenu en sortie de chacune des deux étapes du schéma de détection proposé en figure 1.4. L'image résiduelle en sortie de l'étape de suppression de fond (cf. figure 1.5c) présente des valeurs importantes au niveau des zones de bords et des zones texturées

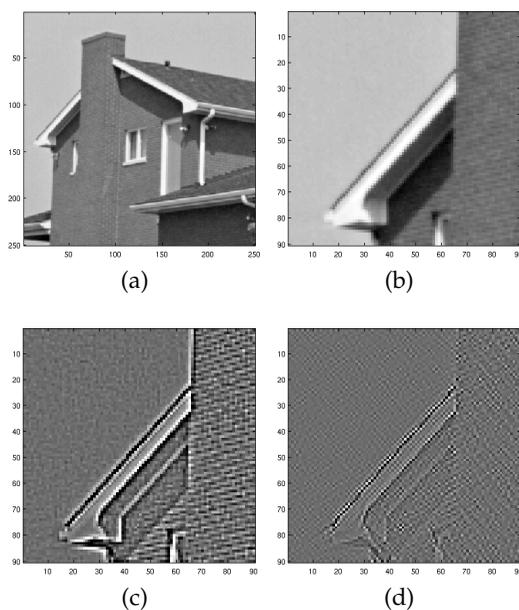


FIGURE 1.5 – Illustration de l'influence des deux étapes du principe de détection issu du GLRT (Fig. 1.4). (a) Image traitée présentant des zones de fortes transitions (bords de toit) et des zones de textures (briques), (b) zoom sur une partie de l'image (a). (c) zoom sur le résidu obtenu par soustraction en chaque pixel de la moyenne estimée par un filtre qui moyenne les pixels inclus dans une fenêtre de 3×3 pixels. (d) zoom sur le résidu obtenu suite à l'application du filtrage adapté sur le résidu obtenu en (c) avec une matrice de covariance estimée à partir de l'ensemble des pixels de l'image.

de l'image. L'application de la deuxième étape du schéma de détection considéré (Fig. 1.4) permet de réduire l'intensité du résidu au niveau des zones texturées, cependant l'intensité au niveau des bords apparaît toujours très élevée et une certaine texture reste toujours visible notamment dans la zone du mur en brique. Le seuillage d'un tel résidu conduit à un fort taux de fausses alarmes au niveau des zones de non-stationnarités de l'image. Il apparaît clairement qu'une simple estimation linéaire des paramètres du modèle de fond n'est pas suffisante face à des fonds non-stationnaires.

Notre intention est de montrer qu'en utilisant des modélisations du fond au premier et second ordre sophistiquées, ce schéma de détection en deux étapes conduit au développement d'algorithmes de détection très performants. Nous allons dans un premier temps voir dans le paragraphe suivant que ce schéma de détection fournit un cadre interprétatif à beaucoup de méthodes de détection existantes, puis dans la suite de ce manuscrit nous proposons de nouvelles approches pour modéliser le fond au premier et second ordre.

1.6 Présentation des principales méthodes de détection

Cette section fait une présentation non exhaustive de méthodes de détection, choisies parmi une littérature extrêmement volumineuse. La plupart des méthodes présentées, exceptées celle de la section 1.6.2, sont des techniques de suppression de fond que nous associons à une modélisation au premier ordre. Les méthodes de détection basées sur une modélisation du fond au second ordre seront présentées à la section 4.1.2 du chapitre 4.

1.6.1 Méthodes par suppression de fond : modélisation du fond au premier ordre

Parmi les méthodes de détection qui se basent sur une estimation du fond au premier ordre, une grande catégorie se détache qui est celle des méthodes par suppression de fond (**BS** : *Background Suppression*). Elles consistent à estimer le fond et puis à soustraire le fond estimé noté \widehat{b} aux données initiales u dans le but de faire ressortir les objets d'intérêt (cf. schéma de principe de la figure 1.6). En reprenant le modèle additif de formation des images donné en équation (1.1), la valeur du pixel en sortie du filtre de détection par **BS** peut s'écrire sous la forme suivante :

$$\widehat{S}(\mathbf{x}) = u(\mathbf{x}) - \widehat{b}(\mathbf{x}) = S(\mathbf{x}) + \epsilon(\mathbf{x}) \quad (1.32)$$

avec $\epsilon(\mathbf{x})$ l'erreur d'estimation du fond. La difficulté est que l'estimation du fond se fait à partir d'images contenant potentiellement des cibles. L'influence de ces dernières doit être diminuée autant que possible. Une illustration du principe de ces approches est présentée en figure 1.7. Idéalement le fond estimé \widehat{b} doit correspondre parfaitement à l'image initiale du fond ; c'est à dire que l'algorithme doit être capable de prédire à la position de la cible, la valeur observée correspondant à la structure du fond en l'absence de cibles, tout en préservant au mieux le reste de l'image. On peut noter que les méthodes de détection par **BS** rentrent parfaitement dans le schéma de détection en deux étapes de la figure 1.4, en considérant qu'une seule matrice de covariance est valable pour l'ensemble des pixels \mathbf{x} de l'image et qu'elle est proportionnelle à l'identité ($\Sigma_{\mathbf{x}} = \Sigma, \forall \mathbf{x} \in \mathcal{X}$ avec $\Sigma \propto I$).

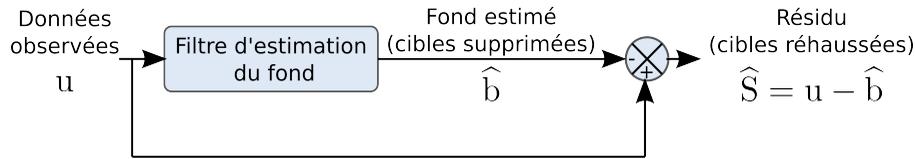


FIGURE 1.6 – Principe des approches par suppression de fond.

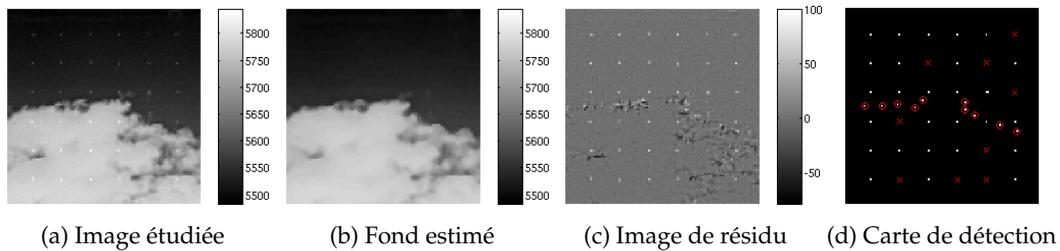


FIGURE 1.7 – Illustration de la suppression du fond sur une partie de l'image de Fig.2. Une matrice d'objets non résolus avec des positions subpixelliques aléatoires est incrustée pour constituer l'image étudiée. Le fond est estimé par un filtre médian [Barnett, 1989]. Le seuil utilisé pour obtenir la carte de détection correspond à une Pfa fixée à 10^{-3} sur toute l'image. Les cercles rouges correspondent aux fausses alarmes et les croix rouges aux objets non détectés.

Filtres d'estimation linéaire locale de la moyenne

Les méthodes BS les plus simples et les plus anciennes sont celles qui consistent simplement à soustraire la moyenne locale (filtrage que nous nommons BSMean dans la suite) [Barnett, 1989, Barnett et al., 1993, Acito et al., 2011]. Cette moyenne peut être estimée soit sur un voisinage spatial soit sur un voisinage spatio-temporel. Certaines méthodes proposent de faire intervenir deux fenêtres de tailles différentes et de soustraire à la moyenne des pixels calculée sur la petite fenêtre la moyenne obtenue sur la couronne définie par la plus grande fenêtre (filtre DWF : Double Window Filter) [Kim et al., 2009, Chen et al., 2007].

Filtres auto-régressifs

Des méthodes d'estimation du fond un peu plus complexe se basent sur une minimisation de l'erreur quadratique du fond (filtre LMS : *Least Mean Square*) [Takken et al., 1979, Soni et al., 1993]. Le principe de ces méthodes adaptatives est de mettre à jour les poids de pondération d'un filtre moyenneur en fonction de l'erreur de prédiction obtenue à l'itération précédente. Les pixels étant estimés en fonction de leurs prédécesseurs et le filtre parcourant l'image ligne par ligne, les derniers pixels de l'image sont donc mieux prédits que les premiers. Un tel filtre est capable de s'adapter à des variations lentes du fond, mais sa vitesse d'adaptation est faible et il n'est pas capable de faire face à la présence dans le fond de fortes variations telles que les bords de nuages.

Filtres d'ordre

Les méthodes linéaires ne permettant pas d'obtenir de bonnes performances de détection dans des fonds non-stationnaires, c'est pourquoi en parallèle de celles-ci de nom-

breuses méthodes non-linéaires ont été développées. Un des filtres non-linéaires de détection d'objets ponctuels les plus classiques consiste à soustraire à l'image étudiée une estimée du fond obtenue par filtrage médian (filtrage que nous nommons BSMedian dans la suite) [Robinson, 1959, Barnett, 1989].

Dans la famille des filtres d'ordre, d'autres méthodes de détection très répandues sont celles basées sur des opérateurs morphologiques : l'opérateur d'érosion (respectivement de dilatation) associe à chaque pixel la valeur minimale (respectivement maximale) contenue dans un voisinage défini par l'élément structurant du filtre [Tom *et al.*, 1993, Barnett *et al.*, 1993]. Le filtre morphologique de détection le plus connu est le Top-Hat aussi appelé WTH (*White Top-Hat*), il soustrait à l'image observée un fond estimé en appliquant un opérateur morphologique d'ouverture. Cet opérateur correspond à appliquer un opérateur d'érosion suivi d'un opérateur de dilatation. Le filtre Top-Hat permet de détecter des cibles de contraste positif. De nombreux autres filtres de détection basés sur les opérateurs morphologiques ont été proposés [Zhang *et al.*, 2005, Bai *et al.*, 2009]. On peut par exemple citer : le filtre Bottom-Hat aussi appelé BTH (*Black Top Hat*) qui permet de détecter des cibles de contraste négatif, le filtre CMO (*Close-Minus-Open*) qui fait ressortir les cibles positives et négatives sans conserver l'information concernant leur signe et le filtre TopBottomHat qui permet de réhausser le niveau des cibles par rapport au fond tout en préservant leur signe [Carnie *et al.*, 2006].

Pour mieux gérer les zones de fortes transitions, des filtres non-linéaires directionnels ont été proposés. Le filtre directionnel le plus connu et le plus utilisé pour estimer le fond est le filtre Max-Median [Deshpande *et al.*, 1999]. Son principe est le suivant : des filtrages médians monodimensionnels sont appliqués sur des directions séparées de 45° dans l'image et, parmi les valeurs ainsi obtenues, le filtre retient la valeur maximale. Pour une fenêtre de dimension $(2F + 1) \times (2F + 1)$ la sortie du filtrage Max-Median peut s'écrire de la façon suivante :

$$\begin{aligned} \widehat{b}_{\text{MaxMedian}}(i, j) &= \max\{z_1, z_2, z_3, z_4\} \\ \text{avec } z_1 &= \text{median}\{u(i, j - F), \dots, u(i, j), \dots, u(i, j + F)\} \\ z_2 &= \text{median}\{u(i - F, j), \dots, u(i, j), \dots, u(i + F, j)\} \\ z_3 &= \text{median}\{u(i - F, j - F), \dots, u(i, j), \dots, u(i + F, j + F)\} \\ z_4 &= \text{median}\{u(i + F, j - F), \dots, u(i, j), \dots, u(i - F, j + F)\} \end{aligned} \quad (1.33)$$

Le succès de cette méthode d'estimation du fond vient du fait qu'elle est simple à mettre en oeuvre et qu'elle permet, tout en éliminant efficacement le bruit impulsionnel, de bien préserver les zones de bords et les discontinuités présentes dans l'image. D'autres méthodes de combinaison de filtrages directionnels ont été proposées. Certaines font intervenir des filtrages morphologiques [Rivest et Fortin, 1996, Carnie *et al.*, 2006, Zhu *et al.*, 2000], une autre fait par exemple intervenir un filtrage LoG (*Laplacian-of-Gaussian*) [Kim, 2011].

Filtres par régression

Des méthodes basées sur des approches de régression non-linéaire ont aussi été proposées dans la littérature. On peut citer l'approche proposée par Samson [Samson, 2002] qui s'inspire de la régularisation bayésienne des problèmes inverses en restauration d'images. Elle a recours à un modèle markovien pour caractériser le fond corrélé, et minimise un critère pénalisé non quadratique pour obtenir une estimée de l'image des cibles au sens du maximum a posteriori. Vasquez *et al.* [Vasquez *et al.*, 2010] proposent une méthode

de détection basée sur une modélisation linéaire ou quadratique de la variation spatiale de la moyenne du fond. Une segmentation de l'image en zones homogènes est proposée et une estimation des paramètres du modèle de variation d'intensité considéré est réalisée pour chacune des zones [Vasquez, 2011]. Cette méthode donne de bons résultats pour des images constituées de régions où les variations d'intensité sont effectivement linéaires ou quadratiques. Cependant elle n'est pas adaptée à des images présentant des zones texturées contenant des hautes fréquences. Boulanger *et al.* proposent une modélisation linéaire de la variation temporelle de la moyenne du fond [Boulanger *et al.*, 2006]. Les paramètres du modèle sont estimés en chaque position de l'image au moyen d'un M-estimateur robuste asymétrique en faisant intervenir une régularisation spatiale des paramètres du modèle [Boulanger *et al.*, 2006, Boulanger, 2007]. On peut noter que suite à cette suppression de fond, deux méthodes de détection sont proposées par Boulanger : un seuillage du résidu faisant intervenir un seuil déterminé automatiquement qui est adapté à chaque signal temporel ou une approche faisant appel à une modélisation *a contrario* [Boulanger, 2007].

Filtres basés sur des approches de débruitage

D'autres approches de détection par BS font intervenir des méthodes de débruitage pour estimer le fond. Certaines [Acito *et al.*, 2011, Bae et Sohng, 2010, Bae, 2011] font intervenir le filtre bilatéral qui réalise une moyenne pondérée des pixels voisins où la pondération correspond au produit de deux noyaux : le premier fait intervenir une mesure de distance spatiale entre les pixels et le second une mesure de similarité de niveau de gris. Bae propose un algorithme adaptatif où les paramètres de ces noyaux vont dépendre de valeurs obtenues par des préfiltrages spatiaux [Bae et Sohng, 2010] ou spatio-temporels [Bae, 2011] définissant une probabilité de présence de cible. Dans les zones où cette probabilité est élevée, des paramètres élevés sont choisis pour bien supprimer la cible et dans les autres zones, des paramètres plus faibles sont sélectionnés afin d'y obtenir une estimation précise du fond. Dans le chapitre 3 nous montrerons que les algorithmes de débruitage récents tels que le NL-means [Buades *et al.*, 2005b] et le BM3D [Dabov *et al.*, 2007b] permettent d'obtenir de très bons filtres BS, moyennant certaines adaptations pour tenir compte de la présence des cibles. Nous avons publié le filtre de détection DNLM fondé sur NL-means en 2011 [Genin *et al.*, 2011]. Plus récemment, nous avons trouvé une méthode de détection elle aussi fondée sur NL-means, proposée par Goudou et publiée, à notre connaissance, uniquement dans son manuscrit de thèse [Goudou, 2007]. La comparaison de ces deux techniques développées indépendamment sera faite au chapitre 3.

1.6.2 Autres méthodes de détection

Parmi la quantité impressionnante de méthodes publiées qui ne font pas partie des catégories présentées ci-dessus, deux familles sont citées dans cette section.

Les méthodes qui détectent les cibles en travaillant dans les domaines des transformées en ondelettes. Ces approches combinent l'information provenant des différents niveaux de décomposition et tirent parti du fait qu'une cible est représentée par un nombre faible de coefficients élevés qui sont corrélés au travers des niveaux de décomposition [Gregoris *et al.*, 1994, Olivo-Marin, 2002, Boccignone *et al.*, 2002].

Les méthodes de détection qui caractérisent le fond par des fractales. Elles se basent sur le fait que les cibles étant le plus souvent d'origine humaine, leur dimension fractale

est différente de celle de milieu naturel (fond nuageux) dans lequel elles se trouvent [Tang et Wang, 2009, Sun *et al.*, 2009].

1.7 Approche considérée

Dans cette thèse nous avons choisi de suivre le schéma de détection en deux étapes proposé en section 1.5 et d’obtenir de nouveaux filtres de détection performants en présence de fonds non-stationnaires en proposant de nouvelles techniques de modélisation au premier et au second ordre. Concernant la première étape, nous nous sommes orientés vers l’utilisation de méthodes de débruitage à patches qui ont l’avantage de bien préserver les zones de transitions et de modéliser les fonds inhomogènes de façon efficace. Concernant la seconde étape, nous nous sommes orientés vers une approche consistant à estimer les matrices de covariance par classes basées sur une classification non supervisée des pixels par mélange de gaussiennes.

Conclusion

Ce chapitre a permis dans un premier temps de formaliser la problématique de la thèse. Les principales méthodes de détection d’objets ponctuels y sont présentées. En nous appuyant sur la théorie statistique, nous avons introduit un schéma de détection en deux étapes. Il est le fil conducteur de notre démarche dans toute la suite de ce manuscrit. Les méthodes basées sur des approches à patches que nous proposons pour la première étape sont présentées dans le chapitre 3. La méthode basée sur une modélisation du fond au second ordre considérée pour la deuxième étape du schéma de détection est présentée au chapitre 4. Avant cela, le chapitre 2 présente la méthode d’évaluation de performances utilisées dans la suite pour comparer les algorithmes ainsi que les données sur lesquelles ces évaluations de performances seront réalisées. Les performances des algorithmes proposés sont présentées dans le chapitre 5.

Chapitre 2

Méthode d'évaluation de performances des algorithmes de détection d'objets ponctuels

Introduction

L'objectif de ce chapitre est de mettre en place une méthode d'évaluation des performances d'algorithmes de détection d'objets ponctuels. Il s'agit à la fois de comparer diverses solutions algorithmiques entre elles et de positionner les algorithmes que nous avons développés par rapport à l'existant.

Organisation du chapitre — La section 2.1 présente tout d'abord les critères de performances utilisés pour comparer les différents algorithmes de détection. En section 2.2, les différentes méthodes envisageables pour évaluer ces critères et la méthode retenue sont ensuite présentées. Une discussion concernant la pertinence et les dangers d'une évaluation de performances par zones est faite en section 2.3. Les deux bases de fonds réels sur lesquels les performances des algorithmes sont évaluées sont présentées en section 2.4. Enfin, la section 2.5 décrit le modèle de cible utilisé et certains détails concernant le mode d'incrustation de la cible dans les images de fond.

2.1 Choix des critères de performances étudiés

Dans cette section nous présentons les principaux critères utilisés dans la littérature pour comparer des algorithmes de détection entre eux, afin de justifier le choix qui a été fait de ne retenir que deux de ces critères. Dans la littérature, on peut distinguer deux grands types de critères : ceux qui se calculent sur le résidu obtenu en sortie de filtrage avant l'étape de seuillage et ceux qui se calculent à partir des décisions binaires obtenues après seuillage de l'image filtrée.

2.1.1 Critères basés sur des mesures de statistiques sur le résidu

Le critère le plus couramment utilisé pour mesurer le contraste d'une cible et qui se trouve à la base de la définition de nombreux autres critères est le Rapport Signal à Bruit

(RSB). Il se définit de la façon suivante pour une cible se trouvant au pixel x :

$$RSB_x = \frac{|t_x - m_x|}{\sigma_x} \quad (2.1)$$

où t_x est l'intensité du résidu au pixel x contenant une cible, m_x et σ_x correspondent respectivement à la moyenne et l'écart-type des résidus au voisinage du pixel x . Notons que certains auteurs choisissent d'estimer la moyenne et l'écart type globalement sur toute l'image [Hilliard, 2000]. Plus le RSB est élevé, plus la cible que l'on cherche à détecter ressort du fond dans l'image des résidus et plus il est facile de détecter l'objet présent. De nombreuses variantes de ce critère existent, ainsi le RSB est parfois défini comme le carré de l'expression (2.1) [Margalit *et al.*, 1985, Samson, 2002].

Le RSB est souvent utilisé pour évaluer l'apport d'un filtre sous la forme d'un gain tel que :

$$RSB_{gain} = \frac{RSB_{out}}{RSB_{in}} \quad (2.2)$$

où RSB_{out} et RSB_{in} correspondent respectivement au RSB de la cible calculée sur l'image originale et sur l'image filtrée [Hilliard, 2000, Yang *et al.*, 2004, Gu *et al.*, 2010]. Le gain de RSB est aussi parfois évalué par rapport au RSB calculé sur l'image en sortie d'un filtre de référence [Margalit *et al.*, 1985].

Avec les critères RSB et RSB_{gain} , le problème est que l'on dispose d'autant de mesures de performances qu'il y a de cibles dans l'image. Pour comparer des algorithmes entre eux, il est préférable d'avoir un indicateur global qui tient compte du contraste de l'ensemble des cibles présentes dans la base de données. Soit \mathcal{X}_1 (respectivement \mathcal{X}_0) l'ensemble des pixels pour lesquels on sait que c'est l'hypothèse H_1 (respectivement H_0) qui est vérifiée, une solution consiste à faire la moyenne des RSB sur toutes les cibles connues :

$$RSB_{moy} = \frac{1}{\text{Card}(\mathcal{X}_1)} \sum_{x \in \mathcal{X}_1} RSB_x. \quad (2.3)$$

Une autre possibilité consiste à évaluer un critère global tel que [Samson, 2002] :

$$RSB_{global} = \frac{(m_1 - m_0)^2}{\sigma_0^2} \quad (2.4)$$

où m_1 , est la moyenne de l'intensité des pixels cible et m_0, σ_0 sont la moyenne et la variance des pixels de fond. On peut aussi utiliser une mesure globale de contraste cible/fond de la forme suivante [Samson, 2002] :

$$C = \frac{\text{Card}(\mathcal{X}_0) \sum_{x \in \mathcal{X}_1} |t_x|^p}{\text{Card}(\mathcal{X}_1) \sum_{x \in \mathcal{X}_0} |t_x|^p}. \quad (2.5)$$

avec $p \in \{1, 2\}$.

Un critère complémentaire est le facteur de suppression de fond (BSF : *Background Suppression Factor*) [Hilliard, 2000, Yang *et al.*, 2004, Gu *et al.*, 2010] :

$$BSF = \frac{\sigma_{0in}}{\sigma_{0out}} \quad (2.6)$$

où σ_{0in} (respectivement σ_{0out}) est l'écart-type des pixels de fond sur l'image originale (respectivement sur l'image filtrée). Remarquons que ce critère ne peut en aucun cas être utilisé seul car il fait intervenir uniquement des caractéristiques de fond et n'apporte donc aucune information concernant la préservation du niveau de la cible par le filtre. Mais, associé à une mesure de gain de RSB, comme c'est le plus souvent le cas dans la littérature, le critère BSF permet de mieux comprendre le comportement des filtres. Par exemple, si un filtre a un fort BSF et un faible gain de RSB, cela indique que la faible performance est liée au fait que le filtre ne préserve pas correctement la cible et non au fait qu'il ne supprime pas correctement le fond.

Pour conclure, il convient à notre avis d'être prudent quant à l'utilisation de tous ces critères basés sur des mesures de statistiques sur le résidu. En effet, mis à part le critère C de l'équation (2.5) avec $p = 1$, tous ces critères ne font intervenir que des moyennes et des variances. Cela implique qu'ils reposent sur un modèle gaussien stationnaire de l'intensité des pixels (fond ou cible). Ces critères sont justifiés uniquement dans le cas où la distribution des pixels correspondant à des cibles et celle des pixels correspondant à du fond sont gaussiennes. Ils ne permettent pas de caractériser précisément des fonds non gaussiens ou non stationnaires tels que les fonds inhomogènes présentant des zones de forts gradients. C'est pourquoi nous n'avons pas retenu ce type de critère dans nos évaluations et nous avons choisi ici de nous restreindre à des critères de performances qui s'évaluent uniquement à partir de la décision binaire de présence ou non d'une cible obtenue en chacun des pixels après seuillage de l'image filtrée.

2.1.2 Critères évalués à partir des décisions binaires obtenues après seuillage

2.1.2.1 Courbes ROC

Parmi ces critères, le plus couramment étudié dans la littérature est la courbe ROC déjà évoquée en section 1.2.2 qui représente la probabilité de détection (Pd) en fonction de la probabilité de fausses alarmes (Pfa).

Sous l'hypothèse que le fond suit une loi de distribution connue, il est souvent possible pour les tests issus de la théorie de la détection (cf. section 1.1) de disposer pour une cible donnée (intensité α_c et signature \mathbf{s}_c) d'une expression théorique de la Pd et la Pfa en fonction du seuil de détection μ et ainsi d'obtenir une expression théorique de la courbe Pd(Pfa). Par exemple, sous l'hypothèse que les composantes de fond des observations associées à chacun des pixels de l'image suivent toute la loi gaussienne $\mathcal{N}(\mathbf{m}, \mathbf{\Sigma})$ considérée dans la construction du test, la Pd et la Pfa pour le filtrage AMF donné par l'équation (1.27) (correspondant au cas où le signe du contraste de la cible est inconnu) s'expriment de la façon suivante :

$$P_{FA}(\mu) = 1 - \phi_{\chi^2}(\mu, 1, 0) \quad \text{et} \quad P_D(\mu) = 1 - \phi_{\chi^2} \left(\mu, 1, \alpha_c^2 \frac{[\mathbf{s}^T \mathbf{\Sigma}^{-1} \mathbf{s}_c]^2}{\mathbf{s}^T \mathbf{\Sigma}^{-1} \mathbf{s}} \right), \quad (2.7)$$

où $\phi_{\chi^2}(x, k, \lambda)$ est la fonction de répartition d'une loi du χ^2 non-centrée à k degrés de liberté et de paramètre de décentralisation λ , \mathbf{s} est la signature *a priori* de la cible dans le test, α_c et \mathbf{s}_c sont les paramètres de la vraie cible. L'expression de la courbe Pd(Pfa)

théorique est alors la suivante :

$$P_D(P_{FA}) = 1 - \phi_{\chi^2} \left(\phi_{\chi^2}^{-1}(1 - P_{FA}, 1, 0), 1, \alpha_c^2 \frac{[\mathbf{s}^T \boldsymbol{\Sigma}^{-1} \mathbf{s}_c]^2}{\mathbf{s}^T \boldsymbol{\Sigma}^{-1} \mathbf{s}} \right). \quad (2.8)$$

On peut noter que pour une **Pfa** fixée, la **Pd** est maximale quand la signature de la cible intervenant dans la définition du test \mathbf{s} correspond à la signature des cibles à détecter \mathbf{s}_c . L'expression de l'équation (2.8) permet ainsi d'analyser la perte de performance liée à la prise en compte d'un mauvais *a priori* sur le modèle de cible dans la construction du test.

Ces évaluations de performances théoriques ont le gros défaut de ne pas tenir compte du fait que le fond ne suit pas forcément la loi de distribution considérée *a priori* pour la construction du test. De plus, pour la majorité des filtres de détection notamment tous les filtres non-linéaires, il est extrêmement difficile d'obtenir des expressions théoriques de la courbe **Pd(Pfa)** sous l'hypothèse que la distribution du fond est connue.

C'est pourquoi la **Pd** et la **Pfa** sont le plus souvent redéfinis de cette façon :

$$P_{FA}(\mu) = \frac{\text{Card}(\{\mathcal{X}_0 | t_x \geq \mu\})}{\text{Card}(\mathcal{X}_0)} \quad \text{et} \quad P_D(\mu) = \frac{\text{Card}(\{\mathcal{X}_1 | t_x \geq \mu\})}{\text{Card}(\mathcal{X}_1)} \quad (2.9)$$

où t_x est la valeur du résidu en sortie de filtrage au pixel de position \mathbf{x} . Ces deux quantités sont calculées expérimentalement à partir de bases de données contenant des pixels pour lesquels on sait qu'une cible est présente (hypothèse H_1 vérifiée) et des pixels pour lesquels on sait qu'aucune cible n'est présente (hypothèse H_0 vérifiée). Dans cette approche, le problème est la représentativité des bases de données utilisées, nous y reviendrons en section 2.2.

A partir de ces courbes de nombreux indicateurs scalaires ont été proposés pour quantifier plus simplement les performances de détection. On peut citer le critère AUC (*Air Under Curve*) qui correspond à l'aire sous la courbe **ROC** [Samson, 2002, Goudou, 2007] :

$$AUC = \int_0^1 P_D(P_{FA}) dP_{FA}. \quad (2.10)$$

Ce critère a le défaut de pondérer de la même façon les **Pd** associées à toutes les **Pfa** comprises entre 0 et 1. Comme dans de nombreuses applications il est spécifié que la **Pfa** doit être beaucoup plus faible que 1, Cafer *et al.* [Cafer *et al.*, 2007] ont proposé, un critère ne prenant en compte que les **Pd** associées aux **Pfa** comprises entre 0 et un seuil à fixer (th) :

$$R_{A_{th}} = \frac{A_{th} - 0.5(th)^2}{1 \times th - 0.5(th)^2} \quad (2.11)$$

où $A_{th} = \int_0^{th} P_D(P_{FA}) dP_{FA}$. Comme illustré en figure 2.1, le critère $R_{A_{th}}$ s'obtient en calculant pour une **Pfa** comprise entre 0 et le seuil th, l'aire de la surface sous la courbe **ROC** moins l'aire de la surface sous la ligne **Pd=Pfa** qui correspond au cas limite où les cibles sont indiscernables du fond. Le critère étant ensuite normalisé par l'aire que l'on obtiendrait dans le cas d'une détection parfaite (zone hachurée).

Cependant en pratique, les algorithmes de détection sont souvent paramétrés pour fonctionner à un niveau précis de **Pfa**. En cherchant à quantifier les performances de détection à l'aide d'un simple scalaire, les critères basés sur des calculs d'aire sous la courbe prennent en compte les performances à des niveaux de **Pfa** qui ne nous intéressent pas.

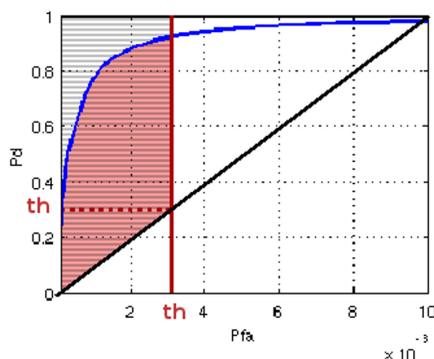


FIGURE 2.1 – Illustration du critère $R_{A_{th}}$ de l'équation (2.11) : $R_{A_{th}}$ correspond au rapport entre l'aire de la surface rouge et l'aire de la partie hachurée.

Pour comparer les algorithmes de détection entre eux, ne disposant pas d'un indicateur scalaire satisfaisant, nous avons choisi d'analyser l'ensemble des courbes ROC pour des valeurs de Pfa comprises entre 10^{-4} et 10^{-2} . La borne supérieure de cette plage de Pfa est nettement supérieure au taux autorisé dans les applications de veille passive. Concernant la borne inférieure il serait intéressant d'étudier des Pfa nettement inférieures à 10^{-4} mais nous sommes ici limités par la quantité de données sur lesquelles l'évaluation des courbes ROC est réalisée.

2.1.2.2 Uniformité de la répartition des Fausses Alarmes : critère DNFA

Notre étude est ici focalisée sur les algorithmes de détection qui ont pour but de tirer parti au mieux de la corrélation spatio-temporelle du fond. Cependant il ne faut pas oublier que ce type d'algorithme intervient dans la pratique dans une chaîne complète de traitement qui fait souvent intervenir une étape de pistage.

Pour qu'un algorithme de pistage soit performant c.a.d. élimine le plus possible de FA il faut idéalement que celles-ci ne soient pas corrélées temporellement. Lorsqu'on utilise une méthode de détection spatiale sur une séquence d'images, la source la plus commune de corrélation temporelle des fausses alarmes est la corrélation temporelle des fonds. En effet, lorsque les FA apparaissent préférentiellement dans certaines zones d'une image, comme les zones de forts gradients, la corrélation temporelle de ces zones dans les images de fonds suivantes de la séquence se transmet aux FA. C'est pourquoi, pour une Pfa fixée, on considère qu'une répartition homogène des FA dans l'ensemble de l'image est plus favorable à l'étape de pistage, que la présence d'amas de FA dans certaines zones.

Afin de quantifier la distribution spatiale des FA, nous proposons d'évaluer la distance moyenne entre une FA et la FA la plus proche. Ce critère est noté DNFA (*Distance to the Nearest False Alarm*). Soit $\mathcal{X}_{FA} = \{\mathbf{x}_1, \dots, \mathbf{x}_n\}$, l'ensemble des pixels de l'image qui correspondent à une FA, on écrit :

$$DNFA = \frac{1}{\text{Card}(\mathcal{X}_{FA})} \sum_{\mathbf{x} \in \mathcal{X}_{FA}} \min(\|\mathbf{x} - \mathbf{x}_i\|, \mathbf{x}_i \in \mathcal{X}_{FA}, \mathbf{x}_i \neq \mathbf{x}). \quad (2.12)$$

Dans la littérature, le modèle aléatoire de référence de répartition homogène de points (*complete spatial randomness*) est le processus de Poisson homogène d'intensité égale à la densité de points. La valeur de DNFA associée au processus de Poisson homogène

d'intensité P_{fa} est [Grégoire, 2011] :

$$DNFA_{\text{poisson}} = \frac{1}{2\sqrt{P_{FA}}}. \quad (2.13)$$

Cette valeur fournit une valeur théorique de référence de $DNFA$ pour une répartition « la plus uniforme possible » des FA lorsque la probabilité de FA vaut P_{fa} . Plus le critère $DNFA$ (équation 2.12) est proche de la valeur $DNFA_{\text{poisson}}$, plus la distribution spatiale des FA est homogène.

2.2 Méthode d'évaluation des critères de performances

Dans la littérature nous pouvons distinguer deux approches d'évaluation de la P_d et de la P_{fa} : l'une évalue ces quantités sur des grandes bases d'images comprenant chacune un nombre limité de cibles réelles, l'autre les évalue sur uniquement une image ou une courte séquence d'images pour laquelle en chaque pixel de l'image une cible simulée est incrustée.

Cibles réelles :

La première approche, assez simple à mettre en oeuvre, est celle qui est la plus utilisée dans la littérature [Goudou, 2007, Gu *et al.*, 2010]. Elle consiste à estimer la P_d et la P_{fa} en comptant le nombre de détections et FA que l'on obtient après filtrage pour des séquences d'images contenant des vraies cibles. La seule difficulté est que le calcul de la P_d et de la P_{fa} nécessite de connaître la vérité terrain. Cela n'est pas toujours possible sur des images réelles : cela suppose l'utilisation d'une technique de détection de référence, par exemple manuelle, ou encore automatique à résolution supérieure. Notamment pour les cibles de faible intensité, il n'est parfois pas possible de distinguer la cible du fond et de déterminer sa position.

De plus, même lorsqu'on utilise des bases importantes, les performances de détection obtenues et surtout le positionnement d'un algorithme par rapport à d'autres, dépendent des caractéristiques des cibles : forme, intensité mais surtout position dans l'image. A part dans des cas très particuliers, la cible ne parcourt pas toutes les positions relativement à une image de fond donnée. On peut donc par exemple évaluer de bonnes performances tout en ayant des zones *aveugles*, c'est-à-dire des zones pour lesquelles les cibles sont non détectables pour l'algorithme considéré. Ces zones n'ont pas été prises en compte dans l'évaluation parce qu'aucune cible n'y est apparue dans la base d'images utilisée. Il est donc nécessaire d'être prudent quant à la généralité des conclusions que l'on pourrait déduire d'une telle évaluation.

Cibles simulées :

Cette seconde approche qui travaille avec des cibles simulées sur des fonds supposés ne contenir aucune cible, a l'avantage que la vérité terrain est connue. Elle est moins répandue dans la littérature sans doute par crainte que la simulation ne soit pas assez réaliste, mais elle permet de prendre en compte le fait qu'une cible peut en pratique se trouver en n'importe quel pixel du fond [Samson, 2002, Caefer *et al.*, 2007, Bar-Tal *et al.*, 1996].

Le travail de Bar-Tal *et al.* tient déjà compte de cela : ils proposent une méthode d'évaluation de performances basée uniquement sur le traitement de fonds sans incrustation de cible. Ils supposent que le filtre atténue les cibles et le fond selon un même facteur :

l'intensité d'un pixel en sortie de traitement est proportionnelle à celle d'entrée. Cette hypothèse est très pratique puisqu'elle permet d'éviter l'incrustation de cible mais elle n'est vraie que pour des filtres linéaires ce qui est beaucoup trop restrictif.

Pour aborder les techniques non-linéaires qui sont étudiées dans ce travail, il convient d'utiliser une méthode d'évaluation basée sur des incrustations de cibles comme utilisée dans [Samson, 2002] et [Cafer et al., 2007]. Ce type de méthode consiste tout d'abord à filtrer le fond en l'absence de cible pour obtenir après seuillage une évaluation de la P_{FA} . Une cible est ensuite incrustée en chacun des pixels et le seuillage de la sortie du filtre permet d'évaluer la P_D (figure 2.2). Cette méthode permet donc de simuler la présence

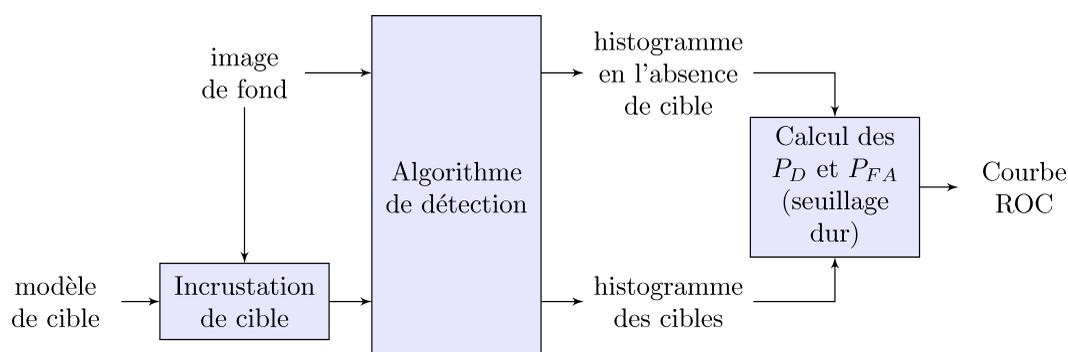


FIGURE 2.2 – Méthode d'évaluation des performances en tout pixel.

d'une cible en tout pixel d'une image. Cependant, comme elle nécessite d'appliquer l'algorithme de détection à chaque fois qu'une cible est incrustée en un pixel du fond, elle est très coûteuse en temps de calcul. De plus il faut que le modèle de cible utilisé corresponde bien aux cibles que l'on rencontre dans la réalité : dans notre contexte de cible non résolue il s'agit surtout d'un problème de modélisation de l'instrument ; nous y reviendrons en section 2.5.

Afin d'obtenir une évaluation des performances pour l'ensemble des pixels de l'image, c'est donc cette seconde méthode d'évaluation que nous utilisons. Pour réduire le temps de calcul nous incrustons des matrices de cibles, c.a.d simultanément plusieurs cibles avant d'effectuer le traitement. A chaque itération des cibles sont incrustées tous les 15 pixels verticalement et horizontalement (ce qui correspond à 0.5% des pixels) de façon à ce que la détection d'une cible ne soit pas perturbée par la présence des autres cibles.

Pour évaluer les performances de détection nous nous fixons des bases de fonds qui contiennent plusieurs images de fonds nuageux acquis par un même instrument (cf. section 2.4). Pour chacune des bases de fonds, nous évaluons une probabilité de détection qui est égale à la moyenne des probabilités de détection obtenues pour chacune des images de la base pour un taux de fausses alarmes donné.

L'évaluation des algorithmes de détection spatio-temporels peut être réalisée uniquement sur les bases pour lesquelles on dispose de plusieurs séquences d'images de fonds et non pas uniquement de plusieurs images de fonds non-corrélées. Cette évaluation nécessite d'incruster des cibles dans les différentes images de chacune des séquences de la base de fonds en considérant la trajectoire et la vitesse des cibles connues *a priori* (cf. section 2.5). Pour chacune des séquences de la base de fonds, on applique un filtrage spatio-temporel avec comme image de référence la même image que celle utilisée pour évaluer les performances des algorithmes spatiaux. On calcule ensuite la P_d et la P_{fa} à

partir de cette unique image filtrée. On a ainsi des valeurs de P_d et P_{fa} pour chacune des séquences constituant la base de fonds. Comme pour les algorithmes de détection spatiaux, pour une P_{fa} fixée, la valeur de P_d retenue est la moyenne des P_d .

Finalement remarquons que le critère $DNFA$ qui ne concerne que la distribution des FA ne pose pas de problème puisqu'il peut être évalué sur des fonds réels sans cible. La valeur de $DNFA$ associée à une base de fonds est la moyenne des $DNFA$ obtenus pour chacun des fonds constituant la base.

2.3 Evaluation par zones

Dans la pratique, il peut être intéressant d'appliquer des filtres de détection différents dans les zones de l'image qui ont des caractéristiques différentes. En effet, d'une part, en ce qui concerne les performances optimales il est plutôt naturel de considérer que le meilleur filtre ne sera pas le même dans les zones homogènes et dans les zones de fortes transitions tels que les bords de nuages. La segmentation peut ainsi permettre d'optimiser les performances en modulant les P_{fa} de chaque zone. D'autre part appliquer des algorithmes différents en fonction des zones de l'image peut permettre d'accélérer le traitement des données sans perdre de performances. En particulier, on peut souvent utiliser dans les zones homogènes des techniques simples et rapides. Des méthodes plus coûteuses mais plus performantes seront alors appliquées uniquement dans les zones où elles permettent d'améliorer les performances. On peut donc se poser le problème de comparer les filtres zones par zones après avoir segmenté l'image.

La segmentation peut par exemple consister à distinguer les portions à basse fréquence spatiale de l'image (zones homogènes ou zones de variations lentes d'intensité) des portions de l'image à haute fréquence comme les zones de fort gradient localisées aux bords des nuages. Sur ce principe, deux exemples de segmentations typiques de l'image de la figure 2 sont présentés en figure 2.3. Ces segmentations sont obtenues en seuillant, l'écart-type local calculé sur une fenêtre de dimension 5×5 pixels, avec un seuil égal à l'écart-type local moyen estimé sur toute l'image pour la segmentation de droite (figure 2.3a) et un seuil deux fois plus élevé pour celle de gauche (figure 2.3b).

Le problème d'une méthode d'évaluation par zones est que le choix de la segmentation, considérée pour définir les zones, influence très fortement les conclusions que l'on obtient concernant les performances des filtres étudiés. La figure 2.4 présentent les courbes ROC évaluées sur l'image de la figure 2 pour quatre filtres de comportements assez différents (cf. chapitre 5 pour plus de détails concernant les filtres évalués), lorsque l'on fait une évaluation globale des performances et lorsque l'on fait des évaluations par zones en faisant intervenir les deux segmentations de la figure 2.3. On constate que les performances de détection par zones, plus particulières pour les zones de basses fréquences dépendent fortement de la segmentation utilisée. Pour la segmentation de la figure 2.3a où le seuil sur l'écart type est bas, les zones de basses fréquences sont très homogènes. Dans de telles zones, comme le montre la figure 2.4b, les filtres de détection simples tels que le $BSMean$ et le $BSMedian$ sont aussi performants que les filtres de détection plus complexes. Pour ces zones, il semble que l'on soit très proche des meilleures performances atteignables. Cette hypothèse est corroborée par une remarquable convergence des performances de toutes les méthodes analysées dans ces zones. Cependant comme le montre la figure 2.4d, si les zones classifiées basses fréquences contiennent des variations d'intensité un peu plus importantes, on a une chute des performances des deux filtres $BSMean$ et $BSMedian$.

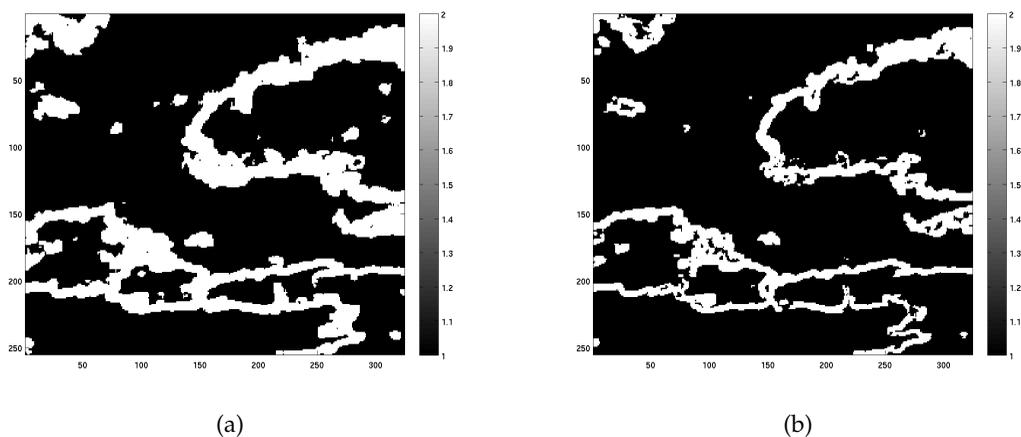


FIGURE 2.3 – Exemple de deux segmentations de l’image présentée en figure 2 que l’on pourrait envisager pour réaliser des évaluations de performances par zones.

A partir du moment où les zones contiennent quelques non-stationnarités, que l’on fasse une évaluation globale des performances ou une évaluation par zone, l’ordre des filtres en terme de performances reste le plus souvent inchangé.

Cherchant, dans cette thèse, à développer de nouveaux algorithmes pour améliorer la détection en présence de fonds non-stationnaires, on choisit pour comparer les algorithmes de détection entre eux, de se restreindre à une évaluation globale des performances. Ce choix nous permet de fournir des évaluations de performance qui ne dépendent pas du réglage d’une méthode de segmentation. Pour les zones très homogènes, nous restons conscient que des algorithmes moins coûteux peuvent donner d’aussi bonnes performances que ceux que nous proposons. On s’autorise bien sûr à proposer des algorithmes de détection qui appliquent des méthodes de filtrage différentes en fonction des zones de l’image, mais dans toute la suite l’évaluation de leurs performances reste une évaluation globale sur toute l’image.

2.4 Images de fond

Nous étudions la détection dans un contexte difficile de fonds inhomogènes. Nous nous intéressons en particulier aux applications de surveillance aérienne du ciel par imagerie infrarouge que se soit aussi bien à partir de systèmes de surveillance terrestre que de systèmes embarqués à bord de satellites. Pour évaluer et comparer les méthodes de détection proposées, nous avons donc constitué deux bases d’images correspondant à ces deux contextes d’applications que nous nommerons respectivement *Base de fonds terrestre* et *Base de fonds satellitaire*.

2.4.1 Base de fonds terrestre

La base de fonds terrestre présentée en figure 2.5, est constituée de neuf images acquises depuis le sol terrestre par nos soins. La caméra utilisée est une caméra non-refroidi Photon 320 de la marque FLIR équipée d’un objectif de focale 35 mm qui travaille

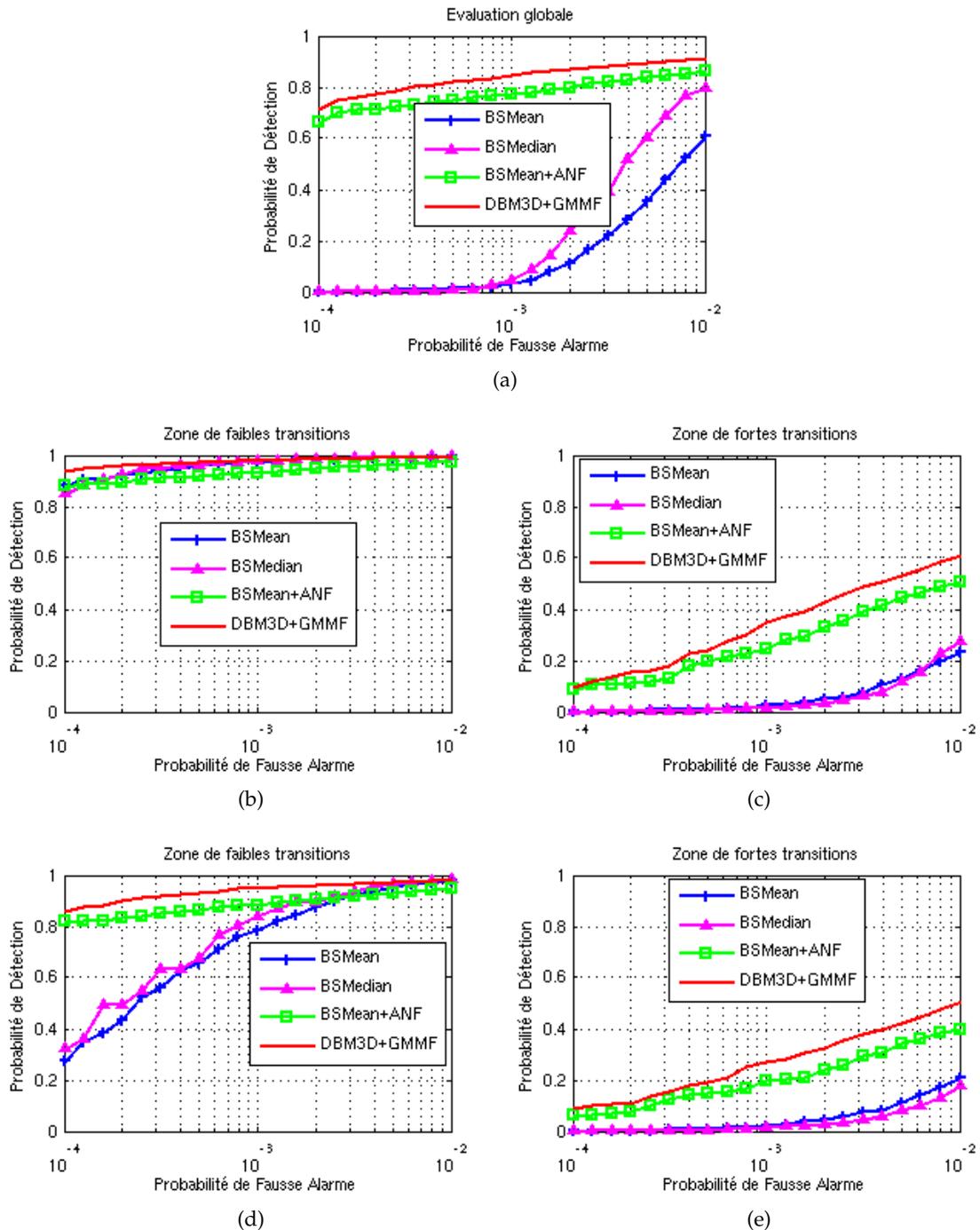


FIGURE 2.4 – Courbes ROC évaluées à partir de l'image de la figure 2 pour différents filtres de détection. En haut, (a) : évaluation globale sur toute l'image ; au milieu : évaluation par zones pour la segmentation de la figure 2.3a ; en bas : évaluation par zones pour la segmentation de la figure 2.3b.

en bande $7.5 - 13.5 \mu\text{m}$, à une fréquence de 25 Hz. Le détecteur est une matrice de 255×324 pixels, sur une dynamique de 14 bits, avec une résolution de 1.086 mrad.

Les images de cette base de fonds présentent essentiellement deux types de fonds bien distincts correspondant respectivement à un ciel clair et à des nuages. De nombreuses zones de forts gradients sont présentes aux bords des nuages. Dans certaines de ces images un dégradé d'intensité vertical est visible dans les zones de ciel. Ce dégradé est dû à une variation d'éclairement en fonction de l'altitude.

Ces neuf images sont extraites de neuf séquences d'images, ce qui nous permet d'évaluer des algorithmes de détection spatio-temporels sur cette base de fonds. Pour évaluer ces algorithmes, on filtre chacune des neuf images de la base de fonds en faisant intervenir pour chaque image, les images de la séquence dont elle est extraite qui la précèdent. La P_d à P_{fa} fixée est ensuite calculée à partir des neuf images filtrées. Les performances de détection sont évaluées pour des séquences acquises à la fréquence nominale de 25Hz. Mais afin d'évaluer les performances des algorithmes en fonction de la corrélation temporelle des fonds, on crée à partir de ces séquences de nouvelles séquences correspondant à une fréquence d'acquisition plus faible (0.5Hz) et donc de corrélation temporelle plus faible en ne retenant qu'une image sur cinquante.



FIGURE 2.5 – Base de fonds terrestre.

2.4.2 Base de fonds satellitaire

La base de fonds satellitaire présentée en figure 2.6, est constituée de neuf images capturées depuis le satellite d'orbite basse MetOp-A (MetOp : *Meteorological Operational satellite*), qui sont téléchargeables depuis le site de EUMETSAT (*European Organisation*

for the Exploitation of Meteorological Satellites) [Eum,]. Ces images ont été acquises par un capteur AVHRR (*Advanced Very High Resolution Radiometer*). Il s'agit d'un radiomètre imageur à six canaux spectraux compris entre 0.6 et 12 μm dont seulement cinq sont transmis au sol à un moment donné. Le canal 3 est un canal dédoublé sur lequel le canal 3a est exploité durant la partie diurne de l'orbite et le canal 3b pendant la partie nocturne. Les images constituant la base de fonds sont des petites images de dynamique 8 bits, de dimension 251×251 pixels extraites des données de la bande 3a qui correspond à une bande spectrale de 1.580 – 1.640 μm . L'instrument AVHRR est un système à balayage caractérisé par un champ d'observation très large. Sa fauchée est de 2940 km et sa résolution de 1.1 km pour la bande 3.

Ce système ayant une fréquence de balayage relativement faible et étant embarqué à bord d'un satellite d'orbite basse n'est pas le plus adapté pour faire de la veille. Cependant les structures nuageuses obtenues sont bien représentatives de celles que l'on pourra observer dans des fonds acquis par de futurs systèmes de veille embarquant des capteurs matriciels à bord de satellites géostationnaires. C'est pourquoi nous avons choisi de nous intéresser à ces fonds.

Le temps entre deux prises de vue par le capteur AVHRR d'un même point terrestre étant très important, le fond nuageux a fortement évolué entre ces deux prises de vue. Ce système d'observation ne permet donc pas de disposer de séquences d'images exploi-

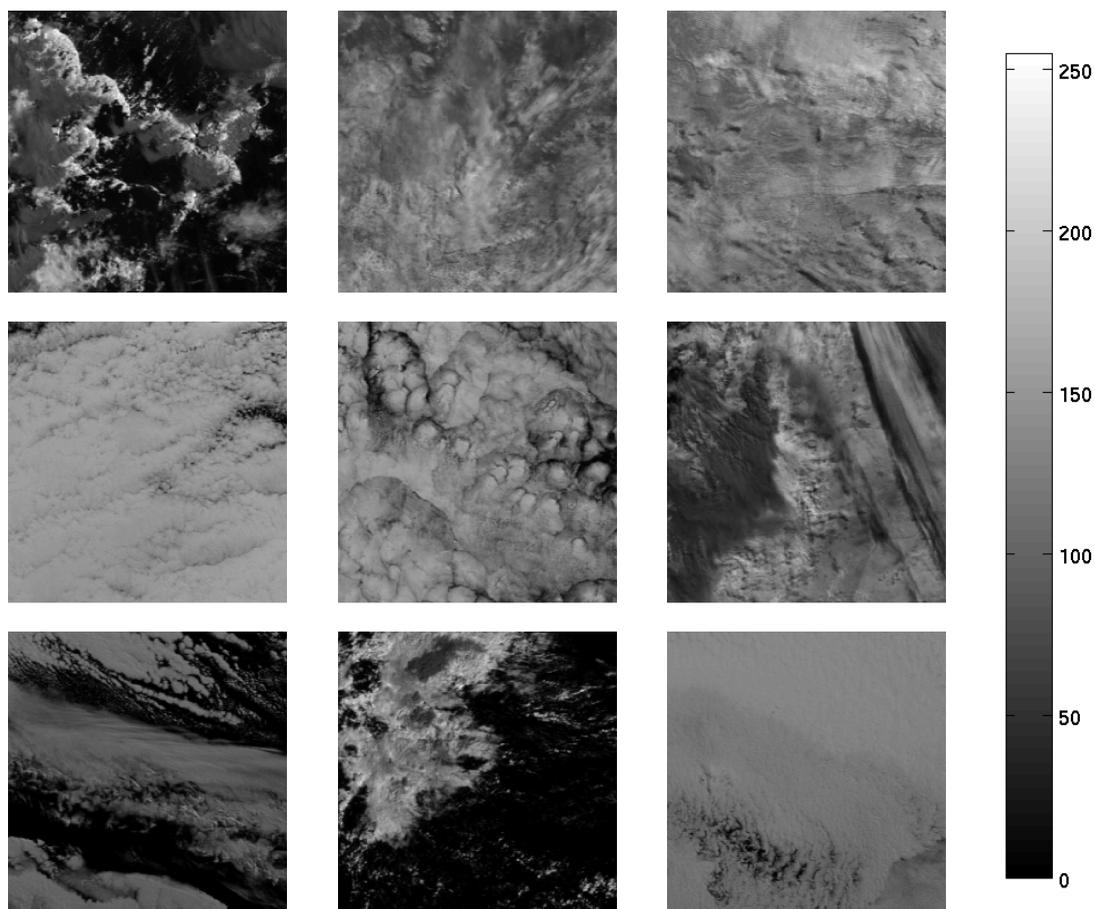


FIGURE 2.6 – Base de fonds satellitaire.

tables temporellement et pour cette base de fonds seules des évaluations de performances dans le contexte mono-image sont réalisées.

Cette deuxième base de fonds contient des fonds nuageux qui présentent différentes zones texturées de diverses fréquences spatiales, on peut considérer qu'elles sont constituées d'une variété de micro-textures. En contraste avec la base de fonds terrestre, il est difficile de segmenter ces images en deux régions labélisées « nuage » et « ciel », séparées par des bords nets.

2.5 Modèle de cible

Afin de pouvoir estimer les performances des filtres de détection en tout pixel de l'image, des cibles simulées sont incrustées sur les images de fonds réels à étudier.

La simulation de cibles se trouvant à une très grande distance du système imageur est d'un premier abord assez simple puisque l'image non échantillonnée de la cible correspond à la réponse impulsionnelle (PSF : *Point Spread Function*) du système optique si l'on considère que les cibles se trouvent à une distance infinie du système imageur. Considérant que dans notre application le système est limité par la diffraction, la PSF de l'optique est supposée être définie par la fonction d'Airy suivante [Samson, 2002] :

$$h_o(v, w) = \frac{1}{\pi} \left[\frac{J_1(\pi r_c \rho)}{\rho} \right]^2, \text{ avec } \rho = \sqrt{v^2 + w^2} \quad (2.14)$$

avec $J_1(\cdot)$ la fonction de Bessel de première espèce, $r_c = f_c/f_e$ la fréquence de coupure normalisée. La fréquence f_c correspond à la fréquence de coupure de l'instrument ($f_c = \frac{D_{\text{pup}}}{\lambda f'}$ avec D_{pup} le diamètre de la pupille, λ la longueur d'onde incidente et f' la focale du système optique) et la fréquence f_e correspond quant à elle à la fréquence d'échantillonnage du capteur.

Cette image de la cible est ensuite échantillonnée par le capteur. La PSF du système échantillonné s'obtient en convoluant la PSF du système optique h_o avec la PSF du détecteur h_d : $h = h_o * h_d$. Comme cela est souvent fait dans la littérature, on modélise la PSF du détecteur par une fonction porte :

$$h_d(v, w) = \begin{cases} 1 & \text{si } |v| > 1/2 \text{ et } |w| > 1/2 \\ 0 & \text{sinon} \end{cases} \quad (2.15)$$

Lors de la modélisation des cibles, il est important faire attention au décalage subpixelique des cibles par rapport au centre des pixels. En effet les cibles ne sont quasiment jamais parfaitement centrées sur un pixel, une cible peut tout à fait se trouver à l'intersection entre deux pixels. On note $\epsilon = (\epsilon_i, \epsilon_j)$, le décalage subpixelique de la cible par rapport au centre du pixel le plus proche : $\epsilon \in [-0.5, 0.5]^2$. Pour les systèmes sous-échantillonnés ($r_c > 0.5$), l'apparence des cibles va très fortement dépendre de leur position subpixelique comme on peut le voir sur la figure 2.7. Cette variation d'apparence des cibles en fonction de leur position subpixelique diminue quand le pas d'échantillonnage du capteur diminue. Cependant, comme pour un grand nombre de système la largeur de la PSF optique est du même ordre de grandeur que le pas du capteur, il est important de prendre en compte ce phénomène de phasage en tenant compte du fait qu'une cible peut tout à fait avoir n'importe quel décalage subpixelique.

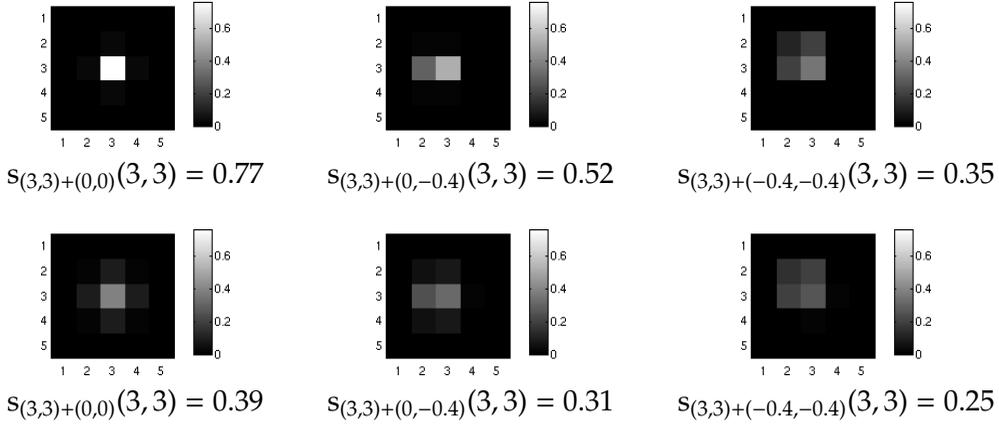


FIGURE 2.7 – Signatures de cibles d'intensité globale : $s_o = 1$ et intensité du pixel central ($\mathbf{x} = (3,3)$), pour différentes positions subpixelliques de cible et différents systèmes optique. En haut : $r_c = 1.5$, en bas : $r_c = 0.8$. De gauche à droite les positions subpixelliques de la cible sont $(0,0)$, $(0,-0.4)$ et $(-0.4,-0.4)$.

La signature des cibles en fonction de leur décalage subpixellique se modélise de la façon suivante [Samson *et al.*, 2004] :

$$s_{\mathbf{x}_c+\boldsymbol{\epsilon}}(\mathbf{x}) = h(\mathbf{x}_c + \boldsymbol{\epsilon} - \mathbf{x}) = s_o \int_{(i_c-i)-0.5}^{(i_c-i)+0.5} \int_{(j_c-j)-0.5}^{(j_c-j)+0.5} h_o(v + \epsilon_i, w + \epsilon_j) dv dw, \quad (2.16)$$

L'intensité globale de la cible est notée s_o et $s_{\mathbf{x}_c+\boldsymbol{\epsilon}}(\mathbf{x})$ correspond à l'intensité du pixel $\mathbf{x} = (i, j)$ pour une cible de position pixellique $\mathbf{x}_c = (i_c, j_c)$ et de décalage subpixellique $\boldsymbol{\epsilon} = (\epsilon_i, \epsilon_j)$.

L'apparence des cibles ainsi que leur niveau d'intensité maximal dépendent fortement de leur position subpixellique et de la fréquence de coupure normalisée du système imageur comme le montre la figure 2.7. La modification de la fréquence de coupure modifie la fonction d'étalement de la cible, plus le rapport f_c/f_e est petit et plus la cible est étalée. La figure 2.7 montre que la valeur associée au pixel le plus intense est toujours inférieure à s_o et dépend du rapport f_c/f_e et de la position subpixellique de la cible (ϵ_i, ϵ_j) . Les cibles que l'on incruste pour évaluer les performances des algorithmes ont une position subpixellique qui est tirée dans l'ensemble $[-0.5, 0.5]^2$ selon une loi uniforme.

Dans le cas de l'évaluation de performances des algorithmes spatio-temporels, la vitesse de la cible est fixée *a priori* et on fait intervenir deux directions de propagation de cibles. On incruste sur la séquence de fonds dans un premier temps des cibles de direction de propagation horizontale (cf. figure 2.8a), puis dans un second temps on incruste des cibles de direction diagonale (cf. figure 2.8b). La Pd à Pfa fixée est évaluée pour chacune des deux directions de propagation, puis on moyenne les deux valeurs de Pd obtenues. Dans nos évaluations nous considérons une vitesse nominale de cible de 1.25 pixels par image, mais nous étudions aussi le contexte plus difficile où les cibles ont une vitesse beaucoup plus faible égale à 0.25 pixels par image.

Pour pouvoir classifier les algorithmes de détection en terme de performances Pd(Pfa), il est nécessaire que l'intensité s_o ne soit pas trop élevée pour que tous les algorithmes ne détectent pas toutes les cibles présentes, ni trop faible pour que tous ne détectent pas quasiment aucune cible. La valeur de s_o doit être choisie de façon à obtenir un problème de

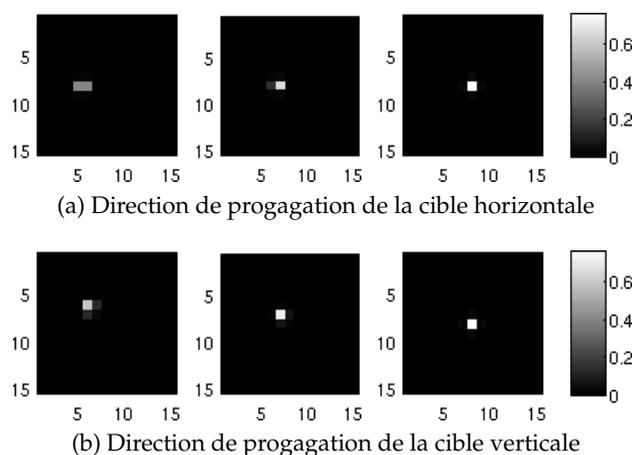


FIGURE 2.8 – Signature temporelle de la cible incrustée pour une séquence de trois images, pour une cible de vitesse 1.25 pixels par image et de direction de propagation soit horizontale, soit diagonale (45° par rapport à la direction horizontale).

détection ni trop difficile, ni trop facile. Dans la suite, pour nos deux bases de fonds, une intensité globale s_o égale à 50 est considérée pour permettre de comparer les algorithmes de détection entre eux pour des Pfa comprises entre 10^{-4} et 10^{-2} . Dans le chapitre 5 des études de performances en fonction de l'intensité de la cible sont réalisées afin de vérifier que les conclusions obtenues n'ont pas une validité qui se limite à ce point de fonctionnement ($s_o = 50$).

En ce qui concerne la caméra Flir, les caractéristiques fournies par le fabricant :

- nombre d'ouverture : $N = 1.4$,
- taille du pixel : $p = 38 \mu\text{m}$,
- bande spectrale : 7.5 à 13.5 μm (soit une bande spectrale moyenne : $\lambda_m = 10.5 \mu\text{m}$),

nous permettent de calculer une valeur théorique de r_c égale à 2.6 sur la base de l'hypothèse que le système est limité par la diffraction ($r_c = \frac{p}{\lambda_m N}$). Des valeurs plus petites de r_c doivent être considérées pour tenir compte des aberrations optiques. En ce qui concerne le capteur AVHRR chargé sur le satellite MetOp-A, la seule information trouvée est que la fonction de transfert de modulation MTF à la fréquence d'échantillonnage MTF_{v_s} , est supérieure à 0.3 [Govaerts, 2006]. En approximant la MTF par une ligne droite passant par les points $(0, 1)$, (v_s, MTF_{v_s}) et $(v_c, 0)$, on obtient que r_c est supérieur à 1.4. Par conséquent, un paramètre r_c égal à 1.5 semble un choix pertinent pour les deux bases de données.

Dans la suite les paramètres dits nominaux associés à la modélisation des cibles sont donc $s_o = 50$, $r_c = 1.5$ et $v_c = 1.25$ pixels par image ; si d'autres paramètres que ceux-ci sont utilisés lors d'une évaluation cela sera précisé.

Conclusion

Dans ce chapitre on a mis en place une méthode d'évaluation des algorithmes de détection d'objets ponctuels. Elle est basée sur des fonds nuageux infrarouges naturels acquis depuis le ciel ou le sol terrestre. Les cibles sont quant à elles modélisées finement en prenant en compte leurs positions subpixelles. Dans la suite et plus particulièrement dans le chapitre 5, on utilise cette méthode d'évaluation pour analyser les performances des algorithmes qui ont été développés au cours de cette thèse.

Chapitre 3

Modélisation du fond au premier ordre : Développement d’algorithmes de détection de cibles ponctuelles basés sur des méthodes de débruitage à patchs

Introduction

Ce chapitre a pour but de présenter de nouvelles méthodes de détection basées sur une modélisation du fond au premier ordre. Ce sont des adaptations de filtres de débruitage dits à patchs dans le cadre d’une approche de détection par suppression de fond. En effet, les méthodes de débruitage à patchs reposent sur une agrégation de patchs similaires trouvés dans l’image elle-même ou dans la séquence d’images et leur adaptation à notre contexte se heurte au fait que la cible brise la similarité. On a donc cherché des moyens de restaurer la similarité entre patchs malgré la présence de la cible.

Organisation du chapitre — En section 3.1, nous introduisons tout d’abord l’idée de base sur laquelle s’appuient les méthodes de détection par suppression de fond proposées dans ce chapitre. Il s’agit d’utiliser la sortie d’un algorithme de débruitage pour modéliser le fond dans le cadre d’une approche de détection par suppression de fond. Dans cette même section 3.1, nous présentons le principe des méthodes de débruitage auxquelles nous nous intéressons qui sont les méthodes à patchs. La section 3.2 présente les deux algorithmes de cette famille que nous adaptons à la détection : **NL-means** (*Non-Local Means*) [Buades *et al.*, 2005b] et **BM3D** (*Block Matching and 3D Filtering*) [Dabov *et al.*, 2007b]. La section 3.3 présente, tout d’abord en partie 3.3.1, l’adaptation du filtre de débruitage **NL-means** à la détection, qui conduit au filtre **DNLM** (*Detection by NLM*). Puis, dans la partie 3.3.2, le nouvel algorithme de détection **DBM3D** (*Detection by BM3D*) basé sur l’algorithme de débruitage **BM3D** est présenté.

3.1 Introduction de l'approche proposée

3.1.1 Approche par suppression de fond

Comme cela a été présenté au chapitre 1 dans la section 1.6.1, les méthodes de suppression de fond (**BS** : *Background Suppression*) consistent à faire ressortir les cibles en soustrayant aux données observées un fond estimé. Cette estimation peut s'interpréter comme une modélisation du fond au premier ordre, première étape du schéma général de la figure 1.4. Dans tout ce chapitre, c'est à ce type de méthodes que nous nous intéressons. Dans l'idéal, l'objectif est de supprimer totalement le fond et de ne pas modifier l'intensité de la cible. On souhaite donc que l'estimation du fond ne soit pas influencée par la présence de cibles, autrement dit qu'elle estime le mieux possible le fond que l'on aurait observé si aucune cible n'avait été présente. On souhaite réciproquement qu'elle modifie le moins possible le fond observé là où aucune cible n'est présente.

3.1.2 Lien avec le débruitage

La problématique de l'estimation du fond malgré la présence de cibles peut être considérée comme un problème de débruitage où les cibles jouent le rôle d'un bruit (bruit impulsionnel de faible fréquence d'apparition) dont il faut s'abstraire.

Le but des méthodes de débruitage est de retrouver l'image originale à partir d'une mesure bruitée. Le modèle de l'image bruitée peut s'écrire sous la forme classique suivante,

$$u(\mathbf{x}) = v(\mathbf{x}) + n(\mathbf{x}) \quad (3.1)$$

où $u(\mathbf{x})$ est la valeur observée, $v(\mathbf{x})$ est la valeur « vraie » et $n(\mathbf{x})$ le bruit au pixel de position \mathbf{x} . On considère couramment que le bruit est gaussien, indépendant et identiquement distribué (i.i.d.). Une cible ne correspond pas à un bruit gaussien mais à un bruit impulsionnel. Cependant, nous verrons qu'en pratique des méthodes de débruitage initialement développées pour éliminer un bruit gaussien peuvent être adaptées pour assurer une bonne élimination des bruits ponctuels tels que les cibles.

La difficulté est d'estimer le fond dans les régions de fortes variations d'intensité du fond telles que les zones de bords de nuages. Dans ces régions, de simples méthodes linéaires d'estimation du fond échouent et le processus de détection par **BS** conduit à un taux élevé de fausses alarmes. Pour améliorer l'estimation du fond dans les zones texturées et les zones de fortes transitions, pour lesquelles les méthodes actuelles échouent, nous nous sommes intéressés aux approches de débruitage à patchs considérées actuellement comme les méthodes de débruitage les plus performantes [Salmon et Strozecki, 2012, Lebrun *et al.*, 2012, Buades *et al.*, 2012]. Ces méthodes à patchs améliorent la qualité du débruitage dans les zones texturées et dans les zones de fortes transitions. De plus les versions spatio-temporelles de ces méthodes [Buades *et al.*, 2005a, Dabov *et al.*, 2007a] présentent l'avantage de ne pas nécessiter d'estimation précise préalable du mouvement du fond. Nous verrons dans la suite que la présence d'une cible modifie le comportement de ces algorithmes et que ceux-ci ne peuvent pas être utilisés tels quels pour faire de la détection par **BS**. Nous montrons comment adapter ces algorithmes pour les rendre robustes à la présence d'une cible et obtenir ainsi un filtre de détection par **BS**.

3.1.3 Principe des algorithmes de débruitage à patches

Les méthodes de débruitage à patches exploitent le caractère répétitif des motifs présents dans une image : toute fenêtre de petite taille (patch) dans une image naturelle a de nombreuses fenêtres similaires dans cette même image. Cette redondance est illustrée en figure 3.1. Cette redondance n'est pas surprenante pour les patches extraits de zone basse fréquence ou à la limite de zones homogènes, en revanche il est intéressant de voir que cette propriété est aussi vraie pour les zones de hautes fréquences.



FIGURE 3.1 – Illustration de la redondance de motifs dans une image naturelle. En gras : patches de référence, en trait simple : patches similaires.

Définition d'un patch :

Définissons de manière plus formelle un patch. Nous reprenons les notations proposées par Salmon *et al.* [Salmon et Strozecki, 2012] qui consistent à indexer un patch par son coin en haut à gauche (comme cela est fait dans la littérature concernant le filtre BM3D [Dabov *et al.*, 2007b]), ce qui permet d'indexer aussi bien des patches de dimension impaire que paire. Dans cette thèse les patches ont une dimension de $N_1 \times N_1$ pixels et le patch extrait de l'image observée bruitée, u , dont l'élément en haut à gauche a pour coordonnée x , est noté :

$$P_x = (u(x + \tau), \tau \in \{0, \dots, N_1\}^2). \quad (3.2)$$

Suivant cette notation le patch dont l'élément en haut à gauche a pour coordonnée $x - \delta$ est noté $P_{x-\delta}$ comme illustré sur la figure 3.2. Et pour les patches de dimension N_1 impaire, celui qui est centré sur le pixel x peut ainsi être noté $P_{x-\delta_{N_1}}$ avec $\delta_{N_1} = (\frac{N_1-1}{2}, \frac{N_1-1}{2})$.

Principe général des approches à patches :

Pour débruiter un pixel, l'idée de base de ces méthodes est de ne pas utiliser uniquement les pixels qui sont proches spatialement mais de tirer parti des pixels qui sont proches d'un point de vue radiométrique. Cette idée de combiner des pixels qui sont proches en termes de radiométrie est déjà présente dans des filtres qui travaillent dans le domaine pixellique ($N_1 = 1$). On peut citer le filtre débruiteur proposé par Yaroslavsky [Yaroslavsky, 1985] qui réalise une moyenne pondérée des pixels se trouvant dans un large voisinage en tenant compte de leur ressemblance radiométrique. Le filtre bilatéral a ensuite été proposé par Tomasi *et al.* [Tomasi et Manduchi, 1998], il combine l'approche proposée par Yaroslavsky avec les filtres moyennant spatiaux. Ce filtre bilatéral réalise une moyenne pondérée où les poids des pixels y intervenant dans la moyenne sont le

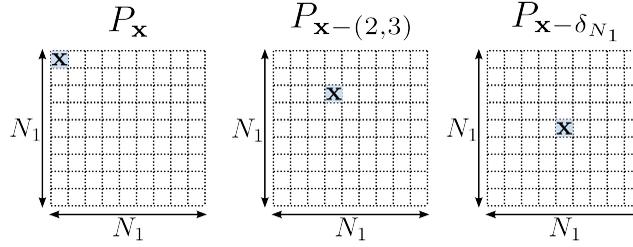


FIGURE 3.2 – Présentation de la convention utilisée pour définir les patches ($N_1 = 9$).

produit d'un terme qui tient compte de la distance spatiale entre les pixels y et x , le pixel à débruiter, et d'un terme qui tient compte de la distance radiométrique entre x et y .

Le filtre **NL-means** (*Non-Local Means*) proposé par Buades *et al.* [Buades *et al.*, 2005b] étend l'idée du filtre de Yaroslavsky et du filtre bilatéral en proposant une nouvelle mesure de distance radiométrique entre les pixels x et y basée sur une mesure de similarité entre les voisinages locaux de chacun des pixels (patches $P_{x-\delta_{N_1}}$ et $P_{y-\delta_{N_1}}$). Le filtrage **NL-means** a constitué un tournant dans le domaine du débruitage dans les années 2000. Il s'interprète comme étant une des toutes premières méthodes de débruitage à patches.

Le principe de base de toutes les méthodes de débruitage à patches est présenté en figure 3.3. Tout d'abord l'image est transformée en une collection de patches. Ensuite chacun des

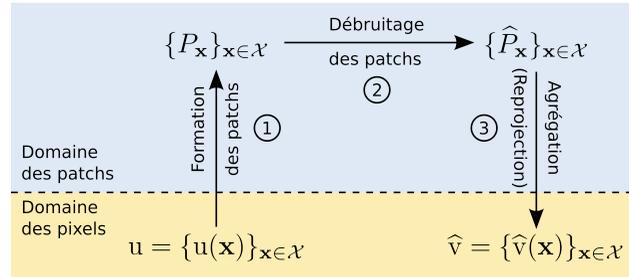


FIGURE 3.3 – Schéma général du principe des approches de débruitage par patches (schéma inspiré de la figure 1 de [Salmon et Strozecki, 2012]).

patches est débruité en tenant compte de la similarité entre les patches. Comme les patches se chevauchent, cette étape conduit à plusieurs estimées de la valeur débruitée de chacun des pixels de l'image. L'information doit donc être combinée pour obtenir une seule valeur débruitée de chaque pixel. Cette dernière étape est l'étape dite d'agrégation des différentes estimées [Dabov *et al.*, 2007b] ou de reprojection [Salmon et Strozecki, 2012].

Le filtre **NL-means** peut-être vu comme une méthode de débruitage à patch qui réalise une simple moyenne pondérée des patches pour les débruiter et qui fait intervenir une étape de reprojection qui consiste simplement à retenir la valeur débruitée des pixels se trouvant au centre des patches [Salmon et Strozecki, 2012] : $\widehat{v}(\mathbf{x}) = \widehat{P}_{\mathbf{x}-\delta_{N_1}}(\delta_{N_1})$.

Depuis cette approche, de très nombreux autres algorithmes de débruitage par patches ont été développés (BM3D [Dabov *et al.*, 2007b], BM3D-SAPCA [Dabov *et al.*, 2009], NLM-SAP [Deledalle *et al.*, 2012], NL-Bayes [Buades *et al.*, 2012], etc. ...) et de nouvelles approches sont développées tous les ans. Ces méthodes diffèrent par l'une des trois étapes du schéma de la figure 3.3.

- Étape de sélection des patches : si la plupart des méthodes considèrent simplement des patches carrés (BM3D, NL-means) comme nous le ferons dans toute la suite de cette thèse, d'autres méthodes proposent d'utiliser des patches de forme adaptative (BM3D-SAPCA, NLM-SAP).
- Étape de débruitage des patches : de nombreuses méthodes de débruitage des patches ont été proposées faisant intervenir des approches variées (simple moyennage [Buades et al., 2005b], DCT [Dabov et al., 2006], ondelettes [Dabov et al., 2007b], PCA [Dabov et al., 2009, Deledalle et al., 2011, Salmon et al., 2012], dictionnaires [Elad et Aharon, 2006, Aharon, 2006], filtrage bayésien [Buades et al., 2012]). Certains algorithmes pour débruiter un patch P_x considèrent l'ensemble des patches P_y et les pondèrent en fonction de la similarité entre le patch P_y et le patch P_x en cours de débruitage (NL-means). D'autres méthodes sélectionnent uniquement un nombre limité de patches : les patches P_y qui sont similaires au patch P_x en cours de débruitage (BM3D).
- Étape d'agrégation : certains algorithmes comme le NL-means [Buades et al., 2005b] retiennent simplement le pixel central de chaque patch débruité. D'autres algorithmes pondèrent les différentes estimées obtenues en fonction d'un critère qui vise à mesurer la qualité du débruitage obtenu pour chaque estimée [Dabov et al., 2007b, Salmon et Strozecski, 2012].

3.1.4 Références sur la détection par méthodes à patches

Le filtre bilatéral, qui peut être considéré comme le précurseur des techniques de débruitage à patches, a déjà été utilisé pour faire de la détection [Pei et al., 2006, Acito et al., 2011, Bae, 2011]. Les méthodes à patches fournissant actuellement des approches très performantes dans le domaine du débruitage, mais également dans d'autres domaines, comme par exemple la super-résolution [Protter et al., 2009] ou la reconstruction de données manquantes (*inpainting*) [Wong et Orchard, 2008], nous avons décidé de les utiliser pour faire de la détection.

Nous avons proposé dans un premier temps d'adapter à la détection le filtre à patches le plus connu : le NL-means. Cela a conduit à un filtre de détection par suppression de fond performant, nommé DNLM qui a fait l'objet de la publication [Genin et al., 2011]. Il est présenté dans ce chapitre en section 3.3.1. Notons que ce travail a été mené indépendamment du travail précédent de J.F. Goudou [Goudou, 2007] qui utilisait déjà NL-means pour la détection. De fait, nous montrons que la méthode DNLM que nous proposons diffère notablement de celle de J.F. Goudou. Au cours de cette thèse nous avons ensuite proposé (section 3.3.2) une deuxième méthode de détection à patches présentée dans [Genin et al., 2012b] et qui se base sur le filtre de débruitage BM3D.

3.2 Présentation de deux algorithmes de débruitage à patches

Nous nous sommes appuyés à des fins de détection sur les deux algorithmes à patches les plus connus : la méthode NL-means qui a constitué un tournant dans le domaine du débruitage et le filtre BM3D. Ce dernier a été considéré comme étant l'état de l'art jusqu'à très récemment et son principe est à la base des méthodes constituant actuellement l'état de l'art : NL-Bayes [Buades et al., 2012] et BM3D-SAPCA [Dabov et al., 2009].

3.2.1 Principe de l'algorithme NL-means

Dans le filtrage NL-means proposé par Buades, la valeur estimée de l'intensité du pixel \mathbf{x} est calculée à partir de l'image bruitée u par une moyenne pondérée de l'intensité des pixels \mathbf{y} se trouvant dans la zone de recherche S_x .

$$\widehat{v}(\mathbf{x}) = \frac{1}{C(\mathbf{x})} \sum_{\mathbf{y} \in S_x} u(\mathbf{y}) w(P_{\mathbf{x}-\delta_{N_1}}, P_{\mathbf{y}-\delta_{N_1}}), \quad (3.3)$$

$C(\mathbf{x})$ est un facteur de normalisation. La zone de recherche S_x est une fenêtre carrée centrée sur le pixel \mathbf{x} de dimension $N_s \times N_s$: $S_x \subseteq \mathcal{I}$ où \mathcal{I} est le domaine de définition de l'image en cours de traitement. La figure 3.4 permet de visualiser les différents voisinages (patches) et supports utilisés par le filtre NL-means pour débruiter le pixel \mathbf{x} . La pondération

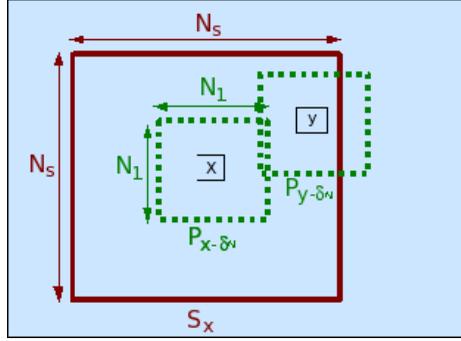


FIGURE 3.4 – Schéma présentant les différents supports et voisinages intervenant dans le débruitage du pixel \mathbf{x} dans le cas spatial. S_x est le support dans lequel on recherche les pixels similaires au pixel \mathbf{x} . $P_{\mathbf{x}-\delta_{N_1}}$ et $P_{\mathbf{y}-\delta_{N_1}}$ sont les patches centrés respectivement sur les pixels \mathbf{x} et \mathbf{y} , utilisés pour calculer la distance entre ces deux pixels.

$w(P_{\mathbf{x}-\delta_{N_1}}, P_{\mathbf{y}-\delta_{N_1}})$ est une fonction décroissante d'une mesure de distance entre les patches $P_{\mathbf{x}-\delta_{N_1}}$ et $P_{\mathbf{y}-\delta_{N_1}}$:

$$w(P_{\mathbf{x}-\delta_{N_1}}, P_{\mathbf{y}-\delta_{N_1}}) = e^{-d^2(P_{\mathbf{x}-\delta_{N_1}}, P_{\mathbf{y}-\delta_{N_1}})/h^2}. \quad (3.4)$$

Cette mesure de distance dépend de la similarité entre les patches centrés sur les pixels \mathbf{x} et \mathbf{y} , elle est calculée par une distance euclidienne avec une pondération gaussienne :

$$\begin{aligned} d^2(P_{\mathbf{x}-\delta_{N_1}}, P_{\mathbf{y}-\delta_{N_1}}) &= \|P_{\mathbf{x}-\delta_{N_1}} - P_{\mathbf{y}-\delta_{N_1}}\|_{2,a}^2 \\ &= \sum_{\mathbf{z} \in Q} G_a(\mathbf{z}) |u(\mathbf{x} + \mathbf{z}) - u(\mathbf{y} + \mathbf{z})|^2, \end{aligned} \quad (3.5)$$

où Q est une fenêtre carrée centrée en $(0,0)$ de dimension $N_1 \times N_1$:

$$Q = \{\mathbf{x} = (i, j) \mid -\frac{N_1-1}{2} \leq i \leq \frac{N_1-1}{2} \text{ et } -\frac{N_1-1}{2} \leq j \leq \frac{N_1-1}{2}\}.$$

G_a est un noyau gaussien qui permet d'accorder une plus grande importance aux pixels qui sont les plus proches du centre des patches : $G_a(\mathbf{z}) = \frac{\exp(-\|\mathbf{z}\|_2^2/2a^2)}{\sum_{\mathbf{z} \in Q} \exp(-\|\mathbf{z}\|_2^2/2a^2)}$ où $\|\cdot\|_2^2$ est la norme euclidienne classique. Dans le calcul du poids donné en équation (3.4), le paramètre h pondère la distance $d(P_{\mathbf{x}-\delta_{N_1}}, P_{\mathbf{y}-\delta_{N_1}})$ et ainsi contrôle l'exigence de similarité du filtre. Quand h est faible, l'exigence de similarité est très stricte et seul un nombre très limité de pixels de la zone de recherche ont un poids significatif dans la moyenne ; dans ce cas

l'image est peu débruitée. Inversement, quand h est élevé, l'exigence de similarité est lâche et tous les pixels de la zone de recherche ont le même poids ; dans ce cas l'image est « lissée », le débruitage est trop important. Le choix de la valeur de ce paramètre est donc particulièrement critique.

Notons de plus que selon l'équation (3.4), $w(P_{x-\delta_{N_1}}, P_{x-\delta_{N_1}}) = 1$ et qu'en pratique ce poids s'avère beaucoup plus élevé que les autres poids. Par conséquent, afin d'éviter la surpondération du pixel x , Buades *et al.* [Buades *et al.*, 2008] choisissent de fixer la valeur de ce poids de la façon suivante : $w(P_{x-\delta_{N_1}}, P_{x-\delta_{N_1}}) = \max_{y \in S_x \setminus \{x\}} w(P_{x-\delta_{N_1}}, P_{y-\delta_{N_1}})$. Présentée comme un détail d'implémentation dans [Buades *et al.*, 2008], cette modification est en fait, selon nos essais, essentielle pour que le filtre soit performant.

Une extension spatio-temporelle de ce filtre a été proposée par Buades *et al.* dans [Buades *et al.*, 2005a] : elle consiste à étendre la zone de recherche à des images temporelles voisines. Les voisinages intervenant dans la mesure de ressemblance entre les pixels restent quant à eux uniquement spatiaux. La figure 3.5 présente une illustration de cette extension du support de recherche. Par rapport aux approches de débruitage par compensation de mouvement, ce filtre de débruitage spatio-temporel a l'avantage de ne pas nécessiter d'estimation précise du mouvement inter-images. En effet tant que l'amplitude du mouvement reste inférieure à la dimension de la fenêtre de recherche, son influence sur la qualité du débruitage est faible.

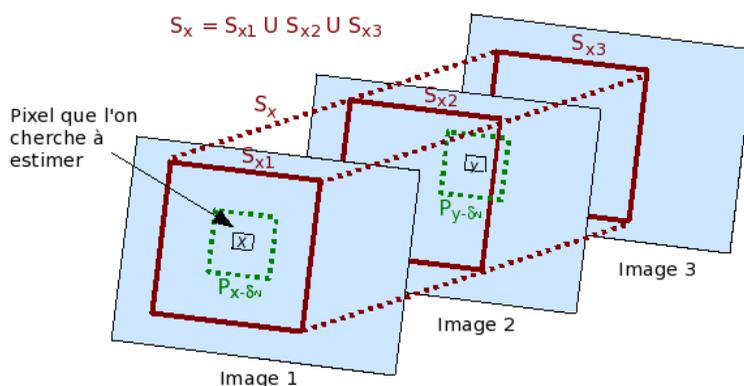


FIGURE 3.5 – Schéma présentant les différents voisinages, supports intervenant dans l'estimation du pixel x dans le cas spatio-temporel à trois images. S_x est le support dans lequel on recherche les pixels similaires au pixel x . S_x^i représente l'intersection entre le support S_x et la $i^{\text{ème}}$ image. $P_{x-\delta_{N_1}}$ et $P_{y-\delta_{N_1}}$ sont les patchs centrés respectivement sur les pixels x et y , utilisés pour calculer la distance entre ces deux pixels. Par convention le pixel en cours d'estimation appartient à l'image numéro 1.

3.2.2 Principe de l'algorithme BM3D

Dans le filtre BM3D [Dabov *et al.*, 2007b], pour débruiter un patch, on utilise uniquement les patchs qui lui ressemblent selon un certain critère. Les patchs similaires au patch de référence sont rassemblés dans un tableau de données 3D nommé groupe (cf. figure 3.6). Le filtrage BM3D propose ensuite une stratégie de débruitage basée sur la décomposition parcimonieuse du signal dans un domaine transformé tridimensionnel. Cette procédure de débruitage des groupes tire parti de la corrélation existant entre les

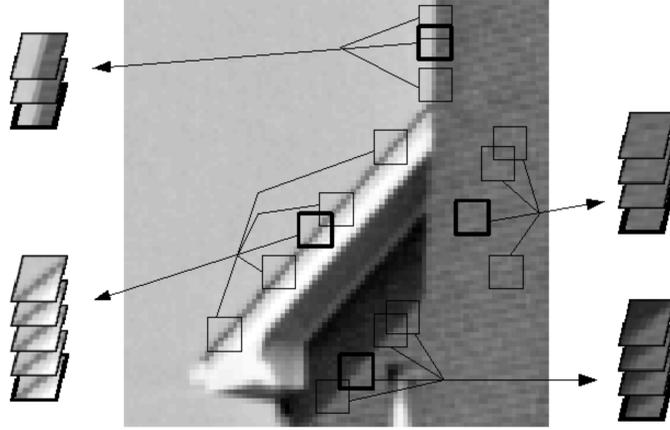


FIGURE 3.6 – Illustration de l'étape de construction des groupes qui consiste à associer à chaque patch de références (carré gras) les patches qui lui sont similaires.

différents patches les constituant. La valeur finale retenue pour un pixel est une moyenne pondérée des différentes estimations provenant de tous les patches débruités contenant ce pixel.

Pour améliorer les performances, le filtre **BM3D** applique successivement deux fois cette procédure générale de débruitage avec des paramètres de réglages spécifiques pour chacune des itérations. Cette deuxième itération du processus fait notamment intervenir un seuillage adaptatif par filtrage de Wiener empirique [Dabov *et al.*, 2007b]. Cependant nous avons observé que la seconde itération conduisait à un gain de performances de détection négligeable tout en augmentant le coût de calcul. Nous avons donc choisi de limiter dans la suite notre étude à l'application uniquement de la première itération de ce processus.

Dans les paragraphes qui suivent nous décrivons rapidement les trois étapes qui constituent la première itération de **BM3D**. Pour connaître plus de détails sur l'algorithme et les spécificités liées à la deuxième itération, le lecteur est invité à se reporter à [Dabov *et al.*, 2007b].

Construction des groupes

Pour construire les groupes, une mesure de dissimilarité entre deux patches de dimension $N_1 \times N_1$ pixels est tout d'abord définie de la façon suivante :

$$d^2(P_x, P_y) = \frac{\|P_x - P_y\|_2^2}{(N_1)^2} = \frac{1}{(N_1)^2} \sum_{z \in Q} |u(x+z) - u(y+z)|^2 \quad (3.6)$$

où $\|\cdot\|_2$ correspond à la norme euclidienne. L'ensemble Q correspond à une fenêtre de dimension $N_1 \times N_1$ dont l'élément en haut à gauche a pour coordonnée $(0,0)$:

$Q = \{\mathbf{x} = (i, j) \mid 0 \leq i \leq N_1 - 1 \text{ et } 0 \leq j \leq N_1 - 1\}$. Les deux patches P_x et P_y sont considérés comme similaires quand la distance $d^2(P_x, P_y)$ est inférieure à une valeur de seuil τ_{match} .

Pour constituer des groupes de patches, Dabov *et al.* commencent par choisir un certain nombre de patches de références, noté P_{x_R} , régulièrement espacés d'un pas $N_{step} \in \mathbb{N}$ dans

l'image. Un groupe sera constitué des patches similaires à un patch de référence. Pour des raisons de temps de calcul il n'est pas possible d'aller chercher des patches similaires dans toute l'image. La recherche de patches similaires au patch $P_{\mathbf{x}_R}$ est donc limitée à un large voisinage local $S_{\mathbf{x}_R}$ de dimension $N_s \times N_s$ centré sur \mathbf{x}_R . On note $C_{\mathbf{x}_R} \subseteq S_{\mathbf{x}_R}$ l'ensemble des coordonnées des patches qui sont similaires au patch de référence $P_{\mathbf{x}_R}$:

$$C_{\mathbf{x}_R} = \{\mathbf{x} \in S_{\mathbf{x}_R} \mid d^2(P_{\mathbf{x}}, P_{\mathbf{x}_R}) < \tau_{match}\}. \quad (3.7)$$

Les patches similaires sont ensuite rassemblés dans un tableau tridimensionnel nommé groupe que l'on note $\mathbf{P}_{C_{\mathbf{x}_R}}$. Toujours pour réduire la complexité algorithmique la taille des groupes est limitée à N_2 patches au maximum. Si il y a plus de N_2 patches qui vérifient la condition de similarité (3.7), le groupe est constitué des N_2 patches les plus similaires autrement dit les N_2 patches $P_{\mathbf{x}}$ qui ont la plus faible mesure de distance $d^2(P_{\mathbf{x}}, P_{\mathbf{x}_R})$ définie en (3.6). La valeur du pas N_{step} est choisie inférieure à la dimension N_1 des patches afin d'avoir un chevauchement des patches. Le chevauchement des patches de référence permet d'assurer l'obtention de plusieurs estimations pour un même pixel qui pourront être combinées dans la dernière étape de l'algorithme.

Pour les images très bruitées l'algorithme de débruitage **BM3D** propose d'utiliser une mesure de distance non pas sur les patches directement extraits de l'image traitée mais sur des patches préfiltrés. Le préfiltrage consiste à réaliser une transformée bidimensionnelle orthonormale du patch fortement bruité suivie d'un seuillage des coefficients obtenus. Cependant dans notre contexte de détection d'objets ponctuels dans des fonds nuageux, les images ne sont que faiblement bruitées et donc l'application d'une mesure de distance directement sur les patches extraits de l'image traitée est pertinente. Dans notre application ce qui peut gêner l'association des patches c'est la présence d'une cible dans un des patches cependant nous verrons au paragraphe 3.3.2 qu'un choix judicieux des paramètres permet de gérer ce cas.

De même que pour le filtrage **NL-means**, il existe une version spatio-temporelle du filtre **BM3D** [[Dabov et al., 2007a](#)]. Celle-ci est basée sur une recherche de patches similaires à chacun des patches de référence $P_{\mathbf{x}_R}$, dans une zone de recherche $S_{\mathbf{x}_R}$ spatio-temporelle. Elle a l'avantage comme le filtre **NL-means** spatio-temporel de ne pas nécessiter d'estimation précise du mouvement du fond.

Débruitage des groupes

Les groupes étant formés de patches similaires, ils présentent un très fort niveau de corrélation dans la dimension que nous nommons par la suite troisième dimension qui correspond à la dimension du groupe (tableau) selon laquelle les patches ont été empilés. L'idée du filtrage **BM3D** est de tirer parti au mieux de cette corrélation. Pour cela l'algorithme fait intervenir une étape de débruitage des groupes composée elle-même de trois étapes : une étape de transformation tridimensionnelle du groupe, une étape de seuillage des coefficients dans le domaine transformé et une étape de transformation tridimensionnelle inverse. Le groupe débruité est noté $\widehat{\mathbf{P}}_{C_{\mathbf{x}_R}}$, sa dimension est la même que celle de $\mathbf{P}_{C_{\mathbf{x}_R}}$. C'est un tableau à trois dimensions qui stocke les patches $\widehat{P}_{\mathbf{x} \in C_{\mathbf{x}_R}}^{\mathbf{x}_R}$ correspondant à des estimées des patches $P_{\mathbf{x}}$ de coordonnées $\mathbf{x} \in C_{\mathbf{x}_R}$.

$$\widehat{\mathbf{P}}_{C_{\mathbf{x}_R}} = \tau_{3D}^{-1}(\Upsilon(\tau_{3D}(\mathbf{P}_{C_{\mathbf{x}_R}}))), \quad (3.8)$$

où τ_{3D} est l'opérateur de la transformée tridimensionnelle, τ_{3D}^{-1} est l'opérateur de la transformée inverse et Υ est un opérateur de seuillage dur avec une valeur de seuil qui dépend de l'écart-type supposé du bruit σ :

$$\Upsilon(\lambda) = \begin{cases} \lambda & \text{si } |\lambda| > \sigma\lambda_{3D} \\ 0 & \text{sinon} \end{cases}, \quad (3.9)$$

où λ_{3D} est un paramètre de réglage du filtre. On peut noter que le paramètre σ qui intervient dans **BM3D** joue un rôle similaire à celui du paramètre h dans le filtrage **NL-means**. Quand σ est élevé, la valeur du seuil appliqué sur les coefficients de la transformée tridimensionnelle est élevée, le filtre a tendance à fortement lisser les différents blocs et donc au final l'image. Inversement, quand le paramètre σ est faible, le filtre **BM3D** débruite très peu les patches et donc l'image.

Agrégation de l'information

Comme les patches se chevauchent et qu'un patch peut se trouver associé à différents groupes, un pixel se retrouve présent dans différents groupes voir même dans différents patches d'un même groupe. On obtient donc de nombreuses estimées de chacun des pixels de l'image et l'information doit donc être combinée judicieusement pour obtenir la meilleure estimée finale possible. L'algorithme réalise une moyenne pondérée de toutes les estimations disponibles. Il associe à chacun des groupes débruités $\widehat{\mathbf{P}}_{C_{\mathbf{x}_R}}$, une pondération qui est fonction de son caractère parcimonieux :

$$w_{\mathbf{x}_R} = \begin{cases} \frac{1}{\sigma^2 N_{\text{har}}^{\mathbf{x}_R}} & \text{si } N_{\text{har}}^{\mathbf{x}_R} > 1 \\ 1 & \text{sinon} \end{cases} \quad (3.10)$$

où $N_{\text{har}}^{\mathbf{x}_R}$ est le nombre de coefficients retenus (non-nuls) après le seuillage dur.

3.3 Adaptation du **NL-means** et **BM3D** à la détection

La principale difficulté rencontrée dans les méthodes de détection par suppression de fond (cf. principe décrit en section 1.6.1) est de parvenir à obtenir une estimation précise du fond à partir de l'image observée tout en étant robuste à la présence de cibles. Idéalement nous souhaitons que dans le fond estimé les structures et zones de fortes transitions de niveaux de gris telles que les bords de nuages soient conservées et qu'en même temps les cibles soient supprimées.

Les méthodes de débruitage à patches ont l'avantage de bien préserver les zones de hautes fréquences de l'image. Cependant, le risque est que comme les cibles sont haute fréquence, celles-ci soient préservées par ces méthodes et donc soient fortement atténuées après l'étape de suppression de fond. Dans un premier temps, en section 3.3.1, on cherche à adapter à la détection le filtrage **NL-means** puis dans la section 3.3.2 on cherche à utiliser le filtre **BM3D** pour faire de la détection.

3.3.1 Algorithme de détection DNLM

3.3.1.1 Adaptation du NL-means à la détection mono-image

Nous analysons d'abord l'impact d'une cible ponctuelle sur le résultat de débruitage par NL-means et montrons son inadaptation pour la détection d'objets ponctuels de forte intensité. Nous montrons ensuite comment corriger ce comportement.

a) Analyse de l'influence de la présence d'une cible

Le filtre NL-means repose sur une mesure de distance quadratique (3.5) qui n'est pas robuste à la présence d'un objet de forte intensité. Comme illustré sur la figure 3.7, la présence d'une cible augmente la mesure de distance (dans le cas particulier où le paramètre a du noyau gaussien G_a tend vers l'infini, la mesure de distance est augmentée de la valeur $2c^2$ avec c le niveau de la cible). Plus le niveau de l'objet ponctuel est élevé et plus l'algorithme a des difficultés à trouver des patches qu'il considère comme similaires au patch centré sur la cible.

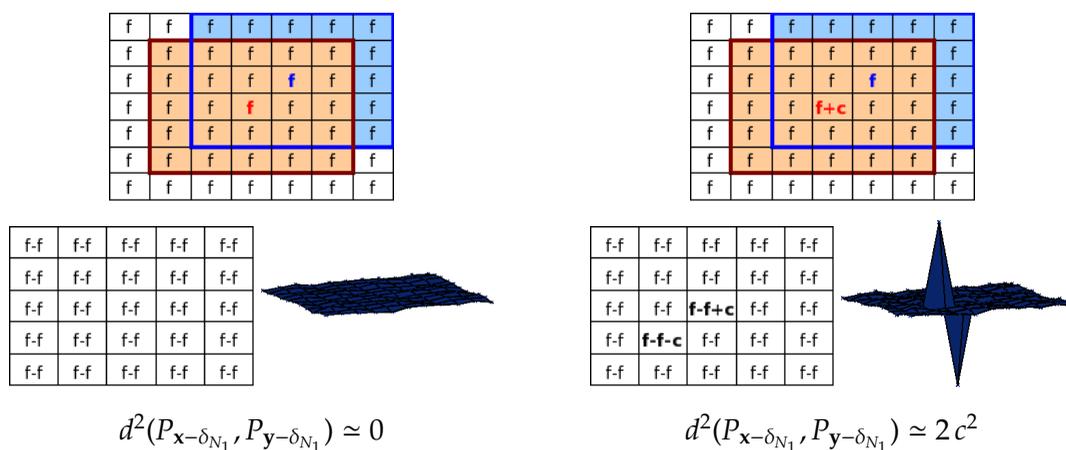


FIGURE 3.7 – Influence de la présence d'une cible sur la mesure de distance (3.5), dans le cas d'un fond homogène. À gauche, cas où aucune cible n'est présente dans le voisinage et à droite, cas où une cible pixelique est présente. De haut en bas : les deux patches pour lesquels on cherche à mesurer la distance ($N_1 = 5$); le résultat de la soustraction des patches (expression littérale et représentation tridimensionnelle); la mesure de distance approchée obtenue avec c le niveau de la cible.

Comme l'estimation du fond est perturbée par la présence d'une cible, la suppression du fond atténue le niveau de la cible. Cette atténuation peut être mesurée pour différents niveaux d'intensité de l'objet afin d'obtenir la « fonction de transfert de l'intensité » (ITF : Intensity Transfer Function) de la méthode BS. L'ITF représente le niveau maximum local d'une cible en sortie du filtre de détection en fonction du niveau du pixel central de la cible incrustée dans le fond.

L'ITF pour la méthode BS utilisant le NL-means pour estimer le fond est présentée en figure 3.8. Elle est comparée à l'ITF du filtre BS le plus simple, noté BSMean, qui consiste à simplement soustraire en chaque pixel la moyenne estimée sur un voisinage de 3×3 pixels. Les paramètres du NL-means utilisés sont les suivants : paramètre d'exigence de similarité h égal à 9, taille du voisinage Q intervenant dans la mesure de similarité

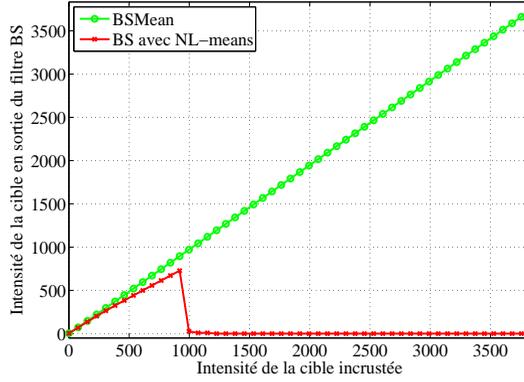


FIGURE 3.8 – Courbes ITF pour différentes méthodes de détection par BS. La cible est incrustée dans la zone de ciel relativement homogène indiquée en figure 2 et ses caractéristiques sont $r_c = 1.5$ et $(\epsilon_i, \epsilon_j) = (0, 0)$ (notations de la section 2.5).

entre deux pixels égale à 5×5 pixels ($N_1 = 5$) et taille de la zone de recherche S_x égale à 11×11 pixels ($N_s = 11$). La cible est incrustée dans la zone de ciel, relativement homogène, présentée sur la figure 2. Cette courbe présente un seuil à partir duquel l'intensité de la cible en sortie du filtre devient nulle. Ce seuil vient du fait que la mesure de distance augmente avec le niveau de la cible, et devient trop grande pour que l'algorithme puisse trouver des patches similaires au patch centré sur la cible en cours de traitement. Dans ce cas, le seul patch $P_{y-\delta_{N_1}}$ qui a un poids $w(P_{x-\delta_{N_1}}, P_{y-\delta_{N_1}})$, défini par l'équation (3.4), non nul est le patch $P_{x-\delta_{N_1}}$ lui-même et son poids est égal à 1. L'algorithme ne trouvant aucun patch similaire autre que lui-même, la cible n'est pas moyennée et apparaît donc dans le fond estimé. Dans ce cas, en sortie du filtre BS, la cible a disparu.

b) Détection par filtrage NL-means

Pour faire face à ce phénomène de seuil, nous adaptons la mesure de similarité à notre problématique de détection. En effet nous voulons qu'elle soit insensible à la présence d'une cible, c'est à dire que lors de la mesure de distance entre deux patches nous obtenions la même valeur qu'une cible soit présente ou non dans un des patches.

Mesure de distance robuste utilisant la médiane :

Un choix populaire de distance robuste est la médiane :

$$d_{Med}^2(P_{x-\delta_{N_1}}, P_{y-\delta_{N_1}}) = med\left(\left\{|G_a(\mathbf{z})|u(\mathbf{x} + \mathbf{z}) - u(\mathbf{y} + \mathbf{z})|^2\right\}_{\mathbf{z} \in Q}\right). \quad (3.11)$$

Elle est robuste à la présence de petits objets, mais elle n'évalue pas correctement le fond dans les zones de fortes variations d'intensité. En effet, le choix d'une distance en médiane revient à considérer comme similaires deux patches qui peuvent avoir jusqu'à 50% de leurs pixels qui sont différents.

Mesure proposée par J.F. Goudou :

Pour rendre l'estimation du fond robuste à la présence de cibles ponctuelles, J.F. Goudou [Goudou, 2007] propose de ne pas prendre en compte le pixel central des patches

dans la mesure de distance entre les patches. En d'autres termes, il utilise une fenêtre Q qui a un trou au centre. La nouvelle distance ainsi définie est donc :

$$d_{\text{Goudou}}^2(P_{\mathbf{x}-\delta_{N_1}}, P_{\mathbf{y}-\delta_{N_1}}) = \sum_{\mathbf{z} \in Q \setminus \{(0,0)\}} G_a(\mathbf{z}) |u(\mathbf{x} + \mathbf{z}) - u(\mathbf{y} + \mathbf{z})|^2. \quad (3.12)$$

Cette mesure de distance entre deux patches a l'avantage d'être robuste à la présence d'une cible ponctuelle au centre de l'un des deux patches. Cependant, il s'avère que le pixel central omis ici, joue un rôle important dans la mesure de distance entre les patches en l'absence de cible. La modification proposée par Goudou dégrade l'estimation du fond d'une façon non négligeable. Elle permet donc d'obtenir une méthode de détection par BS qui préserve bien les cibles (comme le montrent les ITF présentées en figure 3.10) mais elle a le défaut d'augmenter le niveau du résidu correspondant à des pixels de fond.

Mesure robuste proposée :

Lors de l'analyse précédente, nous avons vu que lorsque l'on mesure la distance entre les patches (équation (3.5)), la présence d'une cible conduit à une forte augmentation de deux des valeurs du résultat de la différence entre les patches (cf. figure 3.7). Pour rendre la mesure de distance entre patches plus robuste à la présence de cibles, nous proposons de la modifier en ignorant les deux valeurs maximales de la différence des patches $P_{\mathbf{x}-\delta_{N_1}}$ et $P_{\mathbf{y}-\delta_{N_1}}$ en valeur absolue :

$$d_1^2(P_{\mathbf{x}-\delta_{N_1}}, P_{\mathbf{y}-\delta_{N_1}}) = \sum_{\mathbf{z} \in Q \setminus \{\mathbf{z}_1, \mathbf{z}_2\}} G_a(\mathbf{z}) |u(\mathbf{x} + \mathbf{z}) - u(\mathbf{y} + \mathbf{z})|^2, \quad (3.13)$$

$$\text{avec} \quad \begin{aligned} \mathbf{z}_1 &= \arg \max_{\mathbf{z} \in Q} |u(\mathbf{x} + \mathbf{z}) - u(\mathbf{y} + \mathbf{z})|^2 \\ \mathbf{z}_2 &= \arg \max_{\mathbf{z} \in Q \setminus \{\mathbf{z}_1\}} |u(\mathbf{x} + \mathbf{z}) - u(\mathbf{y} + \mathbf{z})|^2. \end{aligned} \quad (3.14)$$

Le noyau G_a est normalisé en accord avec les pixels intervenant dans la somme.

La présence d'une cible ponctuelle idéale (c.a.d. de dimension exactement un pixel), centrée sur un pixel n'aura pas d'influence sur la mesure de distance entre les patches proposée (d_1) quelle que soit son intensité. La figure 3.9 illustre ce comportement, elle est à comparer à la figure 3.7. Pour des cibles ponctuelles réelles dont la signature dépend de la PSF échantillonnée du système et qui n'ont aucune raison particulière d'être centrées sur un pixel, la correction de la mesure de distance proposée permet de rehausser correctement des cibles plus intenses. En effet sur la figure 3.10, le seuil à partir duquel les cibles sont supprimées en même temps que le fond, est beaucoup plus élevé avec la mesure de distance proposée (éq. (3.13)) qu'avec la mesure du filtre NL-means (éq. (3.5)).

Le réglage du nombre de maxima, N_{max} , que l'on omet volontairement dans la mesure de similarité est une question de compromis : plus on en omet, mieux la cible est conservée mais l'estimation du fond se dégrade. Les ITF présentées en figure 3.10 montrent qu'un choix de N_{max} égal à 4 au lieu de 2 permet de mieux préserver les cibles présentant un décalage subpixelique et qui ont donc leur énergie qui se répartit sur plusieurs pixels (cf. section 2.5). Les courbes ROC présentées en figure 3.11, montrent que pour une cible de faible intensité ($s_o = 50$, cf. 2.5), les performances de détection diminuent quand N_{max} augmente. Cela vient du fait que l'estimation du fond se dégrade quand N_{max} augmente.

Pour les cibles ponctuelles, c.a.d de dimension inférieure ou proche de celle du pixel, ce nombre N_{max} sera usuellement choisi égal à 2. Ce nombre de maxima supprimés

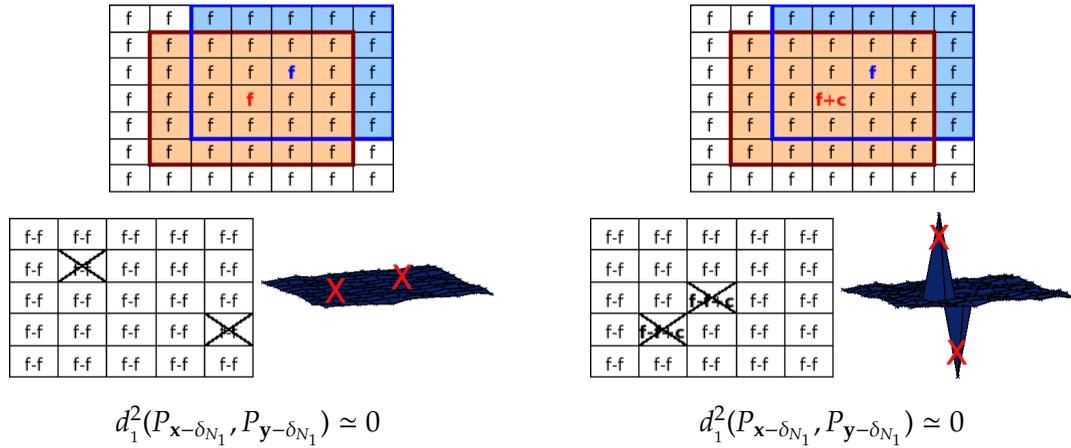


FIGURE 3.9 – Influence de la présence d’une cible sur d_1 (3.13), dans le cas d’un fond homogène. A gauche, cas où aucune cible n’est présente dans le voisinage et à droite, cas où une cible pixellique est présente. De haut en bas : les deux patchs pour lesquels on cherche à mesurer la distance ($N_1 = 5$); le résultat de la soustraction des patchs dans lequel les deux maxima sont marqués d’une croix (expression littérale et représentation tridimensionnelle); la mesure de distance approchée obtenue.

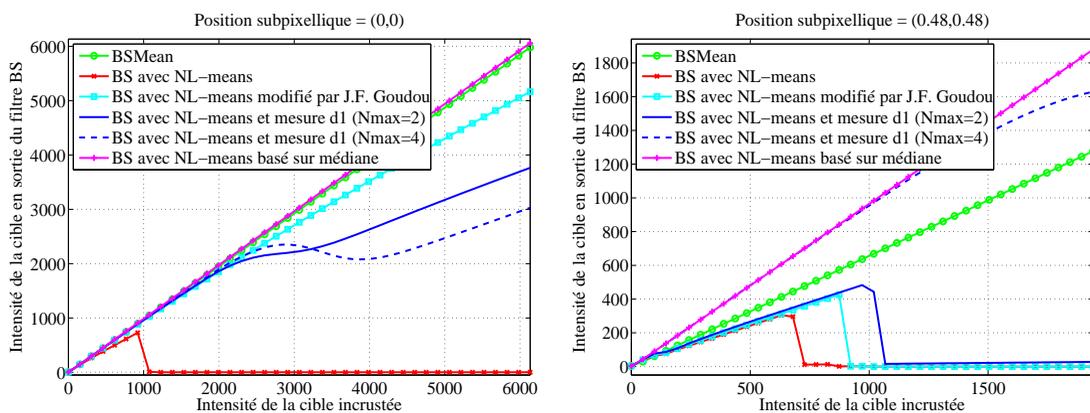


FIGURE 3.10 – Courbes ITF pour différentes méthodes de détection par BS. La cible est incrustée dans la zone de ciel relativement homogène indiquée en figure 2 et ses caractéristiques sont $r_c = 1.5$ et à gauche $(\epsilon_i, \epsilon_j) = (0, 0)$ et à droite $(\epsilon_i, \epsilon_j) = (0.48, 0.48)$ (notations de la section 2.5).

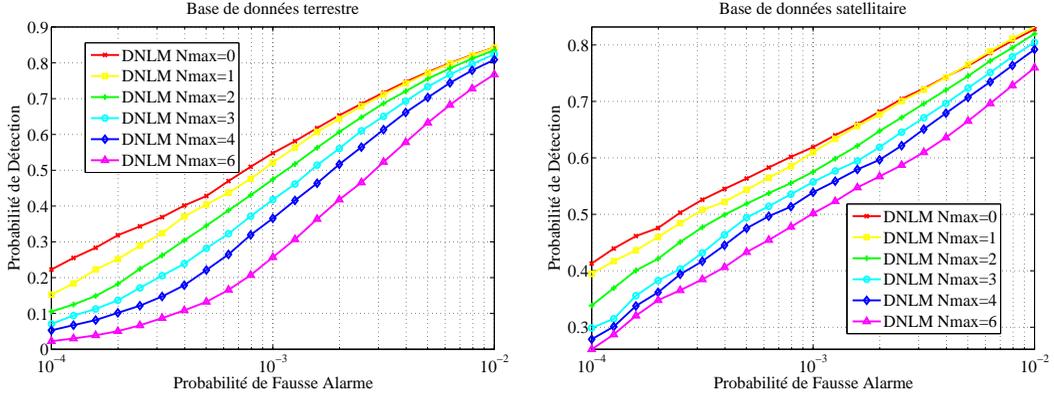


FIGURE 3.11 – Courbes ROC obtenues pour chacune des bases de données (cf. chapitre 2) pour le filtre DNLM faisant intervenir la distance d_1 (c.a.d. un paramètre $\lambda > 1$ dans la mesure d_2 définie en équation (1.28)), en fonction du nombre de maxima, $Nmax$, omis dans le calcul de d_1 . Les paramètres du filtre DNLM autres que $Nmax$ et λ sont ceux donnés par le tableau 3.1.

doit être augmenté quand la dimension spatiale de la cible augmente ; cependant le filtre perd de son efficacité quand $Nmax$ devient trop grand. En effet dans ce cas la mesure de similarité tend à être identique pour tous les pixels, les pondérations intervenant dans la moyenne deviennent toutes équivalentes, le filtre peut alors être assimilé à un simple filtre moyenneur. On en déduit que le filtre de détection développé est performant pour les cibles ponctuelles, par contre il n'est pas pertinent pour les objets étendus.

Mesure de distance adaptative :

Bien qu'elle soit très efficace pour préserver les objets ponctuels de forte intensité, la distance d_1 dégrade la sélection des pixels et l'estimation du fond en l'absence d'objet. Cela a pour conséquence de faire chuter les performances de détection de cibles de faible intensité ($s_o = 50$, cf. section 2.5) comme le montre la figure 3.11. Nous proposons donc de combiner les deux mesures d et d_1 et d'utiliser d_1 uniquement quand la présence d'un objet est suspectée. Le choix est guidé par le rapport entre les valeurs du 1^{er} et du 3^{ème} maxima de la différence des patches :

$$d_2^2(P_{\mathbf{x}-\delta_{N_1}}, P_{\mathbf{y}-\delta_{N_1}}) = \begin{cases} d_1^2(P_{\mathbf{x}-\delta_{N_1}}, P_{\mathbf{y}-\delta_{N_1}}) & \text{if } \frac{u(\mathbf{z}_3)}{u(\mathbf{z}_1)} < \lambda \\ d^2(P_{\mathbf{x}-\delta_{N_1}}, P_{\mathbf{y}-\delta_{N_1}}) & \text{otherwise} \end{cases}, \quad (3.15)$$

$$\text{avec } \mathbf{z}_3 = \arg \max_{\mathbf{z} \in Q \setminus \{\mathbf{z}_1, \mathbf{z}_2\}} |u(\mathbf{x} + \mathbf{z}) - u(\mathbf{y} + \mathbf{z})|^2 \quad (3.16)$$

(dans le cas général où $Nmax$ maxima sont omis dans la mesure de distance d_1 , le choix est guidé par le rapport entre le 1^{er} et le $Nmax + 1$ ^{ème} maximum). Le seuil λ est compris entre 0 et 1 et règle le choix entre d et d_1 . Le choix de λ correspond à un compromis entre les fausses alarmes qui augmentent avec λ et les non-détections qui apparaissent quand λ diminue. Par expérience, $\lambda = 0.08$ semble être un bon compromis. Cette nouvelle distance d_2 est celle que nous faisons intervenir dans le filtre de détection DNLM (*Detection by NLM*) proposé. Le filtre ainsi obtenu dans le cadre spatial fournit un bon compromis, il est à la fois performant pour les cibles ponctuelles de faible intensité et de forte intensité.

Toutes les méthodes modifiées sont plus robustes que la méthode **BS** utilisant directement le filtre **NL-means**. Les modifications apportées arrivent à augmenter l'intensité limite au delà de laquelle les cibles ne sont pas rehaussées. Ces filtres seuls ne pourront pas gérer les cibles de très fortes intensités qui pourront être détectées en parallèle en faisant intervenir un seuil de détection très élevé.

La figure 3.12 présente les courbes **ROC** obtenues pour différentes méthodes de détection basées sur le **NL-means** qui viennent d'être présentées. Les courbes ont été obtenues avec les paramètres de filtrage donnés dans le tableau 3.1. On peut noter qu'une attention particulière a été portée au paramètre h qui a été optimisé pour obtenir les meilleures performances en termes de courbe **ROC** pour chacune des méthodes et pour chacune des bases de données. Le filtre **DNLM** donne des résultats nettement supérieurs à la mesure de distance modifiée proposée par J.F. Goudou. Concernant le filtre utilisant la médiane, il a un comportement inhomogène entre les deux bases de données et nous n'avons pas creusé cette voie là. Une étude paramétrique plus complète du filtre spatial **DNLM** est présentée en annexe B. Dans la suite nous présentons une extension spatio-temporelle de ce filtre.

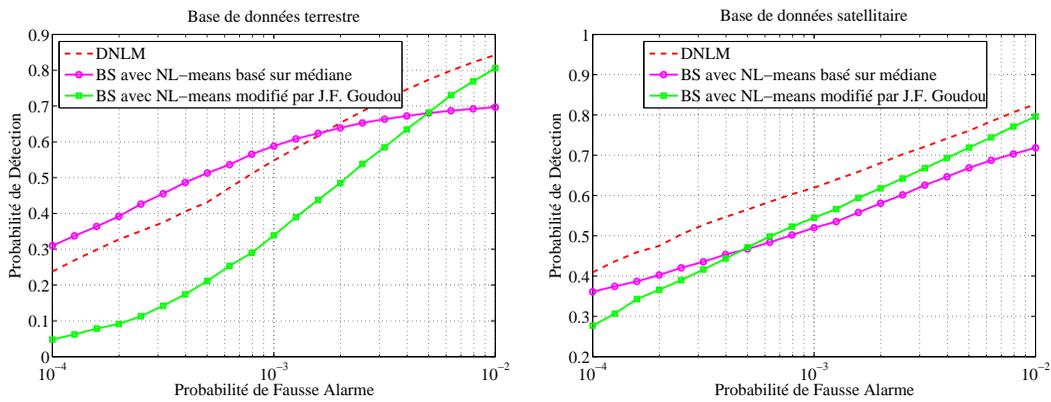


FIGURE 3.12 – Courbes **ROC** obtenues pour chacune des bases de données (cf. chapitre 2) pour les différentes méthodes de détection basées sur **NL-means** qui ont été présentées dans cette partie. Les paramètres utilisés sont ceux donnés par le tableau 3.1.

Paramètres	DNLM	BS avec NL-means basé sur médiane (éq. (3.11))	BS avec NL-means de Goudou (éq. (3.12))
N_1		5	
a		1.5	
N_s		11	
h	Base de données terrestre	9	8
	Base de données satellitaire	6	6
N_{max}	2		-
λ	0.08		-

TABLE 3.1 – Tableau présentant les valeurs des paramètres nominaux utilisées pour les différents filtres de détection basés sur le **NL-means** considérés dans cette section. Dans tout le chapitre ce sont ces valeurs qui sont utilisées si aucune précision particulière n'est indiquée.

3.3.1.2 Adaptation au contexte spatio-temporel

Il est possible de tirer parti de la corrélation temporelle du fond. Pour cela nous étendons notre approche au cas spatio-temporel. Dans le cas multi-images, les modifications de la mesure de distance que nous venons de décrire ne seront pas suffisantes pour rendre le filtre NL-means adapté à la détection. Une deuxième modification est donc apportée, justifiée par une brève étude de l'influence de la présence de cibles.

a) Analyse de l'influence de la présence d'une cible

Dans le cas d'une cible parfaitement pixellique centrée sur un pixel, c'est à dire un unique pixel d'intensité non nulle, le filtre NL-means utilisant la nouvelle mesure de distance d_2 définie ci-dessus permet d'obtenir de très bonnes performances de détection en spatio-temporel. Cependant en réalité les cibles ne sont jamais parfaitement pixelliques (les cibles naturelles forment des tâches) et elles ne sont jamais parfaitement centrées sur un pixel (leur position est subpixellique) [Samson *et al.*, 2004] (cf. section 2.5). En présence de ces phénomènes, les performances de détection du filtre chutent brutalement comme le montrent les courbes ITF présentées en figure 3.13.

Nous rappelons que pour le filtre NL-means, le poids du pixel y intervenant dans l'estimation du pixel x se calcule de la façon suivante :

$$\forall y \in S_x \setminus \{x\} \quad w(P_{x-\delta_{N_1}}, P_{y-\delta_{N_1}}) = e^{-\frac{d(P_{x-\delta_{N_1}}, P_{y-\delta_{N_1}})}{h^2}}$$

$$\text{et } w(P_{x-\delta_{N_1}}, P_{x-\delta_{N_1}}) = \max_{y \in S_x \setminus \{x\}} w(P_{x-\delta_{N_1}}, P_{y-\delta_{N_1}}) \quad (3.17)$$

avec S_x le support dans lequel on recherche des pixels similaires ; il est ici spatio-temporel (cf. figure 3.5).

Les modifications de la mesure de distance entre les patches présentées ci-dessus permettent pour des cibles de relativement faible intensité, d'augmenter le niveau de la cible en sortie du filtre. Cependant pour les cibles de forte intensité le niveau de celles-ci en sortie de filtre chute à partir d'un certain seuil (cf. les ITF présentées en figure 3.13). Cela vient du fait que lorsque le niveau de la cible augmente le filtre a tendance, pour estimer le pixel, à sélectionner uniquement les pixels qui correspondent à la cible présente dans les autres images de la séquence. Dans ce cas la cible apparaît dans le fond estimé et le filtre ne permet donc pas de conserver correctement la cible, celle-ci est supprimée en même temps que le fond, son intensité chute. La figure 3.14 illustre ce phénomène en présentant les cartes des pondérations des pixels du support S_x .

Même si une cible peut avoir son intensité principalement concentrée en un pixel, elle a une certaine influence sur les autres pixels qui l'entourent. Les différents pixels correspondant à une même cible (présente dans les différentes images de la séquence) vont donc avoir des voisinages assez similaires. Quand le niveau de la cible est élevé, le filtre malgré la modification de la mesure de distance va, pour estimer le fond, faire une moyenne uniquement des pixels correspondant à la cible. Le filtre n'est donc pas correctement adapté à la détection dans un cadre spatio-temporel.

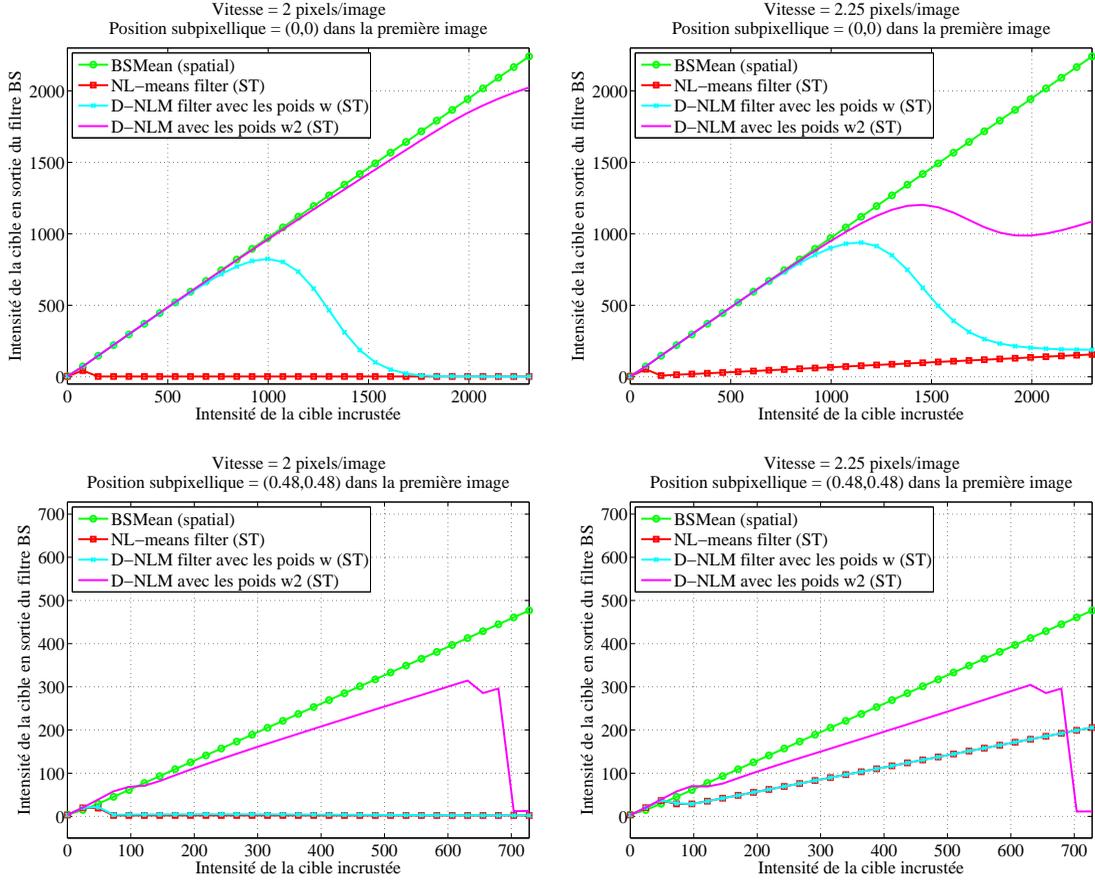


FIGURE 3.13 – Courbes ITF pour différentes méthodes de détection par BS. La cible est incrustée dans la zone de ciel relativement homogène indiquée en figure 2 et a pour paramètre : $r_c = 1.5$. La vitesse de la cible est de 2 pixels/image à gauche et de 2.25 pixels/image à droite. La position subpixelique de la cible à la position du pixel étudié est $\epsilon = (0, 0)$ en haut et $\epsilon = (0.48, 0.48)$ en bas. Les filtres étudiés sont : pour la courbe verte, le BSMean spatial qui estime le fond sur une fenêtre de 3×3 pixels ; pour la courbe rouge, le filtre NL-means spatio-temporel (ST) développé par Buades *et al.* [Buades *et al.*, 2005a] ; pour la courbe cyan, le même filtre mais utilisant la nouvelle mesure de distance d_2 (éq. (3.15) avec $N_{max} = 2$) ; pour la courbe magenta, la nouvelle méthode d'estimation de fond proposée dans le cas spatio-temporel (qui utilise la nouvelle mesure de distance d_2 (éq. (3.15) avec $N_{max} = 2$) et les nouveaux poids w_2 (éq. (3.19)).

b) Détection spatio-temporelle par filtrage NL-means

Pour obtenir un filtre performant, la méthode de calcul des poids doit elle aussi être modifiée.

Pour cela le support S_x dans lequel on recherche les pixels similaires est divisé en sous-soutports. S_x^i représente l'intersection entre le support S_x et la $i^{\text{ème}}$ image.

$$S_x = \bigcup_{i=1}^L S_x^i \quad (3.18)$$

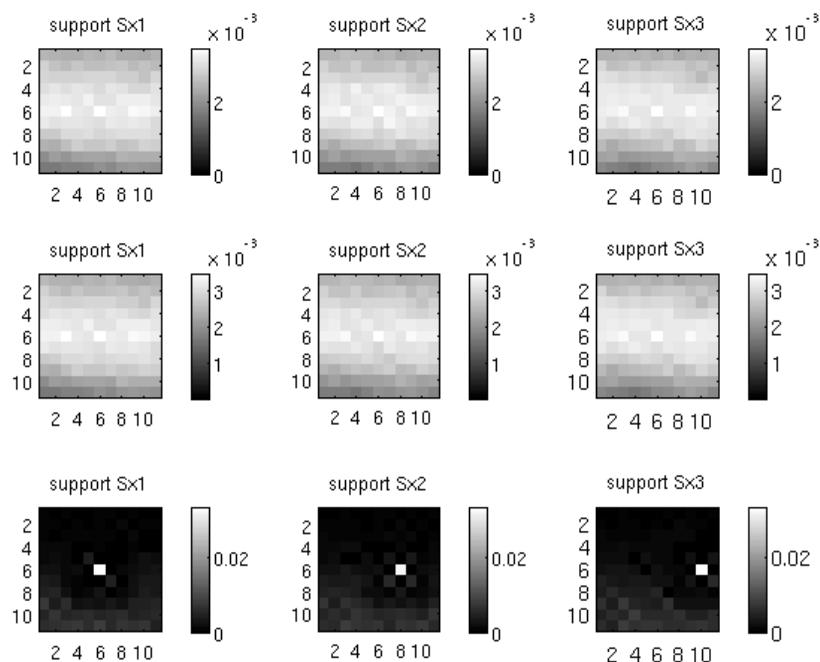


FIGURE 3.14 – Valeurs des poids des pixels du support S_x , intervenant dans l'estimation du pixel du centre de la zone de ciel de la figure 2, par le filtre NL-means spatio-temporel utilisant la nouvelle mesure de distance d_2 (éq. (3.15)). En haut : cas où aucune cible n'est présente ; au milieu : cas d'une cible purement pixellique, de vitesse 2 pixels/image et de niveau 136. en bas : cas d'une cible étalée ($f_c/f_e = 0.8$) centrée sur le pixel ($\epsilon = (0, 0)$), de vitesse 2 pixels/image et d'énergie totale égale à 350 (soit un pixel central de niveau 136).

avec L le nombre d'images de la séquence intervenant dans l'algorithme de détection. La figure 3.5 permet de visualiser ces supports dans un cas spatio-temporel de trois images.

Les poids associés aux pixels \mathbf{y} intervenant dans l'estimation du pixel \mathbf{x} dans le cadre du nouvel algorithme de détection sont calculés de la façon suivante :

$$w_2(P_{\mathbf{x}-\delta}, P_{\mathbf{x}-\delta}) = \max_{\mathbf{y} \in S_x^1 \setminus \{\mathbf{x}\}} e^{-d_2^2(P_{\mathbf{x}-\delta}, P_{\mathbf{y}-\delta})/h^2} \quad (3.19a)$$

$$w_2(P_{\mathbf{x}-\delta}, P_{\mathbf{y}_i-\delta}) = \max_{\mathbf{y} \in S_x^i \setminus \{\mathbf{y}_i\}} e^{-d_2^2(P_{\mathbf{x}-\delta}, P_{\mathbf{y}-\delta})/h^2} \quad \text{pour } i = 1..L, \quad (3.19b)$$

$$\text{avec } \mathbf{y}_i = \arg \max_{\mathbf{y} \in S_x^i} e^{-d_2^2(P_{\mathbf{x}-\delta}, P_{\mathbf{y}-\delta})/h^2}. \quad (3.19c)$$

Le poids associé à \mathbf{x} est remplacé par le second max (éq. (3.19a)) se trouvant dans S_x^1 et pour chacune des autres images de la séquence, le poids maximum est remplacé par le second maximum (éq. (3.19b) et (3.19c)).

La modification apportée écriète la valeur de la pondération maximale de chacun des sous-supports S_x^i . Lors du traitement d'un pixel correspondant à une cible, ces pondérations maximales sont le plus souvent égales aux pondérations associées aux pixels correspondant à la cible elle-même qui est présente dans chaque image du support. L'écriète

tement de ces pondérations permet de corriger le biais introduit par la présence répétitive de la cible dans le support S_x .

Les cartes des poids en l'absence et présence de cibles de la figure 3.15 sont beaucoup plus semblables qu'ils ne l'étaient sur la figure 3.14, les nouvelles pondérations w_2 permettent de rendre l'estimation du fond dans le cadre spatio-temporel moins sensible à la présence d'une cible.

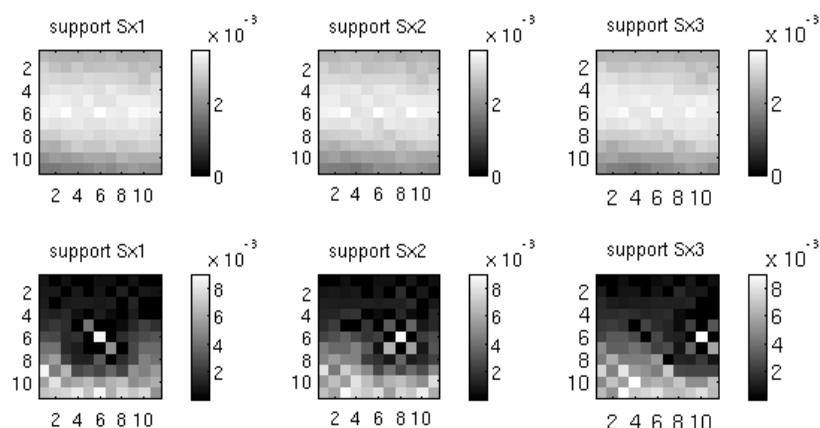


FIGURE 3.15 – Valeurs des poids des pixels du support S_x , intervenant dans l'estimation du pixel du centre de la zone de ciel de la figure 2, par la nouvelle méthode proposée dans le cas spatio-temporel utilisant la nouvelle mesure de distance d_2 (éq. (3.15)) et les nouveaux poids w_2 (éq. (3.19)). En haut : cas où aucune cible n'est présente ; en bas : cas d'une cible étalée ($f_c/f_e = 0.8$) centrée sur le pixel ($\epsilon = (0, 0)$), de vitesse 2 pixels/images et d'énergie totale égale à 350 (soit un pixel central de niveau 136).

La figure 3.16 montre que pour une cible de relativement forte intensité, l'utilisation des poids modifiés w_2 permet à l'algorithme d'être plus robuste à la présence de cibles de faible vitesse. En effet la chute de performance observée quand on passe d'une vitesse de cible de 1.25 pixels/image à une vitesse de 0.25 pixels/image est beaucoup plus importante pour l'algorithme DNLM spatio-temporel faisant intervenir les poids w (DNLM w) que celui faisant intervenir les poids w_2 (DNLM w_2). Cependant pour des cibles de plus faible intensité, la figure 3.17 montre que pour le cas nominal d'une vitesse de cible de 1.25 pixels par images, l'algorithme de détection DNLM w donne des performances de détection qui sont très faibles comparées à celles obtenues avec DNLM w_2 . On peut voir de plus sur cette figure que quand la vitesse de la cible est plus faible (0.25 pixels/image) même si la chute de performance du filtre DNLM w_2 est beaucoup plus faible que celle du filtre DNLM w , le filtre DNLM w_2 n'est pas pour autant plus performant que le filtre DNLM w pour détecter des cibles lentes.

Le contexte qui nous intéresse dans cette thèse étant celui des cibles de faibles vitesses, le filtre DNLM spatio-temporel qui est donc considéré dans toute la suite de ce manuscrit est celui qui fait intervenir les poids non-modifiés w (équation (3.17)) et la mesure de distance entre les pixels d_2 . Le nouveau filtre de détection spatio-temporelle ainsi obtenu garde l'avantage du filtre NL-means initial de ne pas nécessiter d'estimation précise du mouvement. Ses performances sont étudiées dans le chapitre 5.

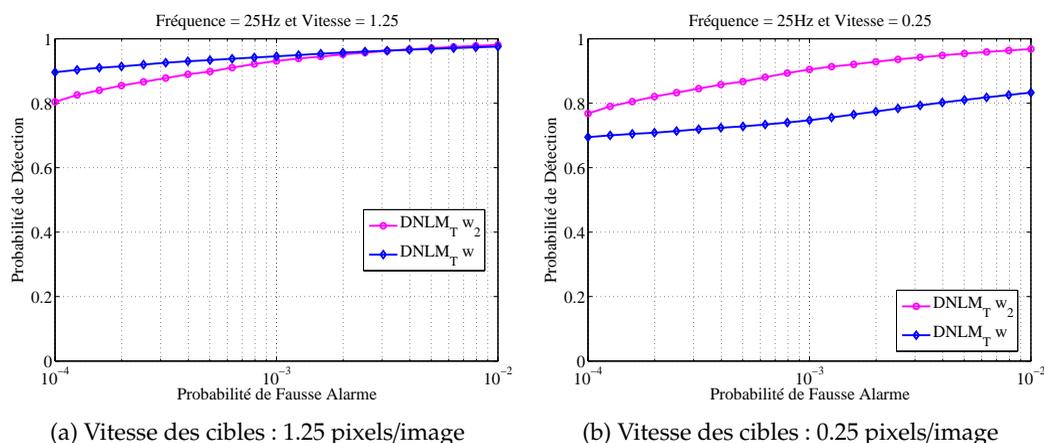


FIGURE 3.16 – Courbes ROC obtenues avec l’algorithme DNLM spatio-temporel faisant intervenir 2 images qui utilise les poids w et avec celui qui utilise les poids w_2 , pour une cible de paramètre $s_0 = 100$ (cf. section 2.5), pour différentes vitesses de cible (le paramètre h a été optimisé pour fournir les meilleures performances : $h = 9$)

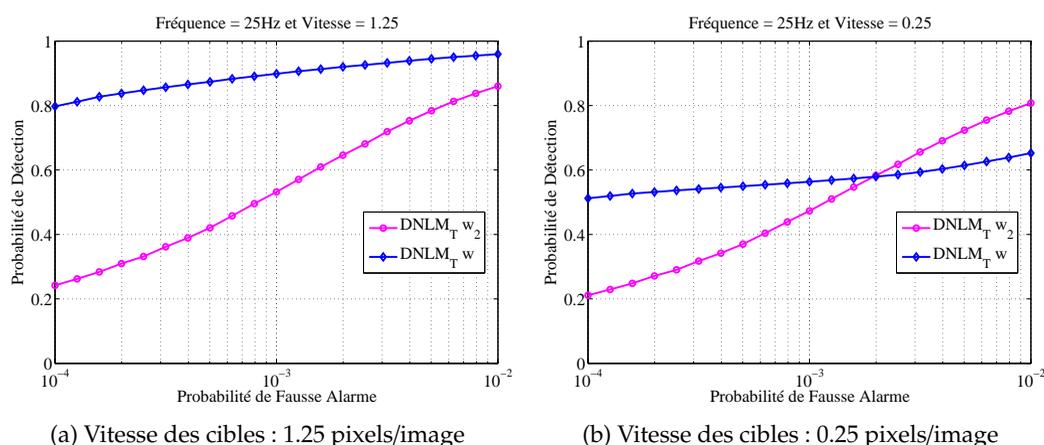


FIGURE 3.17 – Courbes ROC obtenues avec l’algorithme DNLM spatio-temporel faisant intervenir 2 images qui utilise les poids w et avec celui qui utilise les poids w_2 , pour une cible de paramètre $s_0 = 50$ (cf. section 2.5), pour différentes vitesses de cible (le paramètre h a été optimisé pour fournir les meilleures performances : $h = 9$).

Goudou de son côté a aussi proposé une extension spatio-temporelle de son approche de détection [Goudou, 2007]. Il propose d’utiliser des patches spatio-temporels (3D). Pour être robuste à la présence de cibles, il fait la forte hypothèse que les cibles se déplacent dans un plan horizontal et prend donc en compte dans sa mesure de distance uniquement les pixels du patch 3D qui ont un site strictement différent du pixel traité. Cette approche a le désavantage d’être basée sur une hypothèse forte en ce qui concerne la dynamique de la cible. De plus en l’absence de cible, cette approche a le désavantage qu’avait déjà l’approche spatiale de Goudou de dégrader l’estimation du fond en omettant les pixels se trouvant proches du centre des patches.

3.3.2 Algorithme de détection **DBM3D**

3.3.2.1 Adaptation de la méthode de débruitage **BM3D**

Préservant bien les zones de fortes transitions, le filtrage **BM3D** nous permet d'obtenir une très bonne estimation du fond en l'absence de cibles. Cependant pour obtenir de bonnes performances de détection, il faut également s'assurer que **BM3D** est capable de supprimer la cible afin qu'après l'étape de suppression le niveau de celle-ci ne soit pas fortement atténué. Pour que la cible ait bien disparu dans le fond estimé, il est essentiel que l'algorithme puisse trouver des patchs qui sont similaires à des patchs de référence contenant la cible, pour que l'on puisse avoir d'autres pixels que ceux correspondant à la cible qui interviennent dans l'estimation du pixel où se trouve la cible. Si le niveau de la cible est faible, l'algorithme va normalement arriver à trouver des patchs qui sont considérés comme similaires au patch contenant la cible, c.a.d trouver des patchs P_x pour lesquels la distance $d^2(P_x, P_{x_R})$ est inférieure au seuil τ_{match} . Mais pour un certain niveau de cible, l'algorithme ne va plus pouvoir trouver de patchs similaires au patch contenant la cible.

Pour faire face à ce problème, la première approche envisagée a consisté à remplacer la mesure (3.6) par une mesure robuste comme la mesure d_2 (éq. (3.15)) qui a été proposée pour adapter le filtre **NL-means** à la détection. L'intervention de cette nouvelle mesure de distance dans le filtrage **BM3D** a un désavantage. La mesure de distance étant robuste à la présence d'une cible et les patchs se recouvrant, le pixel correspondant à la cible risque de se retrouver plusieurs fois présent dans un même groupe $P_{C_{x_R}}$. La présence répétée du pixel cible dans le groupe dégrade son débruitage, c'est à dire qu'une partie de l'intensité de la cible reste présente dans le groupe débruité $\widehat{P}_{C_{x_R}}$. De plus, la modification de la mesure augmente la complexité algorithmique du filtre.

En fait, l'algorithme peut être rendu robuste aux cibles par un réglage spécifique de ses paramètres. Nous avons observé que pour des paramètres de réglage permettant d'obtenir une bonne estimation de fond nuageux en l'absence de cible, la dimension des patchs était le plus souvent limitée par le paramètre N_2 (nombre de patchs maximum par groupe) et non par le paramètre τ_{match} (valeur de la distance au-delà de laquelle on considère les patchs comme dissimilaires). L'adaptation du **BM3D** à la détection par **BS** que nous proposons consiste simplement à choisir une valeur de τ_{match} très élevée de sorte que l'on trouve des patchs similaires malgré la présence d'une cible relativement intense. Dans ce cas seul le paramètre N_2 est déterminant et l'étape d'association des patchs consiste à assembler dans un bloc les N_2 patchs les plus similaires au patch de référence. Si une cible est présente dans le patch de référence, le nombre de patchs ne comprenant pas la cible dans le groupe est généralement suffisant pour permettre à l'étape de débruitage d'annuler l'influence de la cible. Ce comportement est favorisé par un réglage adapté du paramètre de lissage σ selon le niveau d'intensité des cibles attendues et le SNR. En l'absence de cible, l'estimation du fond n'est que peu dégradée car le nombre de patchs par groupe étant limité à N_2 , peu de patchs parasites viennent perturber l'estimation. Avec ce réglage adapté, on obtient donc un très bon compromis entre estimation du fond et suppression de la cible. Ce nouveau filtre de détection d'objets ponctuels par suppression de fond estimant le fond au moyen du filtre **BM3D** est noté **DBM3D** (pour *Detection by BM3D*).

La figure 3.18 qui présente les courbes ROC obtenues avec les filtres de détection spatiaux DNLM et DBM3D, montre que la méthode de détection DBM3D s'avère beaucoup plus performante aussi bien pour la base de données terrestre que satellitaire.

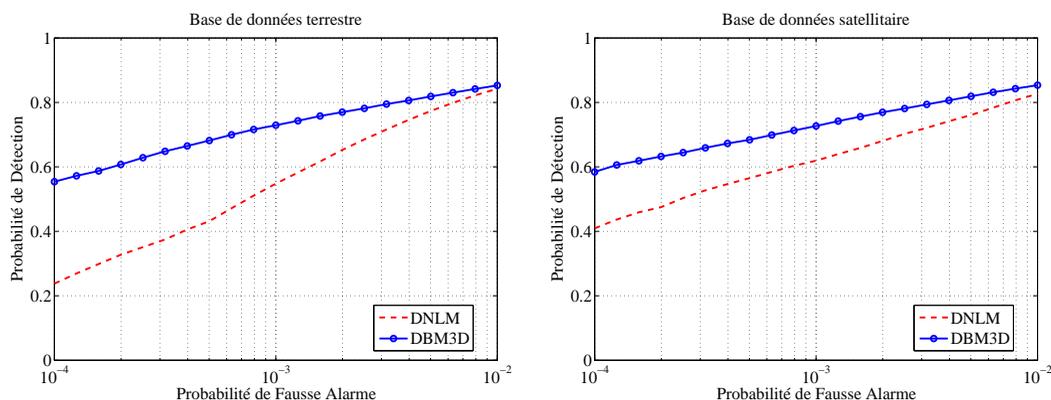


FIGURE 3.18 – Courbes ROC obtenues pour chacune des bases de données pour le filtre DNLM et pour le filtre DBM3D faisant intervenir les paramètres donnés dans les tableaux 3.1 et 3.2.

3.3.2.2 Étude paramétrique

Pour rappel, les paramètres du filtre DBM3D sont les suivants (ce sont exactement les mêmes que ceux intervenant dans BM3D [Dabov *et al.*, 2006]) :

- σ : paramètre de lissage intervenant principalement dans la définition du seuil appliqué sur les coefficients des groupes transformés (éq. (3.9)),
- N_1 : dimension du côté des patches,
- τ_{match} : seuil utilisé pour déterminer si deux patches sont similaires ou non (éq. (3.7)),
- N_2 : nombre de patches maximum par groupe,
- N_s : dimension du côté de la fenêtre carré S_x dans laquelle on recherche des patches similaires à P_x ,
- N_{step} : pas entre les pixels x_R qui définissent les patches de référence,
- λ_{2D} : paramètre qui définit le seuil appliqué sur les coefficients des groupes transformés (éq. (3.9)),
- T_{3D} : transformée tridimensionnelle appliquée sur chacun des groupes (éq. 3.8) ; elle se décompose en une transformée bidimensionnelle (T_{2D}) qui est appliquée sur chacun des blocs et une transformée unidimensionnelle (T_{1D}) sur le groupe contenant les blocs transformés,
- β : paramètre de la fenêtre de Kaiser qui est utilisée pour pondérer les pixels intervenant dans l'étape d'agrégation de l'information pour réduire les effets de bord qui sont liés aux transformations bidimensionnelles.

En accord avec la discussion faite en section 3.3.2.1, le paramètre τ_{match} est augmenté et fixé à 10^{10} . Tous les autres paramètres sont identiques à ceux proposés dans l'article original de Dabov *et al.* concernant le filtre de débruitage BM3D [Dabov *et al.*, 2007b], excepté le paramètre σ qui est optimisé pour obtenir les meilleures performances de détection pour chacune des bases de données : il est fixé égal à 6 (respectivement 5) pour la base de données terrestre (respectivement satellitaire). Les paramètres nominaux considérés pour le filtre DBM3D sont résumés dans le tableau 3.2. Dans cette section on

analyse l'influence de la variation de certains de ces paramètres sur les performances de détection.

Paramètres	DBM3D	
N_1	8	
τ_{match}	10^{10}	
N_2	16	
N_s	39	
N_{step}	3	
σ	Base de données terrestre	6
	Base de données satellitaire	5
T_{2D}	2D-Bior1.5	
T_{1D}	1D-Haar	
λ_{3D}	2.7	
β	2	

TABLE 3.2 – Tableau présentant les valeurs des paramètres nominaux utilisées pour le filtre de détection **DBM3D**. Dans tout le chapitre ce sont ces valeurs qui sont utilisées si aucune précision particulière n'est indiquée.

Influence du paramètre σ

La figure 3.19 présente la variation des performances de détection du **DBM3D** pour différentes valeurs de σ . Ce paramètre a un rôle qui est similaire à celui du paramètre h dans **DNLM**. Quand σ est élevé, les cibles sont très bien rehaussées dans l'image de résidu mais le fond est moins bien supprimé ce qui augmente le nombre de fausses alarmes. Inversement, quand le paramètre σ est faible le fond est parfaitement supprimé mais les cibles le sont en partie aussi et le nombre de non-détections augmente. Le paramètre σ est le paramètre du filtre **DBM3D** le plus délicat à régler, il dépend du niveau d'intensité des cibles attendues et du SNR. La figure 3.19 montre que pour la base de données terrestre (respectivement satellitaire) une valeur de σ égale à 6 (respectivement 5) fournit les meilleures performances de détection.

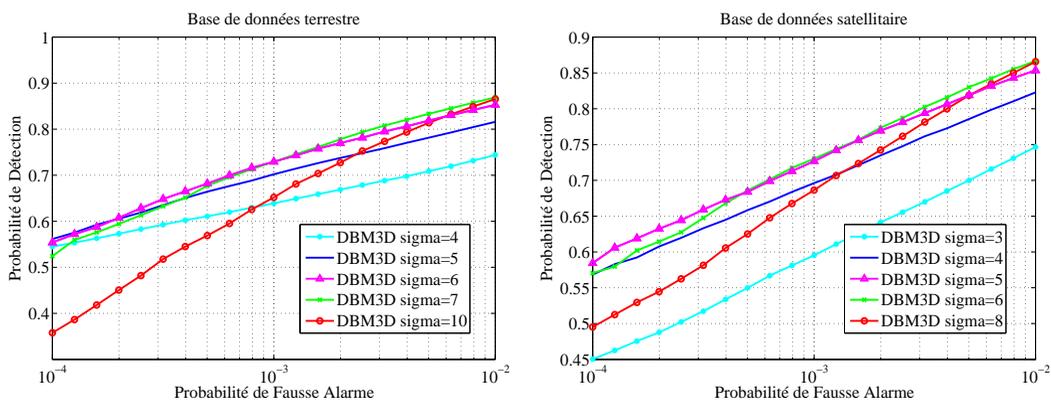


FIGURE 3.19 – Influence du paramètre σ : Courbes **ROC** obtenues avec **DBM3D**, pour chacune des bases de données, pour différentes valeurs de σ .

Influence du paramètre N_1

La figure 3.20 présente les performances de détection obtenues pour différentes valeurs de N_1 , c.a.d. pour différentes dimensions de patches. La dimension des patches qui s'avère le mieux adapté à l'échelle des textures présentes dans les deux bases de données est 8×8 pixels.

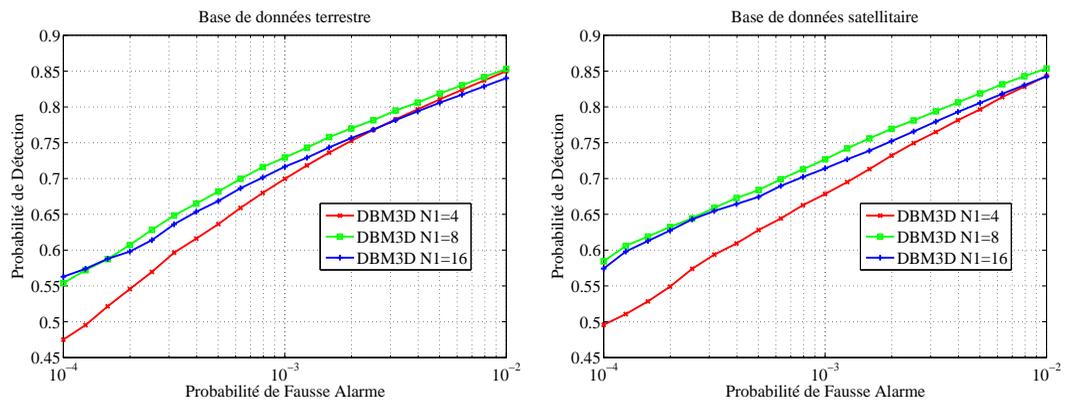


FIGURE 3.20 – Influence du paramètre N_1 : Courbes ROC obtenues avec DBM3D, pour chacune des bases de données, pour différentes valeurs de N_1 .

Influence du paramètre N_2

La figure 3.21 présente les performances de détection obtenues avec DNLM pour différentes valeurs de N_2 . Si la valeur de N_2 est trop petite on risque de ne pas disposer de suffisamment de patches dans les groupes pour correctement estimer le fond malgré la présence d'une cible, inversement si le nombre N_2 est trop grand on risque de rassembler dans un même groupe des patches qui sont dissimilaires et ainsi dégrader l'estimation du fond. En pratique, on constate que seul le choix d'une valeur trop faible ($N_2 = 2$ ou 4) dégrade sensiblement la performance. A partir de $N_2 = 8$ les performances sont très proches. Dans la suite, on choisit la valeur $N_2 = 16$.

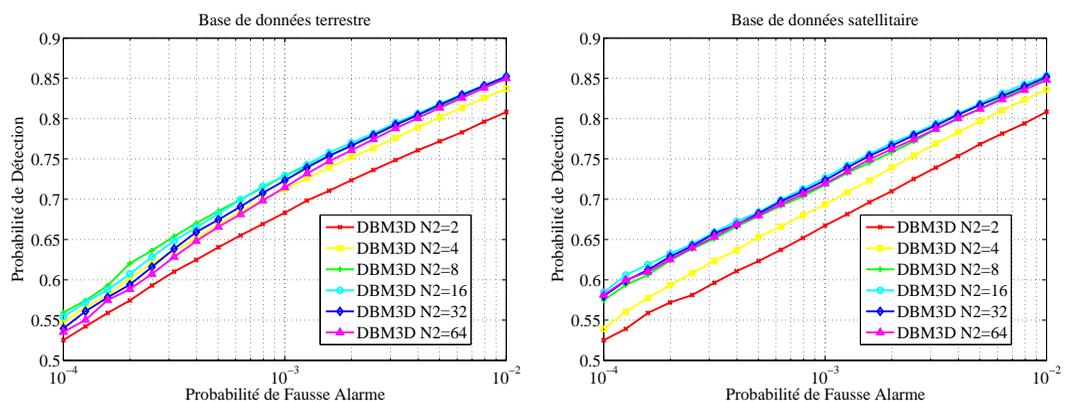


FIGURE 3.21 – Influence du paramètre N_2 : Courbes ROC obtenues avec DBM3D, pour chacune des bases de données, pour différentes valeurs de N_2 .

Influence du paramètre N_s

La figure 3.22 présente les performances de détection obtenues avec DNLM pour différentes valeurs de N_s . On observe que l'on obtient des performances de détection quasiment identiques avec toutes les valeurs de N_s testées. Nous avons choisi comme valeur nominale de N_s celle utilisée par Dabov *et al.* [Dabov *et al.*, 2007b] : 39. Cependant cette étude paramétrique montre qu'en choisissant une valeur inférieure on aurait pu réduire les coûts de calcul sans dégrader les performances de détection.

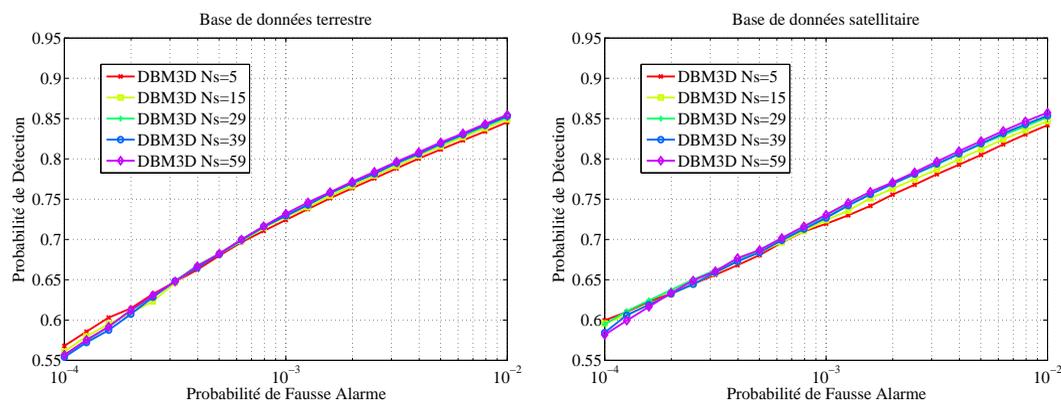


FIGURE 3.22 – Influence du paramètre N_s : Courbes ROC obtenues avec DBM3D, pour chacune des bases de données, pour différentes valeurs de N_s .

3.3.2.3 Cadre spatio-temporel

Comme une version spatio-temporelle de l'algorithme de débruitage BM3D existe [Dabov *et al.*, 2007a], il est tout à fait naturel de proposer une version spatio-temporelle de l'algorithme DBM3D. Le filtre DBM3D spatio-temporel va chercher des patches similaires aux patches de référence non plus uniquement dans l'image en cours de traitement mais aussi dans les autres images de la séquence. Quand la corrélation temporelle de la séquence est forte, pour chaque patch de l'image en cours de traitement il existe un patch très similaire dans chacune des images précédentes. Cependant comme on se fixe une valeur de τ_{match} élevée pour être robuste à la présence de cibles, un groupe est constitué de ces patches très similaires au patch de référence mais aussi d'un nombre important de patches moins similaires ($N_2 - L$, où L est le nombre d'images considérées) qui ont pour conséquence de diminuer l'impact des patches très similaires. De plus quand une cible est présente dans le patch de référence, les patches très similaires associés à la cible présente dans les images précédentes vont se trouver présents dans le groupe. La présence répétée de la cible dans un groupe risque de dégrader l'estimation du fond et donc de faire chuter l'intensité de la cible après l'étape de suppression de fond. Selon nos tests une telle extension spatio-temporelle de l'algorithme DBM3D ne s'avère pas beaucoup plus performante que la version spatiale de celui-ci. La version spatio-temporelle du filtre DNLM (cf. section 3.3.1.2) est plus performante.

Conclusion

Dans le cadre du schéma de détection en deux étapes considéré pour répondre à notre problématique de détection (cf section 1.5, figure 1.4, page 24), nous avons proposé dans ce chapitre de nouvelles méthodes de détection basées sur la modélisation du fond au premier ordre. Nous avons adapté des approches à patches issues du domaine du débruitage qui ont l'avantage de bien préserver les zones de bord et de modéliser les fonds inhomogènes de façon très efficace. Ces approches de débruitage à patches se basent sur une recherche puis une agrégation de patches similaires. Pour pouvoir faire de la détection avec de telles approches, nous avons cherché des moyens permettant de trouver afin de les agréger les patches de fonds similaires en dépit de la présence d'une cible.

Nous avons tout d'abord proposé une nouvelle adaptation du filtre de débruitage NL-means (*Non-Local Means*) introduit par Buades *et al.* en 2005 [Buades *et al.*, 2005b] qui a conduit au filtre de détection DNLM (*Detection by NLM*). Il nous apparaît aujourd'hui qu'une mesure de distance robuste basée sur la norme L_1 pourrait s'avérer être plus pertinente que la mesure de distance d_2 . On obtient ainsi un algorithme moins coûteux (car il ne nécessite pas d'étape de recherche de maxima) pour lequel les résultats préliminaires montrent que l'on obtient des performances de détection identiques à celles obtenues avec d_2 . Cependant, comme l'utilisation du filtre à patches BM3D (*Block Matching and 3D Filtering*) [Dabov *et al.*, 2007b] considéré actuellement comme un des filtres de débruitage les plus performants s'est avéré conduire à une approche de détection nettement plus performante, nous n'avons pas poursuivi dans cette voie. Nous avons montré que le filtre BM3D peut être adapté aisément et conduire à un nouveau filtre de détection DBM3D (*Detection by BM3D*) très performant.

La figure 3.23 présente la sortie du filtre DBM3D (des résultats similaires sont obtenus avec le DNLM). Dans les régions de fortes variations d'intensité, le DBM3D apparaît être plus efficace que les méthodes basées sur un moyennage linéaire et que les méthodes de détection non linéaires par BS comme les filtres max-médian ou morphologiques. En fait, le filtre moyenneur et le filtre top-bottom hat fournissent des intensités résiduelles élevées dans les zones de bord et donc un taux de fausses alarmes élevé. Pour le filtre max-médian, les valeurs résiduelles sont très faibles dans les zones de bord et atteignent des valeurs négatives. Cela implique que l'intensité des cibles dans ces zones est fortement atténuée entraînant une diminution du taux de détection. Le filtre DBM3D proposé est plus efficace dans les zones à fortes variations d'intensité, il fournit des valeurs résiduelles qui ne sont pas significativement supérieures ou inférieures par rapport au reste de l'image. Même si avec DBM3D on obtient un très bon compromis entre estimation du fond et suppression des cibles, des textures de petites échelles sont néanmoins toujours visibles dans le résidu comme on peut le voir sur la figure 3.23(d). Nous montrons dans la suite de ce manuscrit que ces textures peuvent être gérées par des modèles du fond au second ordre.

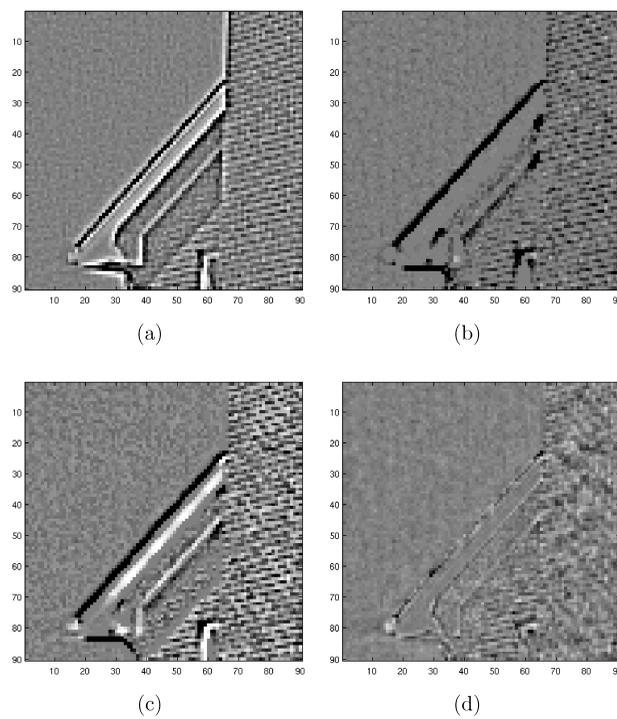


FIGURE 3.23 – Zoom sur les images résiduelles obtenues pour l'image de la figure 1.5a après l'application de diverses méthodes de suppression de fond : (a) BS qui estime le fond avec un filtre moyennneur, (b) BS qui estime le fond avec un filtre max-médian, (c) algorithme top-bottom hat, (d) algorithme DBM3D.

Chapitre 4

Modélisation du fond au second ordre : Développement d'une méthode de détection basée sur une classification des pixels par mélange de gaussiennes

Introduction

Ce chapitre a pour but de présenter une nouvelle méthode de détection basée sur une modélisation du fond au second ordre. La méthode que nous proposons est appelée **GMMF** pour *Gaussian Mixture Matched Filter*. Elle fait tout d'abord intervenir une étape de classification non supervisée des pixels qui permet de faire face à des fonds non-stationnaires. Puis dans un second temps elle fait intervenir un test statistique dans lequel des modèles de cibles plus ou moins complexes peuvent intervenir. Dans le cadre du schéma de détection en deux étapes introduit en figure 1.4, le **GMMF** se place après une méthode de suppression de fond basée sur une modélisation au premier ordre comme par exemple une de celles proposées dans le chapitre 3.

Organisation du chapitre — La section 4.1 rappelle l'état de l'art et introduit le principe de la méthode de détection proposée dans ce chapitre. L'algorithme **CEM** (*Classification Expectation-Maximisation*) de classification des pixels que nous utilisons est ensuite introduit en section 4.2. La section 4.3 présente comment ce principe est appliqué pour conduire à une classification robuste à la présence de cibles. Le filtre de détection proposé, que nous appellerons **GMMF** pour *Gaussian Mixture Matched Filter* est ensuite résumé en section 4.4 et des variantes possibles sont présentées en section 4.5. La section 4.6 discute du paramétrage de l'étape de classification et propose des approches permettant d'ajuster automatiquement certains paramètres du filtre. Dans la section 4.7 on aborde le sujet de la signature des cibles intervenant dans le test statistique. Afin de tenir compte de la position subpixelique des cibles, on propose de reprendre et d'introduire dans notre approche, deux méthodes basées respectivement sur le **LRT** et **GLRT** qui sont proposées dans la littérature dans le cadre de cibles de signe inconnu. Pour finir, en section 4.8, nous proposons des extensions de la méthode au contexte de la détection spatio-temporelle.

4.1 Introduction de la méthode proposée

4.1.1 Approche par filtrage adapté adaptatif

Afin de prendre en compte la texture résiduelle en sortie des filtres de suppression de fond et ainsi pouvoir améliorer les performances de détection, on cherche à appliquer un filtrage adapté correspondant à la deuxième étape du schéma de détection introduit en figure 1.4. On travaille directement sur l'image de résidu, r , issue de l'étape de suppression de fond décrite en section 1.6.1 pouvant faire intervenir par exemple une des approches proposées au chapitre 2. A chacun des pixels $\mathbf{x} \in \mathcal{I}$ où \mathcal{I} est le domaine de définition de l'image de résidu, on associe l'observation vectorielle \mathbf{r}_x constituée des valeurs des pixels inclus dans le patch de dimension $N \times N$ centré sur le pixel \mathbf{x} . Autrement dit, en reprenant les notations utilisées dans le chapitre précédent, \mathbf{r}_x est la version vectorielle du patch $P_{x-\delta}$ extrait de l'image de résidu r . Supposant dans ce chapitre que la moyenne du fond a déjà été correctement éliminée par l'algorithme de suppression de fond, nous considérons donc qu'en l'absence de cible, le vecteur \mathbf{r}_x est la réalisation d'un vecteur aléatoire gaussien de loi $\mathcal{N}(0, \Sigma_x)$. Pour pouvoir appliquer le filtrage adapté défini en équation (1.25), il est nécessaire de connaître la matrice de covariance. Notre problématique de détection se transforme donc ici en un problème d'estimation, en présence éventuelle de cibles, de la matrice de covariance (Σ_x) associée à la composante de fond en chacun des pixels de l'image.

4.1.2 Limitations des méthodes de détection par filtrage adapté existantes

Une méthode très basique, notée **GMF** (*Global Matched Filter*) [Caefer et al., 2008], consiste à estimer pour l'ensemble des pixels de l'image une unique et même matrice de covariance Σ . Cependant, cette méthode n'est pas adaptée à des fonds non-stationnaires pour lesquelles différentes textures peuvent être visibles dans l'image résiduelle. En fonction de la texture, la matrice de covariance varie. Sur l'exemple présenté en figure 4.1, la matrice de covariance calculée à partir des observations vectorielles associées aux pixels se trouvant en bords de nuages diffère de celle obtenue avec les pixels correspondant au ciel. Il est donc nécessaire d'adapter la matrice de covariance à la texture locale pour améliorer les performances de détection [Chan et al., 1998].

Cependant, pour obtenir une estimation précise des matrices de covariance, le nombre d'échantillons indépendants utilisés pour estimer chacune d'elles, doit au minimum être égal à deux fois la longueur du vecteur d'observation [Margalit et al., 1985], c'est-à-dire $2 \times N^2$ pixels. Margalit et al. [Margalit et al., 1985] ont proposé une approche basée sur une estimation locale de la matrice de covariance. Elle consiste à diviser l'image en sous-images non-recouvrantes et à attribuer une matrice de covariance pour chaque sous-image. Comme les sous-images doivent être suffisamment grandes pour que l'estimation des matrices de covariance soit précise (dans l'article [Margalit et al., 1985] : 32×32 pixels), cette méthode est uniquement capable de gérer des variations lentes de la covariance du fond. Elle n'est pas en mesure de faire face à des variations importantes de la covariance sur des zones réduites, comme cela se produit aux bords des nuages

Une autre approche consiste à réduire le nombre de paramètres qui doivent être estimés. Pour cela, on peut approcher les matrices de covariance par des matrices diagonales. Dans ce cas, le **GLRT** (équation (1.25)) faisant intervenir la signature de cibles donnée par la figure 1.3 revient à normaliser l'image de résidu par une estimation de l'écart-type local σ . Ce paramètre σ peut être estimé en chaque pixel sur un voisinage

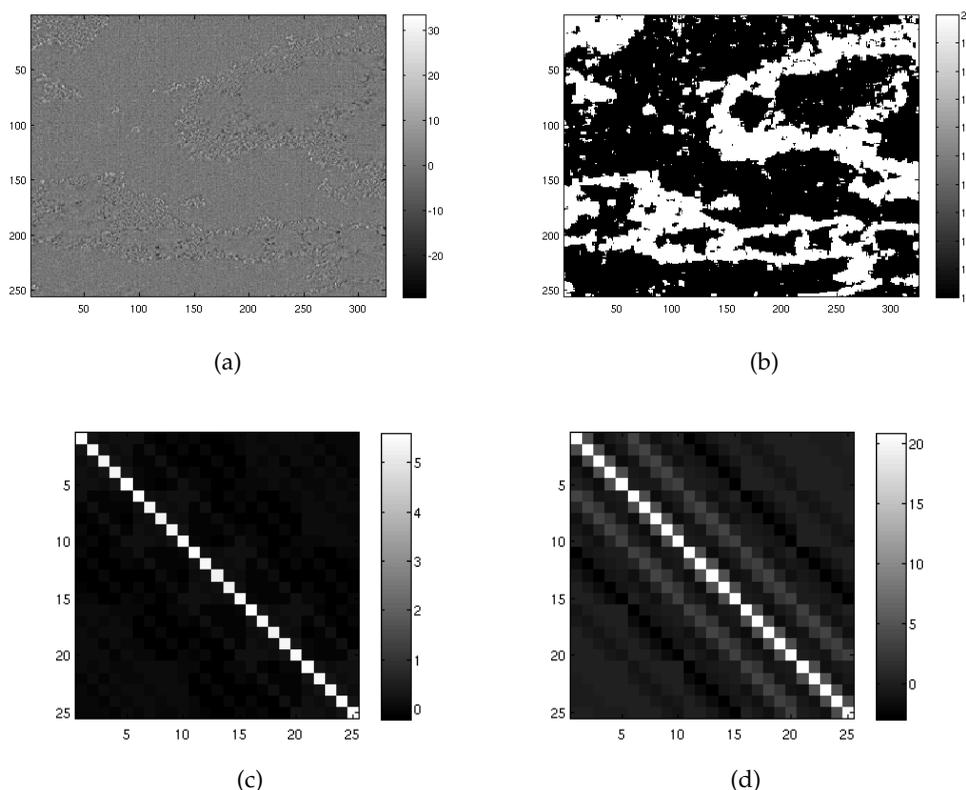


FIGURE 4.1 – Matrices de covariance estimées par zones. (a) Résidu obtenu suite à l’application du filtrage DBM3D sur l’image de la figure 2. (b) Classification des pixels considérée pour estimer les matrices de covariance. (c) Matrice de covariance estimée à partir des observations associées aux pixels en bord de nuage, zones blanches de la classification (b). (d) Matrice de covariance estimée à partir des observations associées aux pixels de ciel, zones noires de la classification (b). Taille du patch utilisé : 5×5 .

restreint. Cette approche nommée « seuillage adaptatif » (*Adapt Tresholding*) dans la littérature [Deshpande *et al.*, 1999], permet de diminuer le taux de fausses alarmes dans les zones de forte transition de l’image mais en contrepartie il conduit à une diminution du taux de détection dans ces même zones. Le modèle de covariance diagonale a le défaut de ne pas permettre de tirer parti de la corrélation potentielle du fond. Dans la suite de ce document, ce filtre est appelé ANF pour *Adaptive Normalized Filter*.

Une approche permettant de collecter un nombre suffisant d’échantillons pour obtenir une estimation précise des matrices de covariance, sans faire aucune hypothèse les concernant, consiste à classifier les pixels de l’image observée en fonction des statistiques au second ordre de leurs composantes de fond, puis à estimer une matrice de covariance par classe. C’est une approche très courante dans le domaine de l’imagerie hyperspectrale pour estimer les matrices de covariance qui interviennent dans le filtrage RX [Matteoli *et al.*, 2010].

Ashton [Ashton, 1998] est l’un des premiers à proposer une telle approche. Il y fait intervenir deux méthodes de classification : une par K-means et une approche bayésienne par maximum de la densité *a posteriori* qui a l’avantage d’être plus rapide et moins sensible

au paramétrage. Stein *et al.* [Stein *et al.*, 2002] classifient les pixels par une procédure SEM (*Stochastic Expectation Maximisation*) et Acito *et al.* [Acito *et al.*, 2005] proposent de sélectionner de façon automatique le nombre de classes à considérer. Ces approches seraient pertinentes dans notre contexte pour des cibles de très faible intensité. Cependant les cibles de plus forte intensité ont de grandes chances de se trouver rassemblées dans une même classe. Ces méthodes manquent de robustesse face à la présence de cibles.

Ce problème peut être résolu en limitant la connexité de la classification, c'est-à-dire en introduisant des contraintes de régularité spatiale des classes dans la classification. Une telle approche est proposée par Caefer *et al.* [Caefer *et al.*, 2008] en imagerie hyperspectrale. Une première classification est effectuée en utilisant la corrélation interbande des pixels. Le plus souvent seuls quelques pixels isolés ont des classes qui sont différentes des classes associées à leurs voisins et généralement ils correspondent à des cibles. Afin de supprimer ces pixels isolés, Caefer *et al.* [Caefer *et al.*, 2008] appliquent une étape de filtrage qui en chaque pixel réévalue la classe qui lui est associée en fonction de son voisinage. Cette approche en deux étapes n'est pas adaptée au contexte mono-bande. En effet, dans un tel contexte, la première classification ne peut être basée que sur des descripteurs spatiaux ou spatio-temporels. La présence d'une cible affecte dans ce cas non seulement la classification du pixel cible, mais aussi celle de ses voisins. Une réévaluation de la classe en fonction du voisinage n'est alors plus efficace.

Dans un contexte de détection mono-bande, Chan *et al.* [Chan *et al.*, 1998] proposent une classification itérative fondée sur la minimisation d'une fonction de coût impliquant huit descripteurs locaux et un terme de pénalisation qui favorise le rattachement de deux pixels voisins à la même classe. L'ajout de cette contrainte complique la minimisation du critère et le résultat est sensible au réglage des poids associés aux huit descripteurs.

4.1.3 Principe général de l'approche proposée

De notre côté, nous proposons une méthode de classification qui, ne faisant intervenir aucune contrainte de connexité, implique beaucoup moins de paramètres que la méthode de Chan *et al.* et qui est entièrement non-supervisée. Comme Stein *et al.* et Acito *et al.*, nous nous appuyons sur un modèle de mélange de gaussiennes, mais notre méthode est robuste à la présence de la cible. Le principe général de cette classification est présenté en section 4.2. L'idée clé est d'utiliser dans l'étape de classification un descripteur spatial conçu pour être robuste à la présence d'une cible. Ce descripteur est décrit en section 4.3. Comme les références précédentes, nous estimons alors une matrice de covariance pour chaque classe et nous appliquons un MF pour chaque pixel en utilisant la matrice de covariance estimée sur la classe à laquelle appartient le pixel. La méthode de détection qui en résulte, nommée GMMF (*Gaussian Mixture Matched Filter*), est résumée en section 4.4. La description de la méthode est tout d'abord présentée dans le cadre de la détection spatiale, où chaque image de la séquence est traitée séparément, des possibilités d'extensions temporelles de la méthode sont ensuite présentées en section 4.8.

Rappelons que le GMMF est appliqué sur l'image de résidu issue d'une technique de suppression de fond, selon le schéma en deux étapes de la figure 1.4. Dans ce chapitre, la méthode de suppression de fond utilisée est soit le filtre DBM3D présenté au chapitre 2, soit un filtre linéaire pour lequel le fond est estimé en moyennant localement les pixels sur un voisinage 3×3 (filtre noté BSMean, pour *Background Suppression by Mean filter*). Si aucune précision n'est faite concernant les paramètres du DBM3D, ce sont les paramètres nominaux donnés au chapitre 5 dans le tableau 5.3 qui sont appliqués.

4.2 Algorithme de classification

Pour classifier les pixels de l'image on utilise une méthode itérative basée sur un modèle de mélange de gaussiennes. On considère, sous l'hypothèse H_0 correspondant à l'absence de cible, que chacun des vecteurs \mathbf{r}_x correspond à une réalisation d'un vecteur aléatoire dont la distribution appartient à un ensemble de K distributions de probabilités gaussiennes centrées : $\{N(0, \Sigma_k)\}_{k \in \{1, \dots, K\}}$ paramétrées par leur matrice de covariance Σ_k . On considère de plus que toutes les distributions gaussiennes ont la même probabilité *a priori* d'être à l'origine des vecteurs \mathbf{r}_x associés aux pixels de l'image, autrement dit que toutes les gaussiennes ont le même poids dans le mélange. La section 4.5, présentant des variantes de la méthode de détection proposée ici, discute de la possibilité de faire d'autres choix concernant les poids du mélange. En notant $k_x \in \{1, \dots, K\}$, l'indice de la distribution associée au vecteur \mathbf{r}_x , la densité de probabilité de \mathbf{r}_x est :

$$p(\mathbf{r}_x | \Sigma_{k_x}) = \frac{1}{(2\pi)^{N^2/2} |\Sigma_{k_x}|^{1/2}} \exp\left(-\frac{1}{2} \mathbf{r}_x^T \Sigma_{k_x}^{-1} \mathbf{r}_x\right), \quad (4.1)$$

avec $|\Sigma_{k_x}|$ le déterminant de la matrice de covariance Σ_{k_x} . Le problème de classification consiste à déterminer conjointement les paramètres $\{\Sigma_k\}_{k \in \{1, \dots, K\}}$ et à identifier pour chaque pixel \mathbf{x} de l'image ($\mathbf{x} \in \mathcal{I}$), l'indice k_x de la distribution qui a généré le vecteur \mathbf{r}_x parmi toutes les distributions du mélange.

Pour ce faire, on utilise un algorithme CEM [Celeux et Govaert, 1992] (*Classification Expectation-Maximisation*) qui est une version de l'algorithme EM [Dempster et al., 1977] (*Expectation-Maximisation*) dédiée à la classification. Ce dernier est une procédure itérative visant à maximiser une vraisemblance, dans le cadre de modèles paramétriques comprenant des paramètres inconnus. Ce principe adapté à la classification conduit, sous l'hypothèse que les poids des gaussiennes dans le mélange sont égaux, à une procédure itérative alternant entre les deux étapes suivantes :

- **Etape E-C (*Expectation-Classification step*)** : Considérant des estimées $\{\widehat{\Sigma}_k\}_{k \in \{1, \dots, K\}}$ des paramètres définissant le mélange de gaussiennes, cette étape identifie l'indice \widehat{k}_x de la distribution qui est la plus probablement à l'origine de l'observation \mathbf{r}_x par maximisation de la vraisemblance $L(k) = p(\mathbf{r}_x | \widehat{\Sigma}_k)$, pour $k \in \{1, \dots, K\}$. Dans le cas gaussien (équation (4.1)), on obtient :

$$\begin{aligned} \widehat{k}_x &= \arg \min_{k \in \{1, \dots, K\}} -\log p(\mathbf{r}_x | \widehat{\Sigma}_k) \\ &= \arg \min_{k \in \{1, \dots, K\}} \left(\log |\widehat{\Sigma}_k| + \mathbf{r}_x^T \widehat{\Sigma}_k^{-1} \mathbf{r}_x \right), \quad \forall \mathbf{x} \in \mathcal{I}. \end{aligned} \quad (4.2)$$

La sortie de cette étape est une segmentation de l'image en K classes. On note C_k l'ensemble des pixels d'indice \mathbf{x} dont l'observation vectorielle \mathbf{r}_x est associée au $k^{\text{ème}}$ modèle de distribution : $C_k = \{\mathbf{x} \in \mathcal{I} | \widehat{k}_x = k\}$.

- **Etape M (*Maximization step*)** : Etant donnée une classification en K classes $\{C_k\}_{k \in \{1, \dots, K\}}$, cette étape consiste à estimer une matrice $\widehat{\Sigma}_k$ par classe par maximum de vraisemblance :

$$\widehat{\Sigma}_k = \arg \max_{\Sigma} p(\{\mathbf{r}_x\}_{\mathbf{x} \in C_k} | \Sigma) = \arg \max_{\Sigma} \prod_{\mathbf{x} \in C_k} p(\mathbf{r}_x | \Sigma), \quad \forall k \in \{1, \dots, K\}, \quad (4.3)$$

Dans le cas gaussien multivarié, la matrice de covariance empirique donnée par le maximum de vraisemblance est [Anderson, 1984, theorem 3.2.1 p.63] :

$$\widehat{\Sigma}_k = \frac{1}{\text{Card}(C_k)} \sum_{\mathbf{x} \in C_k} \mathbf{r}_x \mathbf{r}_x^T, \quad \forall k \in \{1, \dots, K\}, \quad (4.4)$$

avec $\text{Card}(C_k)$ le cardinal de C_k .

A la fin de ce processus itératif on dispose d'une classification des pixels et d'une estimée de l'ensemble des matrices de covariance $\{\widehat{\Sigma}_k\}_{k \in \{1, \dots, K\}}$ associées aux K classes. On peut alors envisager d'appliquer en chaque pixel $\mathbf{x} \in C_k$ le filtrage adapté qui fait intervenir la matrice de covariance $\widehat{\Sigma}_k$ qui a été estimée sur la classe à laquelle appartient ce pixel en cours de traitement. Cependant, l'application tel quel de ce principe n'est pas pertinent car la classification des pixels n'est pas robuste à la présence de cibles, comme montré dans la section suivante.

4.3 Classification robuste

Le filtrage adapté défini par l'équation (1.25) fait intervenir la matrice de covariance associée à la composante de fond de l'observation \mathbf{r}_x . Or le problème est que la méthode de classification proposée ci-dessus classe les pixels \mathbf{x} de l'image en fonction de l'ensemble des composantes de \mathbf{r}_x et non uniquement de sa composante de fond. Dans ce cas, comme le montre la figure 4.2, un pixel contenant une cible sera associé à une classe en général différente de celle qui lui aurait été associée si aucune cible n'avait été présente. En pratique, comme le montre la figure 4.2, les pixels contenant une cible sont classés avec les pixels de forte transition. A l'issue du filtre adapté, la cible n'est alors pas suffisamment rehaussée par rapport au fond, ce qui conduit à une baisse du taux de détection.

Il faut donc que l'algorithme de classification soit robuste à la présence d'une cible. Il doit associer à un pixel cible la classe qui lui serait associée si aucune cible n'était présente. Pour cela on propose une solution simple qui s'avère très efficace. Elle consiste à appliquer le processus de classification CEM décrit en section 4.2 en considérant comme observation associée au pixel \mathbf{x} non pas le vecteur « complet » \mathbf{r}_x mais un vecteur restreint $\mathbf{r}_x^{\text{trunc}}$. Ce vecteur restreint $\mathbf{r}_x^{\text{trunc}}$ collecte uniquement les pixels qui sont inclus dans une couronne centrée sur le pixel \mathbf{x} . Cette couronne est carrée de dimension externe $N \times N$ et comprend un trou de dimension $n \times n$, la figure 4.3 en donne un exemple dans le cas où $N = 9$ et $n = 3$.

L'utilisation de ce vecteur restreint permet de ne pas tenir compte des pixels se trouvant au centre du patch, qui sont susceptibles d'être modifiés par la présence d'une cible qui se trouverait au sein du pixel en cours de classification. Comme nous l'avons dit en introduction, nous cherchons à détecter des cibles éloignées non résolues dont la signature est la PSF du système imageur. Comme nous nous intéressons à des cas où le système est fortement sous échantillonné ($r_c > 0.5$, cf. section 2.5), l'énergie de la cible reste bien concentrée en un nombre limité de pixels. Cependant il ne faut pas oublier que du fait du repliement, plus le système est sous-échantillonné et plus la signature de la cible varie en fonction de la position subpixelique de son image géométrique (phénomène illustré dans la section 2.5 en figure 2.7). Pour tenir compte de ces effets, la dimension du trou choisie ici est de 3×3 pixels ($n = 3$). Concernant le diamètre externe de la couronne, celui-ci est égal à la dimension du patch servant à construire le vecteur \mathbf{r}_x qui intervient

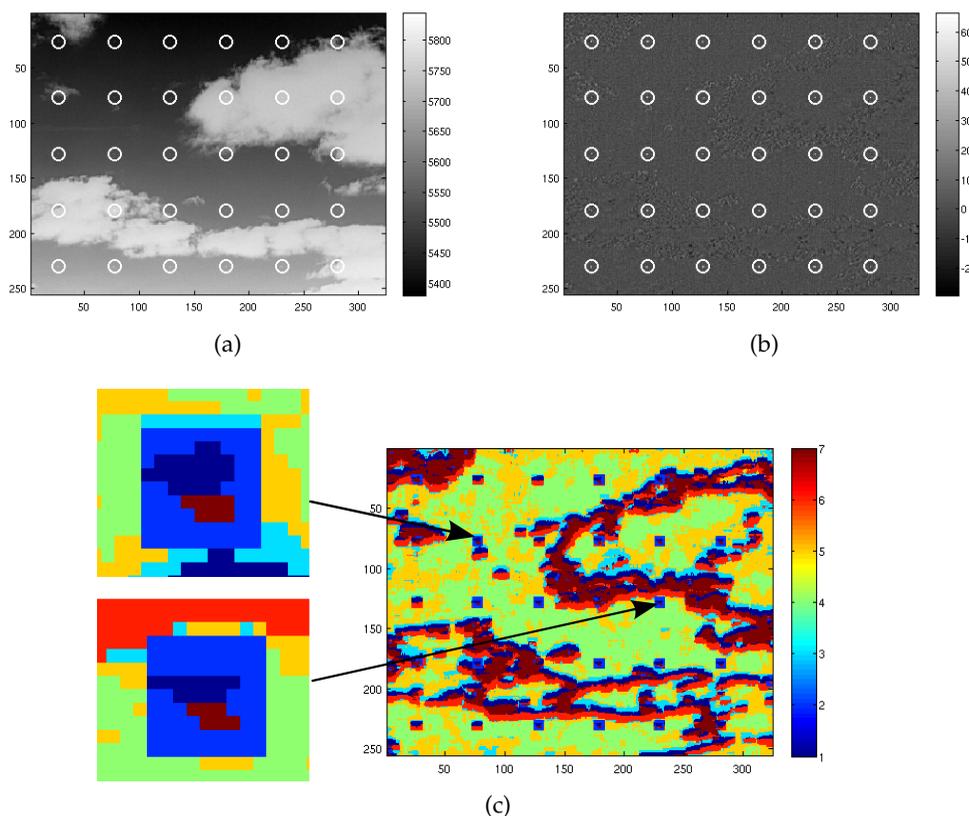


FIGURE 4.2 – Robustesse à la présence de cibles de la classification obtenue à partir des observations $\{\mathbf{r}_x\}_{x \in \mathcal{I}}$. a) Fond nuageux de la figure 2 dans lequel des cibles ont été incrustées : il y a une cible au centre de chaque cercle. b) Résidu obtenu suite à l'application du filtrage DBM3D. c) Classification des pixels obtenue après 9 itérations pour $N = 9$ et $K = 7$ sur l'image de résidu b). d) Zoom sur la segmentation au niveau de deux cibles.

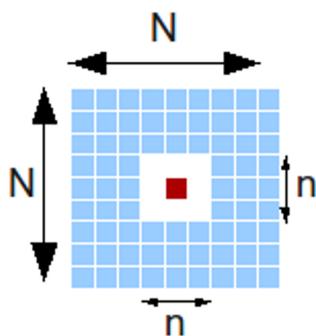


FIGURE 4.3 – Couronne utilisée pour construire l'observation vectorielle $\mathbf{r}_x^{\text{trunc}}$ associée au pixel x (pixel rouge) dans le cas où $N = 9$ et $n = 3$: le vecteur $\mathbf{r}_x^{\text{trunc}}$ collecte les valeurs de tous les pixels bleus.

dans le filtrage adapté. La figure 4.4 présente un exemple de classification des pixels que l'on peut obtenir avec cette approche pour une image sans cible.

La figure 4.5 présente un exemple de classification en présence de cibles. On peut y voir que la classe associée au pixel cible correspond bien à celle qui lui aurait été associée en l'absence de cible. La comparaison des figures 4.5 et 4.4 montre également que la présence d'une cible affecte la classification de ses pixels voisins, mais cela ne nuit pas à la détection. Dans la pratique, cet effet secondaire se révèle en fait être bénéfique en termes de FA. En effet, les pixels dont la classe est modifiée par la présence de cibles dans l'image ont bien dans leur voisinage une forte variation d'intensité (la cible) et celle-ci doit bien être considérée comme faisant partie de la composante de fond pour ne pas augmenter le taux de fausses alarmes dans les voisinages de vraies cibles.

Itérer les équations (4.2) et (4.4) avec les observations $\{\mathbf{r}_x^{\text{trunc}}\}_{x \in \mathcal{I}}$ conduit à une classification robuste des pixels. Par contre les matrices de covariance associées $\widehat{\Sigma}_k^{\text{trunc}}$ que l'on obtient en sortie du processus CEM sont inappropriées pour appliquer le filtrage adapté qui nécessite de disposer des matrices de covariance associées aux observations complètes \mathbf{r}_x . Ainsi, à la fin du processus itératif CEM, on ne retient que la classification et on évalue pour chaque classe de pixels la matrice de covariance associée aux vecteurs complets \mathbf{r}_x en appliquant une nouvelle fois l'étape M mais cette fois en considérant $\{\mathbf{r}_x\}_{x \in \mathcal{I}}$. Afin que les matrices de covariance obtenues puissent bien être inversibles, on impose un nombre minimum de pixels par classe égal à cinq fois la longueur du vecteur \mathbf{r}_x soit : $5 \times N^2$. Faisant l'hypothèse que le nombre de cibles présentes dans l'image est faible, les vecteur \mathbf{r}_x perturbés par la présence d'une cible ne représentent qu'une portion minime des échantillons d'une même classe k . L'estimation n'est donc pas perturbée par la présence de quelques cibles, les $\{\widehat{\Sigma}_k\}_{k \in \{1, \dots, K\}}$ correspondent bien aux matrices de covariance associées aux composantes de fond des vecteurs \mathbf{r}_x .

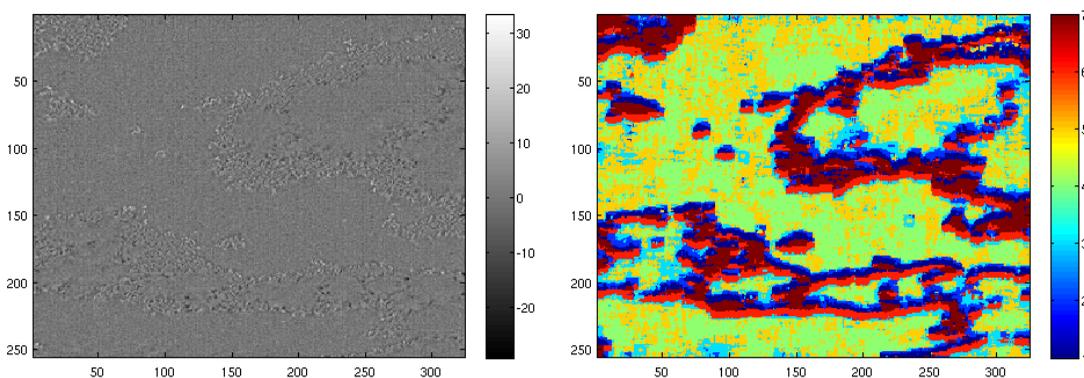


FIGURE 4.4 – Classification des pixels que l'on obtient avec l'observation réduite $\{\mathbf{r}_x^{\text{trunc}}\}$, pour l'image de figure 2 sans ajout de cible, prétraitée par un filtrage DBM3D, après 9 itérations dans le cas où $N = 9$, $n = 3$ et $K = 7$. A gauche : sortie du filtrage DBM3D. A droite : segmentation obtenue.

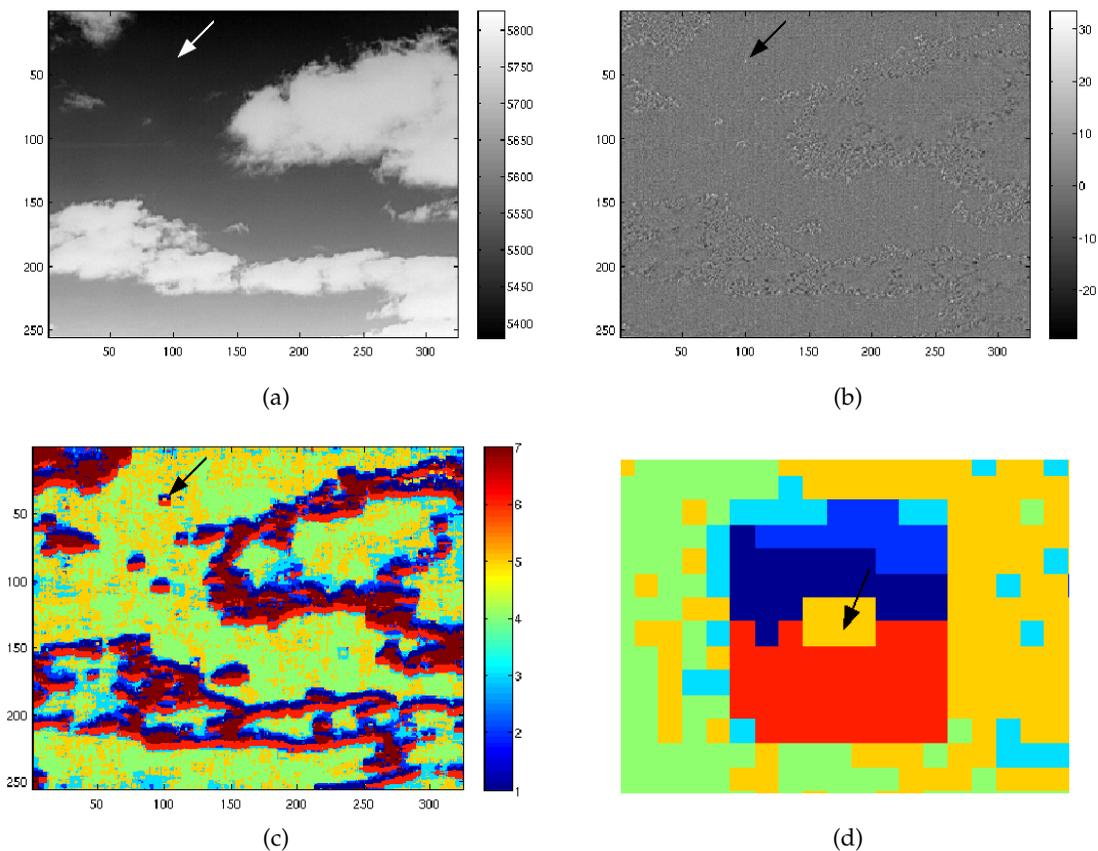


FIGURE 4.5 – Robustesse à la présence de cibles de la classification obtenue à partir des observations restreintes $\{\mathbf{r}_x^{\text{trunc}}\}_{x \in \mathcal{I}}$. (a) Fond nuageux de la figure 2 dans lequel une cible a été incrustée à la position pointée par la flèche. (b) Résidu obtenu suite à l’application du filtrage DBM3D. (c) Classification des pixels obtenue après 9 itérations pour $N = 9$, $n = 3$ et $K = 7$ sur l’image de résidu (b) et zoom de la segmentation au niveau de deux cibles.

4.4 Filtrage GMMF

Pour résumer, le filtrage GMMF proposé consiste à appliquer itérativement les équations (4.2) et (4.4) des étapes E-C et M du processus CEM avec l’ensemble des vecteurs d’observations restreints associés aux pixels de l’image, $\{\mathbf{r}_x^{\text{trunc}}\}_{x \in \mathcal{I}}$. Ensuite, pour chaque classe de pixels k , la matrice de covariance $\widehat{\Sigma}_k$ est estimée à partir des observations « complètes » associées à tous les pixels x qui appartiennent à cette classe k ($\{\mathbf{r}_x\}_{x \in C_k}$) en appliquant l’équation (4.4). Au final, on dispose pour chaque pixel x , d’un label \widehat{k}_x et d’une estimée de la matrice de covariance associée aux pixels ayant ce label $\widehat{\Sigma}_{\widehat{k}_x}$; on peut donc appliquer le filtrage adapté AMF défini par l’équation (1.25) à ce pixel :

$$\frac{\mathbf{s}^T \widehat{\Sigma}_{\widehat{k}_x}^{-1} \mathbf{r}_x}{[\mathbf{s}^T \widehat{\Sigma}_{\widehat{k}_x}^{-1} \mathbf{s}]^{1/2}} \underset{H_0}{\overset{H_1}{\geq}} \text{seuil}. \quad (4.5)$$

Un schéma de principe du filtre GMMF est donné en figure 4.6.

On peut noter que le filtre **GMMF** dans le cas où $K = 1$, correspond exactement au filtrage **GMF** (cf. section 4.1.2).

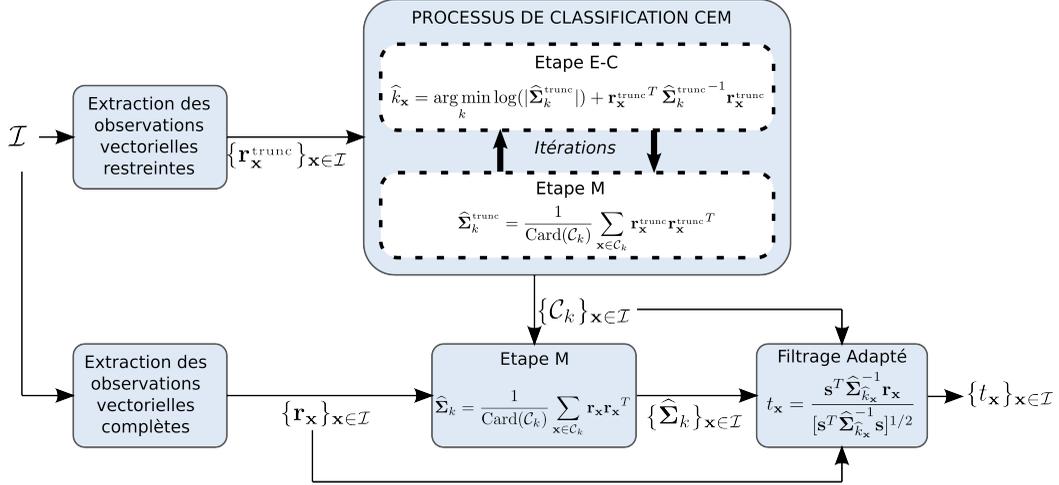


FIGURE 4.6 – Schéma de principe du filtrage **GMMF**.

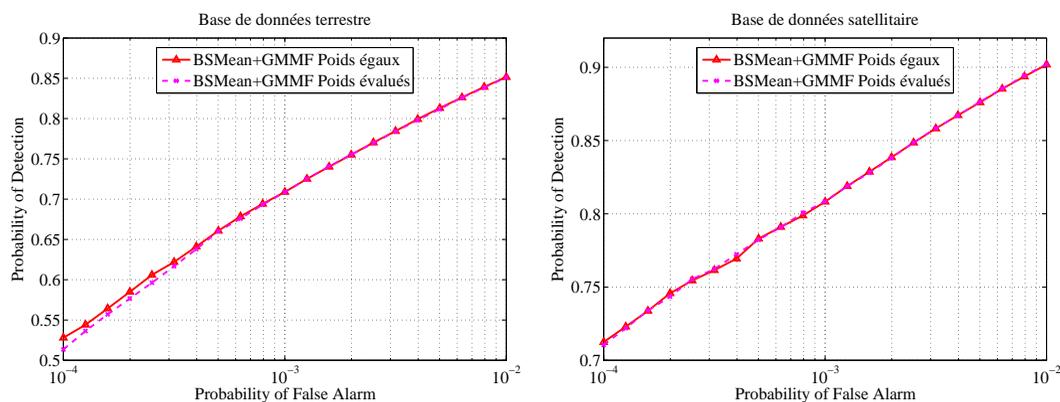
4.5 Discussion : variantes de l'algorithme **GMMF**

4.5.1 Variantes possibles de la classification des pixels par **CEM**

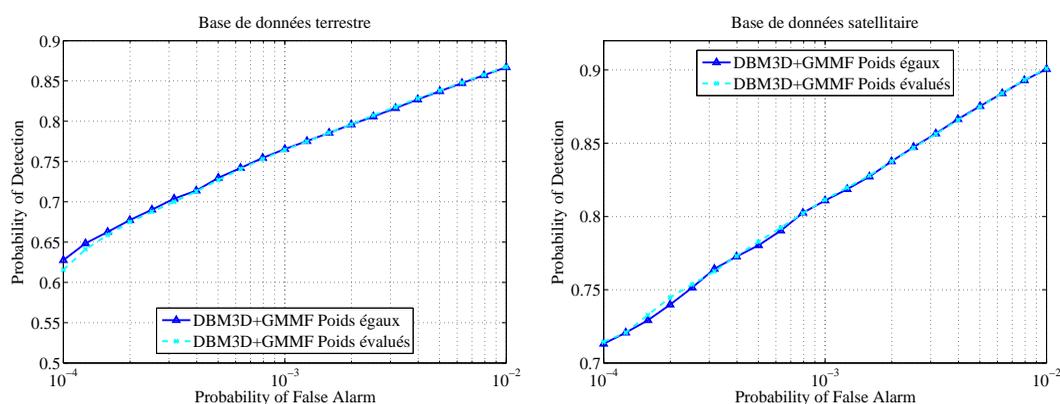
Dans l'approche proposée ci-dessus on classe les pixels en appliquant l'algorithme **CEM** sur les observations $\{\mathbf{r}_x^{\text{trunc}}\}_{\mathbf{x} \in \mathcal{I}}$ en considérant un mélange de gaussiennes de proportions égales, autrement dit un mélange pour lequel toutes les gaussiennes ont le même poids.

On aurait pu envisager que toutes les gaussiennes du mélange n'aient pas la même probabilité *a priori* d'être à l'origine d'une observation \mathbf{r}_x . Dans ce cas ces probabilités *a priori*, $\{\pi_k\}_{k \in \{1, \dots, K\}}$, sont estimées par maximum de vraisemblance dans l'étape M du processus **CEM**. A chaque itération du processus, l'étape M consiste alors à estimer les matrices de covariance associées à chaque classe en utilisant l'équation (4.4) et à estimer les probabilités *a priori* de chaque distribution gaussienne par : $\hat{\pi}_k = \frac{\text{Card}(\mathcal{C}_k)}{\text{Card}(\mathcal{I})}$. L'étape E consiste, quant à elle, à estimer quel est l'indice de la distribution la plus probablement à l'origine de l'observation \mathbf{r}_x en tenant compte de la probabilité *a priori* associée à chaque distribution : $\hat{k}_x = \arg \min_k -2 \log(\hat{\pi}_k) + \log(|\hat{\Sigma}_k|) + \mathbf{r}_x^T \hat{\Sigma}_k^{-1} \mathbf{r}_x$.

Comme le montrent les courbes de performances présentées en figure 4.7, réaliser un **CEM** dans le **GMMF** en estimant les poids du mélange de gaussiennes à chaque itération ou en imposant des poids égaux conduit au final à des performances de détection quasiment identiques. Estimer les poids augmente un petit peu le coût de calcul et pour la base de données terrestre où une classe (celle correspondant au ciel) est souvent nettement plus représentée que les autres, les performances obtenues en fixant les poids sont légèrement supérieures à celles obtenues en estimant les poids. C'est pourquoi nous avons choisi de fixer ces poids.



(a) Préfiltrage par BSMean



(b) Préfiltrage par DBM3D

FIGURE 4.7 – Courbes ROC obtenues, pour chacune des bases de données, pour la version de référence du GMMF qui impose les poids du mélange de gaussiennes égaux (GMMF Poids égaux) et sa variante qui estime ces poids au cours du processus CEM (GMMF Poids évalués). Ces deux filtres sont appliqués à la suite d'un BSMean pour les courbes du haut et à la suite d'un DBM3D pour les courbes du bas. Ces courbes ont été obtenues pour un paramètre σ du filtre DBM3D égal à 8 pour la base de données terrestre et 6 pour la satellitaire.

4.5.2 Possibilité d'utiliser d'autres méthodes de classification robustes

Cette approche de détection (basée sur une classification des pixels et une estimation de matrices de covariance par classe) peut techniquement fonctionner avec n'importe quelle méthode de classification. Mais pour obtenir de bonnes performances de détection, l'étape de classification doit être robuste à la présence de cibles et idéalement rassembler dans une même classe les pixels x pour lesquels les composantes de fonds des vecteurs r_x qui leur sont associées suivent une même loi de distribution. La méthode de classification proposée a l'avantage d'être fondée sur un modèle de mélange de gaussiennes impliquant des statistiques du second ordre ayant la particularité d'être un sous-ensemble de celles qui sont nécessaires pour calculer le filtrage adapté (équation (4.5)). Cependant on peut envisager d'autres méthodes de classification robustes moins coûteuses en calcul. On propose d'étudier une méthode qui consiste à calculer les écart-types locaux associés à

chacun des pixels de l'image et à appliquer un algorithme K-means [MacQueen, 1967] prenant en entrée ces écarts-types locaux. De même que l'algorithme EM, l'algorithme K-means est itératif mais fait intervenir comme statistique uniquement la moyenne des classes. Il peut être vu comme un cas particulier des algorithmes CEM, et conduit à rassembler dans une même classe les pixels qui ont des écarts types locaux de même valeur. Afin que cette classification soit robuste à la présence d'une cible, l'écart-type local du pixel x est calculé à partir des pixels inclus dans une couronne centrée en x (comme celle utilisée pour constituer $\mathbf{r}_x^{\text{trunc}}$: cf. figure 4.3). Le filtrage faisant intervenir cette méthode de classification est dans la suite noté GMMF_0 .

Les courbes ROC obtenues pour chacune des bases de données, pour différentes méthodes de pré-filtrage, avec le filtre GMMF de référence qui fait intervenir un CEM et avec cette variante GMMF_0 moins coûteuse en temps de calcul, sont présentées en figure 4.8. L'algorithme GMMF reste globalement toujours le plus performant. Notamment quand les données ne sont pas prétraitées, il fournit des performances de détection nettement supérieures à celles obtenues avec le filtre GMMF_0 . Cependant, quand les images sont prétraitées par un filtre BSMean ou un filtre DBM3D, les performances du filtre GMMF_0 se rapprochent des performances obtenues avec le GMMF. L'avantage de l'utilisation d'une classification par K-means, c'est qu'elle est beaucoup moins coûteuse en temps de calcul qu'un CEM. Sur des données prétraitées, le filtre GMMF_0 permet d'obtenir un très bon compromis coût de calcul / performances.

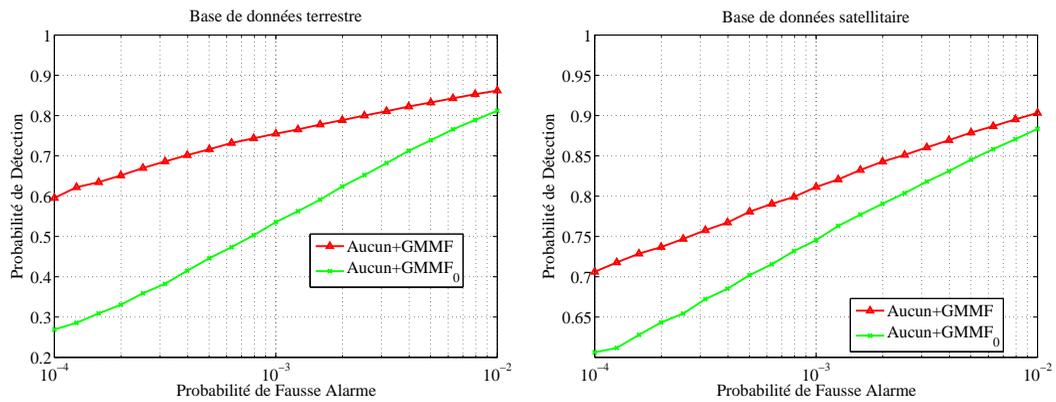
4.5.3 Une approche de filtrage basée sur un mélange de gaussiennes

Toujours basée sur l'idée d'une modélisation du fond par un mélange de gaussiennes, une autre approche a été proposée dans le domaine de l'imagerie hyperspectrale [Stein et al., 2002, Matteoli et al., 2010]. Elle consiste comme précédemment à appliquer en chaque pixel un test du rapport de vraisemblance généralisé (GLRT) mais considère que la composante de fond est une variable aléatoire distribuée directement selon l'ensemble du mélange de gaussiennes et non plus selon une unique gaussienne de ce mélange (celle associée à la classe à laquelle appartient le pixel).

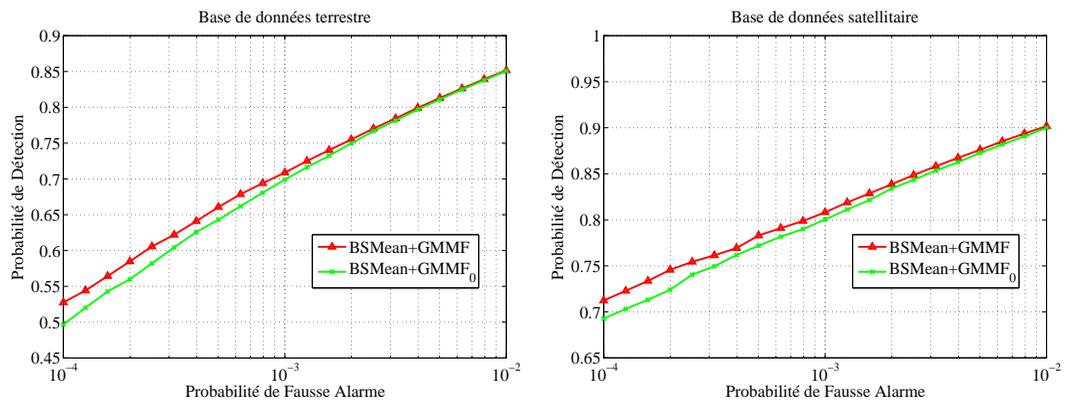
A notre connaissance, le GLRT qui fait intervenir comme distribution du fond un mélange de gaussiennes, n'a été appliqué qu'en imagerie hyperspectral, dans le cas où la signature de la cible est supposée inconnue. On propose dans l'annexe C une adaptation de cette approche à notre contexte de détection, autrement dit au cas où la signature de la cible est considérée connue. Le filtre de détection ainsi obtenu est nommé GLRT-GM (*GLRT-Gaussian Mixture*).

On peut noter que les cibles présentes dans l'image étant supposées peu nombreuses, celles-ci ne vont pas perturber l'estimation du mélange de gaussiennes. De plus, comme il n'y a plus aucune étape de classification des pixels dans la constitution du test à appliquer, il n'est pas nécessaire de faire intervenir l'observation restreinte $\mathbf{r}_x^{\text{trunc}}$. On peut directement estimer le mélange de gaussiennes sur les observations complètes $\{\mathbf{r}_x\}_{x \in I}$. Cherchant à estimer le mélange de gaussiennes et non plus à classifier les pixels, on fait intervenir un processus itératif EM et non plus CEM pour estimer les paramètres du mélange.

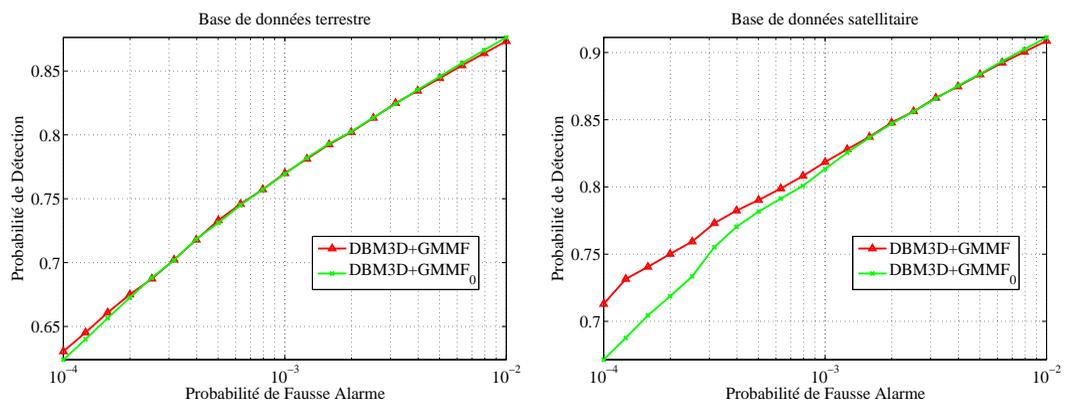
De plus, avec cette approche GLRT-GM, si l'on a des pixels de fond qui ont une structure qui est assez similaire à celle des cibles, on risque de ne plus détecter les cibles quelle que soit leur position dans l'image. Dans les zones de l'image où le fond présente



(a) Aucun préfiltrage



(b) Préfiltrage par BSMeans



(c) Préfiltrage par DBM3D

FIGURE 4.8 – Courbes ROC obtenues, pour chacune des bases de données, pour la version de référence du GMMF qui classe les pixels par CEM et sa variante GMMF₀, de coût de calcul plus faible, qui classe les pixels par K-means. Ces deux filtres sont appliqués sur les images non pré-traitées pour les courbes du haut, à la suite d'un BSMeans pour les courbes du milieu et à la suite d'un DBM3D pour les courbes du bas.

une structure proche de celle des cibles, le **GMMF** (approche classifiante) risque de ne fournir aucune détection ; par contre il aura l'avantage de toujours très bien détecter les cibles se trouvant dans les autres types de fond. C'est pourquoi le **GMMF** est plus performant que cette approche **GLRT-GM**, comme le montrent les courbes **ROC** de la figure 4.9.

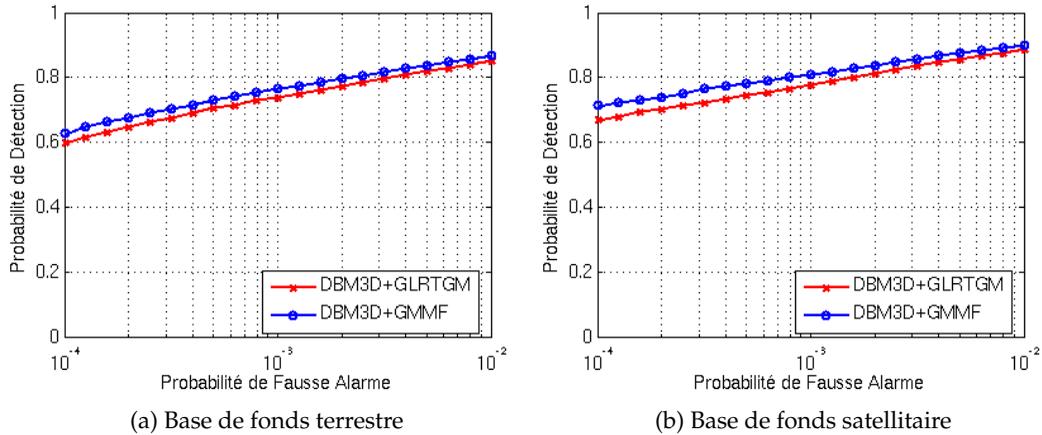


FIGURE 4.9 – Courbes **ROC** obtenues, pour chacune des bases de données, pour le **GMMF** et le **GLRT-GM** appliqués à la suite d'un **DBM3D**.

4.6 Paramètres et détails d'implémentation du processus de classification

Les paramètres du filtrage **GMMF** sont : (i) la dimension du patch ($N \times N$) servant à former le vecteur \mathbf{r}_x , (ii) la dimension du trou ($n \times n$) utilisé pour définir la couronne servant à construire le vecteur $\mathbf{r}_x^{\text{trunc}}$, (iii) le nombre de classes K et (iv) le nombre d'itérations. Ces deux derniers paramètres peuvent être ajustés de façon automatique comme expliqué ci-dessous en sections 4.6.3 et 4.6.4. Notons que ce processus de classification itératif nécessite d'être initialisé. Nous verrons en section 4.6.2 comment il peut l'être. Mais tout d'abord nous étudierons en section 4.6.1, l'influence sur les performances de détection de la variation du paramètre N . Si aucune précision n'est donnée, dans toute la suite les paramètres du filtre **GMMF** utilisés sont ceux donnés dans le tableau 4.1.

4.6.1 Paramètre N : dimension des patches

La figure 4.10 présente la variation des performances de détection du filtre **GMMF** en fonction de la dimension des patches ($N \times N$) servant à former les vecteurs \mathbf{r}_x . On observe que pour les valeurs de N testées, les performances de détection augmentent avec la dimension des patches. Une valeur de N égale à 9 fournit le meilleur compromis coût de calcul / performances de détection.

Paramètres	GMMF
N	9
n	3
K	7
Méthode d'initialisation du CEM	Segmentation obtenue en appliquant un K-means sur les écarts-types locaux.
Nombre d'itérations du processus CEM (<i>ite</i>)	9
Signature des cibles \mathbf{s}	$\mathbf{s} = \begin{pmatrix} 0 & \dots & 0 & \dots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & 1 & 0 & \dots & 0 \\ & & 0 & 0 & 0 & & \\ \vdots & & \vdots & & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & \dots & 0 \end{pmatrix}$

TABLE 4.1 – Tableau présentant les valeurs des paramètres nominaux utilisées pour le filtre de détection GMMF. Dans tout le chapitre ce sont ces valeurs qui sont utilisées si aucune précision particulière n'est indiquée.

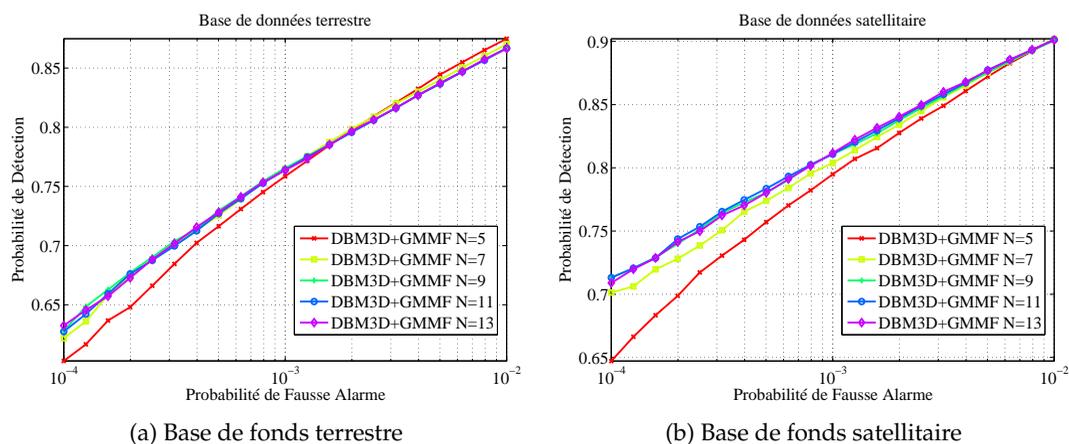


FIGURE 4.10 – Influence de la dimension des patches : Courbes ROC obtenues avec le filtre GMMF sur les données prétraitées par DBM3D (avec $\sigma = 8$ pour la base de donnée terrestre et $\sigma = 6$ pour la satellitaire), pour différentes valeurs de N .

4.6.2 Initialisation de l'étape de classification

Le processus itératif CEM utilisé pour classifier les pixels de l'image nécessite d'être initialisé. Il peut l'être soit par une famille de gaussiennes de paramètres fixés $\{(0, \Sigma_k)\}_{k \in \{1, \dots, K\}}$, soit par une classification des pixels qui permet de définir les ensembles $\{C_k\}_{k \in \{1, \dots, K\}}$. Selon les résultats des tests que nous avons menés sur ce sujet, la méthode d'initialisation ne joue pas beaucoup sur les résultats obtenus mais a une influence sur la vitesse de convergence de la procédure de classification CEM.

Initialisation par une segmentation

Une des méthodes pour initialiser le processus de classification itératif proposé ci-dessus consiste à faire une pré-segmentation des images qui attribue un label à chaque pixel de l'image. On associe ensuite à chaque vecteur d'observation $\mathbf{r}_x^{\text{trunc}}$ issue d'une couronne centrée sur le pixel x un numéro de modèle gaussien k_x qui correspond au label

attribué au pixel x par l'algorithme de pré-segmentation. Dans ce cas, c'est l'étape M de la procédure décrite ci-dessus qui est appliquée en premier.

Comme méthode de pré-segmentation on peut par exemple utiliser un algorithme K-means [MacQueen, 1967] qui classe les pixels en fonction de leurs écart-types locaux, identique à celui présenté dans la section 4.5. Afin d'être robuste à la présence d'une cible dès l'initialisation de notre procédure CEM, l'écart-type local du pixel x est calculé à partir des pixels inclus dans une couronne centrée en x . On choisit la même couronne que celle intervenant dans notre processus de classification CEM.

Notons enfin que dans le cas du traitement d'une séquence d'images, pour initialiser le processus CEM appliqué à l'image courante, on peut utiliser la segmentation que l'on a obtenue à la fin du processus CEM sur l'image précédente.

Par une famille de gaussiennes

L'autre méthode d'initialisation possible du CEM consiste à fixer un modèle initial de mélange de gaussiennes. Dans ce cas c'est l'étape E-C de la procédure CEM qui est appliquée en premier.

Pour générer le mélange de gaussiennes, nous reprenons ici une méthode proposée par Yu [Yu et al., 2012] dans le cadre de la résolution de problèmes inverses pour le traitement des images. Cette méthode consiste à créer une famille de gaussiennes à partir de motifs synthétiques. En se fondant sur l'observation que des pixels associés à des bords (zones de fortes transitions) d'orientations différentes ont des matrices de covariance différentes, Yu propose d'apprendre des matrices de covariance sur des motifs synthétiques qui correspondent à des bords d'orientations différentes uniformément échantillonnés entre 0° et 180° . Un dernier ensemble d'apprentissage est ajouté pour capter les motifs non directionnels. Pour plus de détails le lecteur est invité à se reporter à l'article de Yu [Yu et al., 2012]. Cette méthode fournit des matrices de covariance $\{\Sigma_k\}_{k \in \{1, \dots, K\}}$, associées aux vecteurs complets \mathbf{r}_x de dimension N^2 construits à partir des patches complets. De ces matrices on extrait les composantes qui nous permettent de former les matrices de covariance $\{\Sigma_k^{\text{trunc}}\}_{k \in \{1, \dots, K\}}$ associées aux vecteurs restreints $\mathbf{r}_x^{\text{trunc}}$.

De même que dans le cas de l'initialisation par une pré-segmentation, dans le cas du traitement d'une séquence d'images, on peut utiliser le résultat obtenu sur l'image précédente pour initialiser le processus CEM ; soit ici le mélange de gaussiennes que l'on a obtenu à la fin du processus CEM sur l'image précédente. Cette approche, dans le cadre d'une initialisation par un mélange de gaussiennes pourrait être intéressante en cas de mouvement du fond contrairement à l'approche utilisant une pré-segmentation,

Il est important de noter que la méthode d'initialisation ne joue pas un rôle fondamental si l'on réalise suffisamment d'itérations de notre processus de classification EM. Si l'initialisation de l'algorithme n'est pas adéquate, il faudra simplement plus de temps à l'algorithme pour converger. Disposer d'une initialisation pertinente de l'algorithme peut permettre de réduire le nombre d'itérations et donc le temps de calcul de l'algorithme.

La figure 4.11 présente les classifications obtenues au moyen de ces deux méthodes d'initialisation (aucune itération du processus CEM n'est appliquée) pour l'image présentée en figure 2 pré-traitée par deux filtres différents : le filtre DBM3D et le BSMean. L'initialisation par une famille de gaussiennes obtenue à partir de motifs synthétiques est une méthode d'initialisation qui est originale ; cependant c'est l'initialisation basée sur la méthode de classification par K-means sur les écart-types locaux qui donne la

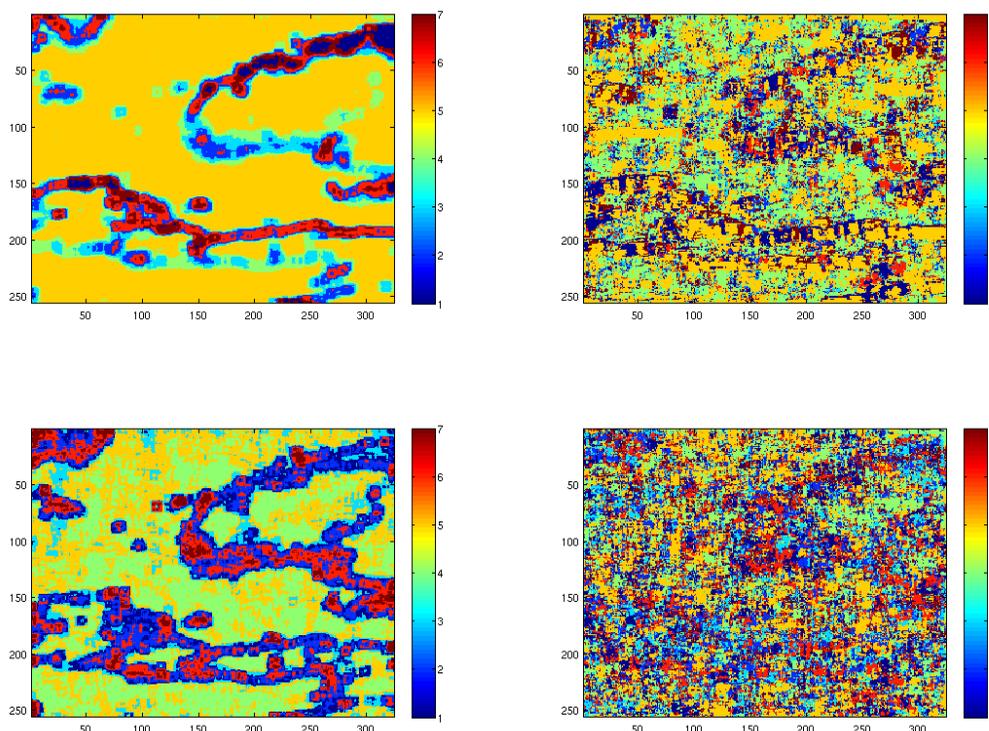


FIGURE 4.11 – Classifications des pixels obtenues (sans itération) par les deux méthodes d'initialisation proposées pour l'image présentée en figure 2 pré-traitée par BSMean (en haut) et par DBM3D (en bas). A gauche : initialisation par une segmentation par K-means sur les écarts-types locaux des pixels. A droite : initialisation par un modèle de mélange de gaussiennes.

classification la plus proche de celle fournie par notre processus CEM après un grand nombre d'itérations. C'est donc la méthode d'initialisation par Kmeans qui est utilisée dans toute la suite. On peut noter que le filtre GMMF, dans le cas où aucune itération n'est appliquée, correspond exactement au filtrage noté GMMF₀ présenté dans la section 4.5.2. La segmentation par K-means classe les pixels en fonction de leurs écart-types locaux, l'application du processus CEM permet d'affiner la classification en prenant notamment en compte la corrélation entre les pixels et ainsi d'augmenter les performances de détection comme le montrent les résultats présentés en figure 4.8.

4.6.3 Nombre d'itérations

L'objectif d'un algorithme CEM est de maximiser le critère CML (*Classification Maximum Likelihood*) [Celeux et Govaert, 1992], aussi nommé *complete log-likelihood* (CL) [Biernacki et al., 2000], qui est égal à :

$$CL(\{C_k\}_{k \in \{1, \dots, K\}}, \{\widehat{\pi}_k\}_{k \in \{1, \dots, K\}}, \{\widehat{\Sigma}_k^{\text{trunc}}\}_{k \in \{1, \dots, K\}}) = \sum_{k=1}^K \sum_{\mathbf{x} \in C_k} \log(\widehat{\pi}_k p(\mathbf{r}_{\mathbf{x}}^{\text{trunc}} | \widehat{\Sigma}_k^{\text{trunc}})). \quad (4.6)$$

Celeux et Govaert [Celeux et Govaert, 1992] ont montré que chaque itération du processus CEM augmente ce critère CML. On propose ici d'utiliser ce dernier pour définir un critère d'arrêt automatique de la procédure itérative CEM. On stoppe les itérations quand la différence relative entre deux valeurs successives de ce critère est inférieure à un certain seuil, fixé ici égal à 10^{-4} , c.a.d. quand :

$$\left| \frac{CL_i - CL_{i-1}}{CL_i} \right| < 10^{-4} \quad (4.7)$$

où i est le numéro de l'itération courante.

Si les poids des gaussiennes sont imposés égaux dans la procédure CEM, on a : $\widehat{\pi}_k = \frac{1}{K}$ pour $k = 1, \dots, K$. Si ce n'est pas le cas, les $\{\widehat{\pi}_k\}_{k \in \{1, \dots, K\}}$ sont ceux obtenus en sortie du CEM. Dans le cadre de la variante proposée qui fait intervenir un processus EM on peut envisager de proposer un critère d'arrêt similaire à celui de l'équation (4.7) qui ne sera pas basé sur le critère CML mais directement sur la vraisemblance observée (*observed log-likelihood*) égale à :

$$L(\{\widehat{\pi}_k\}_{k \in \{1, \dots, K\}}, \{\widehat{\Sigma}_k\}_{k \in \{1, \dots, K\}}) = \sum_{\mathbf{x} \in \mathcal{I}} \log \left(\sum_{k=1}^K \widehat{\pi}_k p(\mathbf{r}_x | \widehat{\Sigma}_k) \right). \quad (4.8)$$

La figure 4.12 présente le nombre d'itérations réalisées avant l'arrêt du processus itératif CEM, dans le cas du filtrage GMMF standard qui considère que les poids des gaussiennes sont égaux, pour les différentes images des bases de fonds satellitaire et terrestre prétraitées par le filtre BSMean et le DBM3D. On peut noter que le nombre d'itérations réalisées est en moyenne supérieur pour la base de données satellitaire que pour la terrestre, et qu'il est en moyenne supérieur pour un prétraitement par BSMean que pour un prétraitement par DBM3D.

Dans la suite de la thèse, le nombre d'itérations est fixé égal à 9. Avec un tel choix on obtient des performances de détection similaires à celles obtenues en choisissant automatiquement le nombre d'itérations, ce qui sera confirmé par la figure 4.14 de la section suivante.

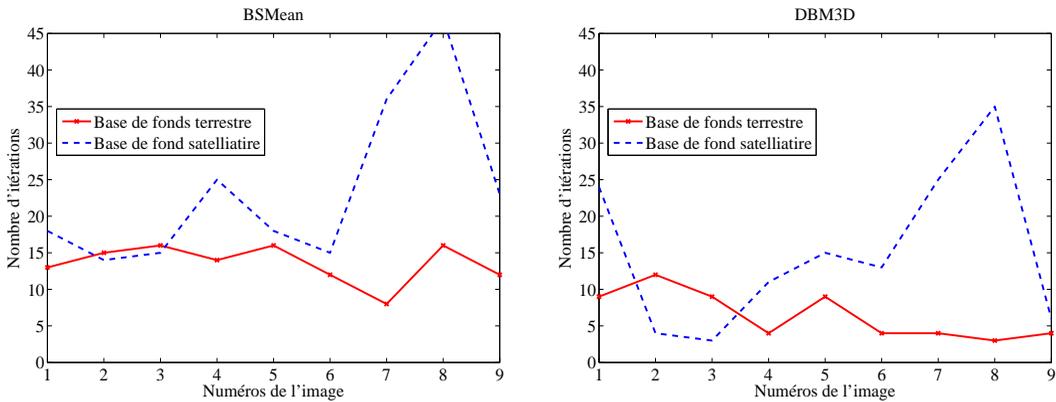


FIGURE 4.12 – Nombre d'itérations réalisées avant l'arrêt du processus itératif CEM inclus dans le filtrage GMMF pour chacune des neuf images des deux bases de fonds, pré-filtrées par BSMean (à gauche) et par DBM3D (à droite), avec un nombre de classes sélectionné de façon automatique (cf. section 4.6.4).

4.6.4 Choix du nombre de classes

Le nombre de classes est relié à la variété des textures présentes dans une image. Cependant celui-ci ne doit pas être trop grand pour la raison suivante. Il est nécessaire que chaque classe contienne un nombre suffisant de pixels pour que l'estimation des matrices de covariance soit précise et robuste à la présence éventuelle de cibles. De plus, comme les matrices de covariance nécessitent d'être inversées lors de l'application du filtrage adapté (éq. (4.5)), celles-ci ne doivent pas être singulières.

Pour ce faire, on impose un nombre minimum de $5 \times N^2$ pixels par classe. Un nombre maximum de classes K est fixé initialement, mais si la classification conduit à des classes trop petites qui ne respectent pas cette contrainte sur le nombre minimum de pixels par classe alors la classification est réinitialisée avec un nombre de classes K inférieur.

Il est possible d'ajuster automatiquement le nombre de classes, autrement dit le nombre de gaussiennes utilisé pour constituer le modèle de mélange; pour cela nous proposons de maximiser le critère **ICL** (*Integrated Classification Likelihood*) [Biernacki *et al.*, 2000]. Ce critère a été proposé par Biernacki *et al.* pour répondre au problème de surestimation du nombre de classes qui apparaît lors de l'utilisation du critère usuel **BIC** (*Bayesian Information Criterion*) dans une approche de classification par **CEM**. Le critère **ICL** est défini pour un nombre de classes K de la façon suivante :

$$ICL(K) = \sum_{k=1}^K \sum_{\mathbf{x} \in C_k} \log(\widehat{\pi}_k p(\mathbf{r}_{\mathbf{x}}^{\text{trunc}} | \widehat{\Sigma}_k^{\text{trunc}})) - \frac{\nu_K}{2} \log(NO), \quad (4.9)$$

où $\widehat{\pi}_1, \dots, \widehat{\pi}_K$ sont les poids estimés du mélange de gaussiennes. NO est le nombre total d'observations restreintes à classifier, il est égal au nombre de pixels dans l'image : $NO = \text{Card}(I)$.

La variable ν_K est le degré de liberté du modèle de mélange de gaussiennes impliquant K gaussiennes. Considérant dans le processus **CEM** que les poids des gaussiennes sont égaux, ces poids sont aussi imposés égaux pour le calcul du critère **ICL** : $\widehat{\pi}_k = \frac{1}{K}$ pour $k = 1, \dots, K$. Sous l'hypothèse que les gaussiennes sont centrées et que leurs poids sont égaux, ν_K est égal à :

$$\nu_K = K \times d \times \frac{(d+1)}{2} \quad (4.10)$$

où $d = N^2 - n^2$ est la dimension du vecteur $\mathbf{r}_{\mathbf{x}}^{\text{trunc}}$.

Dans le cas de la variante de la méthode de classification (cf section 4.5) où les poids du mélange de gaussiennes sont considérés inconnus et sont estimés par **ML**, on peut noter qu'il faut faire intervenir dans le critère **ICL** les poids estimés obtenus à la fin du **CEM** et que le degré de liberté du modèle est augmenté : $\nu_K = K \times d \times \frac{(d+1)}{2} + K - 1$. Dans le cas où l'on considère non plus l'approche classifiante mais celle consistant à prendre en compte l'ensemble du mélange de gaussiennes dans le **GLRT**, ce dernier est estimé par un processus **EM**. Dans ce cas, c'est théoriquement le critère **BIC** qu'il faut utiliser pour déterminer le nombre optimum de gaussiennes K à faire intervenir dans le mélange :

$$BIC(K) = \sum_{\mathbf{x} \in I} \log\left(\sum_{k=1}^K \widehat{\pi}_k p(\mathbf{r}_{\mathbf{x}} | \widehat{\Sigma}_k)\right) - \frac{\nu_K}{2} \log(NO), \text{ avec } \nu_K = K \times N^2 \times \frac{(N^2+1)}{2} + K - 1$$

La figure 4.13 présente le nombre de classes qui est retenu par la méthode proposée, dans le cas du filtrage **GMMF** standard qui considère que les poids des gaussiennes sont égaux, pour les images des bases de fonds satellitaire et terrestre prétraitées par différents filtres : le **BSMean** et le **DBM3D**. Le critère $ICL(K)$ est évalué ici pour un nombre de classes

variant entre 1 et 15. On peut noter que le nombre de classes retenu est en moyenne un peu plus élevé pour la base de données satellitaire que terrestre, cela s'explique par le fait qu'il y a une variété plus grande de textures présentes dans la base de données satellitaire que dans la base de données terrestre. De même on peut noter que le nombre de classes retenu est plus faible après un filtrage **DBM3D** qu'après le **BSMean**. Cela s'explique par le fait que le résidu présente une texture plus importante après le simple filtre linéaire **BSMean** qu'après le filtre non-linéaire **DBM3D**.

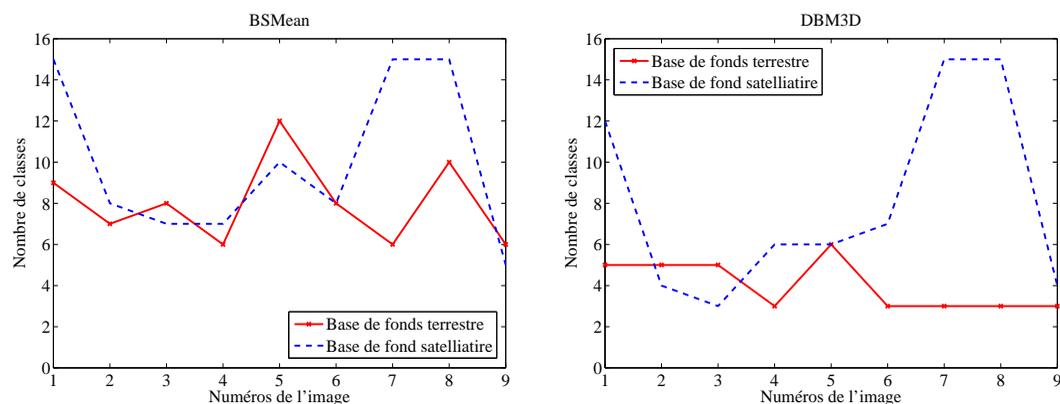


FIGURE 4.13 – Nombre de classes automatiquement sélectionnées dans notre filtrage **GMMF** pour chacune des neuf images des deux bases de fonds pré-filtrées par **BSMean** (à gauche) et par **DBM3D** (à droite) avec un nombre d'itérations déterminé de façon automatique (cf. section 4.6.3)

Cette procédure de sélection automatique du nombre de classes est coûteuse en temps de calcul, car elle nécessite d'appliquer l'ensemble du processus de classification par **CEM** pour chacun des nombres de classes (K) envisagés. En effet c'est une fois seulement que toutes les classifications sont réalisées que la procédure détermine le nombre de classes selon le critère **ICL**. Dans des applications concrètes de détection qui le plus souvent doivent pouvoir fonctionner en temps réel, il n'est donc pas envisageable de sélectionner automatiquement le nombre de classes pour chaque image rencontrée en maximisant le critère **ICL**. C'est pourquoi le plus souvent le nombre de classes est fixé à l'avance. Ce choix du nombre de classes pourra se baser sur l'étude de ce critère **ICL** sur une base de fonds représentative. Dans ce travail, pour l'ensemble des images des deux bases de données on se fixe un nombre de classes K égal à sept ce qui permet d'obtenir de très bonnes performances, comme le montrent les courbes **ROC** de la figure 4.14. La comparaison des deux courbes obtenues dans le cas où le nombre d'itérations est déterminé de façon automatique, montre que les performances de détection obtenues en fixant le nombre de classes égal à sept sont identiques, et même plutôt légèrement supérieures, à celles obtenues quand le nombre de classes est sélectionné de façon automatique pour chaque image des bases de données.

Comportement du **GMMF** face à des fonds homogènes

Dans le cas d'un fond homogène comme celui présenté en figure 4.15a, le filtre le plus performant est le **GMF** et on peut penser que proposer un filtre qui fait intervenir un modèle de fond plus complexe (c.a.d. un modèle avec K classes) va conduire à un surapprentissage et à une dégradation des performances : ce n'est pas le cas. Pour montrer

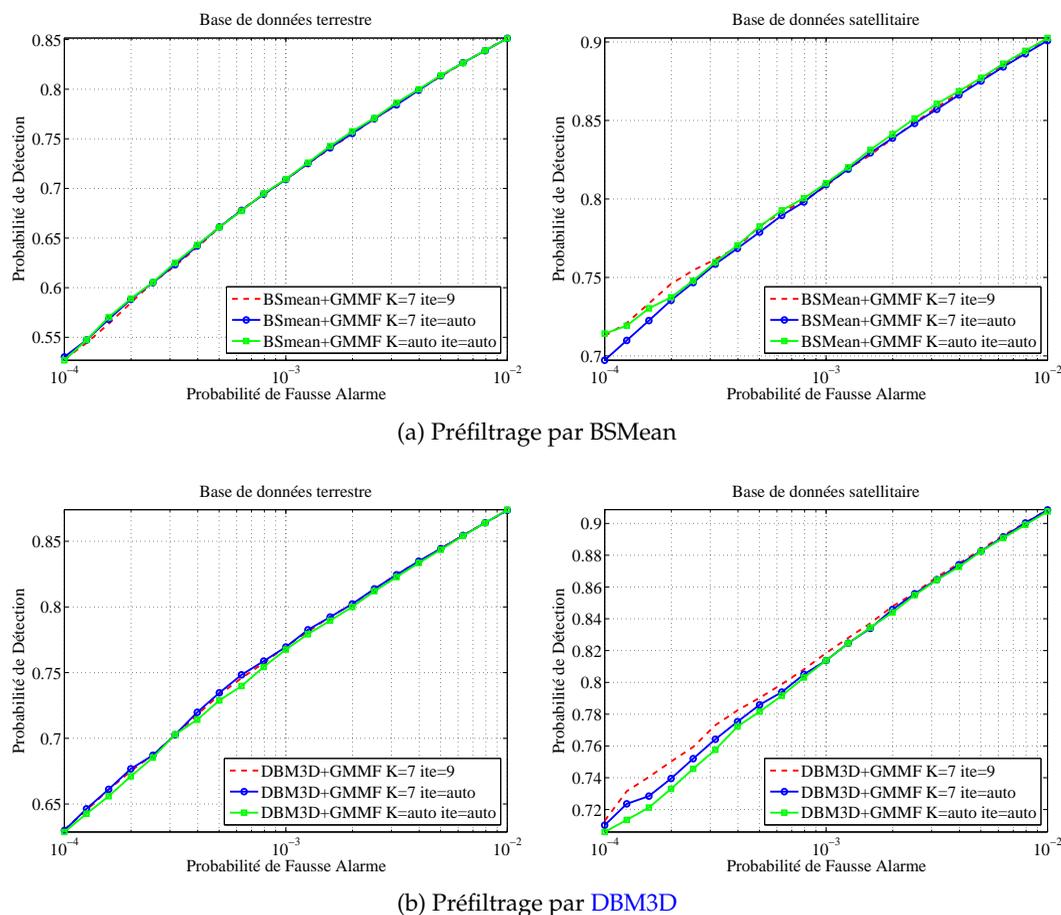


FIGURE 4.14 – Courbes ROC obtenues avec le filtre GMMF i) cas où le nombre de classes est fixé égal à 7 et le nombre d'itérations égal à 9, ii) cas où le nombre de classes reste fixé égal à 7 mais où le nombre d'itérations est déterminé de façon automatique pour chacune des images des bases de données et iii) cas où à la fois le nombre d'itérations et le nombre de classes sont déterminés de façon automatique.

cela, nous avons appliqué l'algorithme GMMF sur un fond homogène naturel (cf. figure 4.15a) acquis avec la caméra Photon utilisée pour construire la base de fond terrestre (cf. section 2.4.1). Après l'étape de suppression de fond (cf. figure 4.15b) l'image résiduelle a l'apparence d'un bruit blanc. L'étape de classification du filtre GMMF conduit comme on peut le voir en figure 4.15c à une répartition aléatoire des classes dans toute l'image. Le nombre de pixel par classes est suffisant pour obtenir des estimations précises des matrices de covariance, une telle classification des pixels ne s'avère pas pénalisante en termes de détection. Les courbes ROC obtenues sur cette image avec les filtres GMF et GMMF ($K = 7$) sont présentées en figure 4.16 et elles montrent que les performances obtenues avec ces deux filtres sont quasiment identiques. On peut noter que si la contrainte de $5N^2$ n'avait pas été respectée, le nombre de classes aurait automatiquement été réduit jusqu'à ce qu'elle le soit. C'est cette contrainte qui permet de maintenir la performance de détection si le nombre initial de classes K est surestimé.

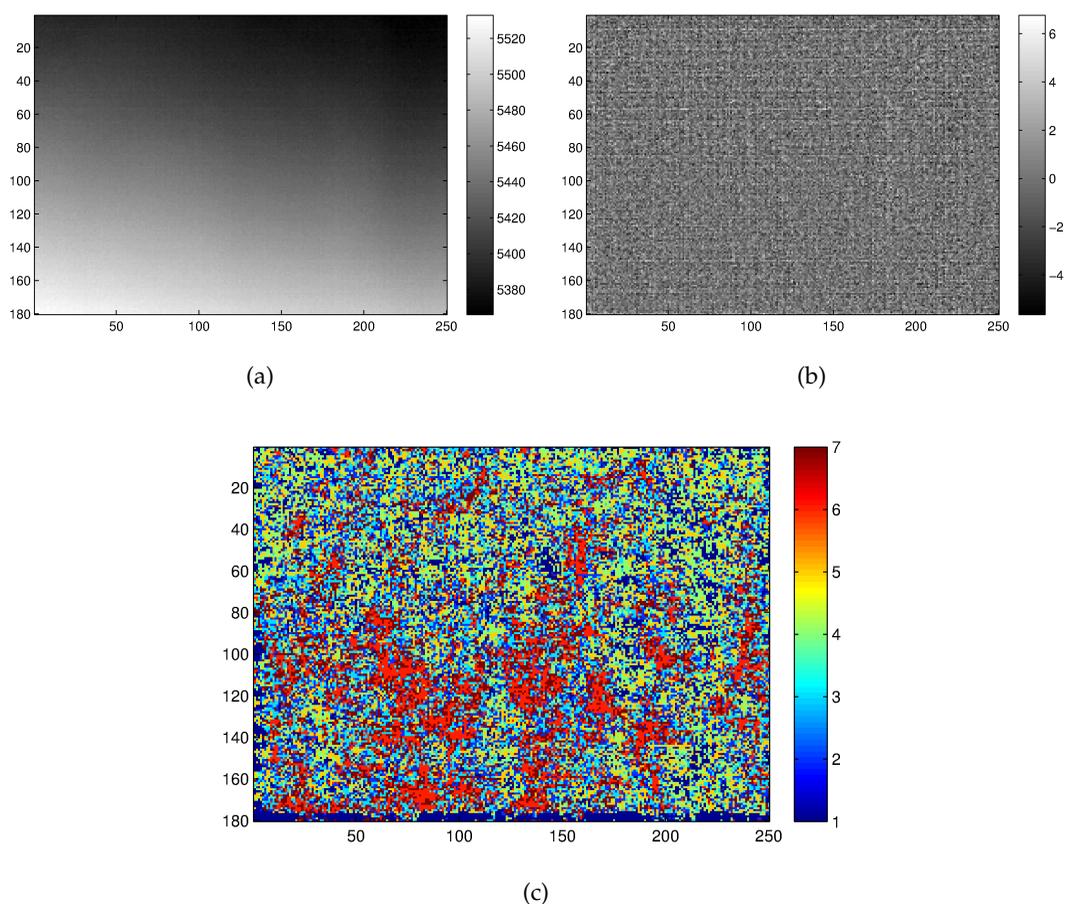


FIGURE 4.15 – (a) Fond homogène naturel. (b) Image résiduelle obtenue en sortie du **DBM3D**. (c) Classification de l’image résiduelle (b) obtenue après 9 itérations avec $N = 9$ et $K = 7$. La classe la plus petite contient 4912 pixels.

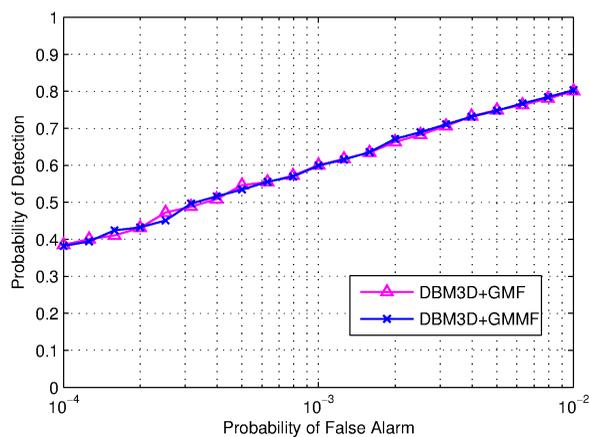


FIGURE 4.16 – Courbes ROC obtenues avec les filtres **DBM3D+GMF** et **BM3D+GMMF** ($K = 7$) sur l’image de la figure 4.15a (intensité des cibles : $s_0 = 10$).

4.7 Modélisation de la signature des cibles

L'approche de détection par modélisation du fond au second ordre (cf. étape 2 du schéma de détection en deux étapes de la figure 1.4, page 24) nécessite de connaître la matrice de covariance Σ_x associée au fond et la signature s des cibles que l'on cherche à détecter. Pour l'instant dans ce chapitre nous avons discuté de variantes de notre approche **GMMF** et du réglage de ses paramètres qui influencent l'estimation de Σ_x mais en ce qui concerne la prise en compte de la signature de la cible des variantes sont aussi possibles.

Nous proposons deux voies pour affiner le modèle de cible. La première consiste à tenir compte du préfiltrage intervenant avant l'application du **GMMF**, l'autre voie consiste à tenir compte du fait qu'une cible a une position subpixelique aléatoire et donc qu'elle n'est quasiment jamais parfaitement centrée sur un pixel.

4.7.1 Prise en compte de la déformation de la cible liée au prétraitement des images

La signature de cibles considérée jusqu'à présent (cf. figure 1.3) correspond à un modèle de l'apparence de la cible observée par le système d'imagerie considéré. Or, la méthode de détection **GMMF** est précédée d'une étape de suppression de fond (cf. schéma 1.4). En pratique, il s'avère que cette étape de suppression de fond modifie le plus souvent la signature de la cible.

Pour les filtres linéaires la déformation des cibles est connue, la figure 4.17 donne un

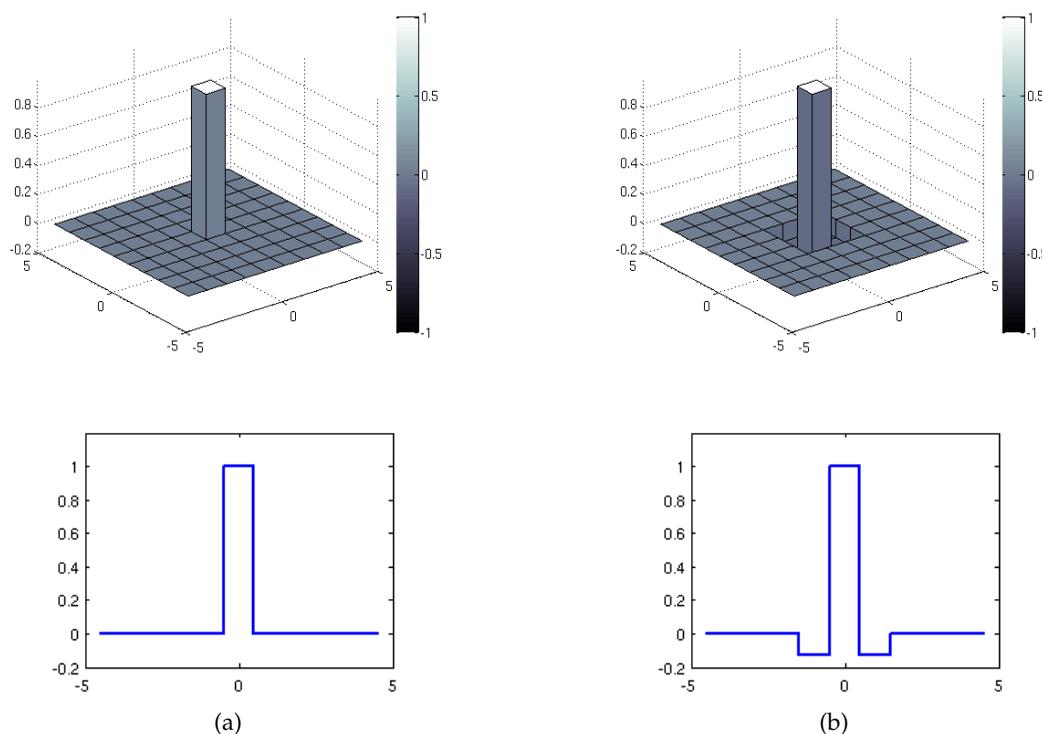


FIGURE 4.17 – Déformation de la signature de la cible liée à l'application d'un préfiltrage BSMean : à gauche signature de la cible dans les données observées, à droite signature de la cible dans l'image de résidu en sortie de BSMean (cas d'une cible centrée sur un pixel).

exemple dans le cas d'un prétraitement par BSMean. Le gain de performances obtenu en considérant la signature de la cible déformée par le BSMean est présenté en figure 4.18. Cette figure montre que associé à un BSMean, le filtrage GMMF faisant intervenir la signature de cible modifiée donnée en figure 4.17b (filtre noté GMMFbs en figure 4.18) est bien plus performant que le GMMF faisant intervenir la signature donnée en figure 4.17a (filtre noté GMMF en figure 4.18). Dans toute la suite, c'est donc la signature de la figure 4.17b qui est utilisée quand le GMMF intervient à la suite d'un BSmean (il reste simplement noté GMMF).

Pour les méthodes de suppression de fond non-linéaires telles que celles proposées dans le chapitre 2, la déformation d'une cible n'est pas prévisible et va dépendre du fond qui l'entoure. On peut penser qu'en moyenne la déformation de la cible va être similaire à la déformation engendrée par le filtre BSMean, mais les résultats d'évaluation de performance de détection présentés en figure 4.18, montrent qu'avec cette hypothèse, on n'améliore pas les performances de détection. Nous considérons donc que la modification de la signature de la cible engendrée par les préfiltrages non-linéaires est négligeable et dans toute la suite, c'est la signature de la figure 4.17a qui est utilisée quand le GMMF intervient à la suite d'un DBM3D.

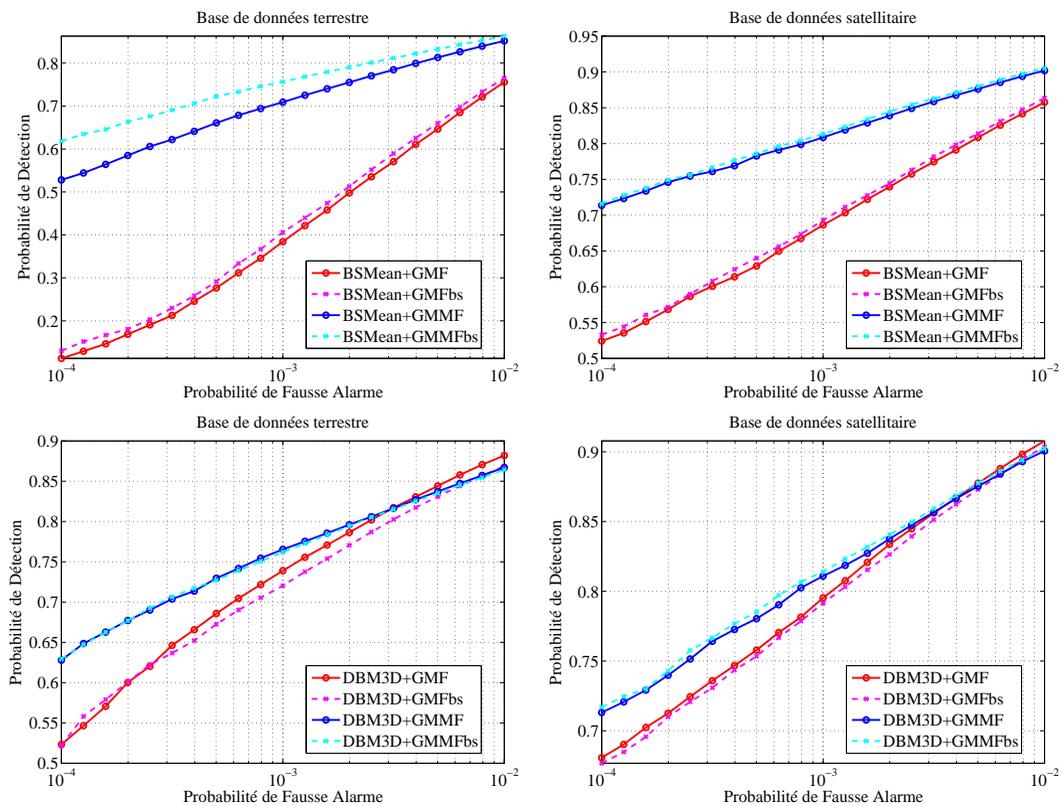


FIGURE 4.18 – Courbes ROC obtenues avec les filtres GMMFbs (filtre GMMF qui fait intervenir la signature de la figure 4.17a) et GMMF (filtre GMMF qui fait intervenir la signature de la figure 4.17b) à la suite d'un BSMean et DBM3D.

4.7.2 Prise en compte de la position subpixellique des cibles

Dans le filtre **GMMF** proposé jusqu'à présent, on considère que la signature des cibles recherchées est unique et ne fait intervenir aucun paramètre inconnu. Or en pratique on sait que la signature d'une cible observée par un système sous-échantillonné dépend très fortement de sa position subpixellique (cf. figure 2.7 de la partie 2.5) et que la position subpixellique d'une cible est parfaitement imprédictible. Quand on considère que la signature des cibles recherchées est unique, comme c'est le cas quand on fait du **MF**, la signature choisie est généralement celle qui correspond au cas où la cible est parfaitement centrée sur un pixel. C'est ce que l'on a fait jusqu'à présent.

On cherche dans cette partie à prendre en compte le fait que la signature de la cible varie en fonction de la position subpixellique de la cible. On suppose que l'on connaît la signature de la cible en fonction de son décalage subpixellique ϵ , celle-ci est notée \mathbf{s}_ϵ . Les deux hypothèses intervenant dans le test d'hypothèse sont alors [Samson *et al.*, 2004] :

$$\begin{cases} H_0 & : \mathbf{r}_x = \mathbf{b}_x & (\text{Fond uniquement}) \\ H_1 & : \mathbf{r}_x = \mathbf{b}_x + \alpha \mathbf{s}_\epsilon & (\text{Présence d'une cible de décalage } \epsilon). \end{cases} \quad (4.11)$$

Samson *et al.* [Samson *et al.*, 2004] ont proposé à partir de ce test différents filtres de détection. On reprend ici deux de leurs propositions pour tenir compte du décalage subpixellique de la cible ϵ et l'intensité de la cible α . Parmi les méthodes qu'ils proposent, ce sont les deux plus compétitives en termes de compromis coût de calcul / performances de détection. La première méthode consiste, sous l'hypothèse qu'une cible est présente (hypothèse H_1), à déterminer par maximum de vraisemblance la position subpixellique de celle-ci. La seconde consiste à accumuler les probabilités de présence de cibles associées à un ensemble fini de positions subpixelliques bien représentatif de l'ensemble des positions subpixelliques que peut avoir la cible.

4.7.2.1 Cible de signe inconnu

Dans un premier temps on considère que le signe de l'intensité de la cible n'est pas connu comme cela est fait dans [Samson *et al.*, 2004].

a) Filtrage **GLRT** (*Generalized Likelihood Ratio Test*) :

Pour faire face à la présence de paramètres inconnus, une première stratégie consiste à appliquer le **GLRT**, soit [Samson *et al.*, 2004] :

$$\Lambda_{GLRT}(\mathbf{r}_x) = \frac{\max_{\alpha \in \mathbb{R}, \epsilon \in [-0.5, 0.5]^2} p(\mathbf{r}_x | H_1, \alpha, \epsilon)}{p(\mathbf{r}_x | H_0)} = \frac{p(\mathbf{r}_x | H_1, \hat{\alpha}_{ML}, \hat{\epsilon}_{ML})}{p(\mathbf{r}_x | H_0)} \underset{H_1}{\overset{H_0}{\leq}} \text{seuil} \quad (4.12)$$

Sous l'hypothèse que $\mathbf{b}_x \sim \mathcal{N}(0, \Sigma_x)$, cela conduit au test de détection suivant [Samson *et al.*, 2004, Samson, 2002] :

$$\frac{[\hat{\mathbf{s}}_{\epsilon_{ML}}^T \Sigma_x^{-1} \mathbf{r}_x]^2}{\hat{\mathbf{s}}_{\epsilon_{ML}}^T \Sigma_x^{-1} \hat{\mathbf{s}}_{\epsilon_{ML}}} \underset{H_0}{\overset{H_1}{\geq}} \text{seuil} \quad \text{avec} \quad \hat{\epsilon}_{ML} = \arg \max_{\epsilon \in [-0.5, 0.5]^2} \frac{[\mathbf{s}_\epsilon^T \Sigma_x^{-1} \mathbf{r}_x]^2}{\mathbf{s}_\epsilon^T \Sigma_x^{-1} \mathbf{s}_\epsilon}. \quad (4.13)$$

Ce test nécessite une maximisation par rapport à ϵ . En pratique, une approximation numérique est faite : on considère un ensemble fini de positions subpixelliques \mathcal{E} et

l'estimateur du ML est calculé sur cet ensemble. Le test à appliquer au final est donc :

$$\max_{\epsilon \in \mathcal{E}} \frac{[\mathbf{s}_\epsilon^T \Sigma_x^{-1} \mathbf{r}_x]^2}{\mathbf{s}_\epsilon^T \Sigma_x^{-1} \mathbf{s}_\epsilon} \underset{H_0}{\overset{H_1}{\gtrless}} \text{seuil}. \quad (4.14)$$

b) Filtrage ALRT (Approximate Likelihood Ratio Test) :

Le test du rapport de vraisemblance n'étant pas optimum au sens de Neyman Pearson, une autre approche consiste à appliquer le ELRT (cf section 1.2.4.1) qui pour le test d'hypothèse donné par l'équation (4.11) correspond à [Samson *et al.*, 2004, Samson, 2002] :

$$\Lambda_{LRT}(\mathbf{r}_x) = \frac{p(\mathbf{r}_x | H_1)}{p(\mathbf{r}_x | H_0)} = \frac{\int_{[-0.5, 0.5]^2} \int_{\mathbb{R}} p(\mathbf{r}_x | H_1, \alpha, \epsilon) p(\alpha) p(\epsilon) d\alpha d\epsilon}{p(\mathbf{r}_x | H_0)} \underset{H_1}{\overset{H_0}{\gtrless}} \text{seuil} \quad (4.15)$$

Dans [Samson *et al.*, 2004], ϵ est distribué selon une loi uniforme sur $[-0.5, 0.5]^2$ et une loi *a priori* non-informative est choisie concernant la distribution de α . Sous l'hypothèse que $\mathbf{b}_x \sim \mathcal{N}(0, \Sigma_x)$, le test à appliquer se reformule alors [Samson *et al.*, 2004] :

$$\int_{[-0.5, 0.5]^2} \underbrace{\frac{1}{[\mathbf{s}_\epsilon^T \Sigma_x^{-1} \mathbf{s}_\epsilon]^{1/2}} \exp\left(\frac{[\mathbf{s}_\epsilon^T \Sigma_x^{-1} \mathbf{r}_x]^2}{2\mathbf{s}_\epsilon^T \Sigma_x^{-1} \mathbf{s}_\epsilon}\right)}_{f(\epsilon|\mathbf{r}_x)} d\epsilon \underset{H_1}{\overset{H_0}{\gtrless}} \text{seuil}. \quad (4.16)$$

En pratique l'intégration par rapport à ϵ doit être approximée numériquement. Samson *et al.* approximent cette intégrale par une quadrature trapézoïdale ne faisant intervenir que les positions demi-pixelliques, notées ϵ_k pour k variant de 0 à 8 (cf. figure 4.19). Le filtre ainsi obtenu est nommé ALRT (Approximate Likelihood Ratio Test) et s'écrit :

$$\Lambda_{ALRT}(\mathbf{r}_x) = \frac{1}{4} \left[f(\epsilon_0|\mathbf{r}_x) + \frac{1}{2} \sum_{k=1}^4 f(\epsilon_k|\mathbf{r}_x) + \frac{1}{4} \sum_{k=5}^8 f(\epsilon_k|\mathbf{r}_x) \right] \quad (4.17)$$

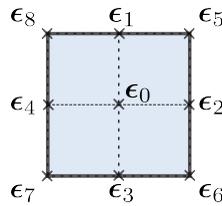


FIGURE 4.19 – Positions subpixelliques considérées pour l'approximation du calcul de l'intégrale (4.21).

4.7.2.2 Cible de signe connu

Les deux approches présentées ci-dessus peuvent être étendues sans difficulté au cas où le signe de l'intensité de la cible est supposé connu.

a) Filtrage GLRT (Generalized Likelihood Ratio Test) :

Concernant l'approche par **GLRT**, quand le signe de α est supposé connu, α_{ML} est l'estimée de α par max de vraisemblance pour $\alpha \in \mathcal{R}$ avec $\mathcal{R} = \mathbb{R}^+$ pour les cibles de contraste positif et $\mathcal{R} = \mathbb{R}^-$ pour celles de contraste négatif. Sous l'hypothèse que $\mathbf{b}_x \sim \mathcal{N}(0, \Sigma_x)$, cela conduit au test de détection suivant :

$$\text{sign}_\alpha \frac{\mathbf{s}_{\hat{\epsilon}_{ML}}^T \Sigma_x^{-1} \mathbf{r}_x}{[\mathbf{s}_{\hat{\epsilon}_{ML}}^T \Sigma_x^{-1} \mathbf{s}_{\hat{\epsilon}_{ML}}]^{1/2}} \underset{H_0}{\overset{H_1}{\geq}} \text{seuil} \text{ avec } \hat{\epsilon}_{ML} = \arg \max_{\epsilon \in [-0.5, 0.5]^2} \text{sign}_\alpha \frac{\mathbf{s}_\epsilon^T \Sigma_x^{-1} \mathbf{r}_x}{[\mathbf{s}_\epsilon^T \Sigma_x^{-1} \mathbf{s}_\epsilon]^{1/2}}, \quad (4.18)$$

avec $\text{sign}_\alpha = +1$ pour les cibles de contraste positif et $\text{sign}_\alpha = -1$ pour les cibles de contraste négatif.

De même que précédemment (cf section 4.7.2.1) la maximisation est approximée numériquement. Le filtre de détection au final s'écrit :

$$\max_{\epsilon \in \mathcal{E}} \text{sign}_\alpha \frac{\mathbf{s}_\epsilon^T \Sigma_x^{-1} \mathbf{r}_x}{[\mathbf{s}_\epsilon^T \Sigma_x^{-1} \mathbf{s}_\epsilon]^{1/2}} \underset{H_0}{\overset{H_1}{\geq}} \text{seuil}. \quad (4.19)$$

b) Filtrage ALRT (Approximate Likelihood Ratio Test) :

Concernant l'approche par **ALRT**, dans le cas où la signature de la cible est connue, le test à seuiller est alors :

$$\Lambda_{LRT}(\mathbf{r}_x) = \frac{p(\mathbf{r}_x | H_1)}{p(\mathbf{r}_x | H_0)} = \frac{\int_{[-0.5, 0.5]^2} \int_{\mathcal{R}} p(\mathbf{r}_x | H_1, \alpha, \epsilon) p(\alpha) p(\epsilon) d\alpha d\epsilon}{p(\mathbf{r}_x | H_0)} \quad (4.20)$$

avec $\mathcal{R} = \mathbb{R}^+$ pour les cibles de contraste positif et $\mathcal{R} = \mathbb{R}^-$ pour celles de contraste négatif.

Considérant les mêmes lois *a priori* pour α et ϵ que celles considérées dans la section 4.7.2.1, sous l'hypothèse que $\mathbf{b}_x \sim \mathcal{N}(0, \Sigma_x)$, le test à appliquer se reformule alors :

$$\Lambda_{LRT}(\mathbf{r}_x) \propto \int_{[-0.5, 0.5]^2} \underbrace{\frac{1}{[\mathbf{s}_\epsilon^T \Sigma_x^{-1} \mathbf{s}_\epsilon]^{1/2}} \exp\left(\frac{[\mathbf{s}_\epsilon^T \Sigma_x^{-1} \mathbf{r}_x]^2}{2\mathbf{s}_\epsilon^T \Sigma_x^{-1} \mathbf{s}_\epsilon}\right) \left[1 + \text{sign}_\alpha \text{erf}\left(\frac{\mathbf{s}_\epsilon^T \Sigma_x^{-1} \mathbf{r}_x}{[2\mathbf{s}_\epsilon^T \Sigma_x^{-1} \mathbf{s}_\epsilon]^{1/2}}\right)\right]}_{g(\epsilon|\mathbf{r}_x)} d\epsilon, \quad (4.21)$$

où $\text{erf}(\cdot)$ est la fonction d'erreur de Gauss.¹

On peut ensuite approximer l'intégrale exactement comme cela est fait dans la section 4.7.2.1, et le filtre à seuiller est alors :

$$\Lambda_{ALRT}(\mathbf{r}_x) = \frac{1}{4} \left[g(\epsilon_0|\mathbf{r}_x) + \frac{1}{2} \sum_{k=1}^4 g(\epsilon_k|\mathbf{r}_x) + \frac{1}{4} \sum_{k=5}^8 g(\epsilon_k|\mathbf{r}_x) \right]. \quad (4.22)$$

4.7.2.3 Signature subpixellique des cibles

Les méthodes de détection tenant compte de la position subpixellique des cibles considèrent connue la signature des cibles en fonction de leur position subpixellique. Les cibles étant non résolues, leurs signatures en fonction du décalage subpixellique sont données par la **PSF** échantillonnée du système optique. Comme discuté en section 2.5, le système est supposé limité par la diffraction avec un paramètre $r_c = 1.5$. La **PSF**

1. $\text{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt$

échantillonnée de ce système est représentée en figure 4.20. On peut noter que cette PSF échantillonnée est celle que l'on utilise pour modéliser les cibles. Le fait d'utiliser exactement la même fonction pour simuler les cibles et les détecter au moyen d'un GLRT risque de fournir des performances de détection supérieures à celles que l'on obtiendrait avec des cibles réelles. C'est pourquoi, nous proposons d'approximer la PSF échantillonnée du système par la fonction « tipi » donnée en figure 4.21. La figure 4.22 présente les signatures de cibles que l'on obtient avec cette fonction tipi pour les neuf positions subpixelliques définies en figure 4.19 intervenant dans le filtrage ALRT. On peut noter que dans le cas d'une cible de position subpixellique $\epsilon = (0, 0)$, on retrouve la signature de cibles utilisée jusqu'à présent quand on ne tient pas compte de la position subpixellique éventuelle des cibles (cf. motif donné en figure 1.3).

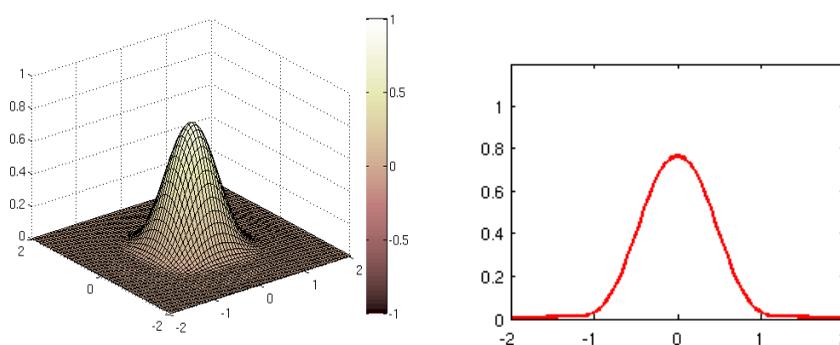


FIGURE 4.20 – PSF du système échantillonnée ($r_c = 1.5$).

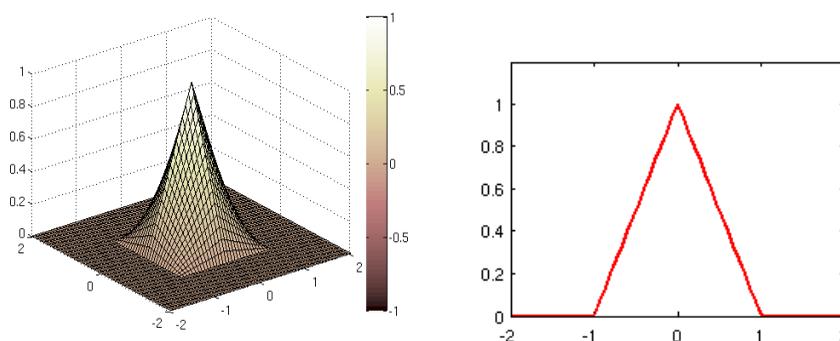


FIGURE 4.21 – Fonction utilisée pour approximer la PSF du système.

Dans la section suivante, nous étudions l'apport de la prise en compte de la position subpixellique des cibles au sein du test d'hypothèse en termes de compromis performances de détection / coût de calcul. On peut noter qu'il est tout à fait possible de prendre aussi en compte la déformation engendrée par le préfiltrage comme proposée en section 4.7.2 même si aucune évaluation n'a été menée sur le sujet.

4.7.2.4 Résultats obtenus

Le contexte qui nous intéresse étant celui de la détection de cibles de contraste positif (autrement dit : signe de l'intensité de la cible connu), nous évaluons les performances des deux approches décrites en section 4.7.2.2. Le filtre de détection qui suit le schéma de

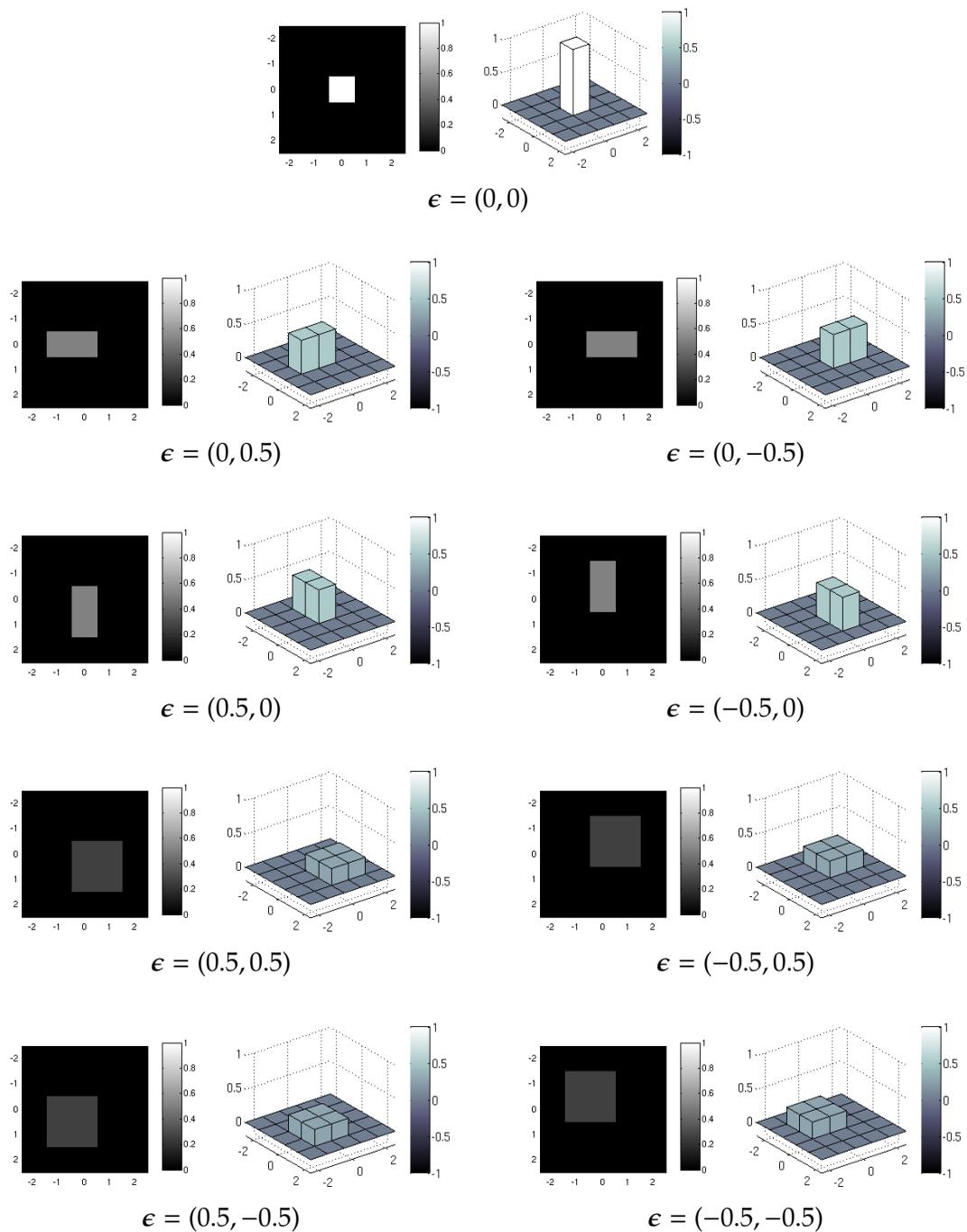


FIGURE 4.22 – Les neuf signatures de cibles associées aux neuf décalages subpixelliques définis en figure 4.19.

principe du **GMMF** donné en figure 4.6 mais qui fait intervenir le test de décision défini en équation (4.19) à la place du filtrage adapté donné en équation (4.5) est noté **GMMF-PSGLRT** (**GMMF** avec prise en compte de la Position Subpixellique par **GLRT**). Celui qui fait intervenir le test défini en équation (4.22) est noté **GMMF-PSALRT** (**GMMF** avec prise en compte de la Position Subpixellique par **ALRT**). Pour les deux filtres **GMMF-PSGLRT** et **GMMF-PSALRT**, on fait intervenir les neuf signatures de cibles définies en figure 4.22.

a) Analyse de cas idéaux où les cibles incrustées ont toute la même position subpixellique

Pour mieux comprendre le comportement de ces filtres, on évalue tout d'abord leurs performances dans le cas de cibles ayant toute la même position subpixellique (soit $\epsilon = (0.5, 0.5)$, soit $\epsilon = (0.5, 0)$, soit $\epsilon = (0, 0)$). Des évaluations ont été réalisées dans le cas où les filtres **GMMF**, **GMMF-PSGLRT** et **GMMF-PSALRT** sont appliqués directement sur les données brutes.

Au vue des performances obtenues dans les différents cas testés (cf. figure 4.23), le premier résultat à retenir est que les deux filtres considérés cherchant à tirer parti de la position subpixellique des cibles (**GMMF-PSGLRT** et **GMMF-PSALRT**) donnent des résultats équivalents. Maintenant analysons leurs performances vis à vis de celles du **GMMF**.

Dans le cas où les cibles sont très fortement décentrées par rapport au centre du pixel, les filtres cherchant à tirer parti de la position subpixellique sont nettement plus performants que le filtre **GMMF** qui considère que la cible est toujours centrée, comme on peut le voir sur la figure 4.23a. Dans ce cas où les cibles sont moins décentrées par rapport au centre du pixel ($\epsilon = (0.5, 0)$), le gain de performance apporté par la prise en compte de la position subpixellique des cibles est plus faible mais les filtres **GMMF-PSGLRT** et **GMMF-PSALRT** restent tout de même nettement plus performants que le filtre **GMMF**. Par contre pour les cibles qui sont que très faiblement décentrées, la prise en compte de la position subpixellique dégrade les performances. Sur la figure 4.23c, on voit que dans les cas où les cibles sont centrées sur les pixels, le filtre **GMMF** est nettement plus performant que les filtres **GMMF-PSGLRT** et **GMMF-PSALRT**.

b) Cas pratique : Cibles de position subpixellique aléatoire

Dans la pratique la position subpixellique des cibles est aléatoire, une cible n'a aucune raison de se trouver en une certaine position plutôt qu'en une autre position. La figure 4.24 montre que pour des cibles ayant des positions subpixelliques aléatoires, la prise en compte de la position subpixellique dégrade les performances de détection. Cela s'explique de la façon suivante. Les positions subpixelliques pour lesquelles on obtient un vrai gain de performances représentent en réalité qu'une portion très faible de toutes les positions subpixelliques pouvant être associées à une cible. L'histogramme de la figure 4.25b, montre que l'intensité du pixel central de la cible chute de façon importante à cause de la position subpixellique des cibles dans bien moins de la moitié des cas. La prise en compte de la position subpixellique, même si elle améliore les performances de détection d'un nombre limité de cibles (les plus décentrées), va dégrader les performances de détection de la majorité des cibles. De plus cette prise en compte de la position subpixellique comme cela est fait dans les méthodes **GMMF-PSGLRT** et **GMMF-PSALRT** a le défaut d'augmenter le coût de calcul des algorithmes. Dans le cadre d'une problématique de détection, la prise en compte de la position subpixellique des cibles n'est pas pertinente,

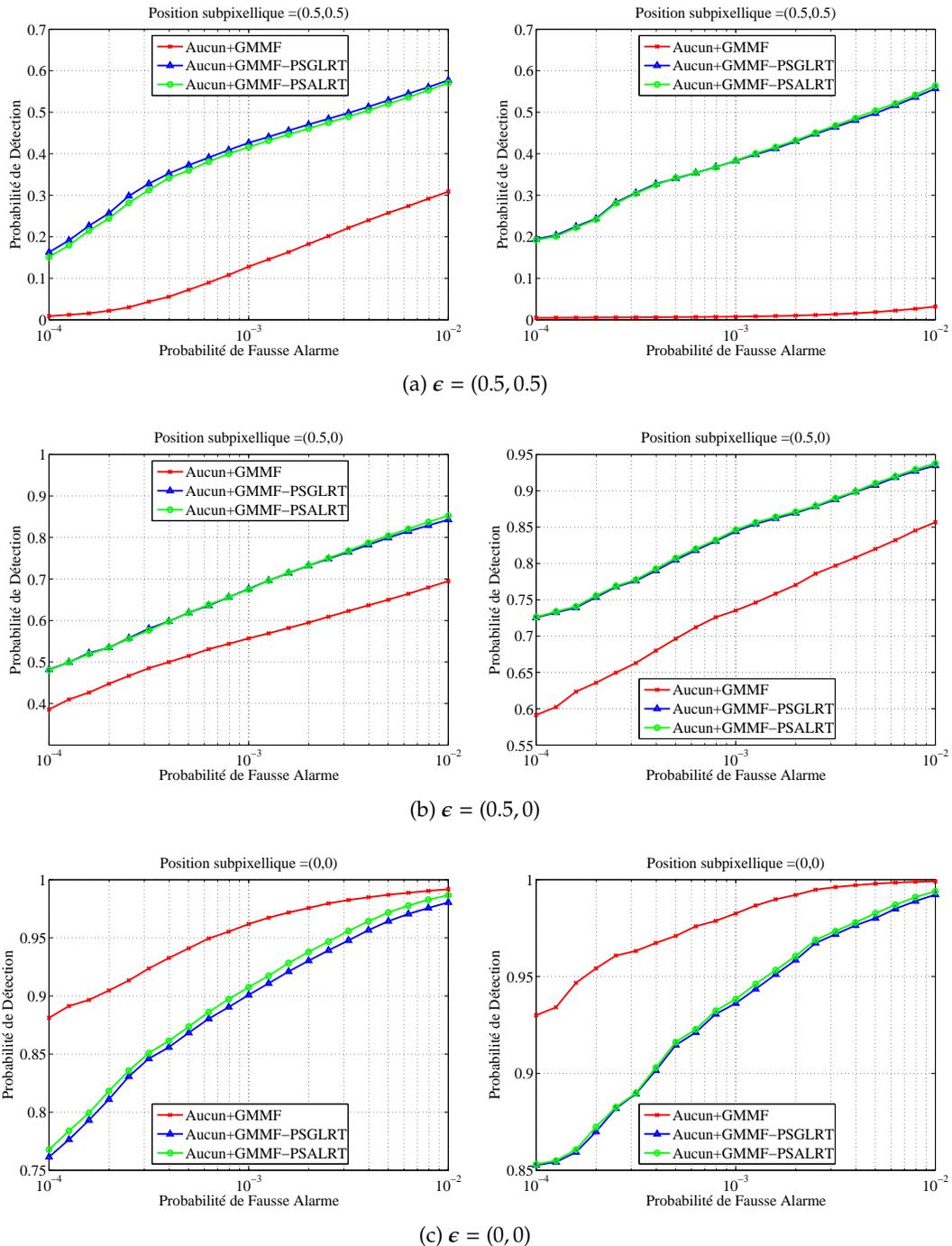


FIGURE 4.23 – Courbes ROC obtenues pour les filtres GMMF, GMMF-PSGLRT et GMMF-PSALRT, dans le cas où toutes les cibles incrustées ont une même position subpixelique. A gauche : base de données terrestre. A droite : base de données satellitaire.

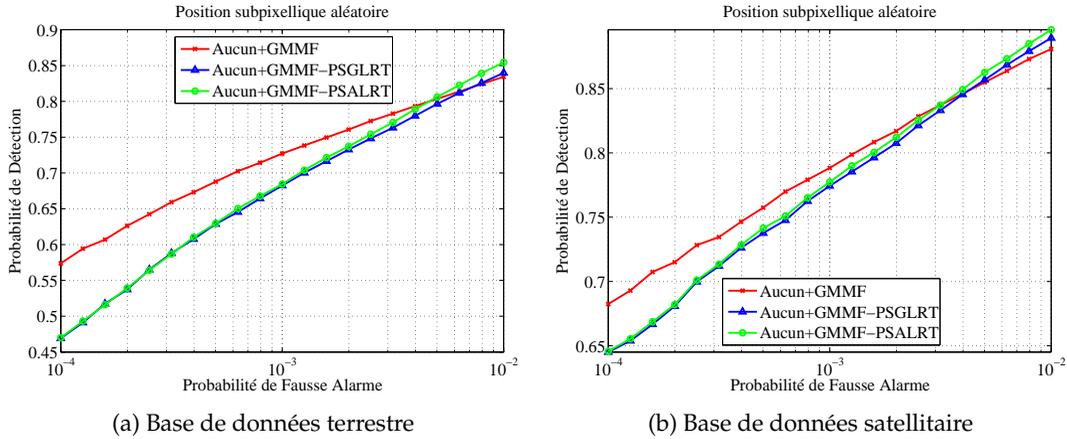


FIGURE 4.24 – Courbes ROC obtenues pour les filtres GMMF, GMMF-PSGLRT et GMMF-PSALRT, dans le cas où la position subpixelique des cibles incrustées est tirée dans un intervalle $[-0.5, 0.5]^2$ selon une loi uniforme.

elle pourra éventuellement être utile dans un second temps pour améliorer la précision de localisation des cibles.

Notons que la conclusion précédente est vraie quelle que soit la dimension de la cible considérée. En effet, plus la cible est étroite et moins fréquents sont les cas où l'intensité du pixel central de la cible chute de façon importante à cause de la position subpixelique (cf. histogramme de la figure 4.25a). D'un autre côté, plus la cible est large et moins sa signature dépend de la position subpixelique, autrement dit moins la plage de variation de l'intensité du pixel central est grande (comparaison des figures 4.26b et 4.26c). Par exemple pour une cible étendue correspondant à un $r_c = 0.8$, cette intensité ne varie que entre 0.2 et 0.4, c'est pourquoi pour les cibles étendues la prise en compte de la position subpixelique ne permet pas d'augmenter les performances de détection.

Pour conclure, le seul cas dans lequel la prise en compte de la position subpixelique est importante est le cas d'une cible fortement décentrée sur plusieurs instants de mesures (cible fixe). Dans les configurations réelles le mouvement de l'objet et de la ligne de visé va rendre l'occurrence de ce cas assez rare.

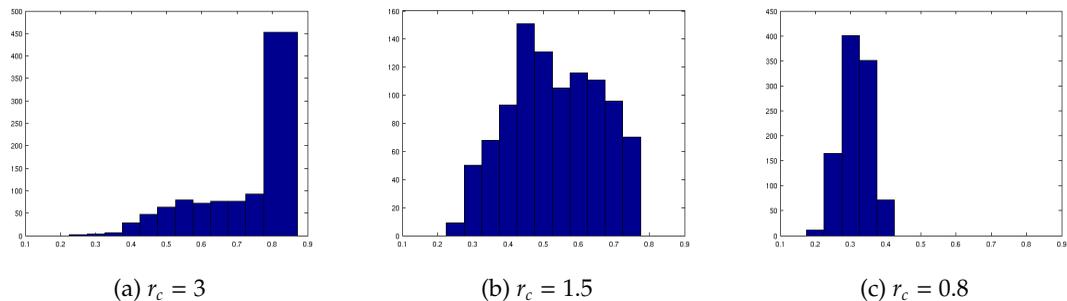


FIGURE 4.25 – Histogramme de l'intensité du pixel central de la cible pour 10^3 cibles de position subpixelique aléatoire, d'intensité totale $s_0 = 1$ pour différentes valeurs de r_c .

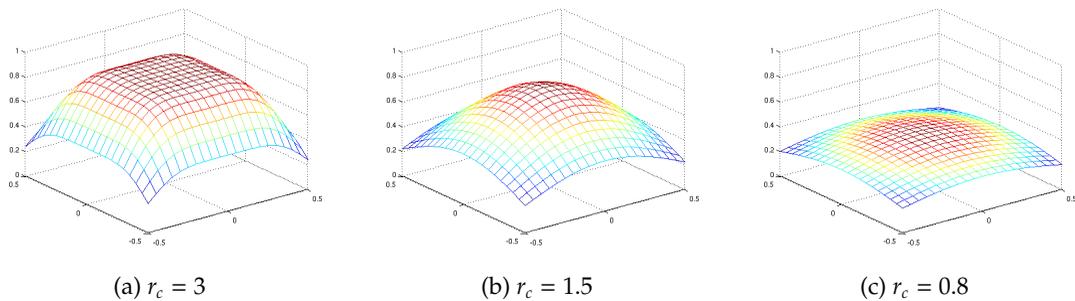


FIGURE 4.26 – Intensité du pixel central de la cible en fonction de sa position subpixelique $\epsilon = (\epsilon_i, \epsilon_j)$ pour différentes valeurs de r_c .

4.8 Extension au contexte spatio-temporel

Le filtre de détection **GMMF** proposé n'a pour le moment été présenté que dans un cadre de détection mono-image où seule l'information contenue dans l'image en cours de traitement est utilisée pour détecter les cibles éventuellement présentes. Dans cette partie nous proposons d'essayer d'améliorer les performances de détection en utilisant l'information contenue dans les images précédentes. L'image dans laquelle on cherche à détecter les cibles est notée \mathcal{I}_t et les L images précédentes utilisées : $\{\mathcal{I}_{t-1}, \dots, \mathcal{I}_{t-L}\}$.

Les approches usuelles de filtrage adapté spatio-temporel consistent à considérer, associée à chaque pixel x de l'image \mathcal{I}_t , une observation vectorielle \mathbf{r}_x constituée des pixels contenus dans un voisinage non plus uniquement spatial mais spatio-temporel de dimension $N \times N \times (L + 1)$, comme cela est illustré en figure 4.27 dans le cas où $L = 1$. Dans ce cas, la signature de la cible \mathbf{s} doit elle aussi être étendue au contexte spatio-temporel.

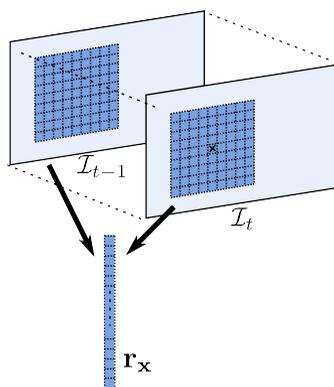


FIGURE 4.27 – Observation vectorielle complète spatio-temporelle associée au pixel x marqué par une croix ($L = 1$ et $N = 9$).

Dans un premier temps, nous considérons simplement que la signature de la cible \mathbf{s} est un vecteur de dimension égale à celle de \mathbf{r}_x dont la composante associée au pixel en cours de débruitage est égale à 1 et dont toutes les autres composantes sont nulles. Une telle signature correspond à faire l'hypothèse que la cible n'est présente qu'une seule fois dans le voisinage spatio-temporel de dimension $N \times N \times (L + 1)$, c'est à dire que sa vitesse est supérieure à $(N + 1)/\sqrt{2}$ pixels par image. Dans le chapitre 5, nous verrons quelles performances de détection on peut obtenir avec une telle signature de cibles pour

des cibles respectant la contrainte sur la vitesse mais aussi pour des cibles de vitesse beaucoup plus faible. Dans les cas où la vitesse et la direction de déplacement de la cible sont connues *a priori*, des signatures de cibles qui tiennent compte de ces *a priori* doivent être considérées. De plus on peut bien entendu chercher à tenir compte de la position subpixelique en appliquant les mêmes méthodes que celles décrites en section 4.7.2 en étendant toutes les signatures de cibles au contexte spatio-temporel ; ou chercher à tenir compte du préfiltrage appliqué comme proposé en section 4.7.1. Comme la problématique qui nous intéresse dans cette section est l'amélioration de l'estimation des matrices de covariance du fond, nous considérons que la signature de la cible spatio-temporelle \mathbf{s} est unique et bien connue.

Comme dans le cas spatial, une méthode simple de détection par filtrage adapté consiste à estimer une unique matrice de covariance à partir de l'ensemble des observations « temporelles » associées à chacun des pixels de l'image $\mathcal{I}_t : \{\mathbf{r}_x\}_{x \in \mathcal{I}_t}$. Cependant celle-ci n'est pas capable de prendre en compte d'éventuelles variations de la corrélation spatio-temporelle du fond. Nous cherchons ici à adapter la méthode de filtrage GMMF proposée pour pouvoir tenir compte de ces variations. Sur le modèle du cas spatial, on classe les pixels par CEM en fonction de leur corrélation spatio-temporelle locale et on estime une matrice de covariance pour chaque classe.

4.8.1 Classification et filtrage adapté spatio-temporels

Une première approche consiste à considérer une observation restreinte spatio-temporelle constituée des pixels inclus dans le même voisinage que celui utilisé pour construire l'observation complète en omettant, comme on le fait dans le contexte spatial, les pixels se trouvant dans un petit voisinage $n \times n \times 1$ centré sur le pixel \mathbf{x} . Ces pixels non pris en compte sont ceux dont la composante de cible risque d'être non nulle, quand une cible est présente en \mathbf{x} . Le principe de constitution de ces observations restreintes temporelles est illustré en figure 4.28 dans le cas où $L = 1$. Le schéma de principe du filtrage GMMF temporel qui découle de cette approche est présenté en figure 4.29. Ce filtrage est pertinent pour les cibles de vitesse très élevée (supérieure à $(N + 1) / \sqrt{2}$ pixels par image) ou dans le cas où l'on cherche à détecter des cibles au moment où elles apparaissent. Comme l'illustre la figure 4.30, la segmentation est robuste à la présence de telles cibles, la

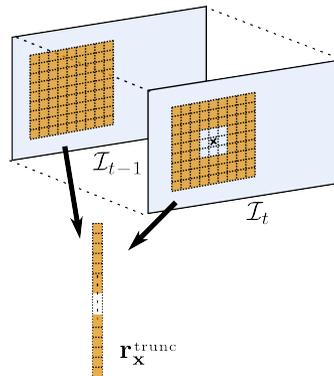


FIGURE 4.28 – Observation vectorielle restreinte temporelle associée au pixel \mathbf{x} marqué par une croix ($L = 1$, $N = 9$ et $n = 3$).

classe associée à un pixel cible est bien celle qui lui serait associée si aucune cible n'avait été présente.

Cependant la segmentation n'est pas robuste à la présence de cibles de vitesse plus faible, comme l'illustre la figure 4.31. En effet dans le cas où une cible de vitesse inférieure à $(N + 1)/\sqrt{2}$ pixels par image est présente au pixel \mathbf{x} , l'observation restreinte $\mathbf{r}_x^{\text{trunc}}$ ne contient plus des pixels qui correspondent uniquement à du fond.

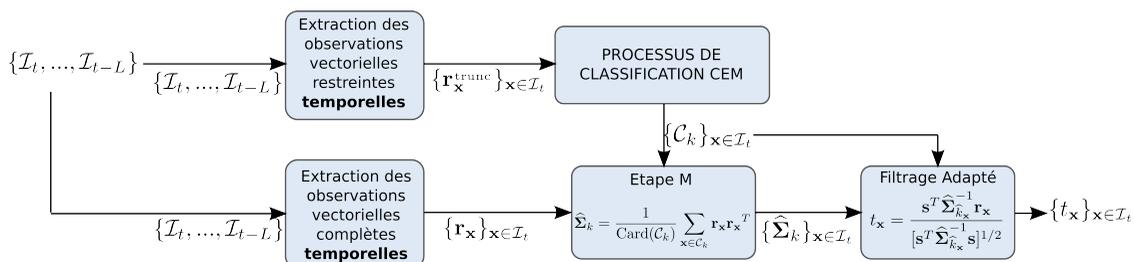


FIGURE 4.29 – Schéma de principe du filtrage GMMF dans le cas où la classification et le filtrage adapté sont spatio-temporels.

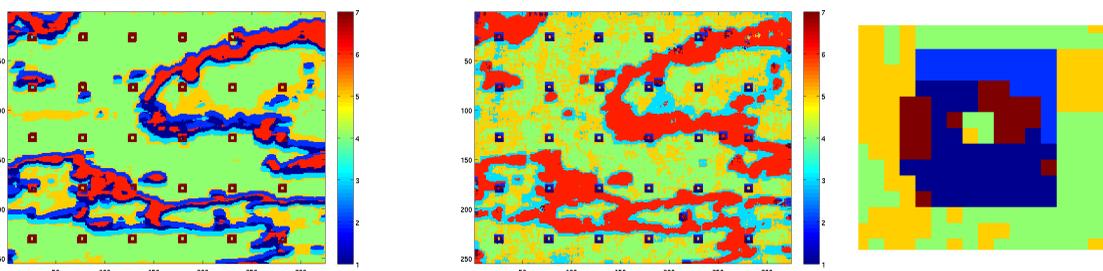


FIGURE 4.30 – Classification obtenue en utilisant une observation restreinte $\mathbf{r}_x^{\text{trunc}}$ temporelle (cf. figure 4.28) dans le cas de cibles naissantes, c.a.d. qui apparaissent dans l'image \mathcal{I}_t alors qu'elles n'étaient pas présentes dans l'image \mathcal{I}_{t-1} ($L = 1, N = 9$ et $n = 3$). A gauche segmentation obtenue directement sur les données brutes (pas de prétraitement), à droite segmentation obtenue suite à un préfiltrage DBM3D et zoom au niveau d'une des cibles.

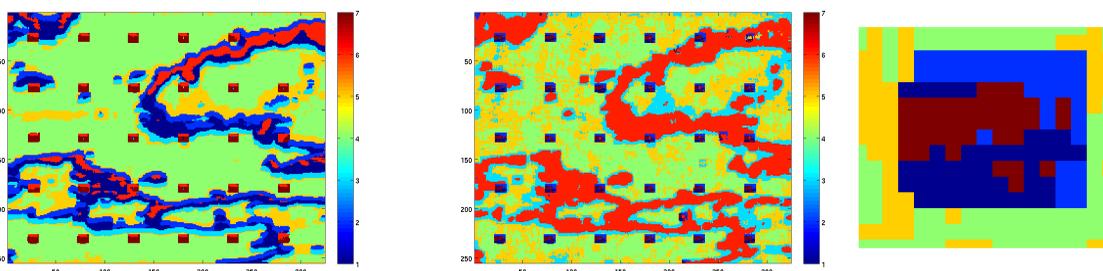


FIGURE 4.31 – Classification obtenue dans le cas de cibles de vitesse 2.25 pixels par image, qui sont donc présentes dans les images \mathcal{I}_t et \mathcal{I}_{t-1} ($L = 1, N = 9$ et $n = 3$). A gauche segmentation obtenue directement sur les données brutes (pas de prétraitement), à droite segmentation obtenue suite à un préfiltrage DBM3D et zoom au niveau d'une des cibles.

4.8.2 Classification spatiale et filtrage adapté spatio-temporel

Il s'avère que pour les données testées, comme le montre la figure 4.32, les segmentations obtenues en l'absence de cible en appliquant le processus de classification CEM sur les observations complètes temporelles (cf figure 4.27) sont très similaires aux segmentations que l'on obtient en utilisant les observations complètes spatiales qui contiennent uniquement les pixels inclus dans un voisinage de dimension $N \times N \times 1$ (soit $L=0$).

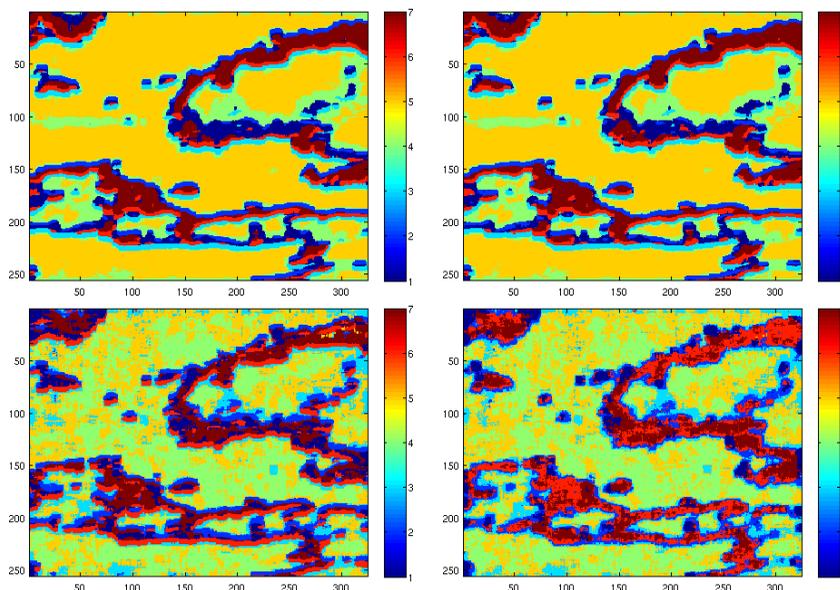


FIGURE 4.32 – Segmentation en l'absence de cible obtenue en classant les pixels à partir des observations complètes \mathbf{r}_x spatiales ou spatio-temporelles (contenant les pixels inclus dans un voisinage de dimension : $N \times N \times 1$ ou $N \times N \times 2$). En haut : aucun prétraitement des données, en bas prétraitement par DBM3D. A gauche : observation spatiale, à droite observation temporelle.

On note que la corrélation spatiale du fond varie au niveau des zones de fortes transitions d'intensité telles que les bords de nuages, et que c'est dans ces mêmes zones que la corrélation spatio-temporelle du fond varie d'après notre méthode de classification des pixels. Cela vient du fait que la corrélation temporelle du fond ne semble pas varier fortement. La corrélation temporelle du fond est liée au mouvement de celui-ci entre les différentes acquisitions constituant la séquence. Le fond est corrélé temporellement ou non, selon que le fond ait bougé ou non entre les acquisitions. Dans certains contextes particuliers comme celui où l'on observe une côte maritime, contenant en même temps une zone de mer et une zone terrestre, on a une variation de la corrélation temporelle du fond en fonction de la position spatiale de celui-ci (selon que l'on se trouve dans la zone terrestre ou la zone de mer). Comme dans notre contexte on observe uniquement des fonds nuageux, la variation de la corrélation temporelle en fonction de la position spatiale du fond est faible et nous considérons qu'elle est négligeable.

Suite à ces observations et afin de disposer d'une classification des pixels qui soit robuste à la présence de cibles de vitesse inférieure à $(N + 1)/\sqrt{2}$ pixels par image on propose de considérer exactement la même classification que celle utilisée dans la version

spatiale de l'algorithme **GMMF**. Le schéma de principe du filtrage **GMMF** temporel qui découle de cette approche est présenté en figure 4.33.

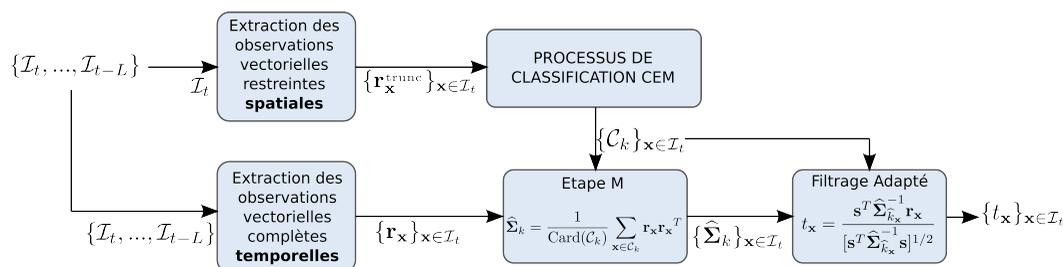


FIGURE 4.33 – Schéma de principe de la version du filtrage **GMMF** temporel retenue : la classification est spatiale et le filtrage adapté est spatio-temporel.

Dans la suite c'est cette méthode de détection qui est retenue comme version temporelle de l'algorithme **GMMF**, car elle est basée sur une classification des pixels robuste à la présence de cibles de vitesse quelconque. Comme nous le verrons au chapitre 5 section 5.2, le filtre **GMMF** temporel proposé a l'avantage d'être capable de tirer parti d'une éventuelle corrélation temporelle des fonds tout en restant très performant face à des fonds de très faible corrélation temporelle. Notons par ailleurs qu'il ne nécessite pas de compensations précises préalables du mouvement du fond.

On peut noter que toutes ces méthodes ont le défaut de ne pas être robustes à la présence de cibles de très faible vitesse. Si on se trouve dans un contexte où les cibles ont un risque non négligeable d'avoir un tel comportement alors on peut le prendre en compte dans la signature de la cible, mais cela va en contrepartie fortement diminuer les performances de détection des cibles ayant une vitesse supérieure. Pour faire face au fait que la vitesse et la direction de propagation des cibles sont inconnues, il est possible de considérer un test **GLRT** au sein duquel ces paramètres inconnus sont estimés par maximum de vraisemblance. Nous reviendrons sur ce point à la fin de ce manuscrit dans la section abordant les perspectives.

Conclusion

Dans le cadre du schéma de détection en deux étapes considéré pour répondre à notre problématique de détection (cf section 1.5, figure 1.4, page 24), nous avons proposé une seconde étape de filtrage adapté basée sur une modélisation du fond au second ordre.

Elle consiste dans un premier temps à regrouper les pixels ayant un comportement au second ordre similaire, en utilisant un algorithme de classification **CEM** (*Classification Expectation-Maximisation*) basé sur une modélisation du fond par un modèle de mélange de gaussiennes centrées. Pour faire face à la présence de cibles dans l'image observée le **CEM** est appliqué sur des observations vectorielles restreintes par rapport à celles intervenant dans le filtrage adapté **MF** appliqué pour détecter les cibles. L'observation restreinte associée à un pixel est constituée de l'intensité des pixels qui sont inclus dans une couronne de pixels centrée sur le pixel à classifier. Ce schéma simple de classification permet d'être robuste à la présence de cibles sans avoir recours à des techniques plus complexes comme par exemple contraindre la connexité des pixels par champs de Markov. Les matrices de covariance nécessaires pour le **MF** sont ensuite estimées à partir de la

classification obtenue en chaque pixel. Pour finir en chaque pixel on applique le MF impliquant la matrice de covariance estimée sur la classe à laquelle il appartient.

De nombreuses variantes de cette approche ont été étudiées dans cette thèse et la solution proposée résulte de choix qui ont été validés par des études de performances. Nous avons entre autre cherché à prendre en compte la position subpixelique des cibles et proposé de ne pas classifier les pixels mais de considérer l'ensemble du mélange de gaussiennes au sein du GLRT. Cependant ces variantes se sont avérées moins performantes que l'algorithme GMMF (*Gaussian Mixture Matched Filter*) retenu plus simple.

L'algorithme de détection résultant GMMF implique un nombre très limité de paramètres et dans ce chapitre des techniques de réglages automatiques de certains d'entre eux ont été proposées. Le GMMF a de plus l'avantage de pouvoir faire face à des fonds non-stationnaires. Il est capable de traiter des textures résiduelles présentes après une simple suppression de fond linéaire et ainsi d'augmenter considérablement la probabilité de détection à taux de fausses alarmes fixe, comme cela est montré dans le chapitre suivant.

Une version spatio-temporelle de ce filtre GMMF a aussi été développée dans ce chapitre. Elle a l'avantage de ne pas nécessiter d'estimation précise préalable du mouvement du fond. Les performances de la version spatiale et de la version spatio-temporelle du filtre GMMF retenues sont analysées dans le chapitre suivant.

Chapitre 5

Évaluation des performances et analyse critique des méthodes de détection proposées

Introduction

Ce chapitre a pour but d’analyser de façon plus approfondie les performances des nouvelles méthodes de détection que nous avons proposées dans les chapitres 3 et 4 et de comparer ces méthodes à l’existant. En s’appuyant sur le schéma de détection en deux étapes de la figure 1.4 présentée dans le chapitre 1, nous allons chercher à déterminer quelle est l’association de filtres basés sur la modélisation du fond au premier ordre d’une part et sur la modélisation du fond au second ordre d’autre part, qui donne les meilleures performances et celle qui donne le meilleur compromis coût de calcul / performances. Les performances des filtres de détection sont évaluées sur des fonds réels dans lesquels des cibles synthétiques sont incrustées comme cela a été présenté au chapitre 2.

Organisation du chapitre — La section 5.1 analyse les performances des filtres de détection spatiaux basés sur une modélisation du fond au premier ordre qui ont été proposés au chapitre 3 (DNLM et DBM3D). La section 5.2 analyse les performances des filtres de détection spatiaux basés sur une modélisation du fond au second ordre qui ont été proposés au chapitre 4 (GMMF et ses variantes). La section 5.3 est consacrée à l’évaluation des performances des algorithmes de détection spatio-temporels. L’influence de la vitesse des cibles et du mouvement du fond y est analysée.

5.1 Algorithmes de détection par suppression de fond

Dans cette section, nous commençons par faire un rappel des valeurs des paramètres des filtres DNLM et DBM3D proposés au chapitre 3. Ceux-ci sont ensuite comparés, dans le cas de cibles dites nominales (c.a.d. modélisées avec les paramètres suivants : $s_o = 50$, $r_c = 1.5$ et ϵ tiré dans un intervalle $[-0.5, 0.5]^2$ selon une loi uniforme), à diverses méthodes de détection spatiales existantes basées sur des approches par suppression de fond BS (*Background Suppression*). Dans un second temps, nous étudions les performances de détection du filtre BS le plus performant, soit DBM3D, en fonction des caractéristiques

des cibles (dimension et intensité). Enfin, pour compléter la caractérisation des méthodes par BS, une analyse de la répartition des fausses alarmes est réalisée.

5.1.1 Rappel des paramètres des filtres étudiés

Les paramètres considérés pour les filtres DNLM et DBM3D dans cette section sont résumés dans le tableau 5.1.

Nom du filtre	Paramètres	Valeurs nominales des paramètres	
DNLM	N_1	5	
	a	1.5	
	N_s	11	
	h	Base de données terrestre	9
		Base de données satellitaire	6
	N_{max}	2	
	λ	0.08	
DBM3D	N_1	8	
	τ_{match}	10^{10}	
	N_2	16	
	N_s	39	
	N_{step}	3	
	σ	Base de données terrestre	6
		Base de données satellitaire	5
	T_{2D}	2D-Bior1.5	
	T_{1D}	1D-Haar	
	λ_{3D}	2.7	
	β	2	

TABLE 5.1 – Tableau récapitulant les paramètres nominaux considérés pour les filtres de détection par suppression de fond proposés au cours de cette thèse. Dans toute cette section 5.1, ce sont ces valeurs qui sont utilisées si aucune précision particulière n'est indiquée.

5.1.2 Comparaison avec l'existant

Les filtres DNLM [Genin et al., 2011] et DBM3D [Genin et al., 2012a] sont comparés avec les filtres de détection par BS suivants :

- une méthode BS où le fond est estimé en moyennant localement les pixels sur un voisinage de 3×3 pixels (BSMean : *Background Suppression by Mean Filter*).
- une méthode BS où le fond est estimé en calculant les valeurs médianes des pixels inclus dans un voisinage de 3×3 pixels (BSMedian : *Background Suppression by Median Filter*).
- une méthode BS où le fond est estimé en utilisant un filtrage médian pondéré qui fait intervenir les quatre plus proches voisins et le pixel central avec un poids égal à 3 (BSMedian Pondéré).

- un filtrage top-hat, qui correspond à une méthode de BS où le fond est estimé en appliquant un opérateur morphologique d’ouverture, avec un élément structurant carré de dimension 3×3 pixels (TopHat).
- une méthode BS où le fond est estimé en appliquant un filtrage max-médian [Deshpande *et al.*, 1999] sur un voisinage de 3×3 pixels (BSMax-median).
- une méthode BS qui prédit le fond au moyen d’une approche régularisée globale non-linéaire (avec les 4 plus proches voisins, $\lambda = 0.3$, $T_\chi = 1$, $T_\phi = 1$) [Samson, 2002] (Samson).

La taille du voisinage impliquée dans chacune des méthodes BS Mean, BS Median, BS Median Pondéré, TopHat et BS Max-median a été optimisée et c’est un voisinage de dimension 3×3 qui donne les meilleures performances de détection pour chacun des filtres.

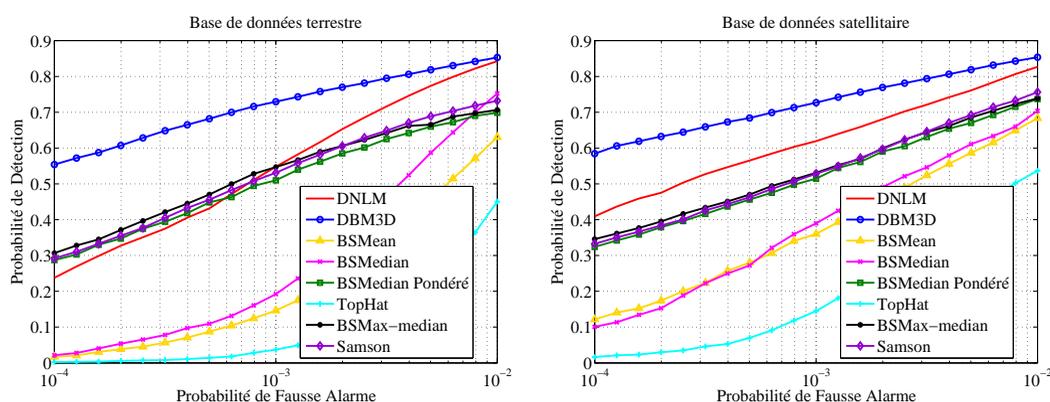


FIGURE 5.1 – Courbes ROC obtenues pour différentes méthodes de détection par BS.

La figure 5.1 présente les courbes ROC obtenues avec les différentes méthodes listées ci-dessus. Nous décrivons les comportements des méthodes, qui sont assez similaires pour les deux bases d’images, en commençant par la moins performante jusqu’à la plus performante. Les méthodes Top-Hat, BS Mean et BS Médian sont les moins performantes. Ensuite on trouve un groupe de 3 méthodes avec des performances moyennes et très similaires : BS Max-median, BS Median Pondéré et Samson. Les deux méthodes basées sur des techniques de débruitage à patchs que nous avons proposées donnent les meilleures performances. Pour la base terrestre, la performance de DNLN est légèrement moins bonne que les 3 méthodes moyennes (BS Max-median, BS Median Pondéré et Samson) pour les faibles Pfa. Mais DNLN donne de meilleurs résultats pour les Pfa plus élevées et pour la base satellitaire, quelque soit la Pfa. Enfin, la méthode DBM3D surclasse nettement toutes les autres méthodes et en particulier DNLN, pour les deux bases d’images.

5.1.3 Étude de l’influence des paramètres associés aux cibles

Dans cette section nous étudions la sensibilité des performances de détection à la variation des caractéristiques des cibles (intensité et dimension). Nous choisissons de présenter cette étude pour deux filtres représentatifs : le filtre BS Mean, filtre linéaire de complexité algorithmique très faible, et DBM3D, le filtre non linéaire le plus performant selon les résultats de la section précédente.

5.1.3.1 Étude de la robustesse à la présence de cibles de forte intensité

Les performances de détection des filtres [DBM3D](#) et [BSMean](#) en fonction de l'intensité des cibles incrustées s_0 sont présentées en figure [5.2](#), pour des valeurs de P_{fa} fixées à 10^{-2} , 10^{-3} et 10^{-4} . Le comportement des filtres [BSMean](#) et [DBM3D](#) est très différent : [DBM3D](#) est toujours bien plus performant que [BSMean](#) pour les niveaux de cible limités (inférieurs à 100) mais ses performances s'effondrent au-delà de 200, alors que [BSMean](#) marche de mieux en mieux.

Le filtre [DBM3D](#) n'est donc absolument pas robuste à la présence d'une cible de forte intensité. Cette dégradation des performances est liée à l'étape de débruitage des blocs par seuillage des coefficients de la transformée en ondelette (cf. section [3.2.2](#), page [55](#)). Quand une cible est très intense le coefficient en ondelette correspondant aux hautes fréquences est très élevé, il est du coup conservé lors du seuillage car il dépasse le seuil qui dépend du paramètre σ . On peut noter que la robustesse à la présence d'une cible du filtre [DBM3D](#) augmente avec σ mais en même temps, en contrepartie, les performances de détection pour les cibles de faible intensité diminuent. Les tentatives de modifications de [DBM3D](#) réalisées jusqu'à présent pour augmenter sa robustesse à l'augmentation de l'intensité des cibles ont toutes conduites à une perte de performances de détection des cibles de faible intensité qui sont notre principal sujet d'intérêt dans cette thèse. Des recherches restent donc à poursuivre sur le filtre [DBM3D](#) pour corriger ce point.

Cependant, la robustesse à la présence de cibles de très forte intensité ne s'avère pas être un véritable problème en pratique. En effet, nos filtres de détection très performants pour les cibles de faible intensité pourront être combinés à un simple filtre linéaire type [BSMean](#). Le seuil de détection associé à ce filtre sera choisi très élevé afin de ne détecter que les cibles de très forte intensité. Il est important d'être conscient de la sensibilité du filtre [DBM3D](#) à l'augmentation du niveau des cibles pour prendre les mesures nécessaires afin de ne rater aucune cible.

5.1.3.2 Étude de la robustesse à l'augmentation de la dimension des cibles

Les performances de détection des filtres [DBM3D](#) et [BSMean](#) en fonction du paramètre r_c intervenant dans la modélisation des cibles (cf. section [2.5](#)) sont présentées en figure [5.3](#), pour des valeurs de P_{fa} fixées à 10^{-2} , 10^{-3} et 10^{-4} . Plus le paramètre r_c est petit et plus les cibles sont larges. Pour s'affranchir du fait que le niveau du pixel central de la cible diminue avec r_c , l'énergie globale de la cible s_0 est ajustée pour chaque paramètre r_c de telle façon que $s_{(0,0)}(0,0) = 38$. La performance de [BSMean](#) est toujours bien plus faible que celle de [DBM3D](#), mais, hormis ce facteur, l'évolution de toutes les courbes est similaire, avec une performance qui s'effondre pour r_c supérieur à 0.75, soit une largeur de cible supérieure à 1.5 pixels (largeur à mi-hauteur de la [PSF](#) échantillonnée). Ces méthodes spécifiques aux objets ponctuels ne sont pas pertinentes pour des objets étendus. Au delà d'une certaine taille de cible la problématique de détection devient différente et des méthodes de détection adaptées à des cibles de plus grande dimension doivent être utilisées

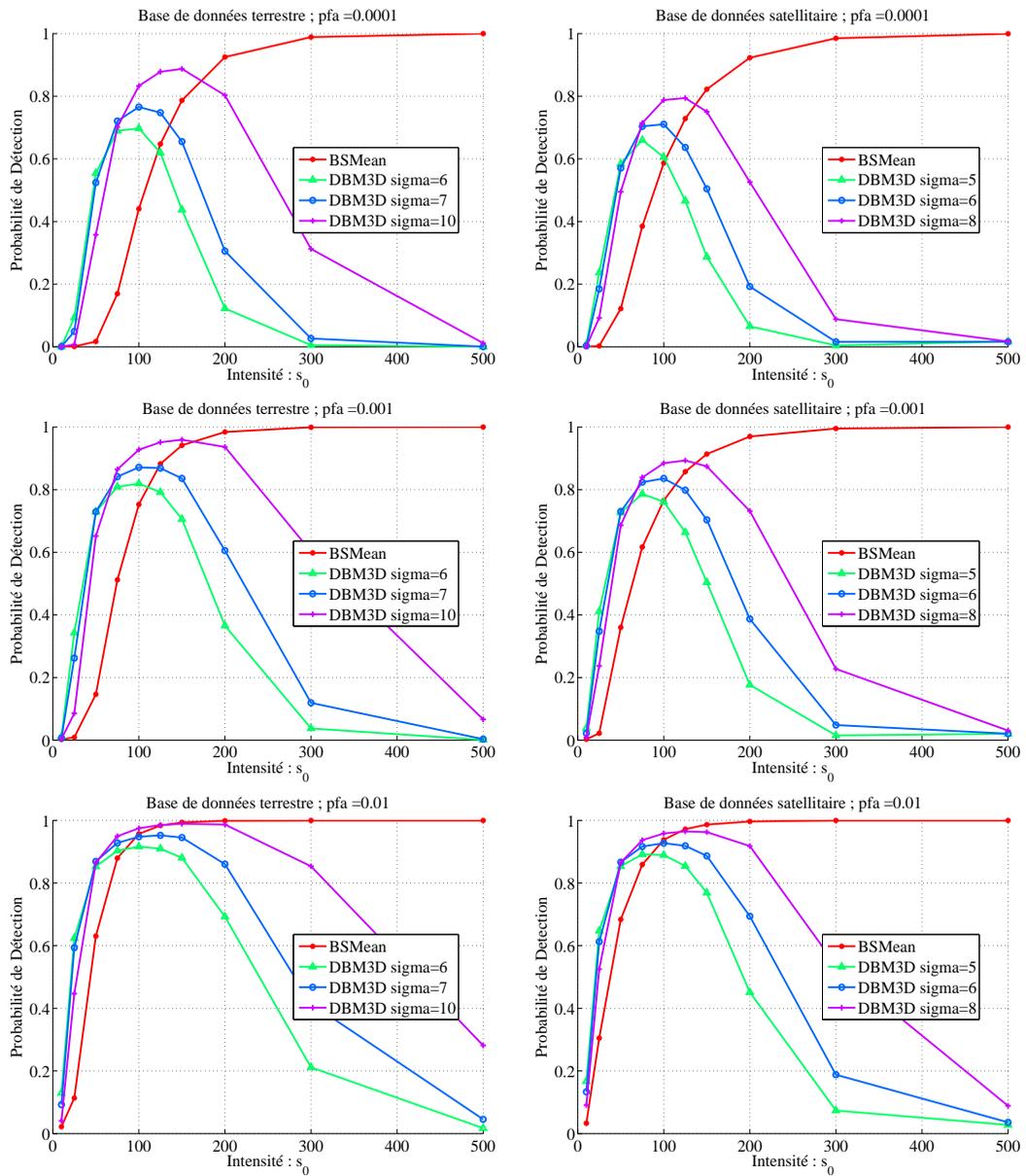


FIGURE 5.2 – Influence de l'intensité de la cible : Courbes ROC obtenues pour des cibles ayant été simulées avec différentes intensités s_0 .

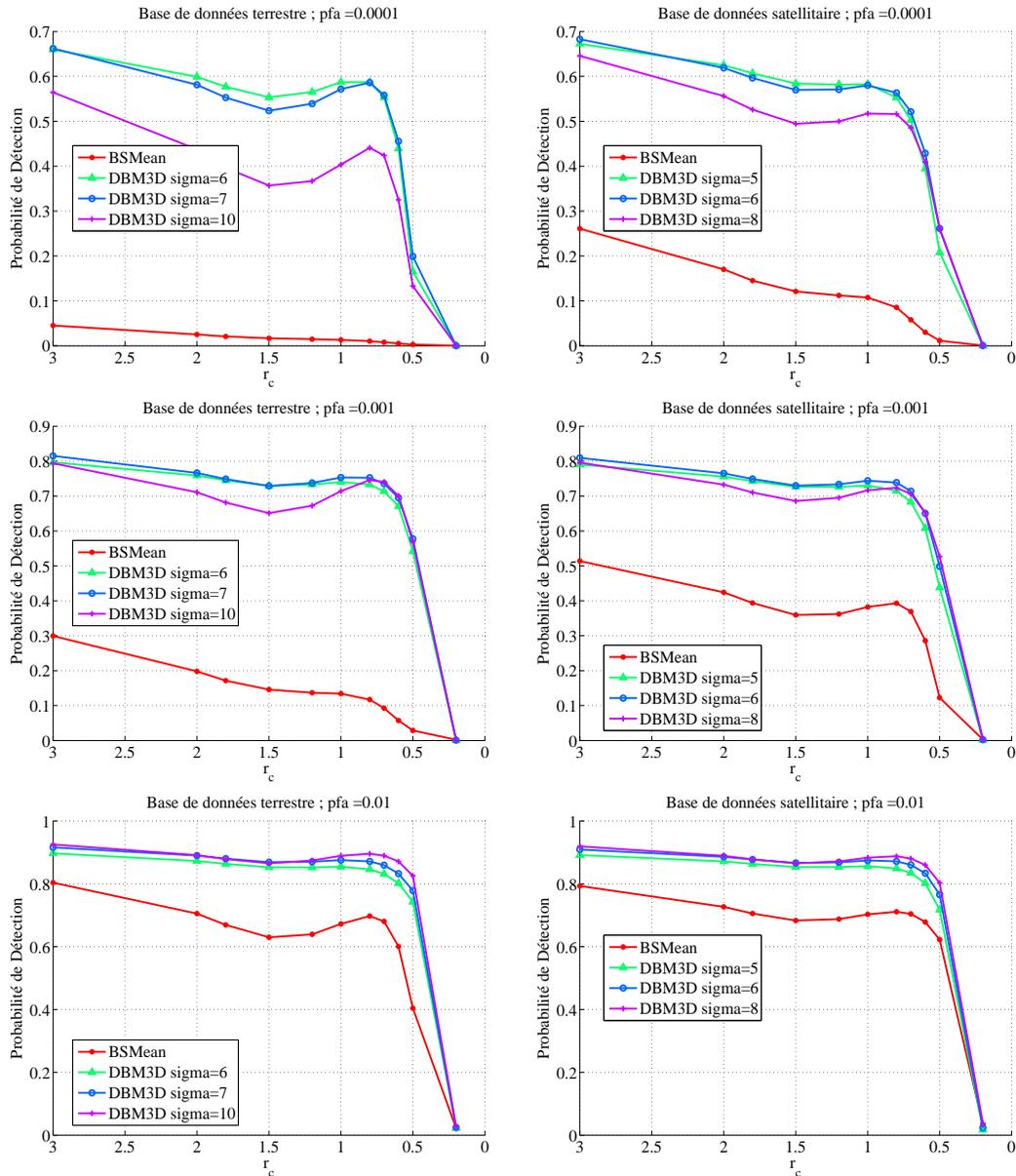


FIGURE 5.3 – Influence de la dimension de la cible : Courbes ROC obtenues pour des cibles ayant été simulées avec différentes valeurs de r_c (avec s_0 ajusté de telle façon que $s_{(0,0)}(0,0)$ soit toujours égal à 38, cf. section 2.5).

5.1.4 Caractérisation complémentaire des algorithmes : analyse de la répartition des fausses alarmes et non-détections

Pour compléter la caractérisation des filtres de détection par **BS**, nous étudions dans cette section la répartition des **FA**. Comme discuté au chapitre 2 en section 2.1.2.2, il est intéressant d'étudier cette répartition en supplément des courbes **ROC** considérant que des **FA** répétées dans des zones difficiles de l'image telles que les bords de nuages vont être corrélées temporellement et vont conduire à un taux élevé de fausses pistes. Les premières lignes des figures 5.4 et 5.5 présentent la répartition des **FA** obtenues pour une image de chacune des deux bases de données considérées jusqu'à présent suite à l'application de méthodes de détection **BS** variées. Ces **FA** sont les pixels dont l'intensité résiduelle est supérieure au seuil de détection réglé ici pour obtenir un taux de fausses alarmes de 10^{-2} . Pour le filtre **BSMean**, on peut voir sur les figures 5.4 et 5.5 que les **FA** sont fortement concentrées dans les zones de bords. Ces groupes de **FA** ont de fortes chances d'apparaître dans les autres images de la séquence et donner lieu à de fausses pistes. Comme on peut le voir, les méthodes **BSMax-Median**, **DNLM** et **DBM3D**, conduisent à une distribution plus uniforme des **FA** que le filtre **BSMean**, elles ont un comportement qui apparaît comme étant avantageux pour les opérations ultérieures de pistage de cibles.

Notons que cette observation doit être reliée à la distribution des non-détections. Cette distribution est obtenue de la façon suivante : on associe à chaque pixel l'intensité de sortie d'une cible incorporée à cette position dans le fond et les non-détections sont ensuite les pixels dont l'intensité associée est inférieure au seuil de détection. Les dernières lignes des figures 5.4 et 5.5 présentent la distribution des non-détections obtenues pour divers filtres **BS**. Comme discuté par Cafer *et al.* dans [Cafer *et al.*, 2007], il apparaît que toute amélioration de l'uniformité de la distribution des **FA** est associée à une augmentation du taux de non-détections dans certaines régions. Il faut donc rester conscient que les filtres qui ont le comportement recherché de conduire à une distribution uniforme des **FA** vont manquer des cibles principalement au niveau des zones de bords de nuages.

Dans le chapitre 2, nous avons proposé le critère **DNFA** (cf. équation (2.12)) pour quantifier l'homogénéité de la répartition des **FA**. Plus la distribution des **FA** est homogène, plus ce critère est proche de la valeur théorique $DNFA_{\text{poisson}}$ (cf. équation (2.13)). Les valeurs de ce critère obtenues pour les filtres **BS** étudiés dans la section 5.1.2 sont présentées dans le tableau 5.2. Deux catégories de filtres se distinguent : ceux pour lesquels les valeurs du critère **DNFA** sont nettement inférieures aux valeurs $DNFA_{\text{poisson}}$ et qui ont leurs fausses alarmes concentrées au niveau des bords (comme par exemple le **BSMean**, cf. figures 5.4 et 5.5) et ceux qui ont des valeurs de **DNFA** qui sont plus proches des valeurs idéales $DNFA_{\text{poisson}}$ qui ont une répartition plus homogène de leurs **FA** (comme par exemple **DBM3D**, cf. figures 5.4 et 5.5). Les filtres **BSMean**, **BSMedian**, **Top-Hat** et **DNLM**, qui ont une valeur de **DNFA** à une P_{fa} de 10^{-3} inférieure à 10 en moyenne pour les deux bases de données (valeur bien inférieure à la valeur $DNFA_{\text{poisson}}$ égale à 15), font partie de la première catégorie (cf. tableau 5.2). Les filtres de **BS** qui appartiennent à la deuxième catégorie sont : **BSMedian pondéré**, **BSMax-Median**, **Samson** et **DBM3D**.

En conclusion, en plus d'être le filtre qui donne les meilleures performances en terme de courbes **ROC**, le filtre **DBM3D** appartient à la catégorie de filtres **BS** permettant d'obtenir les répartitions de **FA** les plus uniformes.

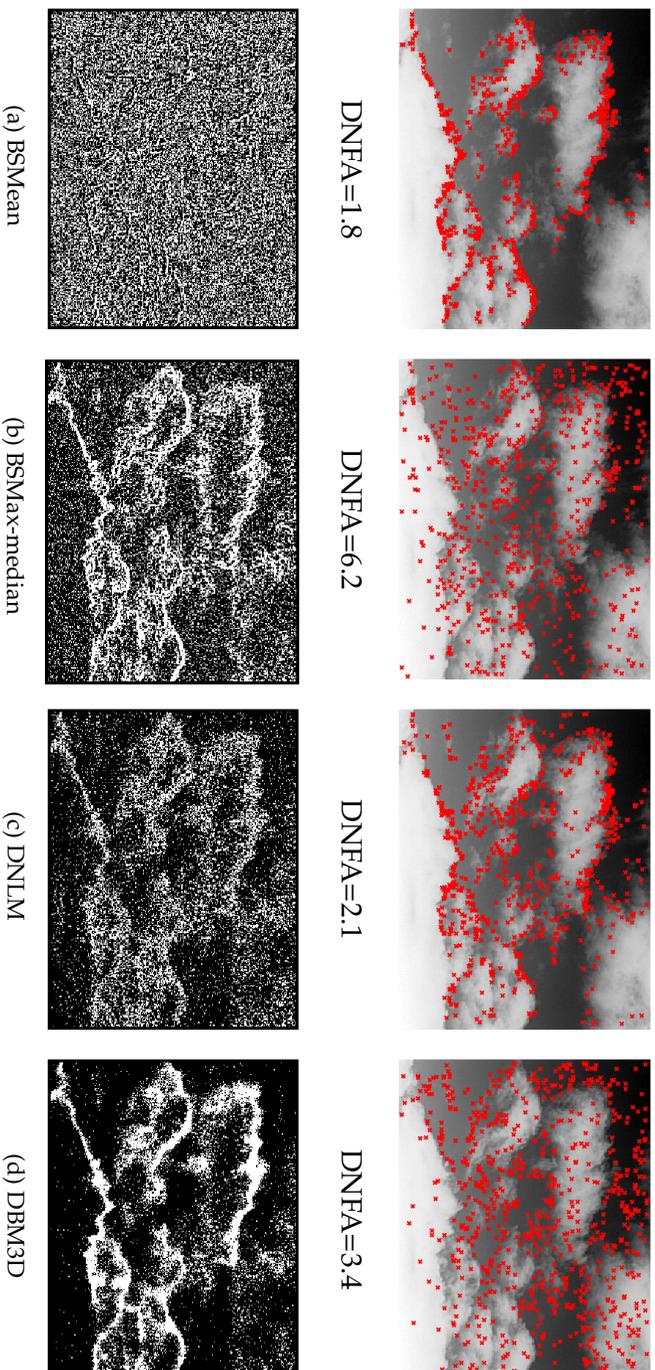


FIGURE 5.4 – Première ligne : positions des fausses alarmes obtenues pour une image de la base de donnée terrestre après l'application de diverses méthodes de BS avec un seuil de détection associé à une P_{fa} de 10^{-2} . Deuxième ligne : critère DNFA associé. Troisième ligne : positions des non-détectés (pixels blancs) obtenues avec le même seuil ($P_{FA} = 10^{-2}$).

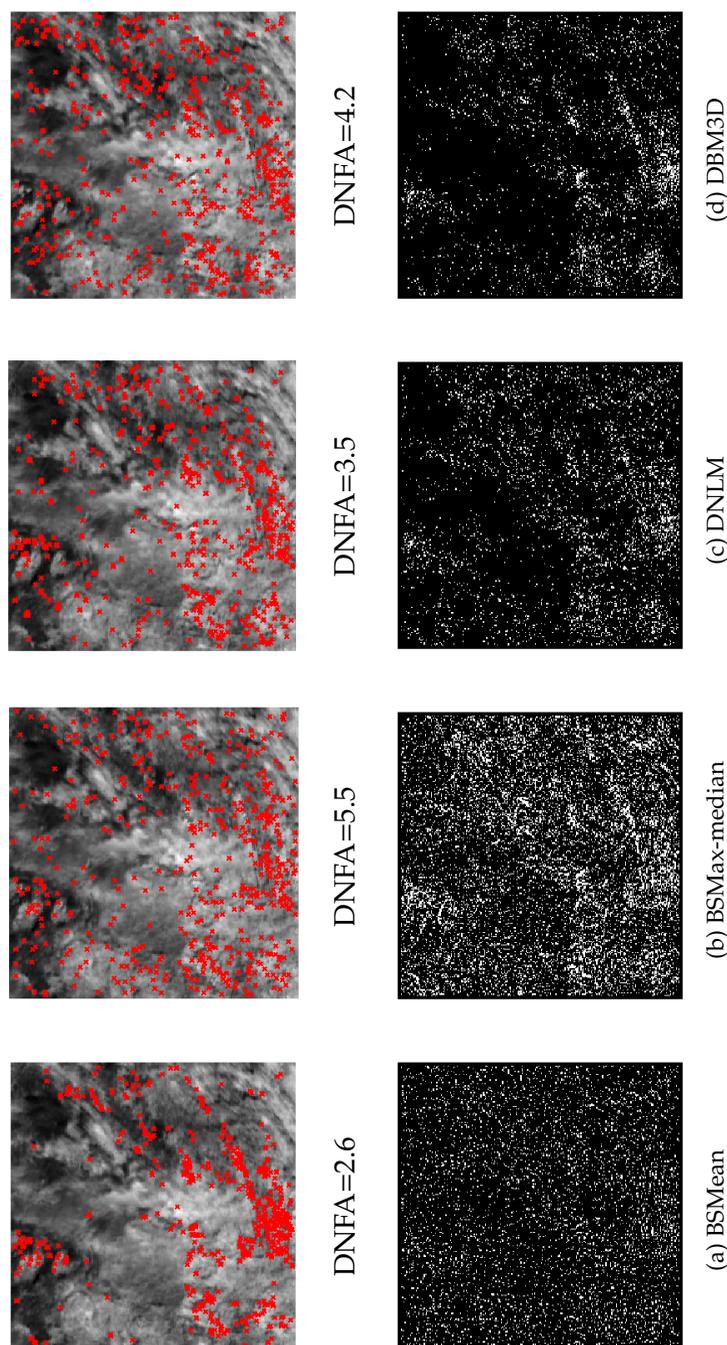


FIGURE 5.5 – Première ligne : positions des fausses alarmes obtenues pour une image de la base de donnée satellitaire après l'application de diverses méthodes de **BS** avec un seuil de détection associé à une **Pfa** de 10^{-2} . Deuxième ligne : critère **DNFA** associé. Troisième ligne : positions des non-détections (pixels blancs) obtenues avec le même seuil ($P_{FA} = 10^{-2}$).

Filtre	DNFA		
	$P_{FA} = 10^{-4}$	$P_{FA} = 10^{-3}$	$P_{FA} = 10^{-2}$
BSMean	29,2	6,9	2,2
BSMedian	40,5	7,9	2,4
BSMedian Pondéré	51,0	12,3	5,4
TopHat	30,1	5,0	1,8
BSMaxmedian	51,6	12,3	6,0
Samson	45,3	10,5	4,8
DNLM	37,4	9,1	2,8
DBM3D	39,0	11,0	3,9

(a) Base de données terrestre

Filtre	DNFA		
	$P_{FA} = 10^{-4}$	$P_{FA} = 10^{-3}$	$P_{FA} = 10^{-2}$
BSMean	28,8	7,8	2,4
BSMedian	34,2	8,2	2,5
BSMedian Pondéré	28,5	12,0	4,9
TopHat	20,2	5,1	2,0
BSMaxmedian	38,0	11,8	5,6
Samson	28,4	10,8	4,4
DNLM	42,5	9,0	3,2
DBM3D	45,5	12,2	3,8

(b) Base de données satellitaire

Filtre	DNFA _{poisson}		
	$P_{FA} = 10^{-4}$	$P_{FA} = 10^{-3}$	$P_{FA} = 10^{-2}$
Distribution de Poisson	50,0	15,8	5,0

(c) Distribution uniforme

TABLE 5.2 – Critère **DNFA** pour les deux bases de données et le critère théorique associé à des nombres pseudo-aléatoires uniformément distribués ($DNFA_{poisson} = 1/(2\sqrt{P_{FA}})$).

5.2 Filtres de détection basés sur une modélisation au second ordre

Maintenant que nous avons vu que le filtrage **DBM3D** était l'approche **BS** la plus performante en spatial, nous étudions dans cette section les performances des approches de détection basées sur une modélisation du fond au second ordre.

5.2.1 Paramètres des filtres étudiés

Les paramètres considérés pour le filtre **GMMF** et sa variante $GMMF_0$ dans cette section sont présentés dans le tableau 5.3 ainsi que les paramètres des filtres utilisés pour prétraiter les données (**DBM3D** et **BSMean**) en accord avec le schéma de détection de la figure 1.4.

Nom du filtre	Paramètres	Valeurs nominales des paramètres		
GMMF et GMMF ₀	N	9		
	n	3		
	K	7		
	Méthode d'initialisation du CEM	Segmentation obtenue en appliquant un K-means sur les écarts-types locaux.		
	Nombre d'itérations du processus CEM (<i>ite</i>)	GMMF	9	
		GMMF ₀	0	
Signature des cibles \mathbf{s}	si aucun prétraitement ou prétraitement par DBM3D	$\mathbf{s} = \begin{pmatrix} 0 & \dots & 0 & \dots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ & & 0 & 0 & 0 \\ 0 & \dots & 0 & 1 & 0 & \dots & 0 \\ & & 0 & 0 & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & \dots & 0 \end{pmatrix}$		
	si prétraitement par BSMean	$\mathbf{s} = \begin{pmatrix} 0 & \dots & \dots & 0 & \dots & \dots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots \\ & & 0 & 0 & 0 & 0 & 0 \\ 0 & \dots & 0 & -0.125 & -0.125 & -0.125 & 0 \\ & & 0 & -0.125 & 1 & -0.125 & 0 & \dots & 0 \\ & & 0 & -0.125 & -0.125 & -0.125 & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & \dots & \dots & 0 & \dots & \dots & 0 \end{pmatrix}$		
DBM3D	N_1	8		
	τ_{match}	10^{10}		
	N_2	16		
	N_s	39		
	N_{step}	3		
	σ	Base de données terrestre	+Aucun Postraitement	6
			+ANF ou +GMF	7
			+GMMF ou +GMMF ₀	10
	Base de données satellitaire	+Aucun Postraitement	5	
		+ANF ou +GMF	6	
		+GMMF ou +GMMF ₀	8	
	T_{2D}	2D-Bior1.5		
	T_{1D}	1D-Haar		
λ_{3D}	2.7			
β	2			
BSMean	Dimension du voisinage	3×3		

TABLE 5.3 – Tableau récapitulant les paramètres des filtres proposés au cours des chapitres 3 et 4 que l'on associe selon le schéma de détection en deux étapes proposé en chapitre 1. Dans toute cette section, ce sont ces paramètres qui sont utilisés si aucune précision particulière n'est indiquée.

Optimisation du paramétrage de DBM3D

Le paramètre σ du filtre DBM3D a été optimisé en fonction de la méthode de détection au second ordre à laquelle il est associé.

Dans la section 3.3.2.2 nous avons vu que σ était le paramètre du DBM3D le plus délicat à régler et que son réglage était une question de compromis (cf. figure 3.19). Quand σ est faible, le fond est très bien supprimé mais les cibles le sont en partie aussi et cela fait

chuter le taux de détection. Inversement, quand σ est élevé les cibles sont préservées mais les textures du fond et les zones de fortes transitions sont très nettement présentes dans le résidu ce qui augmente les fausses alarmes. Au final, comme le montrait la figure 3.19, une valeur de 5 ou 6 (suivant la base) donnait le meilleur compromis, les valeurs plus élevées (8 ou 10) ou plus faibles (3 ou 4) donnant des performances significativement moins bonnes.

La figure 5.6 présente les performances de détection obtenues avec le filtre DBM3D+GMMF pour différentes valeurs de σ . La valeur de σ donnant les meilleures performances de détection est plus élevée pour le filtre DBM3D+GMMF (cf. figure 5.6) que pour le filtre DBM3D seul (cf. figure 3.19). Cela s'explique de la façon suivante. La valeur de σ un peu plus élevée a pour conséquence de mieux préserver les cibles et d'augmenter la structure du fond présente en sortie du filtre DBM3D, comme le montrent les figures 5.7 et 5.8. Mais l'application du filtrage GMMF permet une meilleure modélisation de ces textures résiduelles et une réduction du nombre de fausses alarmes.

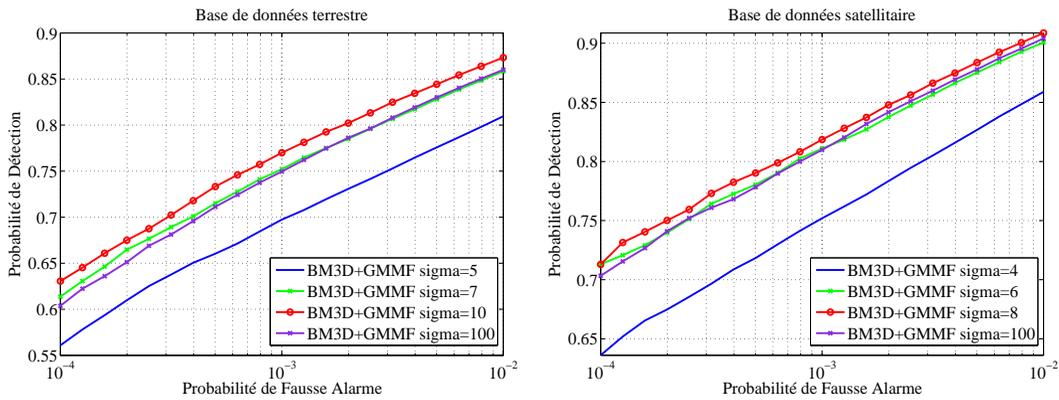


FIGURE 5.6 – Influence du paramètre σ : Courbes ROC obtenues avec DBM3D, pour chacune des bases de données, pour différentes valeurs de σ .

Quand on augmente nettement σ au delà de la valeur optimale, la chute des performances est nettement moins importante que celle observée avec le filtrage DBM3D seul (cf. figure 3.19). Dans ce cas, les performances du DBM3D+GMMF convergent vers les performances que l'on obtient en appliquant directement le filtre GMMF sur les données brutes, cf. figure 5.12.

5.2.2 Comparaison des méthodes de détection au second ordre

La version spatiale du filtre GMMF [Genin *et al.*, 2012a] est comparée aux filtres de détection basés sur une modélisation du fond au second ordre suivants :

- le filtrage adapté donné par l'équation (1.25) où on considère une seule matrice de covariance valable pour tous les pixels de l'image, celle-ci est estimée à partir de l'ensemble de l'image (GMF : *Global Matched Filter*). La signature de cible \mathbf{s} qui intervient est la même que celle utilisée dans le filtre GMMF (cf. tableau 5.3). On peut noter que le filtre GMF correspond au filtrage GMMF dans le cas où le nombre de classes K est égal à 1.
- le filtre couramment utilisé qui consiste simplement à normaliser les pixels par l'écart-type local (ANF : *Adaptive Normalized Filter*). Ce dernier est estimé sur un voisinage de 7×7 pixels en excluant le pixel central. Comme discuté dans le chapitre

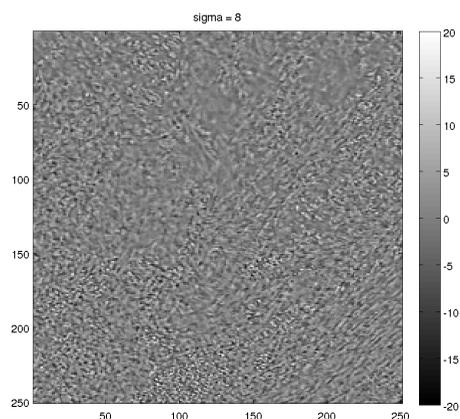
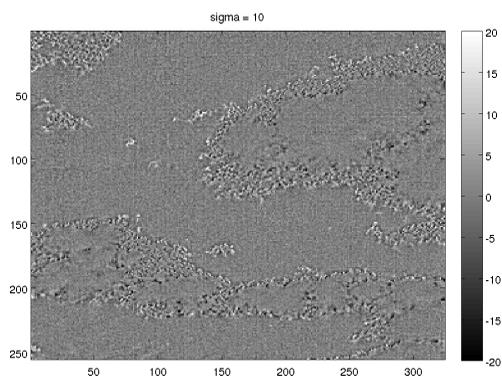
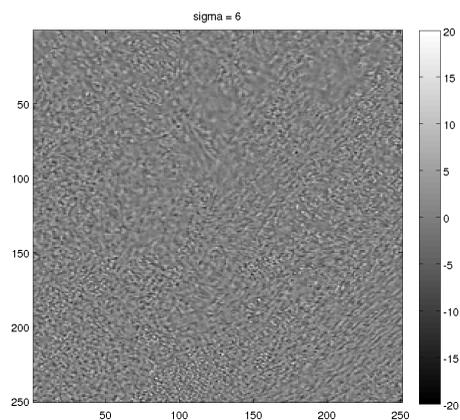
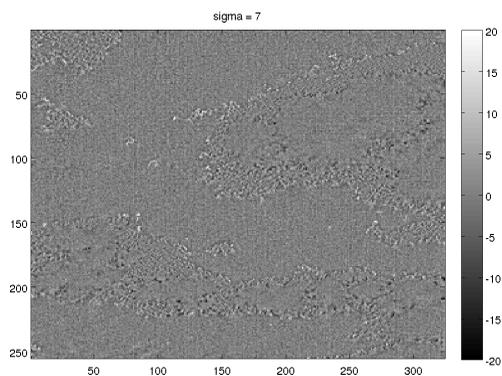
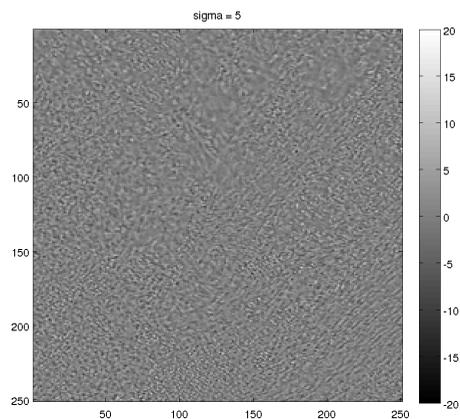
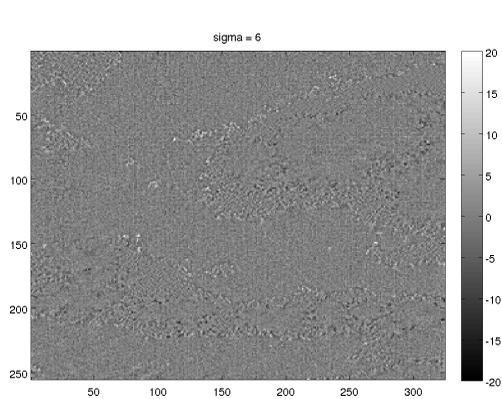


FIGURE 5.7 – Résidu obtenu avec DBM3D pour une des images de la base de données terrestre, pour différentes valeurs de σ (images saturées pour améliorer la visibilité des textures).

FIGURE 5.8 – Résidu obtenu avec DBM3D pour une des images de la base de données satellitaire, pour différentes valeurs de σ (images saturées pour améliorer la visibilité des textures).

4, ce filtre ANF s'interprète comme un filtre GMMF dans le cas où les matrices de covariance sont approximées par des matrices diagonales.

Les courbes ROC obtenues avec ces différents filtres de détection sont présentées en figure 5.9, dans le cas où ils sont appliqués à la suite du meilleur algorithme BS obtenu en section 5.1, soit DBM3D et dans le cas où ils interviennent à la suite du filtre linéaire BSMean. Dans chacun des cas, les performances de l'algorithme de détection BS seul sont aussi incluses dans la comparaison.

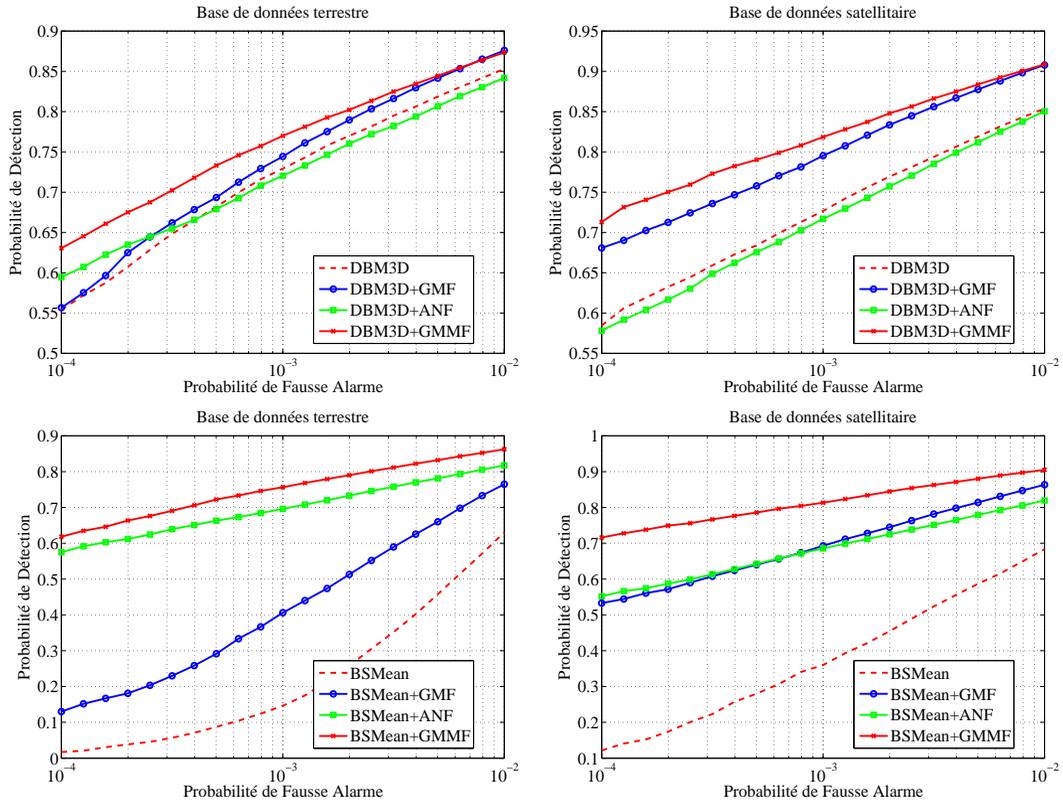


FIGURE 5.9 – Courbes ROC obtenues pour les filtres BSMean et DBM3D associés à différentes méthodes de détection par modélisation du fond au second ordre.

Commençons tout d'abord par commenter les résultats obtenus à la suite d'un DBM3D. Notons que le filtre DBM3D seul et le filtre DBM3D+ANF donnent globalement pour l'ensemble de deux bases de données de moins bonnes performances de détection que les filtres DBM3D+GMF et DBM3D+GMMF. Le filtre ANF utilise uniquement l'information de variance et ne tire pas parti de la corrélation existante entre les pixels du résidu. Or le filtrage DBM3D est capable de bien prendre en main les zones de fortes transitions ce qui a pour conséquence qu'il n'y a quasiment pas de variations de la variance locale dans les images résiduelles comme on le peut le voir sur les résidus présentés en figures 5.7 et 5.8. C'est pourquoi, à la suite d'un filtre DBM3D, le filtre ANF ne permet pas d'améliorer les performances de détection. Avec le filtrage DBM3D même si les bords sont bien supprimés comparé au cas du BSMean par exemple, des textures nuageuses restent toujours en partie visible. C'est pourquoi les filtres de détection qui tiennent compte de la corrélation entre les pixels du résidu comme GMF et GMMF permettent d'améliorer les performances du DBM3D. On peut noter que l'amélioration des performances est plus

importante pour la base de données satellitaire que terrestre, cela s'explique par le fait qu'il y a nettement plus de textures présentes dans le résidu pour la base de données terrestre que satellitaire. Notons que associé au DBM3D c'est le filtre GMMF qui est le plus performant pour les deux bases de données. Il est plus performant que le filtre GMF, ce qui confirme qu'une unique matrice de covariance n'est pas une stratégie très efficace en présence de fonds non stationnaires.

Commentons maintenant les résultats obtenus à la suite d'un BMean. Ces résultats dépendent fortement de la base de données considérée et peuvent s'expliquer en regardant les images de résidus présentées en figures 5.10 et 5.11. Pour la base de données

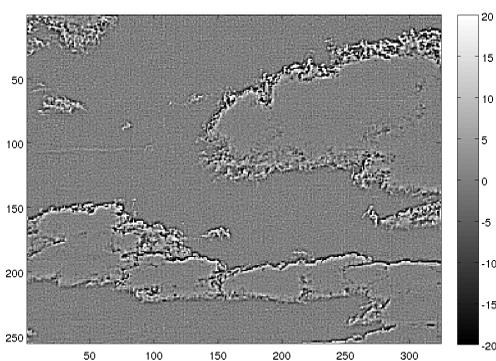


FIGURE 5.10 – Résidu obtenu avec BSMean pour une des images de la base de données terrestre (image saturée pour améliorer la visibilité des textures).

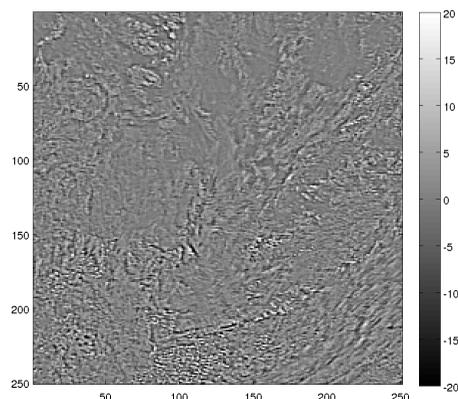


FIGURE 5.11 – Résidu obtenu avec BSMean pour une des images de la base de données satellitaire (image saturée pour améliorer la visibilité des textures).

terrestre, après l'application du BSMean, les zones de bords de nuages sont toujours très marquées. De fortes variations de la variance locale sont donc présentes dans les résidus comme on peut le voir sur la figure 5.10. C'est pourquoi une approche non adaptative telle que le GMF n'est absolument pas efficace et qu'une approche adaptative telle que l'ANF permet de nettement augmenter les performances de détection. Concernant la base de données satellitaire, les résidus sont fortement texturés. Le filtrage ANF a l'avantage d'être adaptatif mais ne tire pas parti de la corrélation entre les pixels. Le filtrage GMF prend en main la texture globalement présente mais n'est pas adaptatif et ne peut donc pas faire face aux variations de texture. Ces deux méthodes qui ont chacune des avantages et des défauts fournissent des performances de détection similaires comme on peut le voir sur la figure 5.9. Le GMMF qui combine les avantages de ces deux méthodes, est donc plus performant que celles-ci pour les deux bases de données (cf. figure 5.9).

Deux conclusions principales se détachent de cette étude. D'une part, en se basant sur la structure de traitement en deux étapes de la figure 1.4, nos résultats montrent que la seconde étape de filtrage fondé sur une modélisation au second ordre améliore toujours les performances par rapport à la première étape seule. Cette amélioration s'observe sur toutes les bases traitées et quelque soient les traitements considérés. D'autre part, le filtre GMMF surclasse nettement tous les filtres au second ordre qui ont été étudiés, car il est capable de prendre en compte la structure de corrélation des textures résiduelles et de s'adapter à la présence de non-stationnarités.

5.2.3 Études des performances en fonction de l'algorithme de BS

La figure 5.12 compare les performances en sortie du filtrage GMMF lorsqu'il est appliqué après différents filtres de BS : aucun filtre (c.a.d. GMMF appliqué directement sur l'image d'entrée), BSMean, DBM3D. Comme on pouvait s'y attendre les trois solutions se classent dans l'ordre d'efficacité du traitement BS, mais les résultats sont assez proches. Il est intéressant de comparer ces performances avec celles du traitement DBM3D sans GMMF, qui est rappelé dans les deux figures (courbe bleue claire). On constate que, quelque soit le traitement BS auquel il est associé (y compris aucun traitement !), tant que les cibles ne sont pas supprimées par ce dernier, le GMMF permet d'obtenir des taux de détection bien plus élevés que le meilleur traitement BS (DBM3D) utilisé seul.

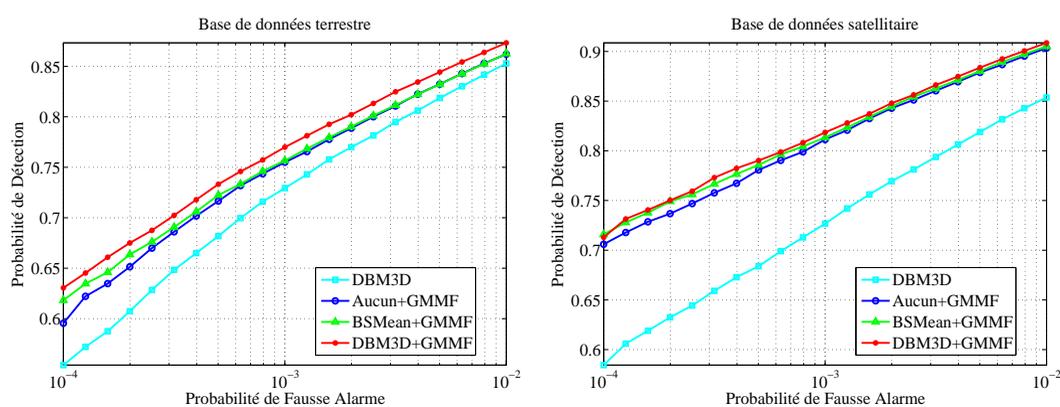


FIGURE 5.12 – Influence de la méthode de BS : Courbes ROC obtenues avec le filtre GMMF associé à différentes méthodes BS

5.2.4 Recherche d'un meilleur compromis coût de calcul / performances

Parmi l'ensemble des combinaisons de filtres testées jusqu'à présent le filtre DBM3D+GMMF est le plus performant, cependant celui-ci a le défaut d'avoir un coût de calcul très élevé (cf. tableau 5.4). Pour pouvoir espérer, après optimisation des codes, disposer d'un algorithme de détection rapide (idéalement en temps réel) et performant, il est nécessaire de regarder dès à présent si il est possible de diminuer le coût de calcul de l'approche en deux étapes proposée sans fortement dégrader les performances de détection. Nous cherchons quelle est l'association des différents filtres étudiés jusqu'à présent dans cette thèse qui permet d'obtenir le meilleur compromis coût de calcul / performances.

Filtre	Temps de calcul	Commentaire
BSMean	1 ms	
DBM3D	0.40 s	1 ^{ère} étape du code mex optimisé fourni par Dabov [BM3,]
BSMean+GMMF ₀	2.3 s	code Matlab non-optimisé
DBM3D+GMMF	18 s	code Matlab non-optimisé

TABLE 5.4 – Temps de calcul obtenus pour différents algorithmes sur un processeur Core 2 Duo 3.0 Ghz pour une image de dimension 255 × 324 pixels.

Dans le chapitre 4 en section 4.5.2, nous avons proposé une variante du filtre GMMF moins coûteuse en temps de calcul où l'étape de classification des pixels fait intervenir

un K-means sur les écarts types-locaux. Nous avons vu (cf. figure 4.8 chapitre 4, page 87) que cette variante notée $GMMF_0$ donne de bonnes performances, proches de celles obtenues avec $GMMF$. C'est pourquoi, nous proposons ici d'étudier les performances de détection obtenues en associant le filtre $GMMF_0$ avec un prétraitement par $BSMean$ (nettement moins coûteux qu'un $DBM3D$, cf. tableau 5.4) en prenant en compte cette fois-ci la déformation de la cible liée à l'application de ce prétraitement (ce qui permet d'augmenter les performances, cf. section 4.7.1). Une méthode encore moins coûteuse consiste à appliquer le filtre 4.5.2 directement sur les données brutes mais elle n'est pas performante (cf. figure 4.8 chapitre 4).

En figure 5.13, les performances obtenues avec ce filtre $BSMean+GMMF_0$ sont présentées et sont comparées à celles obtenues avec $DBM3D+GMMF$ qui est la combinaison de filtres la plus performante, avec le filtre $DBM3D$ seul qui est le filtre BS le plus performant et avec le filtre $BSMean$ qui est le filtre BS le moins coûteux.

Le filtre $BSMean$ est nettement moins coûteux en temps de calcul que les autres approches mais il est aussi nettement moins performant. Le filtre $BSMean+GMMF_0$ fournit des performances de détection supérieures au $DBM3D$ pour un coût de calcul qui est du même ordre de grandeur (le $GMMF_0$ a un temps de calcul plus élevé que $DBM3D$, mais son code n'a pas été optimisé contrairement à celui du $DBM3D$). Ce résultat montre une fois encore, comme nous l'avons vu en section précédente, l'apport essentiel des méthodes au second ordre en termes de performance. Finalement, $BSMean+GMMF_0$ est une solution nettement moins coûteuse en temps de calcul que $DBM3D+GMMF$ et conduit à un taux de détection seulement légèrement plus faible. C'est pourquoi le filtre $BSMean+GMMF_0$ apparaît comme étant le filtre qui donne le meilleur compromis coût de calcul / performances.

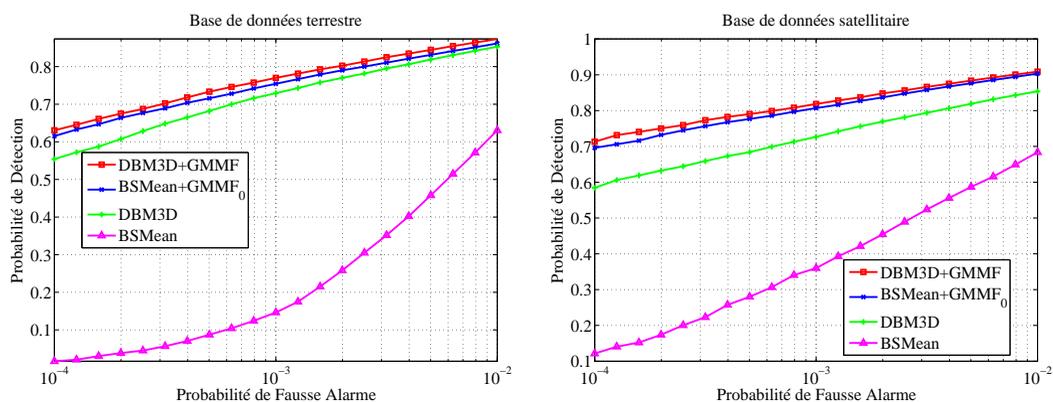


FIGURE 5.13 – Recherche du meilleur compromis coût de calcul / performances : Courbes ROC obtenues avec différentes associations de filtres de coût de calcul plus ou moins élevé.

5.2.5 Caractérisation complémentaire des algorithmes : analyse de la répartition des fausses alarmes et non-détections

Dans cette section, on complète la caractérisation des filtres de détection basés sur une modélisation du fond au second ordre en étudiant la répartition des FA comme cela a été fait en section 5.1.4 pour les filtres BS . Le tableau 5.5 présente les valeurs de $DNFA$ obtenues pour les filtres de détection ANF , GMF , $GMMF$ et $GMMF_0$ précédés soit d'un

BSMean soit d'un DBM3D. Les figures 5.14 et 5.15 présentent la répartition des fausses alarmes et non-détections obtenue avec certains d'entre eux, pour une image de chacune des bases de données.

Parmi l'ensemble des filtres testés, le seul filtre qui a le comportement non souhaité de concentrer les FA au niveau des bords de nuages est le filtre BSMean+GMF. Les filtres basés sur une modélisation du fond au second ordre adaptatifs qui sont le ANF, le GMMF et le GMMF₀ conduisent à une répartition uniforme des FA.

En conclusion, le filtre DBM3D+GMMF qui est le plus performant du point de vue des courbes ROC et le BSMean+GMMF₀ qui donne le meilleur compromis coût de calcul / performances ont tous les deux le comportement souhaité de fournir une répartition uniforme des FA.

Filtre	DNFA		
	$P_{FA} = 10^{-4}$	$P_{FA} = 10^{-3}$	$P_{FA} = 10^{-2}$
BSMean+ANF	53,2	15,6	5,4
BSMean+GMF	28,9	7,9	2,9
BSMean+GMMF ₀	44,8	13,7	4,8
BSMean+GMMF	48,9	14,1	4,9
DBM3D+ANF	49,5	15,0	5,3
DBM3D+GMF	42,4	11,4	4,9
DBM3D+GMMF ₀	47,4	13,3	4,9
DBM3D+GMMF	48,3	14,0	4,9

(a) Base de données terrestre

Filtre	DNFA		
	$P_{FA} = 10^{-4}$	$P_{FA} = 10^{-3}$	$P_{FA} = 10^{-2}$
BSMean+ANF	49,1	14,4	4,8
BSMean+GMF	31,4	7,5	2,8
BSMean+GMMF ₀	46,8	13,6	4,1
BSMean+GMMF	45,2	14,7	4,4
DBM3D+ANF	53,9	14,3	4,7
DBM3D+GMF	33,7	9,9	3,5
DBM3D+GMMF ₀	62,4	12,0	4,3
DBM3D+GMMF	60,1	14,8	4,5

(b) Base de données satellitaire

Filtre	DNFA _{poisson}		
	$P_{FA} = 10^{-4}$	$P_{FA} = 10^{-3}$	$P_{FA} = 10^{-2}$
Distribution de Poisson	50,0	15,8	5,0

(c) Distribution uniforme

TABLE 5.5 – Critère DNFA pour les deux bases de données et le critère théorique associé à des nombres pseudo-aléatoires uniformément distribués ($DNFA_{poisson} = 1/(2\sqrt{P_{FA}})$)

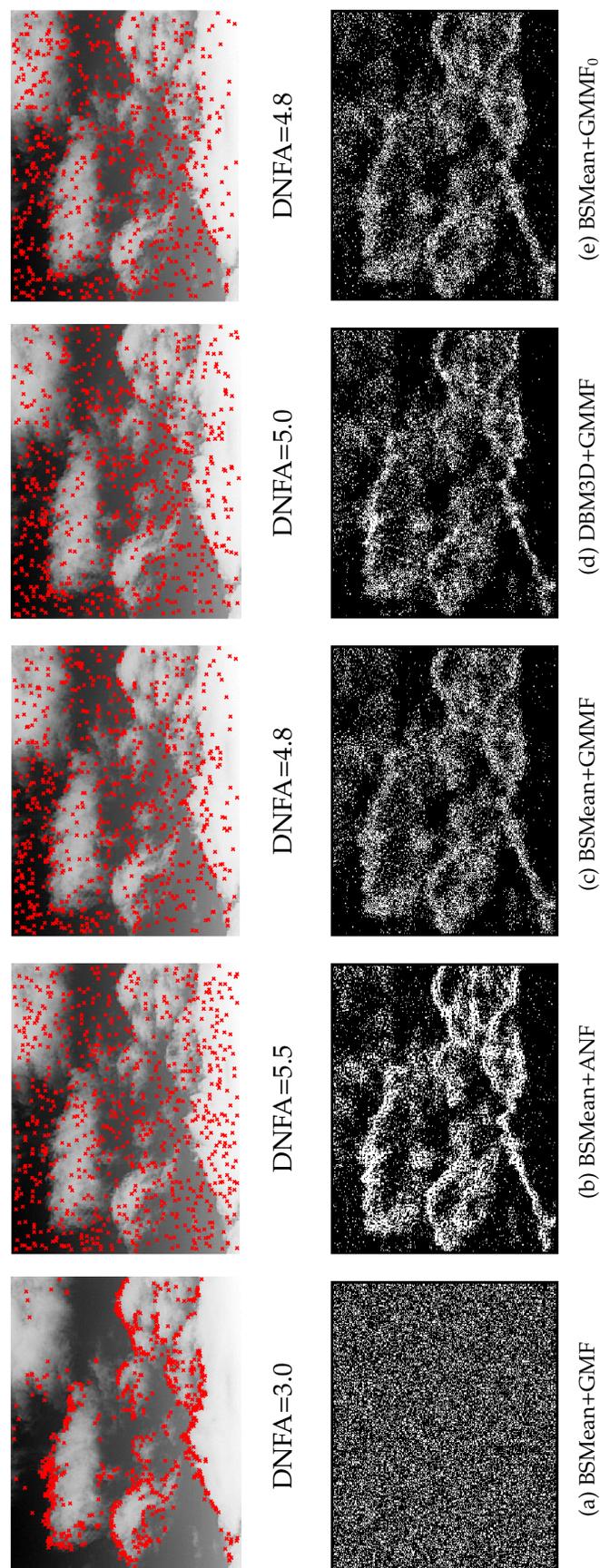


FIGURE 5.14 – Première ligne : positions des fausses alarmes obtenues pour une image de la base de donnée terrestre après l'application de divers filtres avec un seuil de détection associé à une P_{fa} de 10^{-2} . Deuxième ligne : critère DNFA associé. Troisième ligne : positions des non-détections (pixels blancs) obtenues avec le même seuil ($P_{FA} = 10^{-2}$).

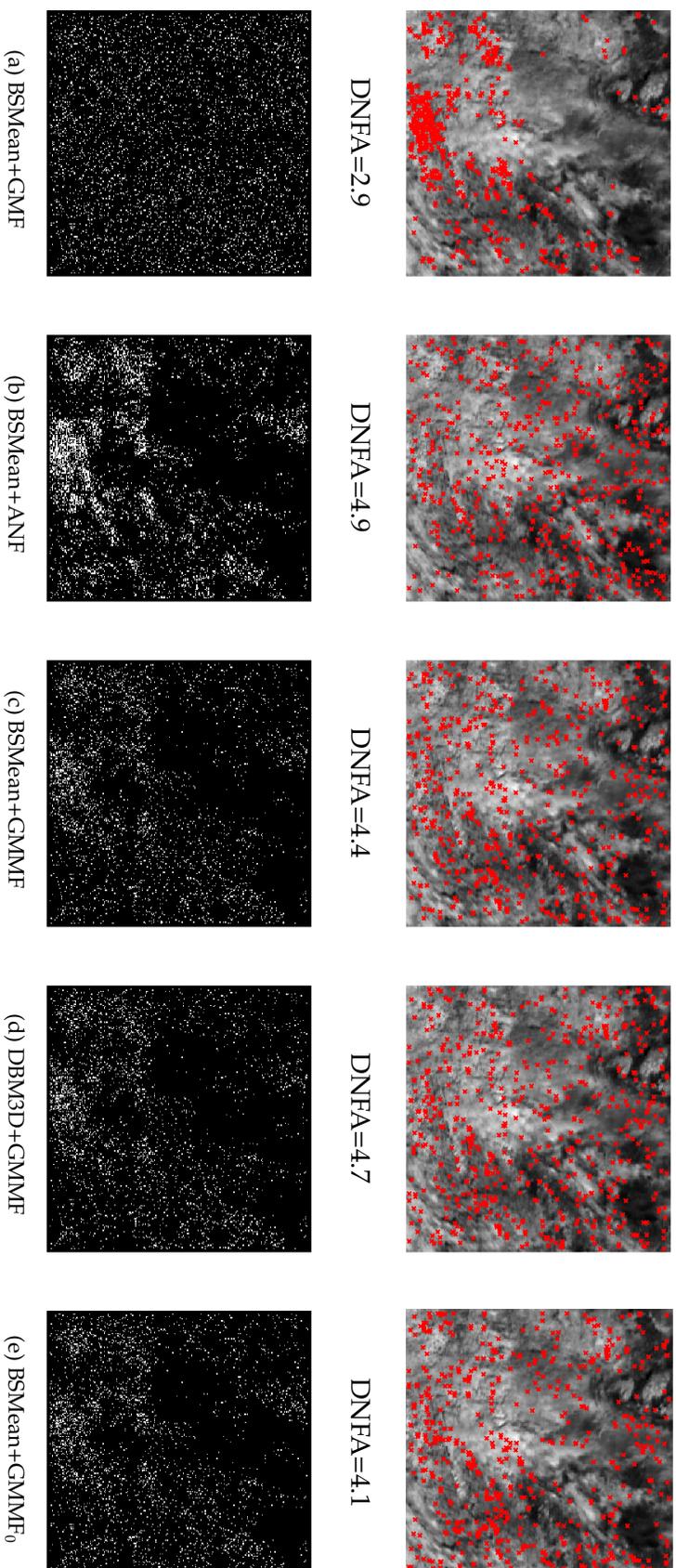


Figure 5.15 – Première ligne : positions des fausses alarmes obtenues pour une image de la base de donnée satellitaire après l'application de divers filtrages avec un seuil de détection associé à une P_{fa} de 10^{-2} . Deuxième ligne : critère DNFA associé. Troisième ligne : positions des non-détections (pixels blancs) obtenues avec le même seuil ($P_{FA} = 10^{-2}$).

5.3 Évaluation des algorithmes spatio-temporels

Dans cette section, nous évaluons les performances de détection des extensions temporelles des algorithmes **DNLM**, **DBM3D** et **GMMF**. Comme pour les versions spatiales, nous commençons par étudier les méthodes par suppression de fond (**DNLM** et **DBM3D**), puis dans un second temps l'approche **GMMF**.

5.3.1 Algorithmes de suppression de fond

Les filtres **DNLM** et **DBM3D** sont comparés avec les filtres de détection spatio-temporels par **BS** suivants :

- la méthode qui consiste à soustraire à l'image courante, l'image précédente (**BS**-temps),
- la méthode **BS** où le fond est estimé en moyennant localement les pixels sur un voisinage spatio-temporel de dimension 3×3 pixels \times 2 images (**BS**spatio-temps).

La version temporelle du filtre **DNLM** intervenant dans cette section (noté **DNLM_T**) est celle qui fait intervenir les poids w (équation (3.17)) qui, comme on l'a montré au chapitre 3, est plus performante que la version faisant intervenir les poids w_2 pour les cibles de faible intensité. Le paramétrage du filtre est modifié par rapport à la version spatiale pour augmenter les performances : le paramètre h est égal à 5 dans toute cette section. Concernant l'algorithme temporel **DBM3D** (noté **DBM3D_T**), on utilise, pour estimer le fond, la première itération du filtre **BM3D** décrit dans [Dabov *et al.*, 2007b] avec un paramètre σ égal à 6 et un paramètre τ_{match} égal à 10^{10} . La version spatiale de l'algorithme **DBM3D**, qui est le meilleur algorithme de détection **BS** spatial, est aussi ajoutée à la comparaison (il est noté **DBM3D_S** pour le différentiel de sa version temporelle).

Pour le cas nominal où la vitesse de la cible est de 1.25 pixels/image, les courbes **ROC** obtenues sont présentées en figure 5.16. Elles montrent que comme prévu dans la

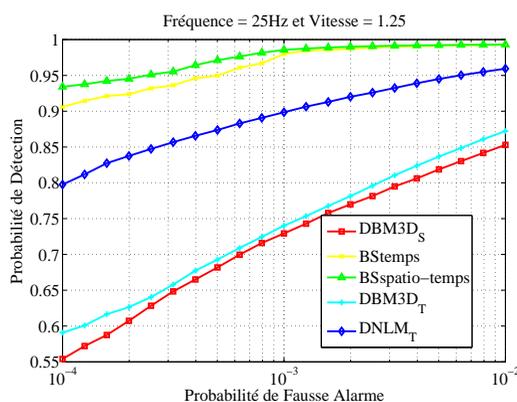


FIGURE 5.16 – Courbes **ROC** obtenues dans le cas dit nominal (vitesse de cible : 1.25 pixels/image et fréquence d'acquisition de la séquence : 25 Hz).

section 3.3.2.3, la version spatio-temporelle du **DBM3D** ne donne pas des performances nettement supérieures à sa version spatiale. La capacité du filtre **DNLM_T** à prendre en compte la corrélation temporelle du fond lui permet d'être plus performant que **DBM3D_S**. Cependant pour ce cas nominal, les algorithmes classiques de détection spatio-temporels **BStemps** et **BSspatio-temps** fournissent d'excellentes performances de détection grâce

à une vitesse d'acquisition des images très élevée (25Hz) impliquant une corrélation temporelle du fond très importante.

Influence du mouvement du fond nuageux

Pour obtenir un fond ayant une corrélation temporelle plus faible, une nouvelle séquence est créée en utilisant uniquement une image sur cinquante (on obtient ainsi une séquence correspondant à une fréquence d'acquisition de 0.5Hz). Les courbes ROC obtenues pour cette séquence sont présentées en figure 5.17a. Une forte dégradation des performances se produit pour l'ensemble des filtres spatio-temporels par rapport au cas nominal présenté en figure 5.16 (fréquence d'acquisition de la séquence 25Hz). Cette dégradation est beaucoup plus importante pour les filtres BStemps et BSspatio-temps que pour les filtres DNLM_T et DBM3D_T. Cela peut s'expliquer par le fait que d'une part ces filtres bénéficient de l'avantage des approches à patches de ne pas nécessiter d'estimation précise préalable du mouvement du fond et d'autre part par le fait qu'ils arrivent à tirer un peu plus parti de la corrélation spatiale des fonds que les filtres BStemps et BSspatio-temps. Cependant il faut noter que toutes les approches temporelles étudiées donnent des performances de détection bien inférieures à celles obtenues avec DBM3D_S. Autrement dit, lorsque la corrélation temporelle diminue, il vaut mieux ne pas utiliser du tout l'aspect temporel et exploiter uniquement la redondance spatiale des images comme le fait DBM3D_S.

Influence de la vitesse des cibles

La figure 5.17b présente les courbes ROC obtenues pour la séquence acquise à la fréquence nominale (25Hz) pour une vitesse de cibles faible (0.25 pixels/image). Une forte dégradation des performances se produit pour l'ensemble des filtres spatio-temporels par rapport au cas nominal présenté en figure 5.16 (vitesse de cible de 1.25 pixels/image) ; et toutes ces approches temporelles étudiées donnent des performances de détection bien inférieures à celles obtenues avec DBM3D_S.

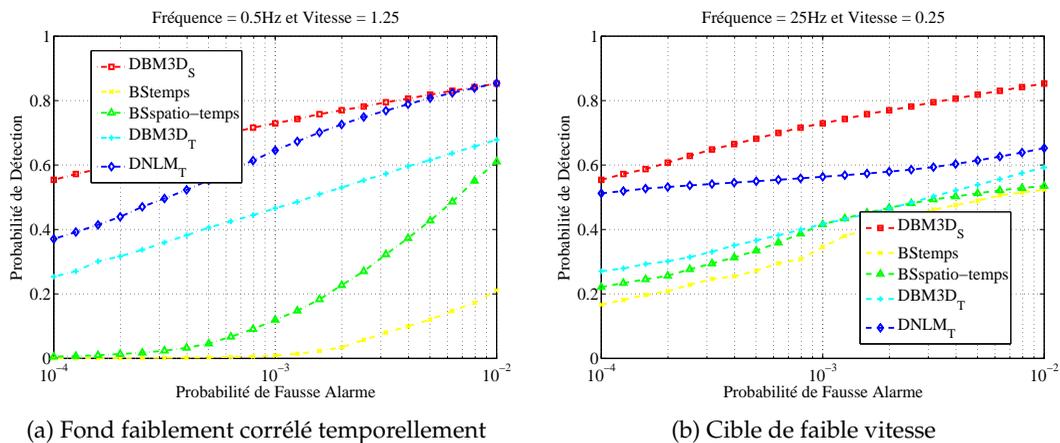


FIGURE 5.17 – Courbes ROC obtenues dans le cas d'une cible de faible vitesse : 0.25 pixels/image (fréquence d'acquisition de la séquence : 25 Hz).

Les résultats obtenus dans cette section montrent que simplement étendre la zone de recherche de patches en temporel des filtres de détection spatiaux DNLM et DBM3D

ne conduit pas à des filtres de détection spatio-temporels intéressants. En effet, ceux ci restent nettement moins performants que les algorithmes BStemps et BSspatio-temporel dans le cas de cibles rapides et de fonds fortement corrélés temporellement et de plus ils sont moins performants qu'une approche spatiale (DBM3D) dans le cas de cibles de vitesse faible et/ou de fonds non-corrélés temporellement.

5.3.2 Algorithme basé sur une modélisation du fond au second ordre : GMMF

Maintenant que nous avons vu que l'extension temporelle des approches BS par patches ne fournit pas de bons résultats, nous analysons les résultats obtenus avec la version spatio-temporelle du filtre GMMF proposée en section 4.8.2. Ce filtre, pour rappel, fait intervenir la même classification des pixels que sa version spatiale mais fait intervenir dans le test de l'équation (4.5) des vecteurs d'observations incluant les pixels compris dans un voisinage spatio-temporel de dimension 9×9 pixels $\times 2$ images et une signature de cible \mathbf{s} dont la composante associée au pixel en cours de débruitage est égale à 1 et dont toutes les autres composantes sont nulles. Ce filtre, noté GMMFspatio-temps, est comparé à sa version spatiale (notée GMMFspatial), aux filtres BSspatio-temps et BStemps (cf. section 5.3.1) et au filtre GMFspatio-temps qui correspond au GMMFspatio-temps dans le cas où le nombre de classes K égal 1. Dans cette section tous les filtres de détection cités sont appliqués directement sur les données brutes.

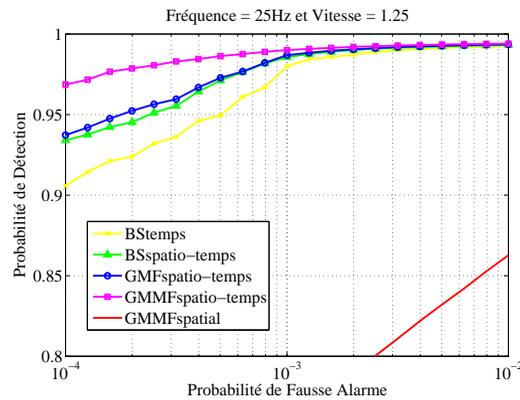
Pour le cas nominal où la vitesse de la cible est de 1.25 pixels/image, les courbes ROC obtenues sont présentées en figure 5.18a. Elles montrent d'abord que les filtres spatio-temporels sont tous plus performants que la méthode spatiale GMMF, ensuite que le filtre GMMFspatio-temporel est le plus performant des filtres spatio-temporels. Notons par ailleurs que les filtres GMFspatio-temps et GMMFspatio-temps ont l'avantage sur les BStemps et BSspatio-temps de ne pas nécessiter d'estimation précise préalable du mouvement du fond. Tant que le pixel de fond fortement corrélé au pixel en cours de test se trouve dans le voisinage considéré pour constituer le vecteur d'observation \mathbf{r}_x l'algorithme est capable de tirer parti de cette corrélation.

Influence du mouvement du fond nuageux

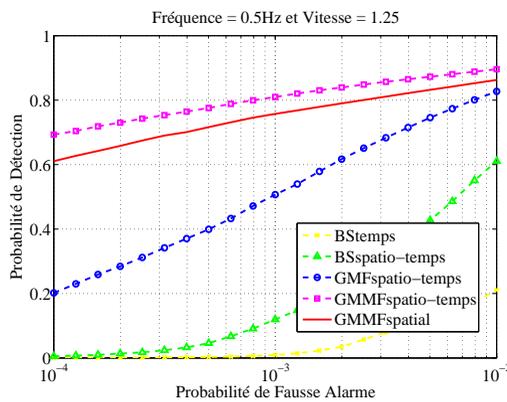
Les courbes ROC obtenues pour la séquence correspondant à une fréquence d'acquisition de 0.5 Hz sont présentées en figure 5.18b. Une forte dégradation des performances se produit pour l'ensemble des filtres spatio-temporels par rapport au cas nominal présenté en figure 5.18a (fréquence d'acquisition de la séquence 25Hz). Comme observé dans la section précédente le filtre spatial surclasse alors tous les filtres spatio-temporels à l'exception notable du filtre GMMFspatio-temps. Parmi les différentes solutions proposées dans cette thèse, le GMMFspatio-temps est donc la seule structure de filtre capable de tirer parti de la faible corrélation temporelle des pixels dans ce cas.

Influence de la vitesse des cibles

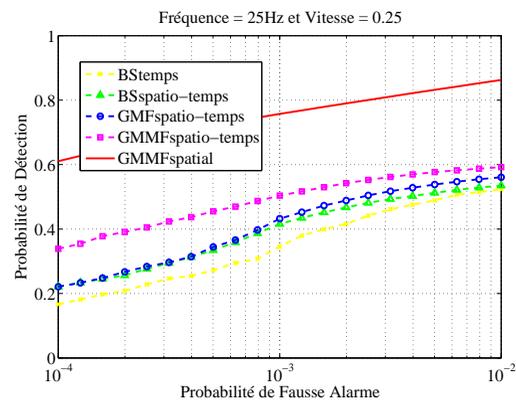
La figure 5.18c présente les courbes ROC obtenues pour la séquence acquise à la fréquence nominale (25Hz) pour une vitesse de cibles faible (0.25 pixels/image). Une forte dégradation des performances se produit pour l'ensemble des filtres spatio-temporels par rapport au cas nominal présenté en figure 5.18a (vitesse de cible de 1.25 pixels/image). Cette fois-ci le GMMF spatial donne des performances bien meilleures que toutes les approches spatio-temporelles. Mais parmi celles-ci, la meilleure performance est obtenue par le filtre GMMFspatio-temps.



(a) Cas nominal



(b) Fond faiblement corrélé temporellement



(c) Cible de faible vitesse

FIGURE 5.18 – Courbes ROC obtenues pour différentes fréquences d'acquisition et différentes vitesses : (a) vitesse de cibles = 1.25 pixels/image et fréquence d'acquisition de la séquence = 25 Hz, (b) vitesse de cibles = 1.25 pixels/image et fréquence d'acquisition de la séquence = 0.5 Hz, (c) vitesse de cibles = 0.25 pixels/image et fréquence d'acquisition de la séquence = 25 Hz.

Le GMMFspatio-temporel apparaît donc très efficace dans tous les cas excepté pour les très faibles vitesses de cible. Dans le dernier chapitre, nous proposerons en perspective une approche qui pourrait permettre de résoudre ce problème.

Conclusion

Dans ce chapitre nous avons analysé les performances des méthodes de détection que nous avons proposées aux chapitres 3 et 4. Le tableau 5.6 récapitule les algorithmes de détection qui, parmi l'ensemble des algorithmes testés, s'avèrent être les plus performants.

Nous nous sommes tout d'abord concentrés sur les méthodes de détection spatiales. La combinaison de filtres basés sur une modélisation du fond au premier et second ordre qui est la plus performante parmi l'ensemble des combinaisons testées est le filtre DBM3D+GMMF. Dans cette combinaison, c'est le filtre GMMF qui s'avère être la clef pour obtenir de si bonnes performances de détection : le filtrage GMMF appliqué directement sur les données brutes est plus performant que le filtre DBM3D seul. L'association qui

Contexte	Mono-image	Multi-images	
		Cibles rapides	Cibles lentes
Filtre le plus performant	DBM3D _S +GMMF _S spatial (Meilleur compromis coût de calcul / performances : BSMean _S +GMMF ₀ spatial)	GMMFspatio-temporel	DBM3D _S +GMMF _S spatial (Meilleur compromis coût de calcul / performances : BSMean _S +GMMF ₀ spatial)

TABLE 5.6 – Synthèse des associations de filtres les plus performantes en fonction du contexte.

donne le meilleur compromis coût de calcul / performances est BSMean+GMMF₀. En effet, elle est beaucoup moins coûteuse en temps de calcul, pour des performances seulement faiblement inférieures à celles obtenues avec DBM3D+GMMF.

Concernant maintenant les approches de détection spatio-temporelles, en reprenant le principe des approches de débruitage spatio-temporelles par patches qui ont l'avantage de ne pas nécessiter d'estimation du mouvement, on espérait obtenir des approches de détection très performantes. Or finalement, on ne retrouve pas le gain observé en spatial avec les méthodes à patches : dans le cas nominal, les traitements simples sont plus performants et dans les cas dégradés, le spatial pur est plus performant. Ces résultats appellent plusieurs commentaires. D'une part, dans le cas nominal les performances obtenues par les approches par soustraction des images précédentes sont excellentes et donc difficiles à améliorer. D'autre part, le principe de la recherche de patches similaires est très perturbé par la présence de la cible dans les images temporellement proches. A cause de ce dernier point, il est selon nous non pertinent de poursuivre dans la voie des approches à patches spatio-temporelles et préférable pour développer des méthodes de détection spatio-temporelles performantes de s'orienter vers d'autres principes.

Par contre l'approche de détection spatio-temporelle **GMMF** proposée en section 4.8.2 apparaît comme étant prometteuse. Cette méthode permet d'exploiter aussi bien la corrélation temporelle que spatiale des fonds, ce qui lui permet d'obtenir les meilleurs résultats des méthodes spatio-temporelles dans tous les cas testés. Son unique défaut à l'heure actuelle est son manque de robustesse aux cibles de faible vitesse, mais nous verrons en conclusion qu'il existe des voies d'améliorations sur ce point.

On peut aussi se poser la question de définir une méthode assurant un bon compromis coût de calcul / performances dans le cas spatio-temporel, comme nous l'avons fait avec BSMean+GMMF₀ dans le cas spatial. Nous n'avons pas eu la possibilité de conduire une étude expérimentale de performance sur ce point dans la durée de ce travail, cependant nous pensons qu'une solution consisterait à associer (1) un filtrage BSMean de chaque image ; (2) une classification spatiale de l'image de référence par K-means et (3) un **GMMF** spatio-temporel utilisant ces classes. Cette solution devrait, à notre avis, fournir de bons résultats pour des cibles rapides. Son évaluation empirique sur les bases d'images est une perspective de notre travail.

Conclusion générale et perspectives

Conclusion générale

Dans cette thèse nous nous sommes intéressés à la détection d'objets ponctuels sur des fonds nuageux dans un contexte de surveillance passive. Les performances de détection par un opérateur humain ayant le défaut de varier en fonction de l'opérateur ou en fonction de l'état de fatigue d'un même opérateur, et la quantité de données à traiter étant de plus en plus grande, il est aujourd'hui indispensable de disposer d'algorithmes performants de détection automatique de cibles. Nous avons concentré nos recherches sur un des contextes les plus difficiles qui est celui de la détection de cibles de faible intensité et de faible vitesse sur des fonds structurés inhomogènes présentant des zones de fortes transitions.

De la théorie de la détection dans un cadre statistique nous avons commencé par faire ressortir un schéma de détection en deux étapes qui consiste à séparer l'estimation des statistiques du fond au premier ordre de l'estimation de celles du second ordre. Ce schéma a été le fil conducteur de notre démarche dans cette thèse. La première étape est une suppression de fond qui repose sur une modélisation du fond au premier ordre, soit une estimation de la moyenne du fond. La deuxième étape est un test du rapport de vraisemblance généralisé (GLRT : *Generalized Likelihood Ratio Test*) basé sur la connaissance des statistiques locales au second ordre du modèle de fond ce qui nécessite d'estimer localement les matrices de covariance. Dans les deux cas le principal problème est d'identifier le comportement statistique du fond en dépit de la présence de cibles dans les données observées.

L'utilisation du principe des méthodes de débruitage à patches s'est avérée être une solution intéressante pour la première étape qui a conduit à des méthodes de détection mono-image par suppression de fond performantes. Nous avons dans un premier temps adapté la méthode de débruitage à patches **NL-means** à la détection mono-image, ce qui a conduit au filtre de détection **DNLM** (*Detection by NLM*) [Genin et al., 2011]. Puis nous avons montré dans un second temps que la méthode de débruitage à patches plus récente **BM3D**, qui est une des méthodes de débruitage actuellement les plus performantes, peut être adaptée aisément à notre problématique de détection et conduire à une méthode de détection spatiale **DBM3D** (*Detection by BM3D*) [Genin et al., 2012a] s'avérant être nettement plus performante que la première approche proposée (**DNLM**).

Concernant la deuxième étape, une approche originale nommée **GMMF** (*Gaussian Mixture Matched Filter*) a été présentée [Genin et al., 2012a]. Les pixels sont regroupés en classes de statistiques du second ordre homogènes en utilisant l'algorithme de classification non-supervisé **CEM** basé sur un modèle de mélange de gaussiennes uniformément pondérées. Nous avons choisi de faire intervenir dans le processus de classification des vecteurs d'observations restreints composés de l'intensité des pixels inclus dans une cou-

ronne centrée sur le pixel à classifier, ce qui permet de traiter de façon simple la présence de cibles dans les données. Pour finir, on applique en chaque pixel le **GLRT** qui fait intervenir la matrice de covariance estimée sur la classe à laquelle il appartient. La solution retenue résulte de l'étude de nombreuses variantes de celle-ci et donc de choix qui ont été validés par des études de performances. En plus de permettre d'augmenter très nettement la probabilité de détection en présence de fonds non-stationnaires, cette approche a l'avantage de ne faire intervenir qu'un nombre très limité de paramètres.

L'association de filtres spatiaux (selon le schéma de détection en deux étapes proposé) qui fournit les meilleures performances de détection est **DBM3D + GMMF**. On peut noter que dans cette combinaison de filtres, c'est la méthode **GMMF** qui semble contribuer majoritairement au rehaussement de cibles. En effet le filtre **GMMF** basé sur une modélisation du fond au second ordre appliqué directement sur les données brutes fournit de meilleures performances de détection que le filtre par suppression de fond **DBM3D**. Une version sous-optimale **BSMean+GMMF₀** permet d'obtenir un très bon compromis coût de calcul / performances et a de plus l'avantage de faire intervenir un nombre de paramètres beaucoup moins important que l'approche **DBM3D+GMMF**.

L'extension spatio-temporelle des approches à patches ne conduit pas à des résultats intéressants, en revanche l'extension temporelle du filtre **GMMF** fournit un filtre de détection très prometteur. Cette extension consiste à faire intervenir dans le test statistique des vecteurs d'observation constitués de l'intensité des pixels inclus dans un voisinage spatio-temporel. La méthode de classification des pixels reste la méthode spatiale qui est utilisée dans la version spatiale du filtre **GMMF**. Le filtre de détection ainsi obtenu a l'avantage pour les cibles rapides de tirer parti à la fois des corrélations temporelles et spatiales des fonds. Il a de plus l'avantage de ne pas nécessiter d'estimation précise préalable du mouvement du fond.

Perspectives

Accélération des algorithmes

Dans cette thèse, nous avons tenu compte du coût de calcul des algorithmes lors de la recherche de l'approche de détection fournissant le meilleur compromis coût de calcul / performances. Cependant les algorithmes de détection ont été développés sous Matlab sans avoir été spécifiquement optimisés en temps de calcul. Des travaux concernant l'accélération des filtres de détection proposés dans cette thèse sont donc à prévoir, en envisageant par exemple une implémentation parallélisée sur une architecture dédiée (FPGA, GPU).

Robustesse de l'algorithme **GMMF** spatio-temporel à la variation de la vitesse des cibles

L'approche de détection spatio-temporelle **GMMF** proposée en section 4.8.2 a le défaut de ne pas être robuste à la présence de cibles de faible vitesse. Dans ce paragraphe nous présentons des résultats préliminaires d'une approche permettant d'améliorer cette robustesse. Rappelons que le **GMMF** présenté en 4.8.2 fait intervenir la signature de cibles donnée en figure 5.19a. Pour être robuste à la présence de cibles de faible vitesse, on peut envisager d'utiliser à la place de cette signature celle donnée en figure 5.19b mais dans ce cas les performances de détection pour les cibles de vitesse plus élevée sont

nettement dégradées. Pour faire face au fait que la vitesse et la direction de propagation des cibles sont inconnues, il est possible de considérer un test **GLRT** au sein duquel, sous l'hypothèse qu'une cible est présente (hypothèse H_0), ces paramètres inconnus sont estimés par maximum de vraisemblance. Plus concrètement, cela consiste à reprendre le test donné par l'équation (4.19) qui a été présenté pour tenir compte de la position subpixelique des cibles, en cherchant la signature temporelle qui maximise le test. Cette maximisation est conduite parmi un ensemble restreint de signatures temporelles de cibles représentatif de l'ensemble des signatures possibles.

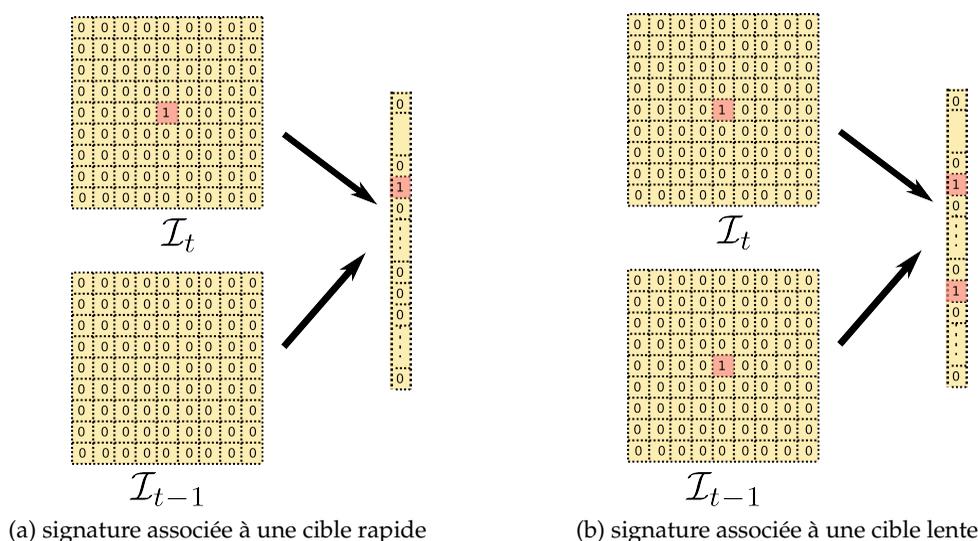


FIGURE 5.19 – Signatures de cibles intervenant dans le test statistique du filtre **GMMF**.

Des résultats présentant les performances obtenues dans le cas où cet ensemble est constitué des deux signatures données en figure 5.19 sont présentés en figure 5.20. Ces premiers tests montrent que cette approche de sélection par maximum de vraisemblance permet de nettement augmenter les performances de détection de cibles lentes (0.25 pixels/image) sans dégrader les performances de détection des cibles plus rapides (1.25 pixels/images ou 8.25 pixels/images). On obtient ainsi, quelque soit la vitesse des cibles, un filtre de détection qui a l'avantage de tirer parti de la corrélation temporelle du fond éventuellement présente tout en étant capable de tirer parti de la corrélation spatiale des fonds aussi bien que la version spatiale du **GMMF** quand la corrélation temporelle des fonds est faible (cf. résultats de la figure 5.20).

Cet algorithme est donc très prometteur et une étude plus approfondie de cette approche reste à réaliser en intégrant par exemple plus de signature de cibles dans l'ensemble dans lequel on recherche la signature de la cible.

Détection directe dans le domaine des patches

Dans cette thèse, nous avons repris le principe des méthodes de débruitage à patches uniquement dans le cadre d'une approche de détection par suppression de fond. Or nous pourrions sortir de cette approche et chercher à obtenir un algorithme de détection basé sur les approches à patches, qui ferait directement ressortir les cibles par rapport au fond dans l'espace des patches (cf. schéma de la figure 3.3). Dans le domaine du débruitage de

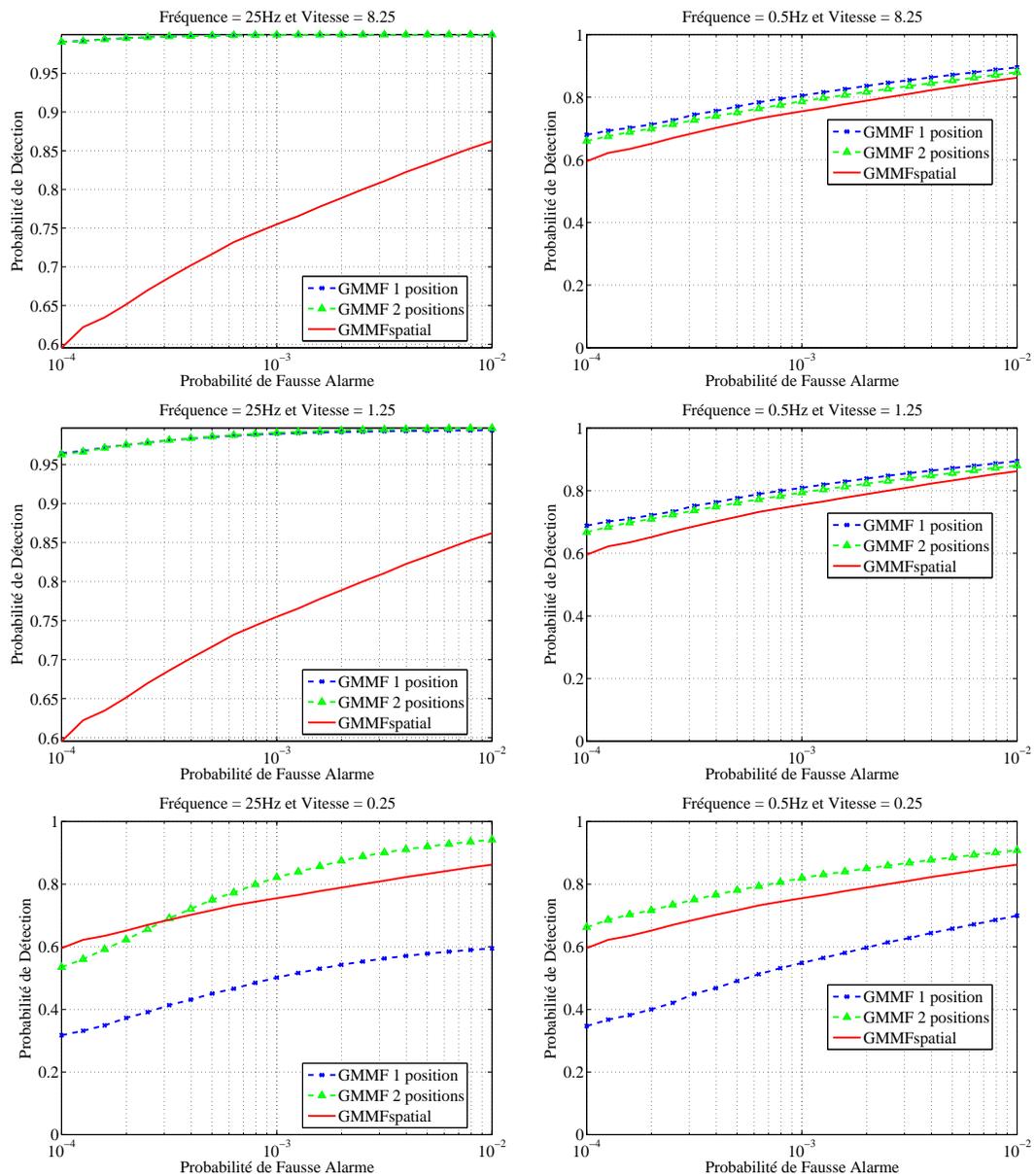


FIGURE 5.20 – Courbes ROC obtenues avec le filtre GMMF spatio-temporel sur 2 images faisant intervenir uniquement la signature de cibles donnée en figure 5.19a (GMMF 1 position), le filtre GMMF qui sélectionne par maximum de vraisemblance la signature de cibles parmi les deux signatures données en figure 5.19 (GMMF 2 positions) et la version spatiale du filtre GMMF (GMMF spatial). A gauche : forte corrélation temporelle des fonds, à droite : faible corrélation temporelle des fonds. De haut en bas : la vitesse des cibles est 8.25 pixel/image, 1.25 pixel/images et 0.25 pixels/image.

nombreux travaux de recherche ont été menés sur les approches à patchs conduisant au développement de nouvelles approches dont une nommée NL-Bayes [Buades *et al.*, 2012]. Cette dernière reprend le principe de construction de blocs proposé dans DBM3D mais les débruite non pas en passant par une décomposition en ondelette mais en se basant sur une modélisation statistique du fond au second ordre. On pourrait donc envisager de reprendre une telle approche et reprendre l'idée de la modélisation statistique du fond dans chacun des blocs pour y faire ressortir les cibles. Ces méthodes de débruitage faisant intervenir des matrices de covariance dans chacun des blocs, on peut envisager de s'en servir pour appliquer des filtres adaptés (équation (1.25)). Dans ce cadre, des travaux de recherche restent à faire pour trouver une méthode efficace de combinaison de l'information provenant des différents patchs et pour évaluer les performances de détection que l'on pourrait obtenir avec une telle approche.

Association avec des méthodes de pistage

Les performances de détection évaluées au cours de cette thèse sont celles obtenues directement en sortie des algorithmes de détection spatiaux ou spatio-temporels basés sur une modélisation de fond que nous proposons. En pratique, ces algorithmes de détection sont le plus souvent suivis d'algorithmes de pistage permettant d'éliminer un certain nombre de fausses alarmes. La caractérisation complémentaire des algorithmes par analyse de la répartition des fausses alarmes et calcul du critère DNFA nous a permis de vérifier que les algorithmes proposés les plus performants n'avaient pas le comportement défavorable de concentrer les fausses alarmes au niveau des bords. Maintenant, il serait intéressant d'associer ces algorithmes de détection à des techniques de pistage et d'évaluer les performances de détection de l'ensemble.

Conception conjointe du système imageur et des traitements pour améliorer la détection

Pour pouvoir améliorer davantage les performances de détection, il serait intéressant d'optimiser l'ensemble : système imageur et algorithme de détection. Comme nous l'avons dit au cours de cette thèse, les cibles étant non résolue au moment de leur détection, leur signature correspond à la PSF échantillonnée du système. Nous avons considéré dans cette thèse uniquement des systèmes pour lesquels la PSF du système imageur correspond à une fonction d'Airy. Cependant, on peut par exemple envisager de proposer des systèmes imageur pour lesquels la PSF sera différente et donc la signature de cibles à détecter aussi et chercher les combinaisons : forme de PSF / algorithme de détection qui permettent d'améliorer la détection. On peut par exemple modifier la PSF en introduisant un mouvement du système imageur et chercher à détecter une signature de cibles liée à un flou de bougé comme proposé dans [Sui *et al.*, 2011]. Dans ce cadre on pourrait utiliser sans difficultés l'algorithme de détection GMMF car la variation de la PSF n'est pas un problème pour lui. Il suffirait simplement de considérer dans le test statistique (équation (4.5)) une nouvelle signature de cibles liée à la nouvelle PSF. On a de plus l'avantage de pouvoir associer au GMMF un modèle théorique de performances ce qui est très intéressant en co-conception.

Annexe A

Calcul du GLRT dans le cas où le signe de la cible est connu

Le but de cette annexe est de montrer, en reprenant les notations et le modèle de fond gaussien présentés au chapitre 1, que dans le cas où l'intensité α de la cible est inconnue mais son signe est connu, le test du **GLRT** :

$$\frac{\max_{\alpha \in \mathcal{R}} p(\mathbf{u}_x | H_1, \alpha)}{p(\mathbf{u}_x | H_0)} \underset{H_0}{\overset{H_1}{\geq}} \mu \quad \text{avec} \quad \mathcal{R} = \begin{cases} \mathbb{R}^+ & \text{pour les cibles positives } (\alpha \geq 0) \\ \mathbb{R}^- & \text{pour les cibles négatives } (\alpha \leq 0) \end{cases} \quad (\text{A.1})$$

est équivalent à :

$$\text{sign}_\alpha \frac{\mathbf{s}^T \boldsymbol{\Sigma}_x^{-1} (\mathbf{u}_x - \mathbf{m}_x)}{[\mathbf{s}^T \boldsymbol{\Sigma}_x^{-1} \mathbf{s}]^{1/2}} \underset{H_0}{\overset{H_1}{\geq}} \sqrt{2 \log(\mu)} \quad \text{avec} \quad \text{sign}_\alpha = \begin{cases} 1 & \text{pour les cibles positives} \\ -1 & \text{pour les cibles négatives} \end{cases} \quad (\text{A.2})$$

où \mathbf{m}_x et $\boldsymbol{\Sigma}_x$ sont la moyenne et la matrice de covariance de la distribution $p(\mathbf{u}_x | H_0)$ et \mathbf{s} est la signature normalisée de la cible.

Preuve:

Afin de montrer cette équivalence, notons tout d'abord que :

$$\max_{\alpha \in \mathcal{R}} p(\mathbf{u}_x | H_1, \alpha) \geq p(\mathbf{u}_x | H_1, \alpha = 0) = p(\mathbf{u}_x | H_0). \quad (\text{A.3})$$

Il s'ensuit que le rapport de vraisemblance généralisé de l'équation (A.1) est toujours supérieur à 1. Ainsi, on peut considérer $\mu > 1$, sans perte de généralité. On applique le logarithme des deux côtés de l'équation (A.1), on multiplie par 2 et permute avec le max. On obtient ainsi que l'équation (A.1) est équivalent à :

$$\max_{\alpha \in \mathcal{R}} 2 \log \left(\frac{p(\mathbf{u}_x | H_1, \alpha)}{p(\mathbf{u}_x | H_0)} \right) \underset{H_0}{\overset{H_1}{\geq}} 2 \log(\mu). \quad (\text{A.4})$$

On spécifie maintenant le logarithme du rapport de vraisemblance dans le cas gaussien, en utilisant le fait que \mathbf{u}_x est la réalisation d'une variable aléatoire U_x qui suit la loi $\mathcal{N}(\mathbf{m}_x + \alpha \mathbf{s}, \boldsymbol{\Sigma}_x)$ (resp. $\mathcal{N}(\mathbf{m}_x, \boldsymbol{\Sigma}_x)$) sous l'hypothèse H_1 (resp. H_0). La fonction $P(\alpha)$ définie de la façon suivante :

$$P(\alpha) = 2 \log \left(\frac{p(\mathbf{u}_x | H_1, \alpha)}{p(\mathbf{u}_x | H_0)} \right), \quad (\text{A.5})$$

est un polynôme du second degré en α qui s'écrit :

$$P(\alpha) = -\mathbf{s}^T \Sigma_x^{-1} \mathbf{s} \alpha^2 + 2\mathbf{s}^T \Sigma_x^{-1} (\mathbf{u}_x - \mathbf{m}_x) \alpha. \quad (\text{A.6})$$

Puisque $\mathbf{s}^T \Sigma_x^{-1} \mathbf{s} > 0$, le polynôme atteint sa valeur maximum sur \mathcal{R} ($\mathcal{R} = \mathbb{R}^+$ ou $\mathcal{R} = \mathbb{R}^-$) soit pour $\alpha = 0$, soit pour $\alpha = \mathbf{s}^T \Sigma_x^{-1} (\mathbf{u}_x - \mathbf{m}_x) / \mathbf{s}^T \Sigma_x^{-1} \mathbf{s}$. L'alternative est gouvernée par le signe de $\mathbf{s}^T \Sigma_x^{-1} (\mathbf{u}_x - \mathbf{m}_x)$ et va dépendre de \mathcal{R} ($\mathcal{R} = \mathbb{R}^+$ ou $\mathcal{R} = \mathbb{R}^-$). Dans la suite, on distingue les deux cas suivants : $\mathcal{R} = \mathbb{R}^+$ et $\mathcal{R} = \mathbb{R}^-$.

Cas où $\mathcal{R} = \mathbb{R}^+$ ($\text{sign}_\alpha = +1$) :

Considérons d'abord $\mathbf{s}^T \Sigma_x^{-1} (\mathbf{u}_x - \mathbf{m}_x) > 0$, pour un tel \mathbf{u}_x ,

$$\max_{\alpha \geq 0} P(\alpha) = \frac{[\mathbf{s}^T \Sigma_x^{-1} (\mathbf{u}_x - \mathbf{m}_x)]^2}{\mathbf{s}^T \Sigma_x^{-1} \mathbf{s}}. \quad (\text{A.7})$$

Puisque $\mu > 1$, $\log \mu > 0$, et comme $\mathbf{s}^T \Sigma_x^{-1} (\mathbf{u}_x - \mathbf{m}_x) > 0$, on peut appliquer une racine carrée des deux côtés de l'équation (A.4) ce qui nous donne l'équation (A.2).

Enfin, considérons $\mathbf{s}^T \Sigma_x^{-1} (\mathbf{u}_x - \mathbf{m}_x) \leq 0$, pour un tel \mathbf{u}_x , $\max_{\alpha \geq 0} P(\alpha) = 0$, et le rapport de l'équation (A.1) est donc égal à 1. Puisque $\mu > 1$, le test de l'équation (A.1) est toujours strictement inférieur au seuil μ . De plus, comme $\mathbf{s}^T \Sigma_x^{-1} (\mathbf{u}_x - \mathbf{m}_x) \leq 0$, le test de l'équation (A.2) est toujours lui aussi strictement inférieur au seuil $\sqrt{2 \log(\mu)}$.

Nous avons montré que, pour tout \mathbf{u}_x les tests des équations (A.1) et (A.2) sont équivalents.

Cas où $\mathcal{R} = \mathbb{R}^-$ ($\text{sign}_\alpha = -1$) :

Considérons d'abord $\mathbf{s}^T \Sigma_x^{-1} (\mathbf{u}_x - \mathbf{m}_x) > 0$, pour un tel \mathbf{u}_x , $\max_{\alpha \leq 0} P(\alpha) = 0$, et le rapport de l'équation (A.1) est donc égal à 1. Puisque $\mu > 1$, le test de l'équation (A.1) est toujours strictement inférieur au seuil μ . De plus, comme $\mathbf{s}^T \Sigma_x^{-1} (\mathbf{u}_x - \mathbf{m}_x) > 0$ et $\text{sign}_\alpha = -1$, on a $\text{sign}_\alpha \mathbf{s}^T \Sigma_x^{-1} (\mathbf{u}_x - \mathbf{m}_x) < 0$ et donc le test de l'équation (A.2) est toujours lui aussi strictement inférieur au seuil $\sqrt{2 \log(\mu)}$.

Enfin, considérons $\mathbf{s}^T \Sigma_x^{-1} (\mathbf{u}_x - \mathbf{m}_x) \leq 0$, pour un tel \mathbf{u}_x ,

$$\max_{\alpha \leq 0} P(\alpha) = \frac{[\mathbf{s}^T \Sigma_x^{-1} (\mathbf{u}_x - \mathbf{m}_x)]^2}{\mathbf{s}^T \Sigma_x^{-1} \mathbf{s}}. \quad (\text{A.8})$$

Puisque $\mu > 1$, $\log \mu > 0$, et comme $\mathbf{s}^T \Sigma_x^{-1} (\mathbf{u}_x - \mathbf{m}_x) \leq 0$, $\text{sign}_\alpha \mathbf{s}^T \Sigma_x^{-1} (\mathbf{u}_x - \mathbf{m}_x) \geq 0$, on peut donc appliquer une racine carrée des deux côtés de l'équation (A.4) ce qui nous donne l'équation (A.2).

Nous avons montré que, pour tout \mathbf{u}_x les tests des équations (A.1) et (A.2) sont équivalents.

Nous avons montré que, pour tout \mathbf{u}_x , que \mathcal{R} soit égal à \mathbb{R}^+ ou à \mathbb{R}^- , les tests des équations (A.1) et (A.2) sont équivalents. ■

Annexe B

Étude paramétrique du DNLM

Pour rappel, les paramètres du filtre DNLM sont les suivants :

- h : paramètre du NL-means ajustant l'exigence de similarité entre les patchs intervenant dans le calcul des poids défini par l'équation (3.4),
- N_{max} : nombre de valeurs maximales de la différence en valeur absolue de deux patchs qui sont ignorées pour évaluer la mesure de dissimilarité (eq. (3.13)) entre ces deux patchs quand le test défini en équation (3.15) est validé,
- λ : paramètre utilisé pour décider si l'on omet ou non les N_{max} valeurs maximales dans le calcul de la dissimilarité entre les patchs (eq. (3.15))
- N_1 : dimension du côté des patchs carré $P_{x-\delta_{N_1}}$ et $P_{y-\delta_{N_1}}$ considérés pour évaluer le poids des pixels y intervenant dans l'estimation du pixel x ,
- a : paramètre du noyau gaussien $G(a)$ qui permet de régler l'importance que l'on accorde aux pixels qui sont les plus proches du centre par rapport aux autres (quand a diminue, le poids accordé aux pixels proches du centre augmente),
- N_s : dimension du côté de la fenêtre carré S_x , centrée sur x , qui contient l'ensemble des pixels intervenant dans l'estimation de ce pixel x par une moyenne pondérée.

Les valeurs nominales des paramètres, qui sont les valeurs utilisées par la suite quand aucune précision n'est donnée, sont présentées dans tableau 3.1. Dans cette annexe, l'influence de la variation de ceux-ci sur les performances du filtre de détection est analysée. Elle permet de justifier les valeurs nominales choisies.

Influence du paramètre h

La figure B.1 présente la variation des performances de détection pour différentes valeurs du paramètre h . Les autres paramètres sont ceux donnés dans le tableau 3.1. Le réglage de ce paramètre est une question de compromis. Quand h augmente, l'exigence de similarité diminue, le filtre d'estimation du fond a tendance à lisser l'image, la cible est donc très bien rehaussée dans l'image de résidu mais le fond est moins bien supprimé ce qui augmente le nombre de fausses alarmes. Quand le paramètre h est très élevé le filtre d'estimation du fond devient équivalent à un filtre moyenneur classique sur la fenêtre S_x . Inversement, quand le paramètre h est faible, l'exigence de similarité est très stricte, le filtre d'estimation du fond moyenne très peu l'image (un nombre très restreint de pixels intervient dans la moyenne pondérée). Dans ce cas, dans l'image de résidu, le fond est parfaitement supprimé mais la cible l'est en partie aussi et son intensité chute. Il faut donc choisir un paramètre h qui n'est pas trop élevé pour ne pas avoir trop de fausses alarmes,

ni trop faible pour que la cible puisse être correctement rehaussée. Ce paramètre est le plus souvent fixé de façon empirique et dépend du niveau de bruit dans l'image.

La figure B.1, montre qu'une valeur de h égale à 9 pour la base de données terrestre et égale à 6 pour la base de données satellitaire est optimum. On peut noter que pour une valeur de h comprise entre 8 et 10 pour la base de données terrestre (respectivement entre 5 et 7 pour la base de données satellitaire), les performances ne sont que très faiblement dégradées. Pour des valeurs de h s'éloignant plus des valeurs nominales, les performances de détection chutent fortement. Le paramètre h semble être le paramètre du filtre DNLM le plus difficile à fixer, il demande donc qu'on lui accorde une attention toute particulière.

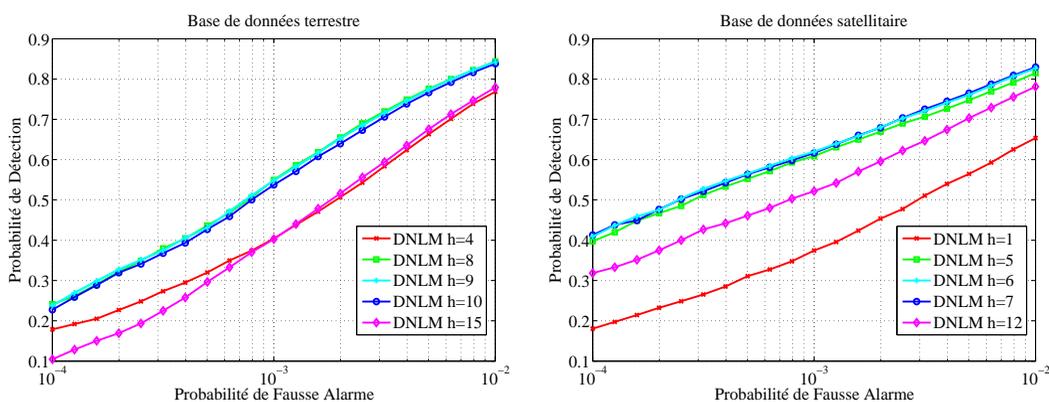


FIGURE B.1 – Influence du paramètre h : Courbes ROC obtenues avec DNLM, pour chacune des bases de données, pour différentes valeurs de h .

Influence du paramètre λ

La figure B.2 présente la variation des performances de détection en fonction du paramètre λ intervenant dans la définition de la mesure de distance entre les patches (éq. (3.15)). Tout d'abord précisons que quand $\lambda < 0$ le filtre DNLM est équivalent au filtre de BS qui estime simplement le fond par NL-means ; il utilise toujours la mesure de distance non-modifiée, d (éq. 3.5). Quand $\lambda > 1$, le filtre DNLM fait toujours intervenir la mesure de distance modifiée d_1 (éq. 3.15). La valeur du paramètre λ ne doit pas être trop grande pour que, en l'absence de cible, on applique bien la mesure de distance non-modifiée (éq. 3.5). Mais d'un autre côté si celle-ci est trop faible on risque d'appliquer aussi la mesure non-modifiée à des cibles intenses qui ne sont pas préservées lors de l'application du filtre de détection qui estime simplement le fond par NL-means. D'après la figure B.2, un paramètre $\lambda = 0.08$ fournit un bon compromis.

Influence du nombre de maxima supprimés

La figure B.3 présente la variation des performances de détection en fonction du nombre de maxima, $Nmax$, intervenant dans la mesure de distance du filtre DNLM (éq. (3.15)). Rappelons tout d'abord que dans le cas où la mesure de distance est toujours la mesure d_1 ($\lambda > 1$), les performances de détection d'une cible de faible intensité sont dégradées quand $Nmax$ augmente comme le montre la figure 3.11. Dans le cas où l'on utilise la mesure de distance d_2 avec $\lambda = 0.08$, le réglage du paramètre $Nmax$ est moins sensible. En effet pour des valeurs de $Nmax$ allant de 1 à 3, les courbes sont quasiment identiques. Le nombre de maxima supprimés doit être adapté à la dimension de la cible

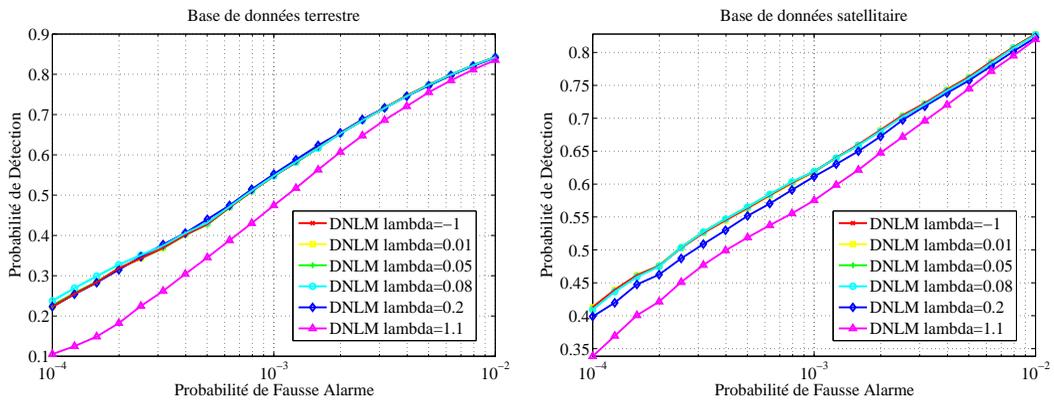


FIGURE B.2 – Influence du paramètre λ : Courbes ROC obtenues avec DNLM, pour chacune des bases de données, pour différentes valeurs de λ .

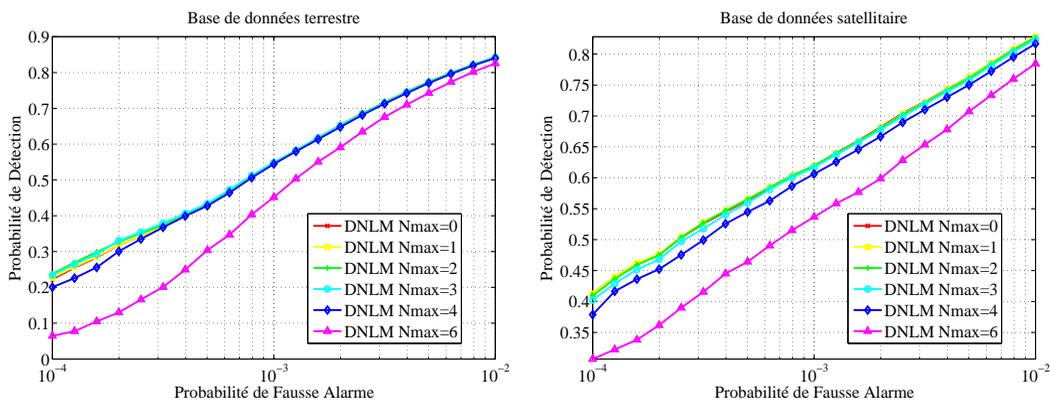


FIGURE B.3 – Influence du paramètre N_{max} : Courbes ROC obtenues avec DNLM, pour chacune des bases de données, pour différentes valeurs de N_{max} .

mais il faut aussi tenir compte de la dimension du voisinage Q utilisé dans la mesure de similarité. Le rapport entre le nombre de maxima supprimés et le nombre de pixels présents dans le voisinage Q doit rester suffisamment grand. Pour les cibles considérées qui sont relativement bien ponctuelles (rapport $r_c = 1.5$, cf. section 2.5), un N_{max} égal à deux apparaît comme étant un bon compromis.

Influence de la dimension du patch

La figure B.4 présente les performances de détection obtenues pour différentes valeurs de N_1 , c.a.d. pour différentes dimensions de patches. On observe que l'on obtient des performances de détection identiques avec quasiment toutes les valeurs de N_1 testées. On observe simplement une chute de performance pour la base de données satellitaire pour $N_1 = 3$. Un patch de dimension 3×3 s'avère être trop petit pour capter les textures présentes dans cette base de données ; des patches de dimension 5×5 sont plus adaptés à l'échelle des textures présentes. Le fait que les performances de détection n'évoluent pas quand on augmente la taille des patches est lié au paramètre a du noyau gaussien $G(a)$ qui a été choisi. En pratique au delà d'une certaine valeur de N_1 c'est le paramètre a qui limite les pixels intervenant dans la mesure de distance entre deux patches et non la taille des patches elle-même. En effet, les pixels se trouvant très éloignés du centre du patch vont

avoir un poids qui tend vers 0. La décroissance des poids avec la distance par rapport au centre du patch va dépendre du paramètre a . Pour le paramètre a considéré ici ($a = 1.5$), un N_1 égal à 5 fournit le meilleur compromis coût de calcul / performances pour les deux bases de données.

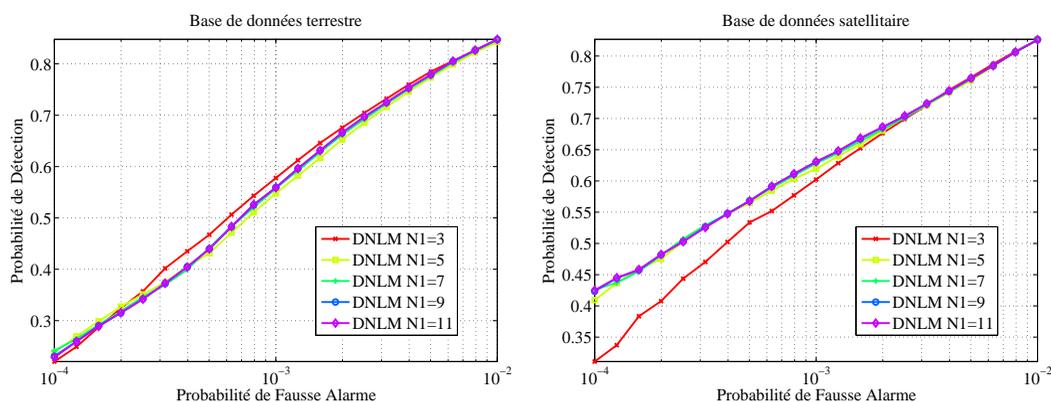


FIGURE B.4 – Influence du paramètre N_1 : Courbes ROC obtenues avec DNLM, pour chacune des bases de données, pour différentes valeurs de N_1 .

Influence du paramètre a du noyau gaussien

La figure B.5 présente les performances de détection pour différentes valeurs du paramètre a . On peut noter que quand a augmente, les courbes convergent toutes vers une même courbe de performances. En effet quand a augmente, tous les pixels d'un patch ont tendance à avoir le même poids : quand a augmente, $G(a)$ de l'équation 3.4 tend vers $1/N_1^2$. Les résultats présentés en figure B.5 montrent que pour nos deux bases de données, accorder une plus grande importance aux pixels qui sont proches du centre des patches permet d'augmenter les performances de détection. La valeur nominale du paramètre a est de 1.5, cependant l'étude paramétrique présentée ici montre qu'une valeur de a un peu plus faible que sa valeur nominale peut permettre d'améliorer les résultats.

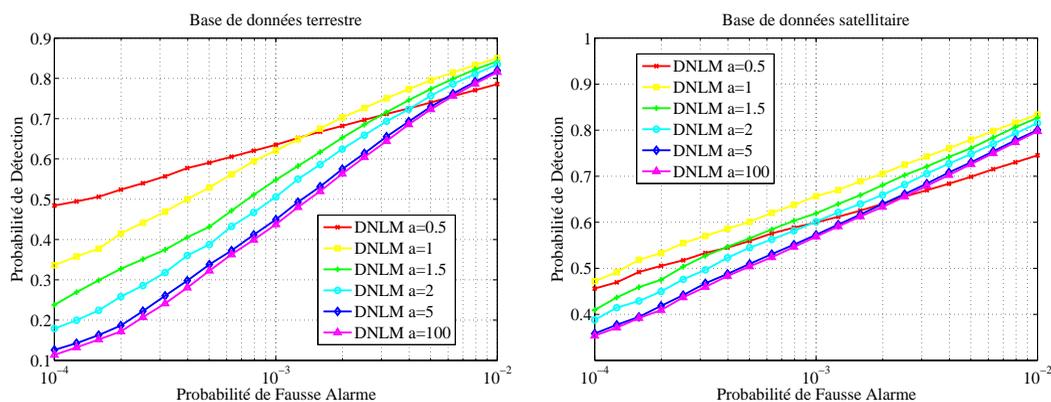


FIGURE B.5 – Influence du paramètre a : Courbes ROC obtenues avec DNLM, pour chacune des bases de données, pour différentes valeurs de a .

Influence de la dimension de la zone de recherche

La figure B.6 présente les performances de détection obtenues avec DNLM pour différentes valeurs de N_s . Elle montre que si le support S_x est très petit ($N_s = 3$), les performances de détection sont dégradées. Quand la dimension du support S_x augmente, les performances augmentent mais au delà d'un certain seuil les performances stagnent et ont même tendance à se dégrader comme on peut le voir sur la figure B.6. On aurait pu penser que le support S_x idéal soit l'image dans sa totalité. Cependant quand on augmente la taille du support S_x , beaucoup de coefficients $w(P_{x-\delta_{N_1}}, P_{y-\delta_{N_1}})$ sont proches de 0 mais non égaux à 0, ce qui a pour conséquence de diminuer l'impact des patchs qui sont les bons candidats c.a.d les patchs pour lesquels $w(P_{x-\delta_{N_1}}, P_{y-\delta_{N_1}})$ est élevé, comme noté par Tasdizen dans [Tasdizen, 2009] et par Salmon et al. dans [Salmon et Strozecki, 2012]. Comme aucun gain en terme de débruitage n'est obtenu en pratique avec un large support S_x , pour nommer la méthode NL-means, il aurait été préférable d'utiliser le terme de « Semi-Local Means » qui apparaît pour la première fois dans [Brox et al., 2008] selon Salmon et al. [Salmon et Strozecki, 2012]. Le support S_x n'est donc pas réduit uniquement pour des questions de coût de calcul, mais aussi pour augmenter les performances d'estimation de la valeur débruitée dans le cadre du débruitage d'image et les performances d'estimation du fond dans le contexte de la détection. D'un autre côté comme la dimension du support S_x doit être suffisamment grande pour trouver un nombre suffisant de patchs similaires au patch centré sur le pixel en cours de traitement, un compromis est donc nécessaire. Dans le cadre du débruitage, Kervrann et al. [Kervrann et Boulanger, 2006] ont proposé une méthode d'optimisation automatique de la zone de recherche basée sur l'application de la méthode de Lepskii. Dans notre problématique de détection, pour les deux bases de données, une valeur de N_s égale à 11, choisie empiriquement, apparaît être un bon compromis.

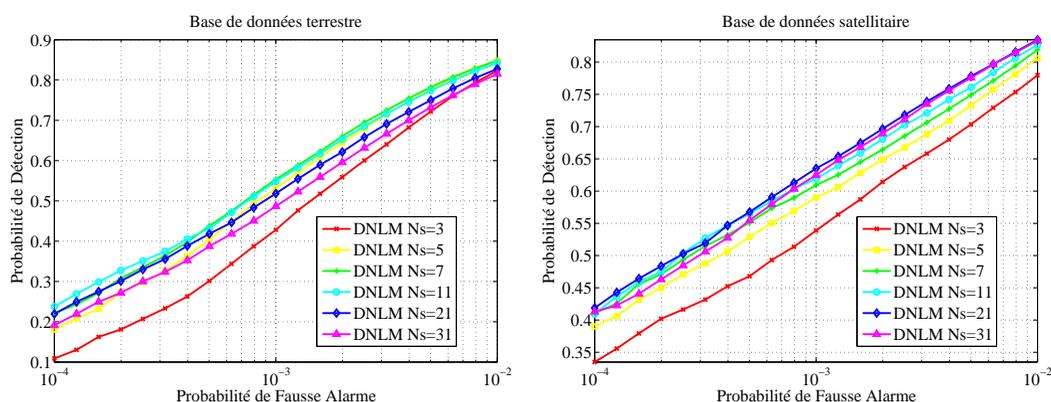


FIGURE B.6 – Influence du paramètre N_s : Courbes ROC obtenues avec DNLM, pour chacune des bases de données, pour différentes valeurs de N_s .

Annexe C

Filtre GLRT-GM

Le filtrage GLRT-GM (GLRT-Gaussian Mixture) abordé au chapitre 4 en section 4.5 est décrit dans cette annexe plus en détail et une comparaison de performances avec le filtre GMMF y est présentée. Ce filtre GLRT-GM considère que le fond peut se modéliser par un unique modèle de mélange de gaussiennes et fait intervenir celui-ci au sein d'un GLRT.

C.1 Principe du GLRT-GM

L'idée de base qui se trouve derrière ce filtre GLRT-GM est la suivante : comme on modélise le fond par un mélange de gaussiennes, plutôt que de classifier les pixels et appliquer un GLRT faisant intervenir comme distribution du fond une unique gaussienne comme le fait le GMMF, ne serait-il pas plus favorable de considérer l'ensemble du mélange de gaussiennes pour modéliser la distribution du fond au sein du GLRT ? L'avantage de conserver l'ensemble du mélange de gaussiennes est que l'on peut ainsi prendre en compte le fait que la texture présente en un pixel peut ne pas correspondre à une unique texture modélisée par une des gaussiennes du mélange mais peut s'avérer être un mélange de deux textures présentes dans l'image et donc avoir un comportement statistique qui se modélise par un mélange de deux gaussiennes.

C.1.1 Expression théorique

Le filtre GLRT-GM consiste à supposer que le fond peut être modélisé en tout point de l'image par le même et unique mélange de gaussiennes. On reprend toutes les notations utilisées au cours du chapitre 4 et on considère comme dans le chapitre 4 que la moyenne du fond a déjà été éliminée par l'application préalable d'un algorithme de suppression de fond. La densité de probabilité de la composante de fond du vecteur \mathbf{r}_x , c.a.d celle de \mathbf{r}_x sous l'hypothèse H_0 est alors :

$$p(\mathbf{r}_x|H_0, \{\widehat{\pi}_k\}_{k \in \llbracket 1, K \rrbracket}, \{\widehat{\Sigma}_k\}_{k \in \llbracket 1, K \rrbracket}) = \sum_{k=1}^K \widehat{\pi}_k g(\mathbf{r}_x, \widehat{\Sigma}_k), \quad (\text{C.1})$$

où K est le nombre de gaussiennes constituant le mélange, $\widehat{\pi}_k$ est le poids estimé associé à la $k^{\text{ème}}$ distribution gaussienne et $g(\cdot, \widehat{\Sigma}_k)$ est la loi normale multidimensionnelle $\mathcal{N}(0, \widehat{\Sigma}_k)$. Les paramètres de ce mélange de gaussiennes, $\{\pi_k\}_{k \in \llbracket 1, K \rrbracket}$ et $\{\Sigma_k\}_{k \in \llbracket 1, K \rrbracket}$, sont estimés au

moyen d'un algorithme **EM** à partir de l'ensemble des vecteurs d'observation $\{\mathbf{r}_x\}_{x \in \mathcal{I}}$ extrait de l'image.

Considérant que la cible a un contraste positif, le **GLRT** s'exprime de la façon suivante :

$$\Lambda(\mathbf{r}_x) = \frac{\max_{\alpha \geq 0} p(\mathbf{r}_x | H_1, \alpha, \{\widehat{\pi}_k\}_{k \in \llbracket 1, K \rrbracket}, \{\widehat{\Sigma}_k\}_{k \in \llbracket 1, K \rrbracket})}{p(\mathbf{r}_x | H_0, \{\widehat{\pi}_k\}_{k \in \llbracket 1, K \rrbracket}, \{\widehat{\Sigma}_k\}_{k \in \llbracket 1, K \rrbracket})} = \frac{\max_{\alpha \geq 0} \sum_{k=1}^K \widehat{\pi}_k g(\mathbf{r}_x | H_1, \widehat{\Sigma}_k)}{\sum_{k=1}^K \widehat{\pi}_k g(\mathbf{r}_x | H_0, \widehat{\Sigma}_k)} \underset{H_1}{\overset{H_0}{\lesssim}} \mu \quad (\text{C.2})$$

Dans notre contexte où le test d'hypothèse se formalise selon l'équation (1.19), le **GLRT** défini ci-dessus est égal à :

$$\frac{\max_{\alpha \geq 0} \sum_{k=1}^K \widehat{\pi}_k g(\mathbf{r}_x - \alpha \mathbf{s}, \widehat{\Sigma}_k)}{\sum_{k=1}^K \widehat{\pi}_k g(\mathbf{r}_x, \widehat{\Sigma}_k)} \underset{H_1}{\overset{H_0}{\lesssim}} \mu \quad (\text{C.3})$$

C.1.2 Approximation

L'application du **GLRT** donné par l'équation (C.3) nécessite de maximiser le terme $\sum_{k=1}^K \widehat{\pi}_k g(\mathbf{r}_x - \alpha \mathbf{s}, \widehat{\Sigma}_k)$ sur $\alpha \geq 0$. Or on ne dispose pas d'expression analytique simple de ce maximum et son calcul nécessite donc de faire une recherche exhaustive très coûteuse en temps de calcul. On choisit donc d'approximer ce maximum sur $\alpha \geq 0$, par le maximum obtenu pour un α sélectionné parmi l'ensemble $\{\widehat{\alpha}_i\}_{i \in \llbracket 1, K \rrbracket}$, où $\widehat{\alpha}_i$ maximise la $i^{\text{ème}}$ distribution gaussienne $g(\mathbf{r}_x - \alpha \mathbf{s}, \widehat{\Sigma}_k)$ sur $\alpha \geq 0$. Le test **GLRT** approximé proposé s'écrit donc :

$$\frac{\max_{\alpha \in \{\widehat{\alpha}_i\}_{i \in \llbracket 1, K \rrbracket}} \sum_{k=1}^K \widehat{\pi}_k g(\mathbf{r}_x - \alpha \mathbf{s}, \widehat{\Sigma}_k)}{\sum_{k=1}^K \widehat{\pi}_k g(\mathbf{r}_x, \widehat{\Sigma}_k)} \underset{H_1}{\overset{H_0}{\lesssim}} \mu \quad \text{avec} \quad \widehat{\alpha}_i = \arg \max_{\alpha \geq 0} g(\mathbf{r}_x - \alpha \mathbf{s}, \widehat{\Sigma}_i). \quad (\text{C.4})$$

D'après les tests réalisés, comme le montre la figure C.1, nous avons une bonne approximation de $\widehat{\alpha} = \max_{\alpha \geq 0} \sum_{k=1}^K \widehat{\pi}_k g(\mathbf{r}_x - \alpha \mathbf{s}, \widehat{\Sigma}_k)$ par $\widehat{\alpha}^* = \max_{\alpha \in \{\widehat{\alpha}_i\}_{i \in \llbracket 1, K \rrbracket}} \sum_{k=1}^K \widehat{\pi}_k g(\mathbf{r}_x - \alpha \mathbf{s}, \widehat{\Sigma}_k)$.

Pour résumer le filtre **GLRT-GM** proposé consiste à estimer par **EM**, le modèle de mélange de gaussiennes permettant de modéliser le mieux possible le fond et à appliquer en chaque pixel le **GLRT** approximé défini par l'équation (C.4).

C.2 Résultats expérimentaux

Dans cette partie, on compare les performances du filtre **GLRT-GM** en termes de courbes **ROC** avec celles du filtre **GMMF** introduit au chapitre 2. Deux cas sont considérés : le cas où les données sont prétraitées par un filtre **BSMean** et celui où elles sont prétraitées par **DBM3D**. Les courbes **ROC** obtenues pour les deux mêmes bases de données que celles utilisées au chapitre 5, sont présentées en figure C.2.

On peut y voir que le plus souvent le filtre **GLRT-GM** reste moins performant que le filtre **GMMF**. De plus, en terme de coût de calcul, le filtre **GLRT-GM** qui prend en compte

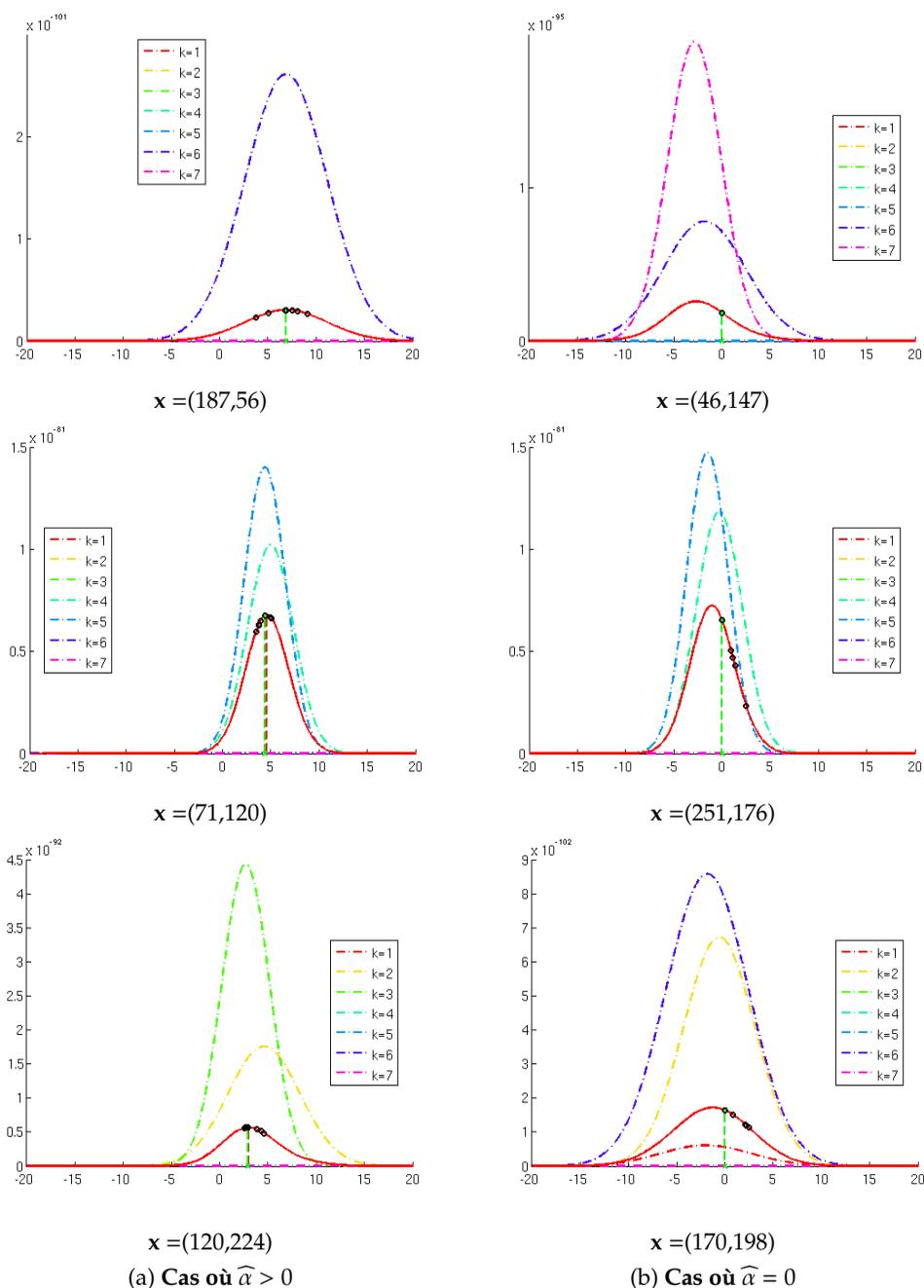
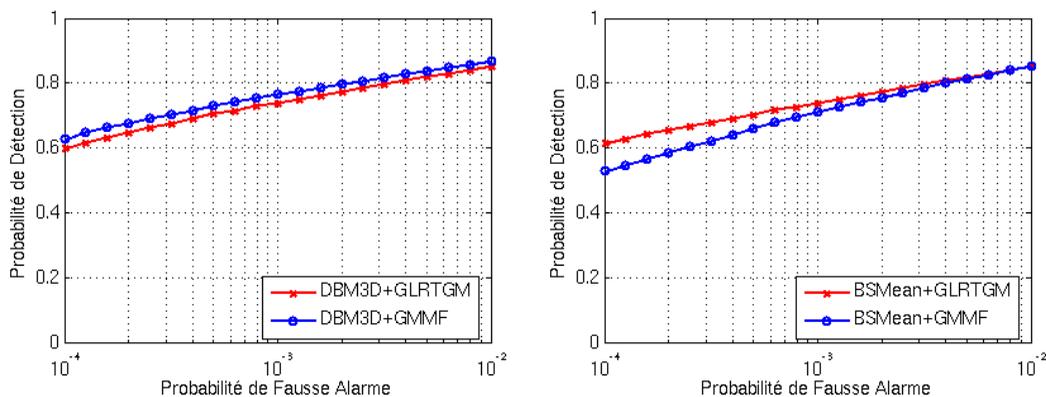
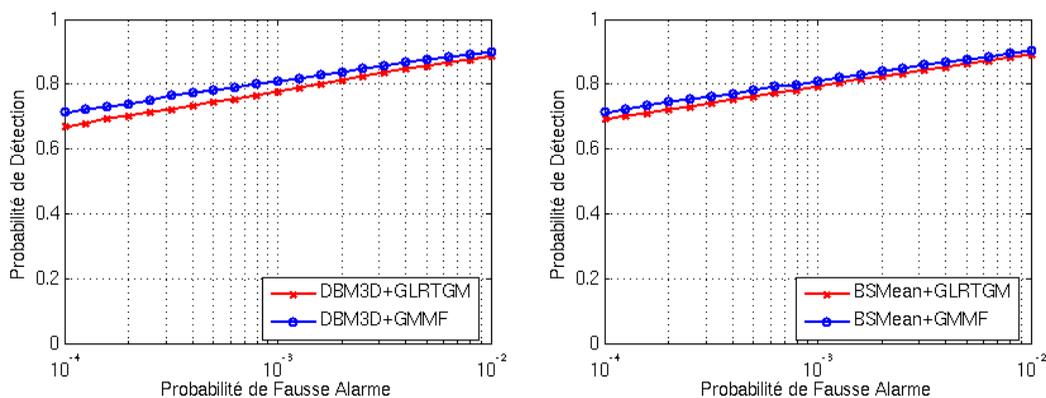


FIGURE C.1 – Illustration de l’approximation pertinente de $\hat{\alpha}$ par $\hat{\alpha}^*$. Pour le mélange de gaussiennes obtenu par EM sur l’image de la figure 2 prétraitée par un filtrage DBM3D, pour différente position x dans l’image, visualisation en fonction de α des $g(\mathbf{r}_x - \alpha \mathbf{s}, \hat{\Sigma}_k)$ (cf. légende du graphique) et de $\sum_{k=1}^K \hat{\pi}_k g(\mathbf{r}_x - \alpha \mathbf{s}, \hat{\Sigma}_k)$ (courbe rouge). Les $\hat{\alpha}_i$ sont indiqués par les ronds noirs, le $\hat{\alpha}^*$ retenu correspond au trait vertical vert et $\hat{\alpha}$ correspond au trait vertical rouge (le plus souvent confondu avec le vert). Des cas où $\hat{\alpha} > 0$ (colonne de gauche) et des cas où $\hat{\alpha} = 0$ (colonne de droite) sont présentés.

l'ensemble du mélange de gaussiennes au sein du GLRT est beaucoup plus coûteux que le filtre GMMF. Le GLRT-GM étant légèrement moins performant et plus coûteux en temps de calcul que le GMMF, chercher à prendre en compte l'ensemble du mélange de gaussienne au sein du GLRT s'avère non pertinent.



(a) Base de fonds terrestre



(b) Base de fonds satellitaire

FIGURE C.2 – Courbes ROC obtenues, pour chacune des bases de données, pour le GMMF et le GLRT-GM. Ces deux filtres sont appliqués à la suite d'un DBM3D pour les courbes de droite et à la suite d'un BSMean pour les courbes de gauche.

Bibliographie

- [Eum,] <https://eoportal.eumetsat.int/userMgmt/protected/dataCentre.faces>. Cité page 42
- [BM3,] <http://www.cs.tut.fi/foi/GCF-BM3D/>. Cité page 128
- [Acito *et al.*, 2005] ACITO, N., CORSINI, G., DIANI, M. et PENNUCCI, G. (2005). Experimental performance analysis of clutter removal techniques in IR images. In *IEEE International Conference on Image Processing (ICIP)*, volume 3, pages III – 561–4. Cité page 78
- [Acito *et al.*, 2011] ACITO, N., ROSSI, A., DIANI, M. et CORSINI, G. (2011). Optimal criterion to select the background estimation algorithm for detection of dim point targets in infrared surveillance systems. *Optical Engineering*, 50(10):107204. Cité pages 26, 28 et 51
- [Aharon, 2006] AHARON, M. (2006). K-SVD : An algorithm for designing overcomplete dictionaries for sparse representation. *IEEE Transactions on Signal Processing*, 54(11):4311–4322. Cité page 51
- [Anderson, 1984] ANDERSON, T. (1984). *An introduction to multivariate statistical analysis, second edition*. Wiley New York. Cité page 80
- [Ashton, 1998] ASHTON, E. (1998). Detection of subpixel anomalies in multispectral infrared imagery using an adaptive bayesian classifier. *IEEE Transactions on Geoscience and Remote Sensing*, 36(2):506–517. Cité page 77
- [Askar *et al.*, 2002] ASKAR, H., LI, X. et LI, Z. (2002). Performance analysis of dim moving point target detection algorithms. In *IEEE International Conference on Communications, Circuits and Systems and West Sino Expositions*, volume 1, pages 605 – 609 vol.1. Cité page 1
- [Aubert *et al.*, 2010] AUBERT, G., BAUDOUR, A., BLANC-FERAUD, L., GUILLOT, L. et LE GUILLOUX, Y. (2010). Detection and tracking of threats in aerial infrared images by a minimal path approach. In *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, pages 1658–1661. Cité page 14
- [Bae et Sohng, 2010] BAE, T. et SOHNG, K. (2010). Small target detection using bilateral filter based on edge component. *Journal of Infrared, Millimeter and Terahertz Waves*, 31(6):735–743. Cité page 28
- [Bae, 2011] BAE, T.-W. (2011). Small target detection using bilateral filter and temporal cross product in infrared images. *Infrared Physics & Technology*, 54(5):403 – 411. Cité pages 28 et 51
- [Bai *et al.*, 2009] BAI, X., ZHOU, F. et JIN, T. (2009). Enhancement of dim small target through modified top-hat transformation under the condition of heavy clutter. *Signal Processing*, 90:1643–1654. Cité page 27
- [Bar-Tal et Rotman, 1996] BAR-TAL, M. et ROTMAN, S. R. (1996). Performance measurement in point source target detection. *Infrared Physics & Technology*, 37(2):231 – 238. Cité page 36
- [Barnett, 1989] BARNETT, J. (1989). Statistical analysis of median subtraction filtering with application to point target detection in infrared backgrounds. In *Proceedings of SPIE*, volume 1050, pages 10–18. Cité pages 26 et 27
- [Barnett *et al.*, 1993] BARNETT, J. T., BILLARD, B. D. et LEE, C. (1993). Nonlinear morphological processors for point-target detection versus an adaptive linear spatial filter : a performance comparison. In *Proceedings of SPIE*, volume 1954, pages 12–24. Cité pages 26 et 27
- [Barniv, 1985] BARNIV, Y. (1985). Dynamic programming solution for detecting dim moving targets. *IEEE Transactions on Aerospace and Electronic Systems*, (1):144–156. Cité page 14
- [Barniv et Kella, 1987] BARNIV, Y. et KELLA, O. (1987). Dynamic programming solution for detecting dim moving targets part II : analysis. *IEEE Transactions on Aerospace and Electronic Systems*, (6):776–788. Cité page 14
- [Barret, 2009] BARRET, M. (2009). *Traitement statistique du signal. Estimation, filtrage de Wiener, méthodes récursives, détection*. Ellipses. Cité page 11
-

- [Biernacki et al., 2000] BIERNACKI, C., CELEUX, G. et GOVAERT, G. (2000). Assessing a mixture model for clustering with the integrated completed likelihood. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(7):719–725. Cité pages 91 et 93
- [Blostein et Huang, 1991] BLOSTEIN, S. et HUANG, T. (1991). Detecting small, moving objects in image sequences using sequential hypothesis testing. *IEEE Transactions on Signal Processing*, 39(7):1611–1629. Cité pages 2, 13 et 14
- [Blostein et Richardson, 1994] BLOSTEIN, S. et RICHARDSON, H. (1994). A sequential detection approach to target tracking. *IEEE Transactions on Aerospace and Electronic Systems*, 30(1):197–212. Cité page 1
- [Boccignone et al., 2002] BOCCIGNONE, G., CHIANESE, A. et PICARIELLO, A. (2002). Small target detection using wavelets. In *IEEE International Conference on Pattern Recognition*, volume 2, pages 1776–1778. IEEE. Cité page 28
- [Boers et Driessen, 2003] BOERS, Y. et DRIESSEN, H. (2003). A particle-filter-based detection scheme. *IEEE Signal Processing Letters*, 10(10):300–302. Cité page 15
- [Boulanger, 2007] BOULANGER, J. (2007). *Estimation non-paramétrique et contributions à l'analyse de séquences d'images : Modélisation, simulation et estimation du trafic-intra cellulaire dans des séquences de vidéo-microscopie*. Thèse de doctorat, Université de Rennes I. Cité page 28
- [Boulanger et al., 2006] BOULANGER, J., KERVRANN, C. et BOUTHEMY, P. (2006). Estimation of dynamic background for fluorescence video-microscopy. In *IEEE International Conference on Image Processing (ICIP)*, pages 2509–2512. Cité pages 1 et 28
- [Brox et al., 2008] BROX, T., KLEINSCHMIDT, O. et CREMERS, D. (2008). Efficient nonlocal means for denoising of textural patterns. *IEEE Transactions on Image Processing*, 17(7):1083–1092. Cité page 151
- [Bruno et Moura, 2000] BRUNO, M. et MOURA, J. (2000). Optimal multiframe detection and tracking in digital image sequences. In *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, volume 5, pages 3192–3195 vol.5. Cité pages 2 et 15
- [Buades et al., 2005a] BUADES, A., COLL, B. et MOREL, J. (2005a). Denoising image sequences does not require motion estimation. In *IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS)*, pages 70–74. Cité pages 48, 53 et 64
- [Buades et al., 2005b] BUADES, A., COLL, B. et MOREL, J. (2005b). A non-local algorithm for image denoising. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 60–65. Cité pages 4, 28, 47, 50, 51 et 73
- [Buades et al., 2008] BUADES, A., COLL, B. et MOREL, J. (2008). Nonlocal image and movie denoising. *International Journal of Computer Vision*, 76(2):123–139. Cité page 53
- [Buades et al., 2012] BUADES, A., LEBRUN, M. et MOREL, J.-M. (2012). Implementation of the « non-local Bayes » image denoising algorithm. *Image Processing On Line*. preprint. Cité pages 48, 50, 51 et 143
- [Cafer et al., 2008] CAEFER, C., SILVERMAN, J., ORTHAL, O., ANTONELLI, D., SHARONI, Y. et ROTMAN, S. (2008). Improved covariance matrices for point target detection in hyperspectral data. *Optical Engineering*, 47:076402. Cité pages 18, 76 et 78
- [Cafer et al., 2007] CAEFER, C., STEFANO, M., NIELSEN, E., RIZZUTO, A., RAVIV, O. et ROTMAN, S. (2007). Analysis of false alarm distributions in the development and evaluation of hyperspectral point target detection algorithms. *Optical Engineering*, 46:076402. Cité pages 34, 36, 37 et 119
- [Carlson et al., 1994] CARLSON, B., EVANS, E. et WILSON, S. (1994). Search radar detection and track with the Hough transform. I. system concept. *IEEE Transactions on Aerospace and Electronic Systems*, 30(1):102–108. Cité page 14
- [Carnie et al., 2006] CARNIE, R., WALKER, R. et CORKE, P. (2006). Image processing algorithms for UAV sense and avoid. In *IEEE International Conference on Robotics and Automation*, pages 2848–2853. Cité pages 2 et 27
- [Celeux et Govaert, 1992] CELEUX, G. et GOVAERT, G. (1992). A classification EM algorithm for clustering and two stochastic versions. *Computational Statistics & Data Analysis*, 14(3):315–332. Cité pages 79, 91 et 92
- [Chan et al., 1998] CHAN, C. L., ATTILI, J. B. et MELENDEZ, K. A. (1998). Image segmentation approach for improving target detection in a 3D signal processor. In *Proceedings of SPIE*, volume 3373, pages 87–94. Cité pages 76 et 78
- [Chapple et al., 2001] CHAPPLE, P., BERTILONE, D., CAPRARI, R. et NEWSAM, G. (2001). Stochastic model-based processing for detection of small targets in non-gaussian natural imagery. *IEEE Transactions on Image Processing*, 10(4):554–564. Cité page 17

- [Chen et Reed, 1987] CHEN, J. et REED, I. (1987). A detection algorithm for optical targets in clutter. *IEEE Transactions on Aerospace and Electronic Systems*, 23(1):46–59. Cité page 17
- [Chen et al., 2007] CHEN, Z., WANG, G., LIU, J. et LIU, C. (2007). Small target detection algorithm based on average absolute difference maximum and background forecast. *International Journal of Infrared and Millimeter Waves*, 28(1):87–97. Cité page 26
- [Dabov et al., 2007a] DABOV, K., FOI, A. et EGAZARIAN, K. (2007a). Video denoising by sparse 3D transform-domain collaborative filtering. In *15th European Signal Processing Conference*, volume 1, page 7. Citeseer. Cité pages 48, 55 et 72
- [Dabov et al., 2006] DABOV, K., FOI, A., KATKOVNIK, V. et EGAZARIAN, K. (2006). Image denoising with block-matching and 3D filtering. In *Proceedings of SPIE*, volume 6064, pages 354–365. Citeseer. Cité pages 51 et 69
- [Dabov et al., 2007b] DABOV, K., FOI, A., KATKOVNIK, V. et EGAZARIAN, K. (2007b). Image denoising by sparse 3-D transform-domain collaborative filtering. *IEEE Transactions on Image Processing*, 16(8):2080–2095. Cité pages 4, 28, 47, 49, 50, 51, 53, 54, 69, 72, 73 et 133
- [Dabov et al., 2009] DABOV, K., FOI, A., KATKOVNIK, V. et EGAZARIAN, K. (2009). BM3D image denoising with shape-adaptive principal component analysis. In *Proc. workshop on signal processing with adaptive sparse structured representations (SPARS)*, volume 49. Citeseer. Cité pages 50 et 51
- [Davey et al., 2008] DAVEY, S., RUTTEN, M. et CHEUNG, B. (2008). A comparison of detection performance for several track-before-detect algorithms. *EURASIP Journal on Advances in Signal Processing*, 2008:41. Cité pages 2 et 13
- [Deledalle et al., 2012] DELEDALLE, C., DUVAL, V. et SALMON, J. (2012). Non-local methods with shape-adaptive patches (NLM-SAP). *Journal of Mathematical Imaging and Vision*, 43(2):103–120. Cité page 50
- [Deledalle et al., 2011] DELEDALLE, C.-A., SALMON, J. et DALALYAN, A. S. (2011). Image denoising with patch based PCA : local versus global. In *Proceedings of the British Machine Vision Conference (BMVC)*. Cité page 51
- [Dempster et al., 1977] DEMPSTER, A., LAIRD, N. et RUBIN, D. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 1–38. Cité page 79
- [Deshpande et al., 1999] DESHPANDE, S. D., ER, M. H., VENKATESWARLU, R. et CHAN, P. (1999). Max-mean and max-median filters for detection of small targets. In *Proceedings of SPIE*, volume 3809, pages 74–83. Cité pages 27, 77 et 115
- [Elad et Aharon, 2006] ELAD, M. et AHARON, M. (2006). Image denoising via sparse and redundant representations over learned dictionaries. *IEEE Transactions on Image Processing*, 15(12):3736–3745. Cité page 51
- [Genin et al., 2012a] GENIN, L., CHAMPAGNAT, F. et LE BESNERAIS, G. (2012a). Background first and second order modeling for point target detection. *Applied Optics*, 51(31):7701–7713. Cité pages 4, 23, 114, 124 et 139
- [Genin et al., 2012b] GENIN, L., CHAMPAGNAT, F. et LE BESNERAIS, G. (2012b). Single frame IR point target detection based on a gaussian mixture model classification. In *Proceedings of SPIE*, volume 8541. Cité page 51
- [Genin et al., 2011] GENIN, L., CHAMPAGNAT, F., LE BESNERAIS, G. et CORET, L. (2011). Point object detection using a NL-means type filter. In *IEEE International Conference on Image Processing (ICIP)*, pages 3533 – 3536. Cité pages 28, 51, 114 et 139
- [Goudail, 2011] GOUDAIL, F. (2011). Introduction au traitement des images. Cours de l’Institut d’Optique Graduate School. Cité page 8
- [Goudou, 2007] GOUDOU, J. (2007). *Apport de la dimension temporelle aux traitements de veille infrarouge marine*. Thèse de doctorat, Telecom Paris. Cité pages 28, 34, 36, 51, 58 et 67
- [Govaerts, 2006] GOVAERTS, Y. (2006). Presentation slides : EUMETSAT mission status, fire products / fire requirements. 2nd Workshop on Geostationary Fire Monitoring and Applications, [http://gofc-fire.umd.edu/products/pdfs/Events/Geo_2006/Govaerts_GOFC\(1\).pdf](http://gofc-fire.umd.edu/products/pdfs/Events/Geo_2006/Govaerts_GOFC(1).pdf). Cité page 45
- [Grégoire, 2011] GRÉGOIRE, G. (2011). Planches de présentation : Introduction aux processus ponctuels spatiaux. Journée satellite STID, 43èmes Journées SFdS. Cité page 36
- [Gregoris et al., 1994] GREGORIS, D. J., YU, S. K., TRITCHEW, S. et SEVIGNY, L. (1994). Wavelet transform-based filtering for the enhancement of dim targets in FLIR images. In Szu, H. H., éditeur : *Proceedings of SPIE*, volume 2242, pages 573–583. Cité page 28

- [Grosjean et Moisan, 2009] GROSJEAN, B. et MOISAN, L. (2009). A-contrario detectability of spots in textured backgrounds. *Journal of Mathematical Imaging and Vision*, 33:313–337. Cité page 1
- [Gu et al., 2010] GU, Y., WANG, C., LIU, B. et ZHANG, Y. (2010). A kernel-based nonparametric regression method for clutter removal in infrared small-target detection applications. *IEEE Geoscience and Remote Sensing Society*, 7:469–473. Cité pages 32 et 36
- [Herranz et al., 2011] HERRANZ, D., ARGUESO, F., SALERNO, E., KURUOGLU, E. et KAYABOL, K. (2011). Bayesian MAP detection of extragalactic point sources in microwave astronomical images. In *IEEE International Conference on Image Processing (ICIP)*, pages 1293–1296. Cité page 1
- [Hilliard, 2000] HILLIARD, C. I. (2000). Selection of a clutter rejection algorithm for real-time target detection from an airborne platform. In *Proceedings of SPIE, Signal and Data Processing of Small Targets*, volume 4048, pages 74–84. Cité page 32
- [Johnston et Krishnamurthy, 2002] JOHNSTON, L. et KRISHNAMURTHY, V. (2002). Performance analysis of a dynamic programming track before detect algorithm. *IEEE Transactions on Aerospace and Electronic Systems*, 38(1):228–242. Cité page 2
- [Kelly, 1986] KELLY, E. (1986). An adaptive detection algorithm. *IEEE Transactions on Aerospace and Electronic Systems (AES)*, 22(2):115–127. Cité page 22
- [Kervrann et Boulanger, 2006] KERVRANN, C. et BOULANGER, J. (2006). Optimal spatial adaptation for patch-based image denoising. *IEEE Transactions on Image Processing*, 15(10):2866–2878. Cité page 151
- [Kim, 2011] KIM, S. (2011). Min-local-log filter for detecting small targets in cluttered background. *Electronics Letters*, 47(2):105–106. Cité page 27
- [Kim et al., 2009] KIM, S., YANG, Y. et LEE, J. (2009). Robust detection of horizontal small targets using synergistic spatial filtering. *Electronics letters*, 45(12):620–621. Cité page 26
- [Lebrun et al., 2012] LEBRUN, M., COLL, B., COLOM, M., BUADES, A. et J.M., M. (2012). L'état de l'art actuel en débruitage. *Image Processing On Line*. preprint. Cité page 48
- [MacQueen, 1967] MACQUEEN, J. (1967). Some methods for classification and analysis of multivariate observations. In *Proceedings of the fifth Berkeley symposium on mathematical statistics and probability*, pages 281–297. California, USA. Cité pages 86 et 90
- [Manolakis et al., 2003] MANOLAKIS, D., MARDEN, D. et SHAW, G. (2003). Hyperspectral image processing for automatic target detection applications. *Lincoln Laboratory Journal*, 14(1):79–116. Cité page 18
- [Margalit et al., 1985] MARGALIT, A., REED, I. et GAGLIARDI, R. (1985). Adaptive optical target detection using correlated images. *IEEE Transactions on Aerospace and Electronic Systems*, 21(3):394–405. Cité pages 16, 17, 32 et 76
- [Matteoli et al., 2010] MATTEOLI, S., DIANI, M. et CORSINI, G. (2010). A tutorial overview of anomaly detection in hyperspectral images. *IEEE Aerospace and Electronic Systems Magazine*, 25(7):5–28. Cité pages 12, 18, 77 et 86
- [Olivo-Marin, 2002] OLIVO-MARIN, J.-C. (2002). Extraction of spots in biological images using multiscale products. *Pattern Recognition*, 35(9):1989–1996. Cité pages 1 et 28
- [Ovarlez et al., 2011] OVARLEZ, J., PANG, S., PASCAL, F., ACHARD, V. et NG, T. (2011). Robust detection using the SIRV background modelling for hyperspectral imaging. In *IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, pages 4316–4319. Cité page 17
- [Pascal, 2006] PASCAL, F. (2006). *Détection et estimation en environnement non Gaussien*. Thèse de doctorat, Univ. Paris X, Nanterre, France. Cité page 17
- [Pascal et al., 2005] PASCAL, F., FORSTER, R., OVARLEZ, J.-P. et ARZABAL, P. (2005). Theoretical analysis of an improved covariance matrix estimator in non-gaussian noise [radar detection applications]. In *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, volume 4, pages iv/69–iv/72. Cité page 17
- [Pei et al., 2006] PEI, J., LU, Z. et XIE, W. (2006). A method for ir point target detection based on spatial-temporal bilateral filter. In *IEEE International Conference on Pattern Recognition (ICPR)*, volume 3, pages 846–849. Cité page 51
- [Protter et al., 2009] PROTTER, M., ELAD, M., TAKEDA, H. et MILANFAR, P. (2009). Generalizing the non-local-means to super-resolution reconstruction. *IEEE Transactions on Image Processing*, 18(1):36–51. Cité page 51
- [Rabaste et al., 2012] RABASTE, O., RICHE, C. et LÉPOUTRE, A. (2012). Long-time coherent integration for low SNR target via particle filter in track-before-detect. In *IEEE International Conference on Information Fusion (FUSION)*, pages 127–134. Cité page 15

- [Reed *et al.*, 1983] REED, I., GAGLIARDI, R. et SHAO, H. (1983). Application of three-dimensional filtering to moving target detection. *IEEE Transactions on Aerospace and Electronic Systems (AES)*, 19(6):898–905. Cité pages 2, 13 et 19
- [Reed *et al.*, 1988] REED, I., GAGLIARDI, R. et STOTTS, L. (1988). Optical moving target detection with 3-D matched filtering. *IEEE Transactions on Aerospace and Electronic Systems*, 24(4):327–336. Cité page 13
- [Reed et Yu, 1990] REED, I. et YU, X. (1990). Adaptive multiple-band CFAR detection of an optical pattern with unknown spectral distribution. *IEEE Transactions on Acoustics, Speech and Signal Processing*, 38(10):1760–1770. Cité page 18
- [Rivest et Fortin, 1996] RIVEST, J.-F. et FORTIN, R. (1996). Detection of dim targets in digital infrared imagery by morphological image processing. *Optical Engineering*, 35(7):1886–1893. Cité page 27
- [Robey *et al.*, 1992] ROBEY, F., FUHRMANN, D., KELLY, E. et NITZBERG, R. (1992). A CFAR adaptive matched filter detector. *IEEE Transactions on Aerospace and Electronic Systems*, 28(1):208–216. Cité pages 21 et 22
- [Robinson, 1959] ROBINSON, D. (1959). Methods of background description and their utility. *Proceedings of the IRE*, 47(9):1554–1561. Cité page 27
- [Salmon *et al.*, 2012] SALMON, J., DELEDALLE, C.-A., WILLETT, R. et HARMANY, Z. T. (2012). Poisson noise reduction with Non-Local PCA. In *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*. Cité page 51
- [Salmon et Strozecski, 2012] SALMON, J. et STROZECKI, Y. (2012). Patch reprojections for non-local methods. *Signal Processing*, 92(2):477–489. Cité pages 48, 49, 50, 51 et 151
- [Samson, 2002] SAMSON, V. (2002). *Approche régularisée pour la détection d'objets ponctuels dans une séquence d'images*. Thèse de doctorat, Université de Paris-Sud, Orsay. Cité pages 1, 19, 27, 32, 34, 36, 37, 43, 99, 100 et 115
- [Samson *et al.*, 2004] SAMSON, V., CHAMPAGNAT, F. et GIOVANNELLI, J. (2004). Point target detection and subpixel position estimation in optical imagery. *Applied Optics*, 43:257–263. Cité pages 11, 19, 44, 63, 99 et 100
- [Soni *et al.*, 1993] SONI, T., ZEIDLER, J. et KU, W. (1993). Performance evaluation of 2-d adaptive prediction filters for detection of small objects in image data. *IEEE Transactions on Image Processing*, 2(3):327–340. Cité page 26
- [Stein *et al.*, 2002] STEIN, D., BEAVEN, S., HOFF, L., WINTER, E., SCHAUM, A. et STOCKER, A. (2002). Anomaly detection from hyperspectral imagery. *IEEE Signal Processing Magazine*, 19(1):58–69. Cité pages 78 et 86
- [Sui *et al.*, 2011] SUI, X., CHEN, Q. et BAI, L. (2011). Detection algorithm of targets for infrared search system based on area infrared focal plane array under complicated background. *Optik - International Journal for Light and Electron Optics*, 123(3):235–239. Cité page 143
- [Sun *et al.*, 2009] SUN, Y.-Q., FENG, X.-Q., LI, L., TIAN, J.-W. et LIU, J. (2009). Fractal-based infrared target detection. In *International Conference on Machine Learning and Cybernetics*, volume 4, pages 2443–2447. Cité page 29
- [Takken *et al.*, 1979] TAKKEN, E., FRIEDMAN, D., MILTON, A. et NITZBERG, R. (1979). Least-mean-square spatial filter for IR sensors. *Applied optics*, 18(24):4210–4222. Cité page 26
- [Tang et Wang, 2009] TANG, Z. et WANG, X. (2009). An Efficient Algorithm for Infrared Small Target Detection. In *IEEE International Conference on Information and Computing Science*, pages 51–54. Cité page 29
- [Tasdizen, 2009] TASDIZEN, T. (2009). Principal neighborhood dictionaries for nonlocal means image denoising. *IEEE Transactions on Image Processing*, 18(12):2649–2660. Cité page 151
- [Tom *et al.*, 1993] TOM, V. T., PELI, T., LEUNG, M. et BONDARYK, J. E. (1993). Morphology-based algorithm for point target detection in infrared backgrounds. In *Proceedings of SPIE, Signal and Data Processing of Small Targets*, volume 1954, pages 2–11. Cité page 27
- [Tomasi et Manduchi, 1998] TOMASI, C. et MANDUCHI, R. (1998). Bilateral filtering for gray and color images. In *IEEE International Conference on Computer Vision*, pages 839–846. Cité page 49
- [Tonissen et Bar-Shalom, 1998] TONISSEN, S. et BAR-SHALOM, Y. (1998). Maximum likelihood track-before-detect with fluctuating target amplitude. *IEEE Transactions on Aerospace and Electronic Systems*, 34(3):796–809. Cité pages 13, 14 et 15
- [Tonissen et Evans, 1996] TONISSEN, S. et EVANS, R. (1996). Performance of dynamic programming techniques for track-before-detect. *IEEE Transactions on Aerospace and Electronic Systems*, 32(4):1440–1451. Cité pages 2, 13 et 14
- [Van Trees, 1968] VAN TREES, H. (1968). *Detection, estimation, and modulation theory : Detection, estimation, and linear modulation theory*. Wiley. Cité pages 10 et 11

- [Vasquez, 2011] VASQUEZ, E. (2011). *Techniques statistiques de détection de cibles dans des images infrarouges inhomogènes en milieu maritime*. Thèse de doctorat, Université d'Aix-Marseille III. Cité page 28
- [Vasquez et al., 2010] VASQUEZ, E., GALLAND, F., DELYON, G. et RÉFRÉGIER, P. (2010). Mixed segmentation-detection-based technique for point target detection in nonhomogeneous sky. *Applied optics*, 49(9):1518–1527. Cité pages 1, 3, 12 et 27
- [Wong et Orchard, 2008] WONG, A. et ORCHARD, J. (2008). A nonlocal-means approach to exemplar-based inpainting. In *IEEE International Conference on Image Processing (ICIP)*, pages 2600–2603. Cité page 51
- [Yang et al., 2004] YANG, L., YANG, J. et YANG, K. (2004). Adaptive detection for infrared small target under sea-sky complex background. *Electronics Letters*, 40(17):1083–1085. Cité page 32
- [Yaroslavsky, 1985] YAROSLAVSKY, L. (1985). Digital picture processing. *Springer Series in Information Sciences*, Springer-Verlag, 9. Cité page 49
- [Yu et al., 2012] YU, G., SAPIRO, G. et MALLAT, S. (2012). Solving inverse problems with piecewise linear estimators : From gaussian mixture models to structured sparsity. *IEEE Transactions on Image Processing*, 21(5):2481–2499. Cité page 90
- [Zhang et al., 2005] ZHANG, F., LI, C. et SHI, L. (2005). Detecting and tracking dim moving point target in IR image sequence. *Infrared Physics and Technology*, 46(4):323–328. Cité pages 2 et 27
- [Zhu et al., 2000] ZHU, Z., LI, Z., LIANG, H., SONG, B. et PAN, A. (2000). Gray-scale morphological filter for small-target detection. volume 4130, pages 28–34. Cité page 27
-

Publications

Article de journal

- Genin, L. ; Champagnat, F. ; Le Besnerais, G. ; "Background first- and second-order modeling for point target detection," *Applied Optics*, Vol. 51, Issue 31, pp. 7701-7713, 2012.

Actes de conférence

- Genin, L. ; Champagnat, F. ; Le Besnerais, G. ; Coret, L. ; "A Non-Local filter for background suppression to detect point objects," *3rd European Workshop on Visual Information Processing (EUVIP)*, Student Poster Session, Paris, France, July 2011
 - Genin, L. ; Champagnat, F. ; Le Besnerais, G. ; Coret, L. ; "Point object detection using a NL-means type filter," *18th IEEE International Conference on Image Processing (ICIP)*, pp.3533 - 3536, Brussels, Belgium, Sept. 2011
 - Genin, L. ; Champagnat, F. ; Le Besnerais, G. ; "Single frame IR point target detection based on a Gaussian mixture model classification," *Proc. SPIE 8541, Electro-Optical and Infrared Systems : Technology and Applications IX*, 854111, Edinburgh, United Kingdom, Sept. 2012
-

Détection d'objets de petite taille sur des séquences aériennes ou satellitaires

Les travaux réalisés dans le cadre de cette thèse ont pour objectif d'améliorer la détection d'objets ponctuels en imagerie optique. Ils se concentrent sur le contexte difficile de la détection d'objets ponctuels de faible vitesse sur des fonds structurés inhomogènes présentant des zones de forts gradients de niveaux de gris. Dans ce contexte, nous proposons des méthodes de détection mono-image cherchant à tirer parti au mieux de la corrélation spatiale du fond. Des extensions spatio-temporelles des méthodes proposées sont étudiées dans un second temps. Basé sur un formalisme du test du rapport de vraisemblance généralisé (TRVG), le problème de détection se ramène à une approche en deux étapes qui consiste à dissocier l'estimation des statistiques locales du fond au premier et second ordre (c.a.d. moyenne et covariance). Pour améliorer les performances des méthodes de détection par modélisation du fond au premier ordre, nous utilisons des approches dites « à patchs » issues du domaine du débruitage. Malgré l'amélioration des performances de détection apportée par ces méthodes à patchs, il s'avère que des textures liées à la structure du fond sont encore visibles après l'étape de suppression de fond. Nous cherchons à améliorer les performances de détection par une modélisation du fond au second ordre. Nous nous intéressons à des méthodes de filtrage adapté par zone basées sur une modélisation par mélange de gaussiennes. Une analyse détaillée des performances des filtres développés est réalisée à partir de fonds nuageux réels sur lesquels des cibles ponctuelles sont incrustées.

Mots-clés : Détection, Cibles ponctuelles, Filtrage adapté, Méthodes à patchs, Classification non-supervisée, Mélange de gaussiennes.

Small object detection on aerial or satellite images sequences

The objective of this thesis is to improve the detection of point objects in optical imaging. They focus on the challenging detection of low velocity point objects on inhomogeneous background including areas of strong gradients of gray levels. In this context, we propose single-frame detection methods trying to take advantage at best of the spatial background correlation. Spatio-temporal extensions of the proposed methods are studied in a second time. Based on a formalism of the generalized likelihood ratio test (GLRT), the problem of detection boils down to a two-step process which consists in separating the first and second order estimation of the local background (i.e. mean and covariance). To improve the performances of the detection methods by first order background modelling, we adapt patch-based denoising method to detection. Despite the improvement of detection performance brought by these patch-based methods, it appears that textures associated with background structures are still visible after the background suppression step. We seek to improve the detection performance by second order modeling. We are interested in matched filter adapted by area based on a Gaussian mixture model. A detailed performance analysis of the developed filters is made from real cloudy background on which point targets are embedded.

Keywords : Detection, Point target, Matched filter, Patch-based methods, Non-supervised classification, Gaussian Mixture Model.

Spécialité : Signaux et Images