

THÈSE

présentée en vue de
l'obtention du titre de

DOCTEUR

de

**L'ÉCOLE NATIONALE SUPÉRIEURE
DE L'AÉRONAUTIQUE ET DE L'ESPACE**

SPÉCIALITÉ : Traitement d'images et télédétection

par

Florence LAPORTERIE

Représentations hiérarchiques d'images avec des pyramides morphologiques.

Application à l'analyse et à la fusion spatio-temporelle de données en observation de la Terre.

Soutenue le 28 mai 2002 devant la Commission d'Examen :

MM.	M. SAMUELIDES	Président
	A. AYACHE	
	G. FLOUZAT	
	M. JOURLIN	Rapporteur
Mmes	M.-J. LEFEVRE	
	C. MERING	Rapporteur

Thèse préparée au sein du Centre d'études spatiales et de la Biosphère (CESBIO).

N° d'ordre:

THÈSE

présentée

devant l'Ecole Nationale Supérieure de
l'Aéronautique et de l'Espace

pour obtenir le grade de :

DOCTEUR DE L'ECOLE NATIONALE SUPÉRIEURE DE L'AÉRONAUTIQUE
ET DE L'ESPACE
Mention TRAITEMENT D'IMAGES ET TÉLÉDÉTECTION

par

Florence LAPORTERIE

Équipe d'accueil : CESBIO / ENSAE
École Doctorale : Informatique et Télécommunications

Titre de la thèse :

**Représentations hiérarchiques d'images avec des pyramides
morphologiques. Application à l'analyse et à la fusion
spatio-temporelle de données en observation de la Terre.**

À soutenir le 28 Mai 2002 devant la commission d'examen

M. :	Manuel	SAMUELIDES	Président
MM. :	Michel	JOURLIN	Rapporteurs
	Catherine	MERING	
	Alain	AYACHE	
MM. :	Marie-José	LEFEVRE	Examineurs
	Guy	FLOUZAT	

Table des matières

Introduction	5
Nomenclature des symboles utilisés	9
I CONTEXTE	11
1 Perception multi-niveau, télédétection et morphologie mathématique	13
1.1 Observation multi-échelle et surfaces terrestres	13
1.2 Perception et télédétection multi-résolution	16
1.2.1 Intérêt et complémentarité des différentes observations	17
1.2.2 Exemple de l'instrument HRVIR sur SPOT 4	19
1.2.3 Exemple de l'instrument VEGETATION sur SPOT 4 et 5	19
1.2.4 Conclusion	19
1.3 Perception et morphologie mathématique	20
1.3.1 Le concept	20
1.3.2 Les filtres morphologiques	23
II LA PYRAMIDE MORPHOLOGIQUE	27
2 Etat de l'art : les algorithmes pyramidaux	29
2.1 Définition	29
2.2 Schéma général d'un algorithme pyramidal	30
2.3 Les principaux algorithmes pyramidaux	31
2.4 Les notions de pyramides morphologiques	33
2.4.1 Pyramides par itération d'un filtrage et d'un échantillonnage	34
2.4.2 Pyramides par itération d'opérations morphologiques	34
2.4.3 Pyramides par méthodes multi-résolution	35
3 La pyramide morphologique	37
3.1 La décomposition	37
3.1.1 Principe	37
3.1.2 Exemple de décomposition morphologique	43
3.2 La recomposition	44

3.2.1	Principe	44
3.2.2	Exemple de recombinaison morphologique	46
3.3	Comparaison entre filtrage linéaire et non linéaire	47
3.3.1	Caractéristiques comparées des filtres linéaire et non linéaire	47
3.3.2	Spécificité de la morphologie mathématique	47
3.3.3	Comparaison entre filtres passe-bas	48
3.4	Synthèse : les familles de pyramides	52
3.5	Propriétés de la pyramide morphologique	53
3.5.1	Filtrage et dualité dans la pyramide	53
3.5.2	Propriétés de la reconstruction	57
3.5.3	Navigation dans les différents niveaux de la pyramide	60
3.6	Conclusion	61

III ANALYSE MULTI-RESOLUTION DES SURFACES TERRESTRES 63

4	Analyse ...	65
4.1	Introduction	65
4.2	Choix des paramètres	66
4.3	Influence des paramètres	67
4.3.1	Evolution des images de détails en fonction de la résolution	67
4.3.2	Influence de la taille de l'élément structurant	69
4.3.3	Influence du pas d'échantillonnage	69
4.3.4	Influence de l'algorithme d'échantillonnage	71
4.4	Extraction d'objets en fonction de leur taille	71
4.5	Extraction et reconstruction d'éléments d'images	75
4.6	Analyse et pré-segmentation d'images	77
4.6.1	Principe	77
4.6.2	Exemples	79
4.7	Analyse et pré-compression de données	83
4.7.1	Principe	83
4.7.2	Exemple	83
4.7.3	Reconstruction « à la carte »	84
4.8	Perspective pour la compression	86
4.9	Conclusion	90

IV FUSION DE DONNEES MULTI-SOURCES 93

5	État de l'art sur la fusion de données multi-résolution	95
5.1	Définition de la fusion de données multi-sources	95
5.1.1	Introduction	95
5.1.2	Les différents types de données en télédétection	97

5.1.3	Les différents types de données en imagerie médicale	101
5.2	Revue des méthodes de fusion multirésolution	101
5.2.1	Méthodes Statistiques	102
5.2.2	Méthodes colorimétriques	104
5.2.3	Méthodes issues du traitement du signal	105
5.3	Synthèse des processus de fusion d'images	109
5.4	Conclusion	112
6	Fusion de données ...	113
6.1	Principe	113
6.2	Mise en oeuvre	114
6.2.1	Calcul théorique de l'image fusionnée	114
6.2.2	Codage des images sur 8 bits	115
6.3	Choix des paramètres de la pyramide	116
6.4	Evaluation de la qualité de la fusion	117
6.4.1	Evaluation des résultats	117
6.4.2	Choix de la répartition pas de décimation / nombre d'étages	132
6.4.3	Comparaison de la pyramide morphologique avec d'autres algorithmes pyramidaux	136
6.5	Recalage de données	136
6.6	Conclusion	138
7	La fusion avec des données multi-temporelles	139
7.1	La problématique	139
7.2	Validation de l'algorithme de fusion temporelle	141
7.2.1	Jeu de données	141
7.2.2	Processus de fusion	142
7.2.3	Résultats	143
7.3	Evolution temporelle d'un paysage terrestre	144
7.4	Suivi temporel à partir d'images multispectrales et grand champ	147
7.5	Conclusion	151
V	CONCLUSION	153
8	Conclusion	155
VI	ANNEXE	157
	La fusion de données Panchromatique / Multispectral	159
	BIBLIOGRAPHIE	165

Introduction

L'arrivée des nouvelles images de télédétection à très haute résolution spatiale va permettre la perception d'une nouvelle gamme d'objets au sol non observables jusqu'à maintenant et pour laquelle on ne dispose pas des moyens d'exploitation adaptés. Cette absence de méthodes de traitement évaluées par les communautés scientifiques et techniques de l'observation de la Terre constitue un obstacle évident au développement de la télédétection dans le milieu professionnel puisqu'il n'existe pas de méthodes transférables. On peut donc imaginer une situation où les bureaux d'étude et les sociétés de services disposeront prochainement d'images d'observation détaillée (semblables aux images aériennes actuelles) mais ne disposeront pas des moyens d'exploitation correspondants.

Ainsi, depuis le lancement du satellite IKONOS en 1999 et celui de QuickBird en 2001, des images satellites à très haute résolution spatiale sont disponibles. Par ailleurs, le lancement de SPOT 5 le 4 mai 2002 et le développement des futurs satellites Pléiades renforcent cette tendance et fourniront sous peu de nombreuses images panchromatiques de résolution infra-métrique et multispectrales de quelques mètres. Ces images permettent la détection d'une gamme d'objets nouvelle et offrent ainsi des perspectives prometteuses pour lesquelles on ne dispose pas des moyens d'exploitation adaptés. Notamment, l'extraction d'objets de ces images devient un élément très important car ceux-ci représentent à ces résolutions des éléments facilement interprétables et exploitables.

Dans cette perspective, l'analyse visuelle doit être vue comme un modèle performant d'extraction d'information. Ainsi, les caractéristiques de la vision primaire reposent sur une perception multi-échelle de l'environnement. Les environnements naturels se prêtent eux aussi à cette décomposition multi-niveaux. Dans ce contexte, l'importance des problèmes d'échelle a donné lieu à de nombreuses études faisant apparaître des hétérogénéités spatiales dépendant de la résolution. La caractérisation des surfaces terrestres par des mesures au sol permet l'accès direct aux variables observées, éventuellement à une fréquence temporelle très élevée. Au contraire, leur étude par télédétection permet de spatialiser les résultats, mais ne permet pas l'accès direct aux variables observées. Deux capteurs complémentaires et cohérents du point de vue spectral et géométrique permettent de travailler dans cette direction : le capteur *grand champ* (VEGETATION) et le capteur multispectral (HRVIR) embarqués sur SPOT 4. En effet, SPOT 4 est la première plate-forme qui intègre simultanément les deux types de charges utiles ouvrant cette possibilité d'observation multi-résolution.

Une approche hiérarchique d'observation est intéressante si elle permet de décrire les continuités et les ruptures des composants de paysage en fonction des échelles permettant leur détection. Suivant cette remarque, une surface décomposée en un ensemble d'objets emboîtés de tailles et formes différentes retranscrit l'organisation spatiale d'un paysage et s'insère bien dans une décomposition pyramidale de données simulant une perception multi-échelle. De plus, les propriétés des filtres non linéaires choisis dans ce travail, combinées à une analyse multi-échelle permettent la séparation à chaque échelle des composantes du paysage représentant les détails et leur répartition en fonction de la variation locale de réflectance sans l'hypothèse implicite d'équi-distribution du filtrage linéaire.

Ainsi, la multiplicité des échelles auxquelles on peut percevoir une information significatives et la diversité des résolutions possibles en télédétection conduisent dans ce mémoire à :

- chercher à décomposer une scène en éléments significatifs à différents niveaux de perception ;
- chercher à composer une observation virtuelle en associant la détection réalisée par plusieurs instruments.

Dans ce contexte, le mémoire présente ainsi le développement d'une représentation multi-échelle d'images par une famille de pyramides morphologiques et ses applications à l'analyse et à la fusion d'images en télédétection. La représentation hiérarchique proposée est basée sur une approche pyramidale utilisant les filtres non linéaires de la morphologie mathématique.

La partie II propose d'abord un état de l'art des transformations pyramidales puis décrit le principe de la pyramide morphologique. Ses propriétés sont étudiées au travers de différents paramétrages et des familles qui en découlent. La pyramide morphologique permet d'une part, de séparer à chaque niveau de résolution les éléments de détails par leur taille et leur réflectance par rapport à l'environnement, et, d'autre part, de représenter les images aux niveaux de perception inférieurs.

La partie III du mémoire est consacrée aux applications des pyramides morphologiques à l'analyse des surfaces observées. La décomposition des éléments imbriqués dans les scènes en signaux séparables à différentes résolutions démontre la capacité de caractérisation multi-échelle. On montre également comment des traitements de reconstruction appliqués aux détails contribuent à cette analyse. Différents exemples de paysages analysés illustrent la méthodologie développée.

La partie IV consiste à fusionner des données de résolutions différentes, notamment celles issues de capteurs à haute résolution et à grand champ de vue. L'approche de fusion par pyramide morphologique crée ainsi des données de synthèse à haute résolution spatiale et à haute fréquence temporelle qui permettent une approche nouvelle du suivi des surfaces terrestres. Les résultats de ce principe de fusion sont présentés en fonction de différents jeux de dates d'acquisition des images.

La conclusion souligne trois perspectives très prometteuses. Premièrement, la pyramide morphologique peut être utilisée comme un navigateur exploitant les différents niveaux de

résolution spatiale permettant l'accès à une information plus ou moins détaillée. Ensuite, la pyramide morphologique ouvre des opportunités intéressantes au recalage d'images de résolutions différentes. Enfin, la pyramide morphologique est un cadre intéressant pour la compression de données par le choix des différents paramètres autorisés.

Nomenclatures des symboles utilisés

Nomenclature de la Pyramide			
$T_{dec,i}$	Transformation réalisant l'étage i de la décomposition	$T_{rec,i}$	Transformation réalisant l'étage i de la recomposition
i	Indice d'étages	N	Nombres d'étages
$FM_{B_i,i}$	Filtre Morphologique	NG	Nombre de niveaux de gris sur lequel l'image est codée
(x_{k_i}, y_{l_i})	Indice de pixels (ligne, colonne) à l'étage i	(K_i, L_i)	Taille des images à l'étage i
Nomenclature des Images basses fréquences			
I_i	Image de base de la décomposition à l'étage i	IF_i	Image filtrée à l'étage i lors de la décomposition
IRF_i	Image Reconstituée mais Filtrée à l'étage i lors de la recomposition	IR_i	Image reconstituée à l'étage i
Nomenclature des Détails			
$D_{sup,filtré,i}$	Détails Sup issus du filtrage à l'étage i	$D_{inf,filtré,i}$	Détails Inf issus du filtrage à l'étage i
$D_{sup,deci,i}$	Détails Sup issus de la décimation à l'étage i	$D_{inf,deci,i}$	Détails Inf issus de la décimation à l'étage i
$D_{filtré,i}$	Détails globaux issus du filtrage à l'étage i	$D_{deci,i}$	Détails globaux issus de la décimation à l'étage i
D_i	Détails globaux obtenus à l'étage i		
Décimation			
$n_i \times m_i$	Pas de décimation <i>ligne</i> \times <i>colonne</i>	$ech_i(n_i, m_i)$	Décimation d'un pas $n_i \times m_i$
Sur-échantillonnage			
$sur - ech_i(n_i, m_i)$	Interpolation générique de taille $n_i \times m_i$, d'un seul étage	Int_k	Interpolation de l'étage $k + 1$ à l'étage k par un pas $n_i \times m_i$
\widetilde{Image}_r^j	Interpolation de <i>Image</i> de la résolution r à la résolution j en passant par toutes les résolutions successives, quelle que soit la résolution de départ		
Termes Mathématiques			
E	Partie entière	ceil	Arrondi par excès

TAB. 1 – Nomenclature des symboles utilisés dans le chapitre 3, « La pyramide morphologique ».

Caractéristiques de la pyramide			
n_i	Pas de décimation carrée ($n_i \times n_i$)	N	Nombre d'étages de la pyramide
Images fusionnées		Indicateurs	
$IFusion_0$	Image fusionnée théorique	IG	Indicateur Global
IHR_{SS}	Image Haute Résolution Sans Saturation ($IHR_{SS} = IFusion_0$)	IL	Indicateur Local
IHR_{AS}	Image Haute Résolution Avec Saturation	p	Indice de voisinage
IHR_{tr+ho}	Image Haute Résolution après Translation et Homothétie	C_p	Valeur de $I_0 - IFusion_0$ sur le palier p

TAB. 2 – Nomenclature des symboles utilisés dans le chapitre 6, « La fusion de données par pyramide morphologique ».

Première partie
CONTEXTE

Chapitre 1

Perception multi-niveau, télédétection et morphologie mathématique

1.1 Observation multi-échelle et surfaces terrestres

Si nos yeux ne voient pas assez précisément, nous prenons une paire de jumelles. Le champ de vision en est réduit, mais la précision augmente, nous percevons plus de détails du paysage observé. Le même phénomène existe pour les images numériques. Un capteur enregistre une image d'une zone de surface S , à une résolution spatiale - précision - r donné. Il faut alors $\frac{S}{r^2}$ pixels pour représenter cette zone.

Supposons maintenant que la résolution soit meilleure, $\frac{r}{2}$ par exemple. Il faut donc $\frac{4S}{r^2}$ pixels pour représenter cette zone, soit 4 fois plus. Dans le cas de l'œil où la taille de l'image est fixe, on ne peut plus représenter S , mais seulement une portion de celle-ci. Dans le cas des images numériques, la même zone géographique est représentée par quatre fois plus de pixels. Inversement, plus la résolution diminue, plus l'observateur semble s'éloigner de la cible.

En fonction de la résolution spatiale à laquelle on discrétise une scène, les résultats des traitements sont totalement différents : la notion de « résolution optimale » en fonction de l'application est alors posée [151]. Si l'on cherche à caractériser des objets (par exemple des parcelles agricoles cultivées) dans leur globalité ou bien à intégrer les sillons du tracteur dans une application, les besoins en résolution spatiale seront totalement différents.

Dans cette perspective, on s'intéresse ici à l'analyse des images aux différentes résolutions. La figure 1.1 illustre cette notion de changement d'échelle dans une image : les images représentées sont à la résolution 1 m, 2 m, 4 m, 8 m, et 16 m. L'image de résolution 1 m possède ainsi quatre fois plus de pixels que celle à 2 m, et ainsi de suite. Dans la représentation de la figure 1.1, quelle que soit la résolution spatiale, le pixel est représenté avec une taille constante : lorsque la résolution s'améliore, la surface contenue dans la représentation de l'image augmente, bien qu'elle caractérise une même aire géographique.

Les différents niveaux de perception apparaissent, ainsi, sur l'image haute résolution, les maisons sont visibles alors que seuls les champs le sont sur les images à basse résolution.

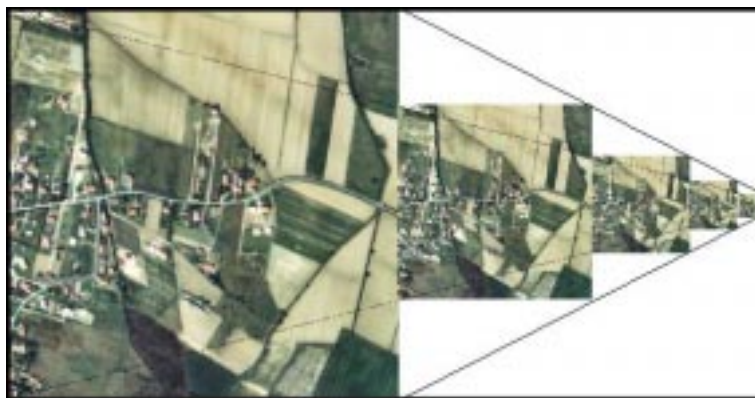


FIG. 1.1 – Différentes échelles pour une même image.

Cette représentation sous forme d'images de plus en plus petites, et ayant de moins en moins de pixels, dessine une sorte de pyramide, à l'origine des termes de décomposition ou d'algorithme pyramidaux. L'image de plus haute résolution est appelée la base de la pyramide, tandis que l'image la plus petite est appelée le sommet. A la limite, celui-ci est de taille 1×1 pixel.

Dans la perspective d'une exploitation des images à très haute résolution spatiale, la fonction image représentant sur un support discret la scène observée doit permettre de s'affranchir d'une description basée uniquement sur la trame des pixels. En particulier, la capacité d'échantillonner une cible à différentes fréquences spatiales et temporelles ouvre la possibilité de « mélanger » des images (on parlera de « fusion ») de résolutions différentes. Cela conduit à une perception composite adaptable.

Par ailleurs, on considère qu'une surface terrestre peut être décrite comme un ensemble d'objets emboîtés de tailles et formes différentes en fonction des attributs que l'on retient comme caractéristiques [3]. Dans ce cadre, la description de l'organisation spatiale d'une scène peut être faite dans une décomposition pyramidale de données simulant une perception multi-échelle [75]. A chaque niveau, ce sont donc des objets différents et des attributs différents qui sont perceptibles et par conséquent séparables. Cette importance des problèmes d'échelle a donné lieu à plusieurs grandes synthèses [115, 90, 3] dans le cas de l'étude des surfaces continentales.

Dans l'étude *in situ* de la structure des surfaces terrestres, la caractérisation au sol est discontinue sur le plan spatial mais elle peut être complète sur le plan des variables observées. Par contre, l'observation par télédétection est continue mais n'accède directement qu'à des caractéristiques spatiales. La paramétrisation de l'organisation des systèmes (distribution 3D des composants) dépend alors beaucoup du niveau de perception, donc de l'échelle d'acquisition ou de synthèse des données.

Une approche hiérarchique pour la description et la quantification de cette structure est donc intéressante si elle permet de préciser la signification des continuités, des ruptures et de leurs composants en fonction des échelles qui permettent de les détecter.

Du point de vue géographique, on admet que les relations structurales et fonctionnelles unissant les constituants d'un système dans un volume donné peuvent être appréhendées par une approche hiérarchique reposant sur la notion de niveau d'intégration [116]. A chacun de ces niveaux, par exemple de l'écosystème local de quelques m^2 jusqu'à la biosphère, l'organisation d'un tel système résulte de l'interaction d'éléments du niveau inférieur, modulée par des processus agissant au niveau supérieur [141, 95]. Ainsi, la superposition de processus agissant à des échelles multiples conduit à un emboîtement de structures spatiales [138] : les processus se déroulant sur de vastes étendues (climat, grandes urbanisations par exemple) se trouvent modifiés au niveau local par d'autres plus limités spatialement (topographie, taille des îlots par exemple). Ceci aboutit généralement à l'accumulation de la variabilité du milieu dans des zones de transition à forts gradients, décelables à certaines échelles seulement.

Dans ces conditions, l'utilisation de la télédétection pour l'étude des surfaces continentales peut être envisagée par une observation multiéchelle selon trois types de méthodologies différentes :

1. Analyse des données permettant de maîtriser le passage d'un niveau de perception à un autre : ce sont les approches multiniveaux et multirésolution.
2. Couplage entre mesures et modèles à des échelles variables : ce sont des méthodes couplant modèles de mesure et modèles spatio-temporels d'objets ou de processus.
3. Intégration d'informations multi-sources pour représenter la surface observée et les mécanismes spatiaux qui s'y déroulent : ce sont les méthodes introduisant des connaissances et leurs représentations.

Les travaux présentés dans ce mémoire se situent dans la première voie. Dans les paragraphes qui suivent, on décrit brièvement deux types d'observations par satellites qui illustrent deux situations de résolutions spatiales et temporelles très différentes. Cela permet de souligner la diversité des mesures qui gagnent à être utilisées conjointement pour décrire une scène le plus complètement possible surtout lorsqu'elle est composée de milieux très différents.

Pour cela, on évoque les deux types d'exemples d'observation par satellites suivants :

1. le cas des couverts végétaux, et, 2. le cas des milieux urbains.

L'étude de la végétation par télédétection repose essentiellement sur l'observation des variations temporelles de la réflexion du rayonnement solaire de courte longueur d'onde. Précisément, la réponse spectrale type d'un couvert végétal montre une forte réflectance dans le proche infra rouge et une faible réflectance dans le rouge. Ces domaines de longueurs d'onde sont donc très utilisés car la différence de réflectances dans ces deux bandes spectrales dépend des caractéristiques chlorophyllienne de la végétation. Ces deux me-

sures permettent de calculer un indice de végétation duquel on déduit un indice foliaire (LAI - Leaf Area Index).

Ainsi, l'étude des couverts végétaux se fait généralement par l'intermédiaire de mesures du LAI (représentant la densité foliaire) pendant un cycle saisonnier, ce qui montre l'évolution de la phénologie. Ces grandeurs ont une évolution rapide notamment en période de croissance et de sénescence. Pour comprendre cette évolution, il est admis que la fréquence temporelle des observations doit être égale ou inférieure à 10 jours.

Par ailleurs, la résolution spatiale doit être suffisante pour permettre de discriminer les différents objets au sol (exemple : les parcelles) comme cela est nécessaire dans de nombreuses applications. Il apparaît qu'une résolution de 20 m, aux latitudes tempérées est suffisante dans la plupart des cas actuels.

Les besoins d'étude des zones urbaines impliquent des approches à plusieurs échelles, compte tenu de la diversité des problèmes posés. Ainsi, une résolution spatiale d'une dizaine de mètres permet d'étudier la croissance des villes, mais pas les variations intra-muros.

Dans les cas d'étude du tissu urbain proprement dit, des résolutions métriques voire infra métrique s'avèrent nécessaires. Par contre, la fréquence temporelle nécessaire dépend de l'application. Généralement, une image par an à la résolution de 10 m s'avère suffisant pour évaluer la croissance urbaine; une image par jour à une résolution métrique serait par exemple nécessaire pour évaluer l'état de saturation d'un parking et en déduire les éventuels aménagements.

Ce très bref aperçu des conditions de mesure en observation de la Terre montre bien que la complexité fréquente de la plupart des scènes de surfaces terrestres observables nécessite différentes résolutions spatiales et temporelles pour être analysées complètement.

1.2 Perception et télédétection multi-résolution

D'une manière générale, il est convenu d'appeler « télédétection » l'ensemble des moyens et méthodes permettant d'obtenir à distance une information sur une scène. En pratique, on ne retient quasi exclusivement que la détection électromagnétique comme moyen de mesure et d'observation. Historiquement, la télédétection a débuté par l'utilisation de ballons (montgolfières) pour observer la Terre depuis le ciel, notamment pour observer les champs de bataille. En 1839, l'invention de la photographie par N. Niepce ouvre de nouveaux horizons. Dès 1855, en effet, Félix Tournachon, surnommé Nadar, prend la première photographie aérienne de Paris. Plus près de nous, les images ont été acquises par avions, ballons stratosphériques puis satellites artificiels orbitant autour de la Terre.

Ces satellites d'observation de la Terre sont des plates-formes orbitant autour du globe et embarquant une ou plusieurs charges utiles à une altitude donnée et une inclinaison

fixée par rapport à l'équateur. Les satellites en orbite basse frottent les hautes couches de l'atmosphère terrestre et ont ainsi une courte durée de vie. Au contraire, les capteurs placés sur des orbites hautes « voient » le sol moins précisément, à capteur identique. Un compromis doit donc être trouvé : en général, les satellites d'observation de la Terre actuels orbitent à environ 800 km. L'optique du capteur et la chaîne d'acquisition définissent la résolution spatiale, la fauchée - largeur - des images acquises, ainsi que leurs résolutions spectrales (assimilables à la largeur des bandes). La période de révolution du satellite et la fauchée du capteur définissent la résolution temporelle des données. Résolution spatiale, spectrale et temporelle définissent donc les caractéristiques des images.

Le développement des technologies numériques et la transmission des images vers le sol ont transformé les satellites en parties très importantes de systèmes d'observations qui nécessitent cependant des segments sol de plus en plus performants pour exploiter les données.

La recherche méthodologique sur les outils correspondants prend actuellement un relief particulier car l'étude de la dynamique des surfaces terrestres est actuellement un thème important motivé par les grands problèmes posés par l'évolution de l'environnement sur Terre. Compte tenu des caractéristiques de résolution des images d'observation de la Terre évoquées plus haut, la télédétection constitue une approche intéressante dans ce type d'objectifs. Toutefois, cela pose les problèmes méthodologiques dus à la multiplicité des échelles qui sont pertinentes pour observer et décrire leur structure et leur fonctionnement. Cela est dû à l'imbrication spatiale et temporelle des constituants et des mécanismes qui rend indispensable leur étude aux différentes échelles où existe une information utile à leur description et à leur modélisation.

Ces quelques remarques montrent donc l'analogie entre :

- le concept décrivant les surfaces comme un ensemble de composants emboîtés ayant leur signification propre à différentes échelles,
- et la télédétection offrant une capacité d'observation continue en surface et à différents niveaux de résolution spatiale et temporelle.

Cela souligne bien le besoin :

1. d'analyse à différents niveaux de résolution, et,
2. de fusion de données ayant des caractéristiques de résolutions complémentaires.

1.2.1 Intérêt et complémentarité des différentes observations

En effet, on doit considérer la surveillance des surfaces terrestres comme un système d'observation mettant en œuvre non seulement des moyens spatiaux d'observation mais également des capacités d'observation aérienne et de mesures *in situ* avec des réseaux de mesures au sol. Cela conduit donc à considérer que cette surveillance peut se réaliser par trois méthodes différentes intervenant dans des proportions variables : les mesures terrains, les capteurs aéroportés, ou les capteurs satellites. Chacune de ces méthodes possèdent un

certain nombre d'avantages et d'inconvénients présentés dans le tableau 1.1.

	Mesures terrains	Images aériennes	Images satellites
Vérité terrain	OUI	NON	NON
Régionalisation	NON, sauf interpolation	OUI, sur partie structurale	OUI, selon la grandeur perçue
Répétitivité	OUI	Faible	Variable

TAB. 1.1 – Avantages et inconvénients comparés des différents types de mesures disponibles.

On comprend qu'aucun type de données ne fournit une information complète au sens du besoin nécessaire à la connaissance des surfaces observées. Les mesures sur le terrain présentent l'avantage d'être des « vérités terrains », c'est-à-dire que la mesure reflète précisément l'état d'une grandeur de la zone d'intérêt au moment de la mesure. De plus, les capteurs correspondant prennent l'information régulièrement et permettent ainsi un suivi temporel performant. Cependant, ces mesures sont très localisées et ne sont significatives que dans une région où leur variabilité spatiale est stationnaire. Cet inconvénient est majeur pour réaliser un suivi sur de grandes régions à partir de mesures locales où ce sont les mécanismes d'interpolation qui induisent la généralisation avec les limitations que cela comporte.

Par contre, les images aériennes ou les images satellites à haute résolution permettent cette régionalisation des données surtout pour ce qui concerne l'organisation spatiale des composants des surfaces. Elles ne fournissent toutefois qu'une mesure de réflectance moyenne sur une zone définie par la résolution spatiale des capteurs. Une analyse ultérieure est donc indispensable pour identifier l'état de la zone observée et l'interpréter comme cela est rappelé plus haut dans le cas des couverts végétaux et des milieux urbains. L'imagerie satellite assure de plus une périodicité des acquisitions, ce qui apporte deux types d'information : 1) sur le type de fonctionnement des surfaces lorsque la nature de celles-ci ne varie pas ; et, 2) sur le type de dynamique spatiale lorsque les composants des surfaces se transforment.

En observation de la terre, trois principaux types de capteurs opérationnels co-existent actuellement, les capteurs grand champ (C-GC), les capteurs multispectraux (C-XS) et les capteurs panchromatiques (C-PAN). De façon résumée, les principales différences entre eux sont :

- la fauchée (environ 2000 km pour C-GC, et moins de 100 km pour C-XS et C-PAN),
- la fréquence temporelle d'acquisition des images (1 jour au moins pour C-GC, et environ 1 mois en moyenne pour C-XS et C-PAN),
- la résolution spatiale (≤ 10 m pour C-PAN, ≤ 20 m pour C-XS et environ 1 km pour C-GC),
- et la résolution spectrale (multibande pour C-GC et C-XS, monobande pour C-PAN).

Depuis peu, sont apparus des capteurs à haute résolution spatiale (de 1 à quelques mètres). Théoriquement, la résolution temporelle des images est très faible, mais ces capteurs sont orientables et programmables, il peuvent être dépointés dans une direction choisie vers l'est ou vers l'ouest pour augmenter le nombre de vues sur un site donné à partir de la même orbite. De ce fait, deux images d'un même site peuvent être obtenues à quelques jours d'intervalle seulement.

1.2.2 Exemple de l'instrument HRVIR sur SPOT 4

L'instrument HRVIR possède 4 bandes spectrales fonctionnant en mode multispectral : 1) la bande XS1, couvrant les longueurs d'onde entre $0,50 - 0,59 \mu m$ (Vert) ; 2) la bande XS2, couvrant les longueurs d'onde entre $0,61 - 0,68 \mu m$ (Rouge) ; 3) la bande XS3, couvrant les longueurs d'onde entre $0,79 - 0,89 \mu m$ (PIR - Proche InfraRouge, ou NIR - Near InfraRed) ; 4) la bande XS4, couvrant les longueurs d'onde entre $1,58 - 1,75 \mu m$ (MIR - Moyen InfraRouge ou SWIR - Short Wave InfraRed).

En mode multibande, la taille d'un pixel est de $20 m \times 20 m$, et de $10 m \times 10 m$ en mode panchromatique ($0,61 - 0,68 \mu m$). Comme la fauchée est de 60 km, les images observées couvrent donc une zone limitée à cette largeur pour une orbite donnée, mais avec une résolution assez élevée. La répétitivité au nadir est de 26 jours mais l'occultation par les nuages ne permet en moyenne qu'un nombre limité d'observations par an si l'on ne fait pas appel au dépointage.

1.2.3 Exemple de l'instrument VEGETATION sur SPOT 4 et 5

L'instrument VEGETATION possède également quatre bandes spectrales, dont 3 sont identiques à celle de HRVIR : 1) la bande b0, couvrant les longueurs d'onde entre $0,43 - 0,47 \mu m$ (Bleu) ; 2) la bande b2 couvrant les longueurs d'onde entre $0,61 - 0,68 \mu m$ (Rouge) ; 3) la bande b3 couvrant les longueurs d'onde entre $0,79 - 0,89 \mu m$ (PIR - Proche InfraRouge, ou NIR - Near InfraRed) ; 4) la bande b4 couvrant les longueurs d'onde entre $1,58 - 1,75 \mu m$ (MIR - Moyen InfraRouge ou SWIR - Short Wave InfraRed).

Ainsi, les bandes XS2 / B2, XS3 / B3 et XS4 / B4 ont les mêmes gabarits spectraux.

La fauchée de l'instrument VEGETATION est de 2250 km. Ainsi, la quasi-totalité de la terre est couverte quotidiennement, seules quelques zones proches de l'équateur ne sont couvertes qu'un jour sur deux.

1.2.4 Conclusion

Les deux types de capteurs multispectraux, haute résolution - du type SPOT 4 HRVIR - et grand champ - du type SPOT 4 VEGETATION - sont complémentaires. Les premiers montrent les éléments fins de la surface observée, tandis que les deuxièmes permettent de suivre l'évolution globale des surfaces. Lorsque les applications font apparaître simultanément des besoins en haute résolution spatiale et haute fréquence temporelle, un choix est nécessaire. Classiquement, l'étude des couverts végétaux se réalise avec des

images grand champ et l'étude des zones urbaines avec des images haute résolution spatiale.

Dans ce mémoire, nous verrons comment combiner ces deux types d'images dans la partie IV. On comprendra également que l'arrivée d'images à très haute résolution spatiale, comme cela est indiqué dans l'introduction, démultiplie fortement l'intérêt de nouvelles approches des images de télédétection.

1.3 Perception et morphologie mathématique

L'algorithme pyramidal que nous proposons dans ce mémoire est basé sur les propriétés des filtres non linéaires de la morphologie mathématique. Ce choix sera illustré à la section 3.3 mais on peut dès maintenant indiquer qu'il a été fait pour respecter les deux propriétés suivantes.

D'une part, sur le plan mathématique, les opérateurs utilisés doivent être invariants en rotation, translation et homothétie pour être indépendants des types de paysages représentés dans les images.

D'autre part, sur le plan physique, les opérateurs utilisés doivent permettre d'extraire des informations à toutes les échelles en postulant la non-équidistribution de la variabilité locale du signal autour d'un plan médian.

Nous allons maintenant rappeler brièvement les principaux concepts de base de la morphologie mathématique indispensables à la compréhension de cette thèse. (le lecteur intéressé par un développement plus complet pourra se référer aux citations de ce chapitre, notamment [123] et [126]).

1.3.1 Le concept

1.3.1.1 Définitions

La morphologie mathématique a été développée pendant les années soixante, à l'Ecole des Mines de Paris par G. Matheron et J. Serra ([123], [126] et [125]). Sa théorie, initialement conçue pour des images binaires, s'est étendue dans les années soixante-dix aux images en niveaux de gris. La morphologie mathématique traite les signaux par une approche géométrique et s'apparente à une méthode de traitement de signal 2D, mais peut néanmoins être utilisée pour traiter des signaux 1D. Cela favorise entre autre la caractérisation des formes et des contours plutôt que d'autres caractéristiques d'une image, comme les variations d'intensité par exemple.

En cela, elle se démarque d'autres techniques usuelles en traitement d'images, comme les décompositions linéaires dans le domaine fréquentiel ou les convolutions dans le domaine spatial. A l'origine, la morphologie mathématique a été conçue pour décrire la

structure des matériaux.

La morphologie mathématique est principalement utilisée en traitement d'images [45], notamment dans quatre domaines : l'extraction d'attributs [50, 62], la reconnaissance et la détection d'objets ou de contours [4, 11, 142, 32, 139], l'organisation spatiale ou volumique d'objets et plus récemment la compression [34, 117]. Elle est également utilisée en télédétection [37].

Les opérations de morphologie mathématique travaillent sur le voisinage d'un point, appelé **élément structurant**. Celui-ci agit en fait comme une sonde qui se déplace sur l'image. Les opérateurs de morphologie mathématique sont non linéaires, et reposent sur une algèbre ensembliste dans laquelle les opérations élémentaires commutent avec *Inf* et *Sup* et non plus selon les opérateurs linéaires habituels.

Dans les paragraphes qui suivent, on rappelle brièvement quelques propriétés sur les opérations de morphologie mathématique. Tout d'abord, la morphologie mathématique se réfère à un treillis complet.

Théorème 1 (Treillis complet) *Soit \mathcal{P} un treillis complet. \mathcal{P} est un ensemble muni d'une relation d'ordre, que nous noterons \leq et tel que toute famille d'élément A_i dans \mathcal{P} admette dans \mathcal{P} un plus petit majorant, que nous noterons $\vee A_i$, et un plus grand minorant, que nous noterons $\wedge A_i$. Ceci implique l'existence dans \mathcal{P} d'un plus grand élément E et d'un plus petit élément \emptyset .*

Dans cet espace, les notions suivantes sont définies :

Définition 1 (la croissance) *Soient X et Y deux ensembles (ou images), et T une transformation morphologique. T est dite croissante si et seulement si :*

$$X \leq Y \Rightarrow T(X) \leq T(Y) \quad (1.1)$$

Définition 2 (l'idempotence) *Soit X un ensemble et T une transformation morphologique. T est dite idempotente si et seulement si :*

$$T[T(X)] = T(X) \quad (1.2)$$

Définition 3 (l'extensivité) *Soit X un ensemble et T une transformation morphologique.*

T est dite extensive ssi :

$$X \subseteq T(X) \quad (1.3)$$

Définition 4 (l'anti-extensivité) *Soit X un ensemble et T une transformation morphologique.*

T est dite anti-extensive ssi :

$$T(X) \subseteq X \quad (1.4)$$

Définition 5 (la dualité) Soit X un ensemble, et T une transformation morphologique. On note $\overline{T}(X)$ la transformation duale de $T(X)$. $\overline{T}(X)$ est calculée comme étant le complémentaire de la fonction T appliqué au complémentaire de X . Mathématiquement, si l'on note X^c le complémentaire de X , et plus généralement c la fonction complémentaire, on a

$$\overline{T}(X) = [T(X^c)]^c$$

Définition 6 (l'invariance en rotation) Soit X un ensemble, T une transformation morphologique et α un angle de rotation. T est dite invariante par translation si

$$T(X_\alpha) = [T(X)]_\alpha$$

En d'autres termes, cela signifie que T est invariante par rotation lorsqu'il est équivalent de faire tourner X par α puis d'appliquer T sur le résultat, ou de commencer par appliquer T sur X et de faire tourner le résultat par α .

Définition 7 (l'invariance en translation) Soit X un ensemble, T une transformation morphologique et h un vecteur de translation. T est dite invariante par translation si

$$T(X_h) = [T(X)]_h$$

En d'autres termes, cela signifie que T est invariante par translation lorsqu'il est équivalent de traduire X par h puis d'appliquer T sur le résultat, ou de commencer par appliquer T sur X et de traduire le résultat par h .

Définition 8 (l'invariance en homothétie) Soit X un ensemble, T une transformation morphologique et λ une constante. T est dite invariante par homothétie si

$$T(\lambda X) = \lambda T(X)$$

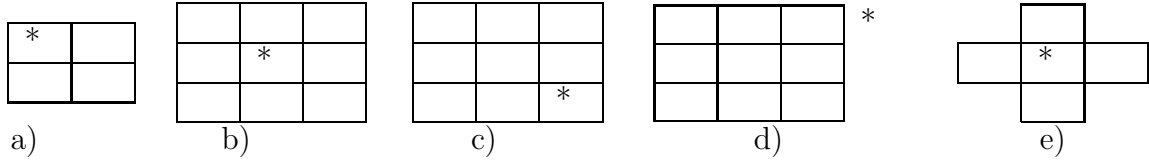
En d'autres termes, cela signifie que T est invariante par homothétie lorsqu'il est équivalent de multiplier X par λ puis d'appliquer T sur le résultat, ou de commencer par appliquer T sur X et de multiplier le résultat par λ .

1.3.1.2 L'élément structurant

Nous avons déjà présenté son rôle : c'est celui d'une sonde qui se déplace sur l'image, afin d'analyser les caractéristiques incluses dans le voisinage qu'il définit.

En continu, l'élément structurant est défini par plusieurs paramètres, la taille, la forme, la connexité et l'origine. L'origine de l'élément structurant correspond au pixel de l'image que l'on traite. Il n'est pas nécessairement symétrique, et son origine ne lui appartient pas forcément.

Voici en numérique des exemples d'éléments structurants : on désigne par le symbole « * » la position de l'origine de l'élément structurant.



La forme la plus régulière d'un élément structurant en numérique est la forme carrée, c'est en effet celle qui se rapproche le plus de la forme ronde ou « boule », naturelle en continu pour un voisinage. Les éléments structurants possibles sont fonction de l'application souhaitée.

En général, un élément structurant est noté B et son symétrique \check{B} .

1.3.2 Les filtres morphologiques

Définition 9 (filtre morphologique) Les applications $\Phi : \mathcal{P} \rightarrow \mathcal{P}$, croissantes et idempotentes sont appelées filtres morphologiques sur \mathcal{P} .

En explicitant, est un filtre toute transformation T qui conserve la relation d'ordre (croissance) et qui est tel que $T[T(X)] = T(X)$ (idempotence). Cette définition est beaucoup plus précise et beaucoup plus stricte que celle d'un filtre en traitement du signal linéaire.

Définition 10 (fermetures, ouvertures) Les filtres extensifs sont appelés fermetures, les filtres anti-extensifs ouvertures.

1.3.2.1 Propriétés des filtres

De nouveaux filtres peuvent être obtenus par composition de filtres. Le théorème suivant est exprimé par G. Matheron :

Théorème 2 Soit ζ et ψ deux filtres tels que $\zeta > \psi$, alors :

1. $\zeta\psi, \psi\zeta, \zeta\psi\zeta, \psi\zeta\psi$ sont des filtres morphologiques (croissants et idempotents)
2. $\psi \leq \psi\zeta\psi \leq \zeta\psi\zeta \leq \zeta$
3. $\zeta\psi\zeta$ est le plus petit filtre plus grand que $\zeta\psi \vee \psi\zeta$ (\vee désigne le sup)
 $\psi\zeta\psi$ est le plus grand filtre plus petit que $\zeta\psi \wedge \psi\zeta$ (\wedge désigne l'inf)
4. On a les équivalences : $\zeta\psi\zeta = \psi\zeta \Leftrightarrow \psi\zeta\psi = \zeta\psi \Leftrightarrow \psi\zeta \geq \zeta\psi$

Le nombre de filtres obtenus par composition est donc assez limité. En effet, composer plus de trois filtres de bases ne fournit pas d'autres filtres à cause de la propriété d'idempotence (par exemple $\zeta\psi\zeta\psi = (\zeta\psi)(\zeta\psi) = \zeta\psi$). Pour une étude plus complète des filtres morphologiques, le lecteur pourra se référer à [125] et [126].

1.3.2.2 Les ouvertures et fermetures morphologiques (ou par adjonction)

Ce sont des ouvertures (resp. des fermetures).

Dilatation / Erosion Les opérateurs de base travaillent sur l'ensemble des pixels inclus dans l'élément structurant. Ainsi, l'origine de l'élément structurant est superposée au pixel à traiter, et tous les points de l'image correspondant aux points inclus dans l'élément structurant sont pris en compte par l'opérateur.

Mathématiquement la dilatation d'une image f par un élément structurant B au point x s'exprime par :

$$(f \oplus B)(x) = \text{Max}\{f(z), z \in B(x)\} \quad (1.5)$$

Dans cette expression, x est l'origine de l'élément structurant, c'est donc le point de l'image que l'on traite. z représente l'ensemble des points inclus dans l'élément structurant. On remarque que si l'origine x n'est pas incluse dans l'élément structurant, alors $x \notin z$.

La dilatation est une opération **croissante et extensive**.

L'érosion est l'opération duale de la dilatation. Ainsi, faire une dilatation sur un objet ou faire une érosion sur son complémentaire donne le même résultat sur une image.

L'expression mathématique de l'érosion de l'image f par l'élément structurant B au point x est :

$$(f \ominus B)(x) = \text{Min}\{f(z), z \in B(x)\} \quad (1.6)$$

L'érosion est une opération **croissante et anti-extensive**.

Les dilatations et les érosions sont invariantes par translation.

Ouverture / Fermeture morphologiques L'ouverture et la fermeture morphologiques sont les deux autres opérateurs de base. Ce sont en fait des compositions de dilatation et d'érosion.

L'**ouverture** morphologique est une transformation composée d'une **érosion puis d'une dilatation** :

$$X \circ B = X_B = (X \ominus \check{B}) \oplus B \quad (1.7)$$

Une ouverture morphologique est une transformation **croissante, anti-extensive et idempotente**.

La **fermeture** morphologique est une transformation composée d'une **dilatation puis d'une érosion**.

$$X \bullet B = X^B = (X \oplus \check{B}) \ominus B \quad (1.8)$$

La fermeture morphologique est une transformation **croissante, extensive et idempotente**.

Une conséquence de ces propriétés est importante pour l'utilisation de la pyramide morphologique : les filtres morphologiques permettent une séparation spatiale de la variabilité locale du signal par les différences de fonctions.

Tout d'abord, si $f_B(x)$ est l'ouvert de la fonction f au point x , la différence $f(x) - f_B(x)$ contient les valeurs élevées de luminance par rapport voisinage dont la période est contenue ou égale à la taille de l'élément structurant. L'inverse est vrai - valeur faible de luminance par rapport au voisinage - pour la différence des fonctions $f^B(x) - f(x)$, où $f^B(x)$ est le fermé de la fonction f au point x .

Ensuite, les parties des signaux extraites par ces différences de fonctions peuvent constituer directement l'information recherchée lorsqu'il s'agit d'éléments apparaissant à une résolution donnée comme des points, des lignes, des textures diverses, etc ... Ils peuvent également constituer des marqueurs d'objets lorsque ceux-ci apparaissent comme des traces schématisées à des niveaux de résolution plus grossiers. La reconstruction des objets marqués pourra alors se faire à un niveau donné indépendamment des autres.

Deuxième partie

**LA PYRAMIDE
MORPHOLOGIQUE**

Chapitre 2

Etat de l'art : les algorithmes pyramidaux

Ce chapitre décrit le principe général des algorithmes pyramidaux et présente également une revue des algorithmes existants.

2.1 Définition

La notion et la caractérisation des échelles dans une image apparaît en 1971 [114] lorsque celles-ci sont étudiées par l'application d'opérateurs de tailles de plus en plus grandes. C'est en 1975 que Tanimoto et Pavlidis proposent les premières pyramides [134] : l'opérateur garde une taille fixe relativement petite et cela améliore les performances en temps de calcul des algorithmes [113]. Dans les années qui suivent, l'approche de Marr [75] est à la base de nombreux développements lorsqu'il s'agit de rechercher une caractérisation multiéchelle de scènes. Le principe de détection dyadique des passages par zéro associés à des différences de filtres passe-bas ([74, 76]) est ainsi particulièrement intéressant : en effet, il semble une bonne simulation des systèmes perceptifs de bas niveau assimilés à des systèmes ayant plusieurs canaux fréquentiels fonctionnant simultanément [150, 149]. De nombreuses pyramides ont alors été développées, la pyramide de Burt et Adelson [8], celle de Crowley et Stern [25], le laplacien pyramidal et le codage en sous bande sont parmi les plus connues. D'une manière générale, ce type d'analyse multirésolution est proche de l'analyse en ondelettes de Grossmann et Morlet [42] et Meyer [82] qui permet de recomposer une image à partir de sa représentation multiniveaux. En 1989, le développement de pyramides à base d'ondelettes par Mallat ([68], [69], [67], [70]) a conduit à un essor important des approches pyramidales.

A l'origine, l'application principale consistait à extraire et à transmettre rapidement l'information contenue dans une image. Ensuite, les pyramides ont été exploitées dans un but de compression de données [51]. Actuellement, les applications sont plus diversifiées avec notamment l'apparition de la fusion de données par algorithmes pyramidaux [104].

Les pyramides à base de morphologie mathématique apparaissent en 1987 [43]. Elles

permettent, grâce au choix de l'élément structurant, une extraction très spécialisée. Toët [135] fait de l'analyse de forme grâce à des pyramides morphologiques en 1989. Sun et Maragos [132] l'utilisent pour des applications de compression de données.

Le principe des algorithmes pyramidaux est de décomposer une image en images de différentes résolutions puis de la recomposer. Dans ce cadre, un algorithme pyramidal se définit comme la décomposition d'une image sous forme d'arbre, de sorte que l'étage supérieur puisse être restitué à partir de l'étage inférieur.

Le développement simultané des pyramides et de leurs champs d'application a conduit à une grande variété d'algorithmes pyramidaux, parfois issus de théories très diverses. Il est de ce fait difficile de donner une définition globale de cette notion. Cependant, Eichmann *et al.* [29] et Haralick *et al.* [43] ont proposé des définitions vers la fin des années 1980. Haralick souligne la nécessité du filtrage et de l'échantillonnage dans chaque algorithme pour supprimer tout effet d'aliasing. Plus récemment, Goutsias et Heijmans [41] se sont essayés à une définition plus globale :

« 1) The pyramid consists of a (finite or infinite) number of levels such that the information content decreases toward higher levels and 2) each step toward a higher level is implemented by an (information-reducing) analysis operator, whereas each step toward a lower level is implemented by an (information-preserving) synthesis operator. »

« 1) La pyramide consiste en un nombre (fini ou infini) de niveaux tel que le contenu en information diminue lorsque le niveau augmente et 2) chaque pas vers un niveau supérieur est implémenté par un opérateur d'analyse (réduisant l'information), tandis que chaque pas vers un niveau inférieur est implémenté par un opérateur de synthèse (préservant l'information). »

D'un point de vue plus général, l'analyse multi-résolution permise par les pyramides considère que l'on dispose d'une mesure spatialement continue de la scène. Cependant, pour acquérir les images, cette mesure est réalisée par un capteur qui délivre une image numérique par échantillonnage de la cible. La donnée obtenue est donc une approximation R_i , de résolution i , de la fonction continue analogique $f(x, y)$ représentée par le paysage. D'une manière générale, l'analyse multi-résolution consiste à traiter l'information présente dans les fonctions spatialement continues $R_i f(x, y)$ pour l'ensemble des niveaux i de résolutions intéressants $i = 1, \dots, r$.

2.2 Schéma général d'un algorithme pyramidal

Dans tous les algorithmes pyramidaux, la décomposition multi-résolution (resp. la reconstruction) se construit par itération d'une transformation $T_{dec, i}(f(x, y))$ (resp. $T_{rec, i}(f(x, y))$) où i est l'indice d'étages. Pour simplifier ce qui suit, on convient d'assimiler $T_{dec, i}(f(x, y))$ à $T_{dec, i}$ et $T_{rec, i}(f(x, y))$ à $T_{rec, i}$. Ainsi, $T_{dec, i}$ fait décroître la résolution et relie les valeurs de la fonction projetée à l'étage i à celles de sa projection à l'étage $i + 1$. Au contraire $T_{rec, i}$ relie les valeurs de la fonction projetée à l'étage $i + 1$ à celles de sa projection à

l'étage i par l'intermédiaire d'images appelées ci-après OI_i .

De façon mathématique, soit I_0 l'image de départ, c'est-à-dire celle au niveau 0 de la pyramide,

$$\{I_1, OI_0\} = T_{dec,0}(I_0) \quad (2.1)$$

où I_1 représente l'image filtrée au niveau 1 de la pyramide, et $\{OI_0\}$ l'ensemble des autres images créées lors de la transformation $T_{dec,0}$. A l'étage suivant, on a,

$$\{I_2, OI_1\} = T_{dec,1}(I_1) \quad (2.2)$$

où I_2 représente l'image filtrée au niveau 2, et $\{OI_1\}$ l'ensemble des autres images créées lors de la transformation $T_{dec,0}$. De façon itérative,

$$\{I_1, I_2, \dots, I_N, OI_{N-1}, OI_{N-2}, \dots, OI_0\} = T_{dec,N-1}[\dots[T_{dec,1}[T_{dec,0}(I_0)]]] \quad (2.3)$$

I_N est l'image basse résolution au $N^{\text{ème}}$ niveau et $\{OI_j\}_{j=0\dots N-1}$ les autres images calculées.

Si la pyramide comporte N niveaux, I_N est alors **l'image basse résolution** de la pyramide.

La reconstruction se base sur la dernière image filtrée. Les opérations suivantes sont ensuite itérées, de :

$$IR_{N-1} = T_{rec}(\{I_N, OI_{N-1}\}) \quad (2.4)$$

jusqu'à

$$\begin{aligned} IR_0 &= T_{rec,0}\{T_{rec,1}\{\dots T_{rec,N-2}\{T_{rec,N-1}\{I_N, OI_{N-1}\}, OI_{N-2}\}, \dots, OI_1\}, OI_0\} \\ &= T_{rec,0}(\{IR_1, OI_0\}) \end{aligned} \quad (2.5)$$

où IR_i représente l'image reconstruite au niveau i .

$T_{rec,i}$ est la transformée inverse de $T_{dec,i}$.

Pour reconstruire une image à la résolution de l'image I_0 , il est nécessaire de faire autant d'étages de reconstruction que l'on en a fait lors de la décomposition. L'image IR_0 est donc la reconstruction de l'image I_0 après passage dans la pyramide :

$$IR_0 = T_{rec,0}\dots T_{rec,N-1}(T_{dec,N-1}\dots T_{dec,1}T_{dec,0}(I_0)). \quad (2.6)$$

2.3 Les principaux algorithmes pyramidaux

De façon générale, plusieurs types d'algorithmes pyramidaux se dégagent :

- les premiers, basés sur l'itération d'un filtre et d'un échantillonnage correspondent plutôt à la définition de modèles de perception ;
- les deuxièmes sont basés sur l'analyse en ondelettes ;
- les troisièmes utilisent les outils de la morphologie mathématique et feront l'objet d'une description plus détaillée dans la section suivante.

Le plus connu des algorithmes pyramidaux basés sur l'itération d'un filtre et d'un échantillonnage, est celui proposé par Burt et Adelson en 1983 [8]. A l'origine, cet algorithme a été développé pour faire du codage le moins redondant possible, tout en transmettant progressivement les niveaux successifs de la pyramide. Il combine le calcul de deux pyramides. La première est appelée pyramide gaussienne et réalise un filtrage et un échantillonnage en un seul processus correspondant à T_{dec} itéré à chaque niveau. T_{dec} ne dépend donc pas du niveau l de la pyramide. La formule de changement de niveau est la suivante :

$$g_l(i, j) = \sum_{m=-2}^2 \sum_{n=-2}^2 w(m, n) g_{l-1}(2i + m, 2j + n) \quad (2.7)$$

Le filtre, de taille 5×5 est une moyenne pondérée par les coefficients du noyau w . L'équation (2.7) est itérée de $l = 0$ à $l = l_{max}$ conduisant à une image $g_{l_{max}}$ de taille unité. L'ensemble des $\{g_l\}_{l=0, \dots, l_{max}}$ forme la pyramide gaussienne et correspond aux images I_0 à $I_{l_{max}}$. La partie reconstruction de la pyramide se fait en moyennant de façon pondérée les pixels de l'image basse résolution les plus proches géographiquement. La pondération utilisée est la même que celle choisie lors de la décomposition.

La deuxième pyramide, la pyramide laplacienne, est construite comme la différence entre deux niveaux de la pyramide gaussienne. Elle se calcule par

$$\begin{cases} L_l &= g_l - EXPAND(g_{l+1}) \\ L_N &= g_N \end{cases} \quad (2.8)$$

où l'opérateur $EXPAND$ est l'interpolation de l'image au niveau supérieur.

Le premier niveau de la pyramide laplacienne, L_N représente la basse résolution de l'image de départ. Ainsi,

$$\begin{aligned} g_l &= L_l + EXPAND(g_{l+1}) \\ g_{N-1} &= L_{N-1} + EXPAND(L_N) \end{aligned} \quad (2.9)$$

On peut reconstruire exactement g_0 à partir de $\{L_i\}_{i=0 \dots N}$ et de l'opérateur $EXPAND$. $\{L_0, L_{N-1}\}$ correspondent aux images $\{OI_0, OI_{N-1}\}$. Dans le cas de la pyramide de Burt et Adelson, il n'y a donc qu'une seule autre image créée par niveau.

Les premiers modèles de pyramides, proposés en 1975 et 1979 par Tanimoto [134, 133] proposent des algorithmes plus simples que celui de Burt et Adelson, consistant par exemple à prendre la moyenne, ou même la somme des pixels (i, j) , $(i + 1, j)$, $(i, j + 1)$, $(i + 1, j + 1)$ avec i et j impairs afin de calculer la valeur du pixel $(\frac{i+1}{2}, \frac{j+1}{2})$ au niveau inférieur, l'échantillonnage et le filtrage sont toujours réalisés en une seule et même étape. Dans le cas de l'algorithme « somme », la reconstruction se base sur la transmission du sommet de la pyramide et de $3/4$ des pixels aux niveaux supérieurs. Les pixels manquants sont obtenus par différence entre la somme codée à l'étage inférieur et les trois pixels de l'étage. Dans la même optique, Wang et Goldberg [146] proposent de ne transmettre que

la différence entre deux étages de la pyramide de somme ou de la pyramide tronquée ainsi que le sommet de la pyramide. Dans la pyramide « S-transform » proposée par Wendler *et al.* [148] et Boucher *et al.* [5], chaque paquet de 4 pixels est transformé en quatre pixels combinaisons linéaires des précédents. C'est cette nouvelle image qui est échantillonnée, en ne gardant qu'un pixel sur 4.

Ces méthodes sont axées vers la transmission progressive de données. Une grande attention est donc portée dans ces publications, à des schémas de reconstruction sans perte, à partir d'un nombre de bits et de pixels transmis minimaux. Le papier de Goldberg et Wang [40] synthétise les avantages et inconvénients de chacune de ces méthodes dans cet objectif. Dans le même esprit, Ho et Gersho [47] itèrent des décimations, et calcule la différence entre deux niveaux successifs. Des contours sont extraits de chaque « détail » (image différence). Crowley et Stern [25] ont également proposé un schéma de pyramide laplacienne en 1984.

En 1990, Jolion [51, 53] propose de construire des algorithmes pyramidaux, soit en fonction de la simplicité d'implémentation, soit en fonction de l'application. Il présente ainsi le principe d'une pyramide stochastique pour laquelle le rapport de décimation et le nombre de descendants sont variables. En 1994, il propose une pyramide de contraste [52] où chaque niveau est constitué d'une mesure de contraste sur un voisinage de taille différente, et qui garde toutefois la propriété de coder exactement l'image de départ. Par la suite, il définit une pyramide basée sur la pyramide gaussienne de Burt et Adelson [72], mais n'ayant pas un nombre de descendants constant sur l'image pour réaliser de la vision exploratoire.

Par la suite, les propriétés des filtres sont mises en valeur notamment afin de supprimer le bruit créé lors de l'échantillonnage. Le filtre peut ainsi inclure un opérateur de détection [131], un opérateur de réduction de bruit [17] ou même être choisi en fonction des propriétés de sa transformée en Z [86].

Certains algorithmes pyramidaux sont également conçus à partir de modèles markoviens hiérarchiques [55].

Les algorithmes basés sur la théorie des ondelettes sont plus généralement appelés analyse multi-résolution et réunissent en une seule étape les opérations de filtrage et d'échantillonnage. Les travaux de Mallat sont parmi les plus connus [68], [69], [67], [70]. Ces méthodes se sont ensuite développées, on peut citer par exemple les travaux de Pichler *et al.* [99] sur l'extraction d'éléments de structure et de Proenca [104] sur la fusion d'images multi-résolution. Les relations entre ces différentes méthodes d'analyse multi-échelle et multi-résolution sont bien décrites par Froment et Morel [38] du point de vue de leurs propriétés en analyse d'images. Une bonne revue des méthodes multi-résolutions linéaires peut être trouvée dans [63].

2.4 Les notions de pyramides morphologiques

Trois notions différentes de pyramides morphologiques coexistent :

1. les pyramides basées sur le principe de l'itération d'un filtrage et d'un échantillonnage,

2. les pyramides construites uniquement par itérations d'opérations morphologiques,
3. les méthodes multi-résolution, souvent dans un but de segmentation.

2.4.1 Pyramides par itération d'un filtrage et d'un échantillonnage

Dans ces méthodes, il y a toujours un opérateur de réduction d'information et un opérateur d'expansion d'information.

Pour Haralick *et al.* [43, 44], l'opérateur de réduction d'information est composé d'une ouverture, d'un échantillonnage, d'une dilatation puis d'un sous-échantillonnage afin de générer l'étage suivant.

Pour Toët [135] en 1989, l'opérateur de réduction d'information est composé d'un filtre (ouverture puis fermeture) suivi d'un échantillonnage (dyadique). Une pyramide d'erreur est ensuite calculée par l'opérateur d'expansion et permet une reconstruction exacte.

Pour Sun et Maragos [132], en 1989 également, la pyramide se construit par itérations d'une ouverture par un élément structurant fixé et d'une décimation. Deux opérateurs d'expansion sont envisagés : le premier calcule la différence entre l'image avant et l'image après ouverture. Dans ce cas, la reconstruction n'est pas exacte. La deuxième calcule la différence entre l'image avant ouverture et celle après interpolation de l'image décimée. La reconstruction est alors exacte. Deux types d'interpolation sont testés (moyenne par une fenêtre 3×3 et dilatation par un élément structurant 3×3).

En 1994, Marshall et Matsopoulos [77] calculent l'opérateur de réduction d'information par itération d'un filtre par un élément structurant symétrique et d'un échantillonnage. L'opérateur expansion d'information se réalise par interpolation. Dans l'exemple cité, le filtre est une ouverture-fermeture par un élément structurant 3×3 et l'échantillonnage est dyadique. Marshall et Matsopoulos appliquent cet algorithme pour réaliser de la fusion de données MR et CT du cerveau humain.

Florencio et Schafer [34] utilisent également ce type d'algorithme pour faire de la compression.

2.4.2 Pyramides par itération d'opérations morphologiques

En 1988, Eichmann *et al.* [29] proposent d'itérer des opérateurs morphologiques (dilatation, érosion, ouverture et fermeture) par des éléments structurants identiques (pour des dilatations et des érosions) ou grossissants. La reconstruction est réalisée par application de la transformation duale, elle n'est donc pas exacte. Ensuite, la différence entre deux étages est calculée.

En 1988 également, Maragos [73] itère n érosions suivies de n dilatations (ou vice-versa) pour définir des ouvertures (resp. des fermetures) multi-échelles. De même, en 1989, Chen et Yan [18] utilisent des éléments structurants grossissants pour étudier l'évolution des contours.

En 1995, Montiel *et al.* [87] itèrent n érosions suivi de m dilatations (et vice-versa) pour aller de haut en bas de sa pyramide.

En 1996, Jackway [50] utilise comme fonction d'échelle une fonction structurante caractérisée par un paramètre d'échelle σ .

En 1998, Chanda *et al.* [11] utilisent des éléments structurants grossissants pour détecter les contours.

En 2001, Mukhopadhyay et Chanda [89] utilisent une pyramide morphologique construite par des éléments structurants grossissants. Cette pyramide est élaborée dans un but de fusion, nous détaillerons donc son principe dans le chapitre 5.

2.4.3 Pyramides par méthodes multi-résolution

D'une façon générale, une étape supplémentaire (extraction d'information, calcul de gradient, ...) est réalisée dans ce contexte à chaque niveau de la pyramide. Les calculs à un niveau donné se décompose donc

- en une décomposition du signal, de façon identique aux schémas vus précédemment,
- en une étape supplémentaire, comprenant par exemple une partie extraction de caractéristique puis une partie décision.

En 1992, Salembier et Serra [118] proposent par exemple une approche multi-échelle morphologique dans un but de segmentation. On peut également citer leur travaux de 1993 [128].

En 1997, Cichosz et Meyer [23] réalisent une segmentation sous forme d'arbres : ils itèrent une étape de filtre et calcul de gradient qui sur-segmente l'image puis une ligne de partage des eaux pour connecter les bassins versants.

Chapitre 3

La pyramide morphologique

La pyramide morphologique permet, en combinant le principe des algorithmes pyramidaux et la théorie de la morphologie mathématique, de **décomposer et recomposer sans erreur** une image quelconque à différentes résolutions en respectant l'hypothèse de non-équidistribution des détails du signal après le filtrage. Ce principe peut avoir de nombreuses applications, mais deux d'entre elles sont particulièrement intéressantes en observation de la Terre et font l'objet du développement de ce mémoire :

- la première est l'analyse des images à différentes résolutions ce qui permet de déterminer à quels niveaux apparaissent différentes catégories d'objets et de les isoler ;
- la seconde est la fusion des données, ce qui permet d'intégrer dans la pyramide et donc de mélanger plusieurs images à des niveaux de résolution différents.

Ces applications seront traitées dans les parties III et IV.

La pyramide morphologique est composée de deux parties, la décomposition et la reconstruction (ou recomposition). Si l'on appelle I_0 l'image de départ et IR_0 l'image reconstruite, pour la pyramide morphologique, alors $IR_0 = I_0$.

La nomenclature des termes utilisés dans ce chapitre est décrite dans le tableau 1.

Comme on ne sait pas « créer de l'information », la partie reconstruction est forcément précédée par une décomposition. Par contre, dans certaines applications, la décomposition peut se suffire à elle-même. Ce principe de la pyramide morphologique fait l'objet de plusieurs publications et communications, en particulier [57, 35].

3.1 La décomposition

3.1.1 Principe

On rappelle qu'une image $I(x, y)$ représente une scène de taille quelconque (*hauteur, largeur*) où (x, y) sont les variables continues de l'espace à deux dimensions de l'image.

Du point de vue numérique, cette image est représentée par son approximation sur un support de trame G formé par un paysage donné [12]. Les coordonnées (x, y) deviennent alors des indices (x_k, y_l) discrets et si K (resp. L) est le nombre de lignes (resp. de colonnes) de la projection de $I(x, y)$ sur G , chaque pixel de l'image projetée correspond alors à une surface physique de la scène selon le support d'homothétie $\frac{\text{hauteur}}{K} \times \frac{\text{largeur}}{L} \text{unité}^2$.

La projection de $I(x, y)$ sur G est donc représentée en espace discret par l'ensemble des pixels $(x_k, y_l)_{k=1\dots K; l=1\dots L}$.

Alors que l'image analogique peut s'écrire comme :

$$I = \{I(x, y) / (x, y) \in \mathbb{R}^2, x \leq \text{hauteur}, y \leq \text{largeur}\} \quad (3.1)$$

où x et y sont des variables physiques exprimées dans la même unité que *hauteur* et *largeur*, la projection de I sur G s'exprime par :

$$\text{Projection}(I(x, y)) = \{I(x_k, y_l) / ((x_k, y_l) \in \mathbb{N}^2, x_k = 1\dots K, y_l = 1\dots L)\} \quad (3.2)$$

où (x_k, y_l) représente le pixel courant et $I(x_k, y_l)$ la valeur de $I(x, y)$ projetée sur la grille en (x_k, y_l) .

Lorsque $(K, L) \rightarrow \text{inf}$, $\text{Projection}(I) \rightarrow I$, ceci signifie évidemment que l'échantillonnage spatial de la scène est de plus en plus fin.

Ici, soit $I_0(x_{k_0}, y_{l_0})_{x_{k_0}=1\dots K_0; y_{l_0}=1\dots L_0}$ l'image numérique de départ de taille K_0 lignes $\times L_0$ colonnes, projetée sur la trame G_0 . (x_{k_0}, y_{l_0}) désigne les coordonnées discrètes et donc le pixel courant des images au niveau 0. Comme au chapitre précédent, i désigne l'indice de niveaux et $(x_{k_i}, y_{l_i})_{x_{k_i}=1\dots K_i; y_{l_i}=1\dots L_i}$ représente le pixel courant au niveau i .

Dans la pyramide morphologique, la transformation $T_{dec,i}$ est en fait composée de quatre étapes distinctes. A chaque étage i de la décomposition, les opérations suivantes s'enchaînent :

1. filtrage morphologique,
2. calcul de la différence Image Initiale / Image filtrée,
3. échantillonnage,
4. calcul de la différence Image avant échantillonnage / Image après échantillonnage.

3.1.1.1 Etape 1 : le filtrage

Le filtre (FM) est un filtre passe-bas, basé sur des filtres morphologiques ou leurs combinaisons. Définir complètement le filtrage implique évidemment de définir un élément structurant par sa forme, sa taille et son origine. Les informations extraite de l'image d'origine représentent les détails contenant les éléments fins ou petits et intégrant les objets de taille liée à la résolution des images à l'étage i .

Par exemple, sur un voisinage 3×3 à la résolution de 1 m, les objets de surface inférieure à 9 m^2 vont être modifiés. A la résolution de 4 m, ce sont les objets de surface inférieure à $9 \times 4^2 = 144 \text{ m}^2$ qui vont l'être.

Mathématiquement, l'étape 1 à l'étage i s'écrit $IF_i(x_{k_i}, y_{l_i}) = FM_{B_i, i}(I_i(x_{k_i}, y_{l_i}))$ où IF_i représente l'image filtrée de I_i , B_i est l'élément structurant au niveau i associé au filtre $FM_{B_i, i}$.

Le filtrage sera pris comme combinaisons d'ouverture et de fermeture. Par exemple, $FM_{B_i, i}(I_i(x_{k_i}, y_{l_i}))$ peut être égal à

$$\begin{aligned} & \gamma_{B_i} \varphi_{B_i}(I_i(x_{k_i}, y_{l_i})), \\ & \varphi_{B_i} \gamma_{B_i}(I_i(x_{k_i}, y_{l_i})), \\ & \gamma_{B_i} \varphi_{B_i} \gamma_{B_i}(I_i(x_{k_i}, y_{l_i})), \\ & \varphi_{B_i} \gamma_{B_i} \varphi_{B_i}(I_i(x_{k_i}, y_{l_i})), \\ & \frac{\gamma_{B_i}(I_i(x_{k_i}, y_{l_i})) + \varphi_{B_i}(I_i(x_{k_i}, y_{l_i}))}{2}. \end{aligned} \quad (3.3)$$

où γ est une ouverture et φ une fermeture par adjonction. Le dernier filtre proposé ne répond pas à la définition des filtres morphologiques, il est toutefois un filtre passe-bas au sens général du terme.

3.1.1.2 Etape 2 : le calcul des détails dus au filtrage

Une fois le filtrage effectué, l'information supprimée par celui-ci est calculée par l'intermédiaire de deux images,

- l'image différence (Image Initiale - Image Filtrée), et les pixels négatifs sont mis à zéro,
- l'image différence (Image Filtrée - Image Initiale), les pixels négatifs sont mis à zéro.

La première image correspond aux pixels dont le compte numérique est abaissé par le filtrage et donc aux zones plus claires sur l'image initiale que sur l'image filtrée. On choisit d'appeler cette image $D_{sup, filtre, i}$ au niveau i . La deuxième image correspond au contraire aux pixels dont le compte numérique est surélevé après le filtrage et donc aux zones plus foncées sur l'image initiale que sur l'image filtrée. On choisit d'appeler cette image $D_{inf, filtre, i}$ au niveau i .

Mathématiquement, à l'étage i , pour le pixel (x_{k_i}, y_{l_i}) , on a donc

$$\begin{cases} D_{sup, filtre, i}(x_{k_i}, y_{l_i}) = Sup(I_i(x_{k_i}, y_{l_i}), IF_i(x_{k_i}, y_{l_i})) - IF_i(x_{k_i}, y_{l_i}) \\ D_{inf, filtre, i}(x_{k_i}, y_{l_i}) = Sup(I_i(x_{k_i}, y_{l_i}), IF_i(x_{k_i}, y_{l_i})) - I_i(x_{k_i}, y_{l_i}) \\ \forall x_{k_i} \in [1; K_i], \forall y_{l_i} \in [1; L_i] \end{cases} \quad (3.4)$$

$D_{sup, filtre, i}$ et $D_{inf, filtre, i}$ ont des valeurs positives, entières et codées sur le même nombre NG de niveaux de gris que les images I_i .

Si l'on appelle $D_{filtre, i}(x_{k_i}, y_{l_i}) = D_{sup, filtre, i}(x_{k_i}, y_{l_i}) - D_{inf, filtre, i}(x_{k_i}, y_{l_i})$, alors

$$D_{filtre, i}(x_{k_i}, y_{l_i}) = I_i(x_{k_i}, y_{l_i}) - IF_i(x_{k_i}, y_{l_i}) \quad (3.5)$$

$D_{filtre, i}$ prend des valeurs positives ou négatives, entières et codées sur $[-NG, NG]$.

Par la définition même des $D_{sup, filtre, i}$ et des $D_{inf, filtre, i}$, on a :

$$D_{sup, filtre, i}(x_{k_i}, y_{l_i}) \text{ non nul} \Rightarrow D_{inf, filtre, i}(x_{k_i}, y_{l_i}) \text{ est nul},$$

et, de façon équivalente,

$$D_{inf, filtre, i}(x_{k_i}, y_{l_i}) \text{ non nul} \Rightarrow D_{sup, filtre, i}(x_{k_i}, y_{l_i}) \text{ est nul}.$$

Les détails filtrés d'un pixel se retrouvent entièrement dans $D_{sup, filtre, i}(x_{k_i}, y_{l_i})$ ou dans $D_{inf, filtre, i}(x_{k_i}, y_{l_i})$, mais jamais mélangés dans les deux.

3.1.1.3 Etape 3 : l'échantillonnage

Vient maintenant l'étape qui consiste à changer de résolution. L'image filtrée est échantillonnée, ce qui réduit la taille de l'image.

L'échantillonnage de l'image IF_i s'exprime par

$$IF_i(x_{k_i}, y_{l_i}) \xrightarrow{ech_i(n_i, m_i)} I_{i+1}(x_{k_{i+1}}, y_{l_{i+1}}) \text{ avec} \quad (3.6)$$

$$\begin{cases} x_{k_i} = [1 \dots K_i], y_{l_i} = [1 \dots L_i] \\ x_{k_{i+1}} = [1 \dots K_{i+1}], y_{l_{i+1}} = [1 \dots L_{i+1}] \end{cases}$$

L'échantillonnage est également modulable. Deux paramètres doivent être choisis.

1. **Le pas d'échantillonnage** représente le premier paramètre. Il est défini par sa taille et sa forme, carrée ou non. En décimant par un carré $n_i \times n_i$, l'image reste homothétique de l'image initiale (*cf.* figure 3.1). Il est également possible de décimer par des rectangles $n_i \times m_i$, et d'anamorphoser l'image par un rapport Nombre de lignes/Nombre de colonnes différent. Le nombre de lignes et de colonnes de l'étage i au suivant varie selon la formule

$$(K_i, L_i) \rightarrow (K_{i+1}, L_{i+1}) = (E(\frac{K_i}{n_i}) + \text{ceil}(\text{reste}_{K_i, n_i}), E(\frac{L_i}{m_i}) + \text{ceil}(\text{reste}_{L_i, m_i})), \quad (3.7)$$

$$(n_i, m_i) \in \mathbb{N}^2, n_i \geq 1, m_i \geq 1$$

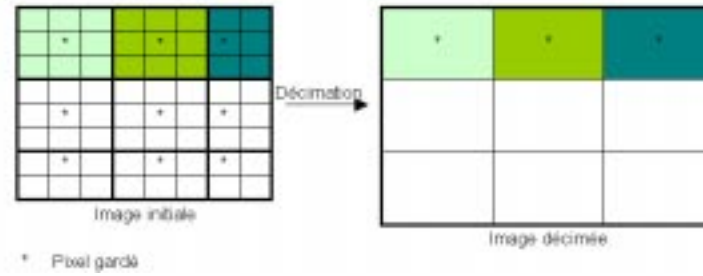
où reste_{K_i, n_i} (resp. reste_{L_i, m_i}) désigne le reste de la division de K_i par n_i (resp. de L_i par m_i); "ceil" arrondit reste_{K_i, n_i} et reste_{L_i, m_i} à l'entier le plus proche par excès et E désigne la partie entière. Les figures 3.1 et 3.2 illustrent le changement de taille des images.

2. La façon dont la décimation est réalisée est également paramétrable. Nous avons évalué trois méthodes différentes. Si le pas de décimation est $n_i \times m_i$:

- i) **L'échantillonnage simple** consiste à ne garder qu'un pixel sur les $n_i \times m_i$ définis par le pas de décimation. A l'étage i , cela correspond à calculer le pixel $(x_{k_{i+1}}, y_{l_{i+1}})$ de I_{i+1} par

$$I_{i+1}(x_{k_{i+1}}, y_{l_{i+1}}) = IF_i(x_{k_{i+1}} \times n_i - E(\frac{n_i}{2}), y_{l_{i+1}} \times m_i - E(\frac{m_i}{2})) \quad (3.8)$$

En choisissant cet algorithme, le pixel gardé est le pixel le plus central possible. Si n_i (resp. m_i) est impair, il sélectionne la ligne (resp. la colonne) milieu, sinon,

FIG. 3.1 – Exemple de décimation simple dans le cas d'un pas de décimation 2×2 .FIG. 3.2 – Exemple de décimation simple dans le cas d'un pas de décimation 3×3 .

il sélectionne la ligne (resp. la colonne) précédent le milieu (*cf.* figures 3.1 et 3.2).

Lorsque le nombre de lignes (resp. colonnes) de l'image n'est pas multiple de celui du pas de décimation, il faut décimer les dernières lignes (resp. colonnes) en prenant en compte un algorithme différent (*cf.* figure 3.2) : dans ces conditions, le nombre de lignes restant (resp. de colonnes) est inférieur à celui du pas de décimation, le pixel gardé par l'algorithme retenu est celui le plus en haut (resp. le plus à gauche). Comme le pas de décimation est petit - rarement plus de 3×3 - le nombre de lignes (resp. de colonnes) restant est faible et cet algorithme n'est pas pénalisant.

- ii) La **décimation moyenne** consiste à garder la moyenne des $n_i \times m_i$ pixels suivant la formule

$$I_{i+1}(x_{k_{i+1}}, y_{l_{i+1}}) = \sum_{p=0}^{(n_i-1)} \sum_{q=0}^{(m_i-1)} \frac{IF_i(x_{k_{i+1}} \times n_i - p, y_{l_{i+1}} \times m_i - q)}{n_i \times m_i} \quad (3.9)$$

- iii) La **décimation médiane** consiste à garder la médiane des $n_i \times m_i$ pixels.

$$I_{i+1}(x_{k_{i+1}}, y_{l_{i+1}}) = \text{Médiane}(IF_i(x_{k_{i+1}} \times n_i - p, y_{l_{i+1}} \times m_i - q)), \quad (3.10)$$

$p = 0 \dots (n_i - 1), q = 0 \dots (m_i - 1)$

I_{i+1} est l'image de base du niveau $i + 1$, c'est donc sur elle que la décomposition est effectuée à nouveau.

3.1.1.4 Etape 4 : calcul des détails de l'échantillonnage

a) Le sur-échantillonnage

Comme pour l'étape 2, la différence entre les images avant et après échantillonnage est calculée afin de ne pas perdre d'information au cours de la décomposition. Pour cela, l'image décimée est sur-échantillonnée du ratio inverse à celui appliqué lors de la décimation. La taille de l'image ainsi obtenue est la même que celle avant décimation. Les deux images-différences calculées sont appelées $D_{sup, dec, i}$ et $D_{inf, dec, i}$.

Mathématiquement, à l'étage i , l'image I_{i+1} est interpolée en $I_{i+1, int.}$ de telle sorte que $I_{i+1, int.}$ et IF_i aient la même taille :

$$I_{i+1}(x_{k_{i+1}}, y_{l_{i+1}}) \xrightarrow{sur-ech_i(n_i, m_i)} I_{i+1, int.}(x_{k_i}, y_{l_i}) \quad \begin{cases} x_{k_i} = [1...K_i], y_{l_i} = [1...L_i] \\ x_{k_{i+1}} = [1...K_{i+1}], y_{l_{i+1}} = [1...L_{i+1}] \end{cases} \quad (3.11)$$

La façon dont le sur-échantillonnage est réalisé peut être choisie sans contrainte, il suffit ensuite de réaliser la même interpolation lors de la reconstruction ; son pas est cependant identique à celui de l'échantillonnage de façon à faire coïncider la taille de $I_{i+1, int.}$ et de IF_i . Le type de sur-échantillonnage constitue par contre un paramètre. Nous avons testé les algorithmes de duplication, interpolation bilinéaire et interpolation bicubique.

b) Calculs des images différences

Ensuite, les images-différences sont calculées par

$$\begin{cases} D_{sup, dec, i}(x_{k_i}, y_{l_i}) = Sup(IF_i(x_{k_i}, y_{l_i}), I_{i+1, int.}(x_{k_i}, y_{l_i})) - I_{i+1, int.}(x_{k_i}, y_{l_i}) \\ D_{inf, dec, i}(x_{k_i}, y_{l_i}) = Sup(IF_i(x_{k_i}, y_{l_i}), I_{i+1, int.}(x_{k_i}, y_{l_i})) - IF_i(x_{k_i}, y_{l_i}) \\ \forall x_{k_i} = [1...K_i], \forall y_{l_i} = [1...L_i] \end{cases} \quad (3.12)$$

$D_{sup, dec, i}(x_{k_i}, y_{l_i})$ et $D_{inf, dec, i}(x_{k_i}, y_{l_i})$ ont des valeurs positives, entières et codées sur le même nombre NG de niveaux de gris que les images de l'étage i .

Comme précédemment, $D_{dec, i}(x_{k_i}, y_{l_i}) = D_{sup, dec, i}(x_{k_i}, y_{l_i}) - D_{inf, dec, i}(x_{k_i}, y_{l_i})$, et donc

$$D_{dec, i}(x_{k_i}, y_{l_i}) = IF_i(x_{k_i}, y_{l_i}) - I_{i+1, int.}(x_{k_i}, y_{l_i}) \quad (3.13)$$

$D_{dec, i}(x_{k_i}, y_{l_i})$ prend des valeurs positives ou négatives, entières et codées sur $[-NG, NG]$. Les deux catégories d'images-différences sont également ajoutées pour créer une image

$$D_{sup, i}(x_{k_i}, y_{l_i}) = D_{sup, filtre, i}(x_{k_i}, y_{l_i}) + D_{sup, dec, i}(x_{k_i}, y_{l_i})$$

et une image

$$D_{inf, i}(x_{k_i}, y_{l_i}) = D_{inf, filtre, i}(x_{k_i}, y_{l_i}) + D_{inf, dec, i}(x_{k_i}, y_{l_i})$$

globales à l'étage i , dont on se servira lors de la recombposition. $D_{sup, i}(x_{k_i}, y_{l_i})$ et $D_{inf, i}(x_{k_i}, y_{l_i})$ ont des valeurs comprises entre $[0, NG]$.

$$D_{sup,i} = Sup(IF_i, I_{i+1,int}) - I_{i+1,int} + Sup(I_i, IF_i) - IF_i$$

4 cas sont à envisager :

- $IF_i > I_i$ et $I_{i+1,int} > IF_i$

$$D_{sup} = 0$$

- $IF_i > I_i$ et $I_{i+1,int} < IF_i$

$$D_{sup} = IF_i - I_{i+1,int} < 255$$

- $IF_i < I_i$ et $I_{i+1,int} > IF_i$

$$D_{sup} = I_i - IF_i < 255$$

- $IF_i < I_i$ et $I_{i+1,int} < IF_i$

$$D_{sup,i} = IF_i - I_{i+1,int} + I_i - IF_i = I_i - I_{i+1,int} < 255$$

La même démonstration peut être conduite pour $D_{inf,i}$.

On définit également

$$\begin{aligned} D_i(x_{k_i}, y_{l_i}) &= D_{sup,i}(x_{k_i}, y_{l_i}) - D_{inf,i}(x_{k_i}, y_{l_i}) \\ &= D_{filtre,i}(x_{k_i}, y_{l_i}) + D_{deci,i}(x_{k_i}, y_{l_i}) \\ &= I_i(x_{k_i}, y_{l_i}) - I_{i+1,int}(x_{k_i}, y_{l_i}) \end{aligned} \quad (3.14)$$

D_i représente l'information perdue lors de la transformation $T_{dec,i}$ et ses valeurs sont comprises entre $[-NG, NG]$.

L'étage i de la décomposition se schématise selon le principe montré en figure 3.3.

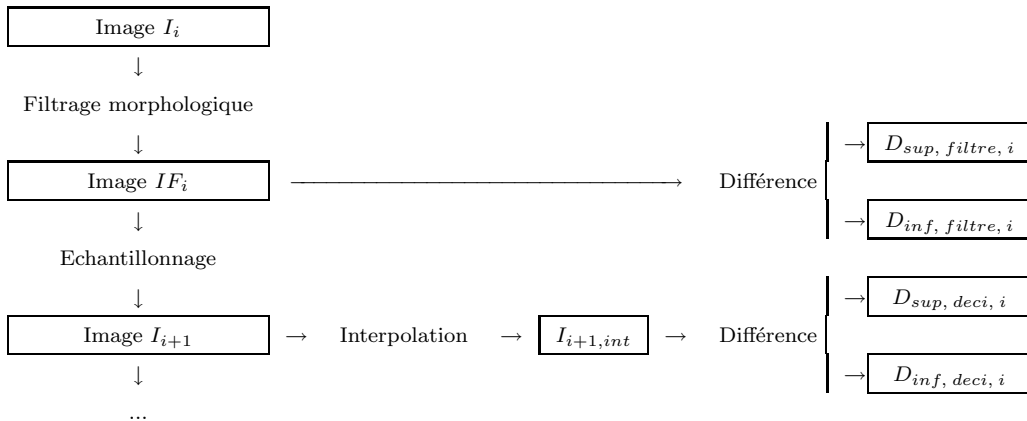


FIG. 3.3 – Schéma de décomposition de la pyramide morphologique.

3.1.2 Exemple de décomposition morphologique

L'image choisie comme exemple de pyramide est typique par sa présence à différentes échelles de détails qui disparaissent après filtrage.

Les figures 3.4 et 3.5 représentent un exemple de décomposition. La pyramide choisie est une pyramide à deux étages. L'élément structurant est un carré 5×5 , d'origine centrée et symétrique. Le filtre est une ouverture suivie d'une fermeture. La décimation est une décimation moyenne 2×2 .

La figure 3.4 représente les « basses fréquences » de la décomposition. Entre l'image 3.4a) et l'image 3.4b), le filtrage se remarque d'une part par l'impression générale classique de flou sur la seconde image et d'autre part par la disparition d'éléments fins. Par exemple, le cordon que tient le cameraman ainsi que le centre du trépied de l'appareil photographique ont disparu sur l'image filtrée.

Sur les images 3.5, les détails sont représentés en négatif : les comptes numériques élevés sont représentés en noir. Le détail $D_{sup, filtre, 0}$ du filtre représente les lignes blanches qui ont disparu entre l'image d'origine et l'image filtrée. Ce sont en fait les zones où le filtrage a abaissé la valeur du compte numérique. Au contraire, sur les détails $D_{inf, filtre, 0}$ on voit les zones noires sur l'image d'origine qui avaient disparu sur l'image filtrée. On retrouve bien sur l'image $D_{sup, filtre, 0}$ le cordon du photographe et les barres transversales du trépied. Au contraire, sur $D_{inf, filtre, 0}$, on retrouve notamment le sommet sombre du building en fond d'image.

3.2 La recomposition

3.2.1 Principe

La transformation $T_{rec,i}$ se décompose en deux opérations (cf. figure 3.6) :

1. Sur-échantillonnage spatial sur la trame de l'image.
2. Ajout de la différence calculée lors des étapes 2 et 4 de la décomposition.

3.2.1.1 Etape 1

Le sur-échantillonnage est le même que celui réalisé à l'étape 4 de la décimation. Cette étape a pour but d'augmenter le nombre de pixels de l'image pour le rendre compatible avec la portée spatiale des détails.

Ainsi, le sur-échantillonnage à l'étage $i + 1$ de l'image IR_{i+1} s'exprime par :

$$\begin{cases} IR_{i+1}(x_{k_{i+1}}, y_{l_{i+1}}) \xrightarrow{sur-ech_i(n_i, m_i)} (IRF_i)(x_{k_i}, y_{l_i}), i \in [0, N - 2] \\ I_N(x_{k_N}, y_{l_N}) \xrightarrow{sur-ech_{N-1}(n_{N-1}, m_{N-1})} (IRF_{N-1})(x_{k_{N-1}}, y_{l_{N-1}}) \end{cases} \quad (3.15)$$

IRF_i représente l'Image Reconstituée Filtrée et IR_i l'Image Reconstituée au niveau i .

3.2.1.2 Etape 2

Cette étape consiste à ajouter à l'image sur-échantillonnée toute l'information perdue lors de la décomposition. Par définition, cette information est contenue dans les $D_{sup, i}$ et les $D_{inf, i}$.

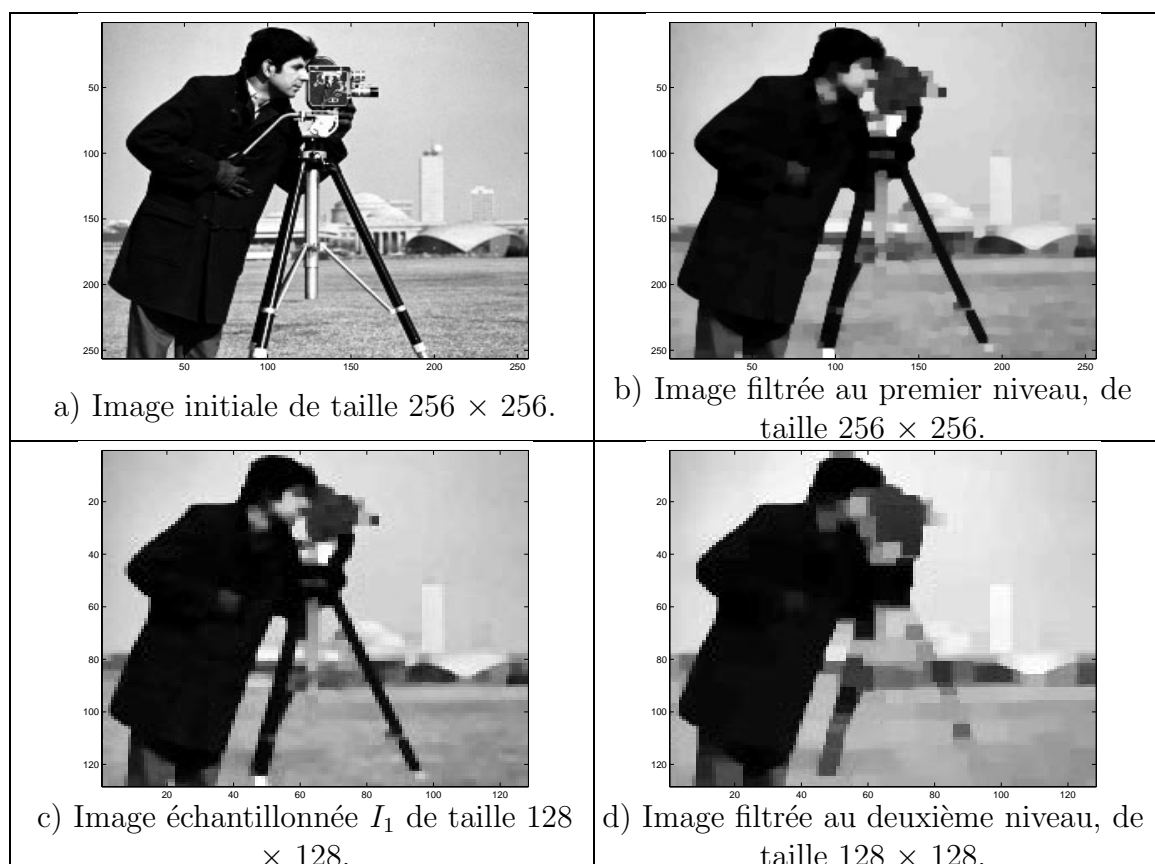


FIG. 3.4 – Images « basses fréquences » obtenues après deux étages de pyramide sur l'image du cameraman. a) Image Initiale I_0 , b) Image Filtrée IF_0 , c) Image Décimée I_1 , d) Image Filtrée IF_1 .

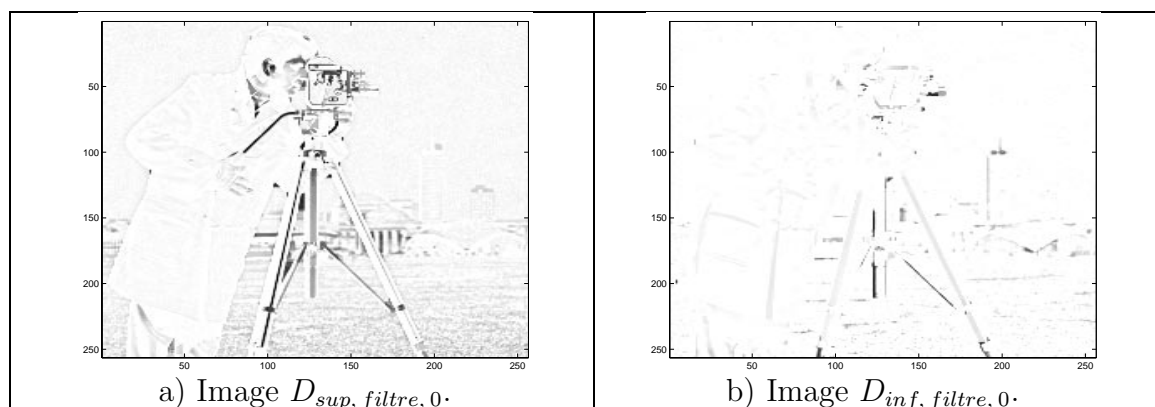


FIG. 3.5 – Images « haute fréquence » obtenues au premier étage de la pyramide sur l'exemple de l'image du cameraman : a) $D_{sup, filtre, 0}$, b) $D_{inf, filtre, 0}$.

Mathématiquement, cela consiste à ajouter les détails calculés lors de la décimation (cf. équation (3.14)).

$$IR_i(x_{k_i}, y_{l_i}) = IRF_i(x_{k_i}, y_{l_i}) + D_i(x_{k_i}, y_{l_i}) \quad (3.16)$$

IR_i représente l'Image Reconstituée au niveau i . C'est à cette image que l'on va reappliquer $T_{rec,i}$.

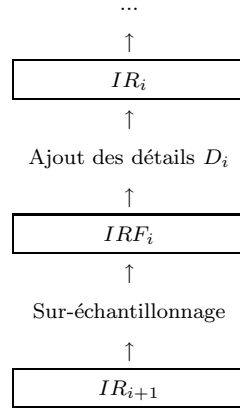


FIG. 3.6 – Schéma de reconstitution de la pyramide morphologique.

La reconstruction commence, pour une pyramide à N étages par le traitement de la dernière image filtrée lors de la décomposition, soit I_N . La première étape correspond à la deuxième ligne de l'équation (3.15). Cette équation est identique à l'équation (3.11) appliquée à $i = N - 1$. Donc $IRF_{N-1}(x_{k_{N-1}}, y_{l_{N-1}}) = I_{N,int.}(x_{k_{N-1}}, y_{l_{N-1}})$.

Combiner les équations (3.14) et (3.16) permet le calcul de IR_{N-1} ,

$$\begin{aligned} IR_{N-1}(x_{k_{N-1}}, y_{l_{N-1}}) &= IRF_{N-1}(x_{k_{N-1}}, y_{l_{N-1}}) + D_{N-1}(x_{k_{N-1}}, y_{l_{N-1}}) \\ &= I_{N,int.}(x_{k_{N-1}}, y_{l_{N-1}}) + I_{N-1} - I_{N,int.}(x_{k_{N-1}}, y_{l_{N-1}}) \\ &= I_{N-1}(x_{k_{N-1}}, y_{l_{N-1}}) \end{aligned} \quad (3.17)$$

En procédant ainsi par récurrence lors de la reconstruction, on a à l'étage i :

$$IR_i(x_{k_i}, y_{l_i}) = I_i(x_{k_i}, y_{l_i}) \quad (3.18)$$

En particulier, pour l'étage 0 et donc pour la reconstruction finale,

$$\boxed{IR_0(x_{k_0}, y_{l_0}) = I_0(x_{k_0}, y_{l_0})}$$

Cette pyramide est donc exacte et permet de décomposer et recomposer à volonté les images.

3.2.2 Exemple de reconstitution morphologique

Voilà la reconstruction (exacte) qui suit la décomposition du paragraphe 3.1.2. La figure 3.7 représente les images reconstituées IR_1 et IR_0 . On a donc $3.7a) = 3.4c)$ et $3.7b) = 3.4a)$.

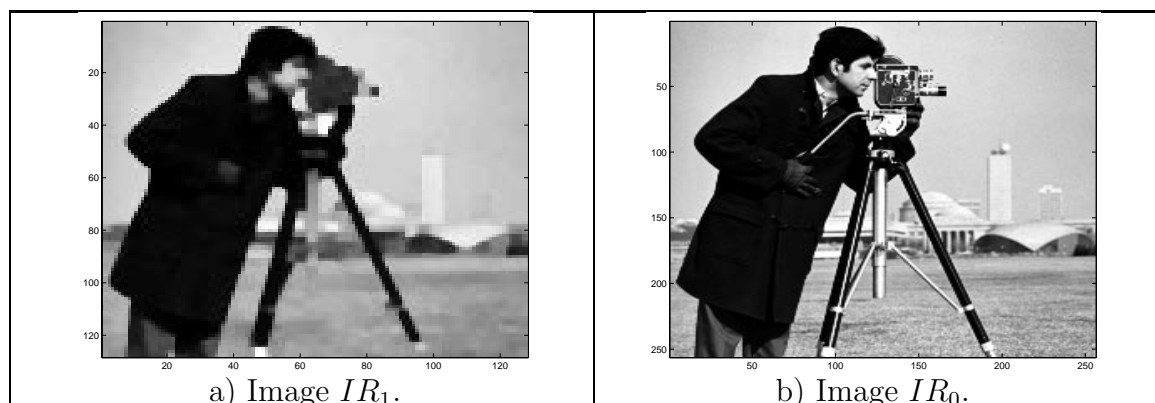


FIG. 3.7 – Etapes de reconstruction appliquées à l'image du cameraman : a) IR_1 , b) IR_0 .

3.3 Comparaison entre filtrage linéaire et non linéaire

Il est toujours tentant de vouloir comparer les propriétés des filtres linéaires et non linéaires. Toutefois, cet exercice est toujours délicat car il fait appel à des principes et des propriétés qui sont différents. Dans cette partie, on choisit simplement de comparer les résultats obtenus par la pyramide morphologique à ceux qui seraient obtenus par application d'autres filtres. Cette étude est divisée en deux parties, la première recense les différences relevées dans la littérature tandis que la deuxième montre des exemples précis de filtrage sur des mires.

3.3.1 Caractéristiques comparées des filtrages linéaire et non linéaire

Selon Maragos [73] les avantages du filtrage linéaire sont sa facilité mathématique et sa relation étroite avec l'analyse de Fourier. Il peut de plus être utilisé aux premières étapes d'un système de modélisation de la vision primaire.

Par contre, les filtres linéaires décalent et rendent flous des caractéristiques importantes comme les bordures, ou les contours [34], [77], [78] et impliquent un paramètre d'échelle σ qui n'est pas directement lié à la définition de l'échelle. Les versions multi-échelles du signal ne correspondent pas à une représentation compacte des formes, sauf dans le cas le plus simple de la différence entre les différentes versions filtrées aux échelles successives. Enfin, les temps de calculs sont en général plus importants lors du filtrage linéaire [77], [78], [29]. Haralick *et al.* [43] préfèrent un filtrage non linéaire pour l'extraction de formes.

3.3.2 Spécificité de la morphologie mathématique

Les filtres linéaires de pas dyadique correspondent aux modèles admis par la voie achromatique de la vision primaire [75]. Sur un autre plan, les filtres non linéaires morphologiques sont adaptés à la détection d'éléments emboîtés de tailles différentes et contrastés

avec leur environnement. Cela constitue une différence majeure entre les deux approches et qui les rend toutes deux intéressantes.

Pour Serra, les intérêts principaux de la morphologie mathématique sont l'irréversibilité des opérations [124] et la suppression des effets d'aliasing [127]. De plus, l'élément structurant permet de choisir ce que l'on extrait, contrairement aux filtres linéaires, notamment gaussiens [18]. Ce sont ces caractéristiques qui donnent leur intérêt aux différences de fonctions.

3.3.3 Comparaison entre filtres passe-bas

Les différences de comportement entre les filtres passe-bas linéaires et non linéaires sont facilement mis en évidence sur des signaux simples à une dimension. Les graphes de la figure 3.8 comparent les effets du filtre linéaire « moyenne » et des filtres non linéaires « médiane », « ouverture » et « fermeture » sur une fonction composée de fonctions portes de tailles grossissantes. Le signal d'origine, représenté par des « o » rouges sur la figure 3.8, est construit par la juxtaposition du schéma de base $(2n + 1)$ pixel(s) de valeur 1 et $(2n + 1)$ pixel(s) de valeur 0, où n croît de 0 à 4. Le signal ainsi créé $I(x_k)$ est donc égal à 101110001111100000... pour $x_k = [1...4]$.

Les filtres morphologiques ouverture et fermeture ne modifient que la portion du signal de plage inférieure à la taille de l'élément structurant, c'est-à-dire 5 dans l'exemple. L'ouverture supprime les pics de longueur 1 et 3, la fermeture les creux de même taille. Par contre, le filtre moyenne modifie tous les contours, y compris ceux des plages de taille supérieure à la longueur du filtre. Le filtre médiane ne modifie pas ce signal, composé uniquement de transition de type « marche ». La propriété de conservation des contours des filtres non linéaires est confirmée sur cet exemple simple.

La première image choisie est une combinaison de sinusoides (*cf.* image figure 3.9). Les filtres sont calculés sur des voisinages 3×3 (*cf.* figure 3.10).

Cette image, composée exclusivement d'éléments fins met en évidence l'extraction des éléments clairs par l'ouverture, sombres pour la fermeture et de tous les éléments pour le filtre moyenne. Le filtre médiane est celui préservant le plus l'image de départ. Les images suivantes (*cf.* figure 3.11) comparent les détails $D_{sup, filtre}$ et les $D_{inf, filtre}$ obtenus à partir de chacun des quatre filtres étudiés. L'image de départ (*cf.* figure 3.9) est régulièrement répartie entre les détails de réflectances claires sur un fond plus sombre et ceux de réflectances sombres sur un fond plus clair. Ce résultat se retrouve facilement sur la figure 3.11. $D_{sup, filtre}$ et $D_{inf, filtre}$ obtenus par le filtre moyenne sont de contenu similaire, la même remarque peut être faite pour le filtre médian. Au contraire, $D_{inf, filtre}$ obtenu par l'ouverture et $D_{sup, filtre}$ obtenu par la fermeture sont nuls par définition. La symétrie de l'image de détails se retrouve dans la ressemblance entre $D_{sup, filtre}$ fourni par l'ouverture et $D_{inf, filtre}$ fourni par la fermeture. Seuls ces filtres morphologiques peuvent discriminer l'extraction des éléments en fonction de leur réflectance par rapport à leur environnement.

L'image 3.12 montre la meilleure conservation des contours avec des filtres non linéaires qu'avec le filtre moyenne. Filtrer avec un élément structurant 7×7 et une ouverture, ou une fermeture, ou un filtre médiane ne modifie pas l'image, tandis qu'un filtre moyenne

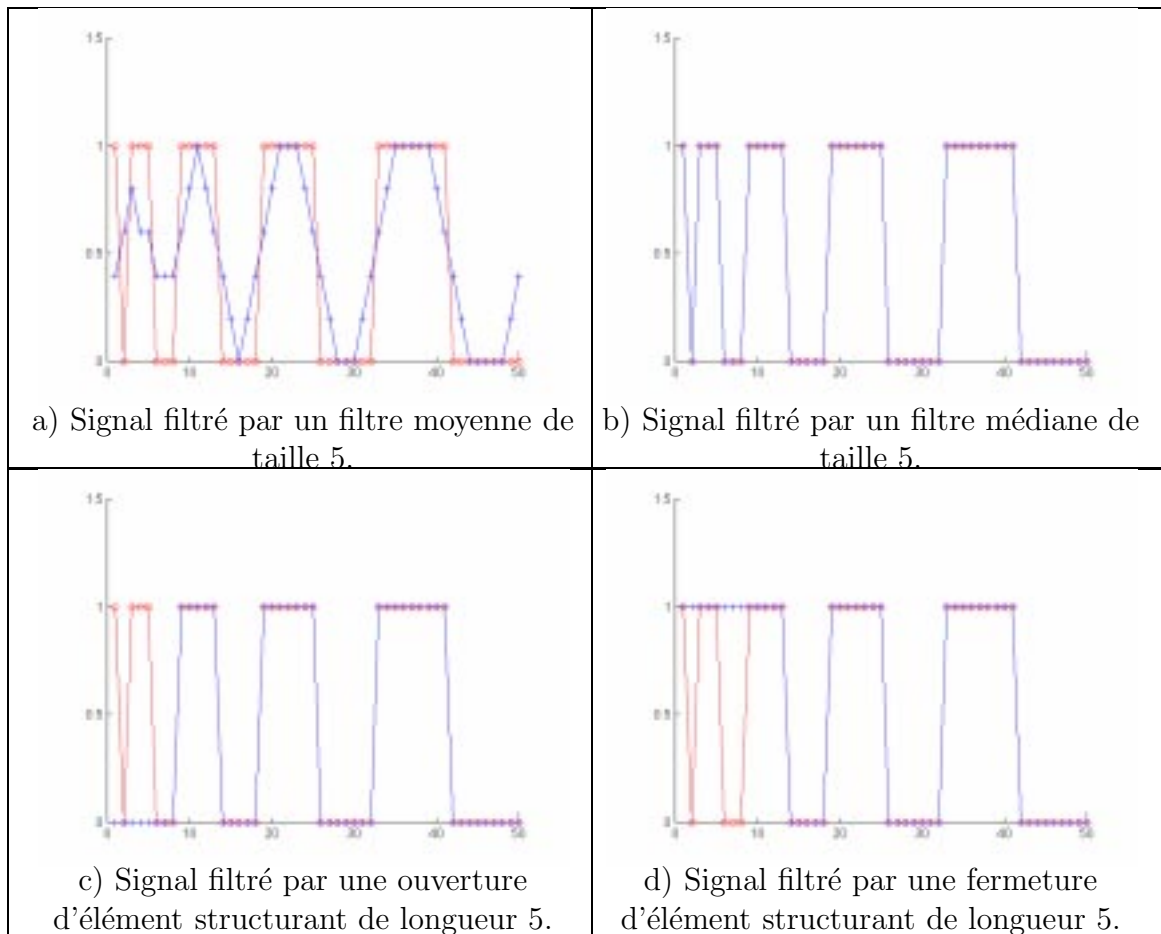


FIG. 3.8 – Le signal d'origine est signalé par des « o » rouges, les signaux filtrés par des « + » bleus.

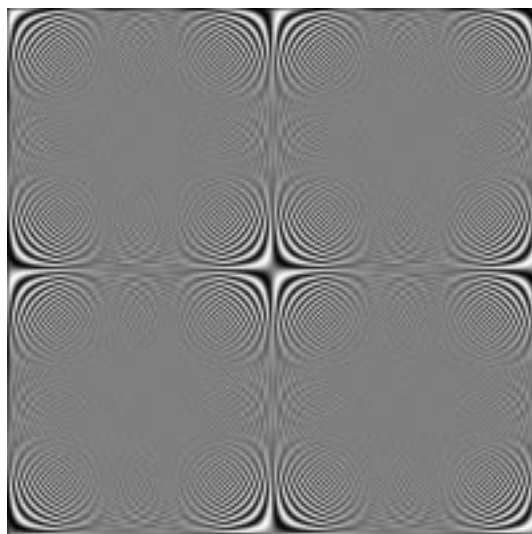


FIG. 3.9 – Image initiale, combinaisons de sinusoides.

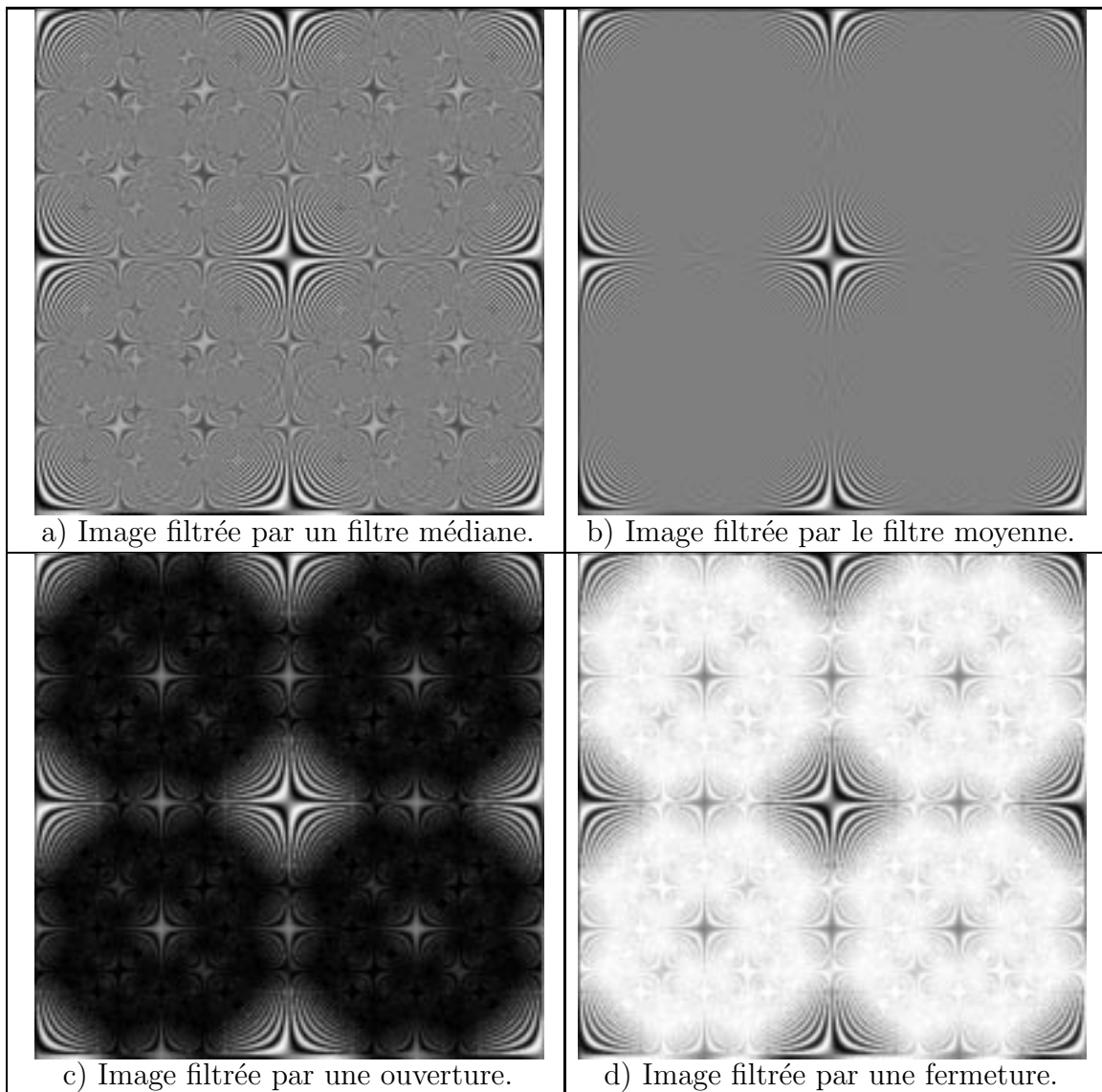
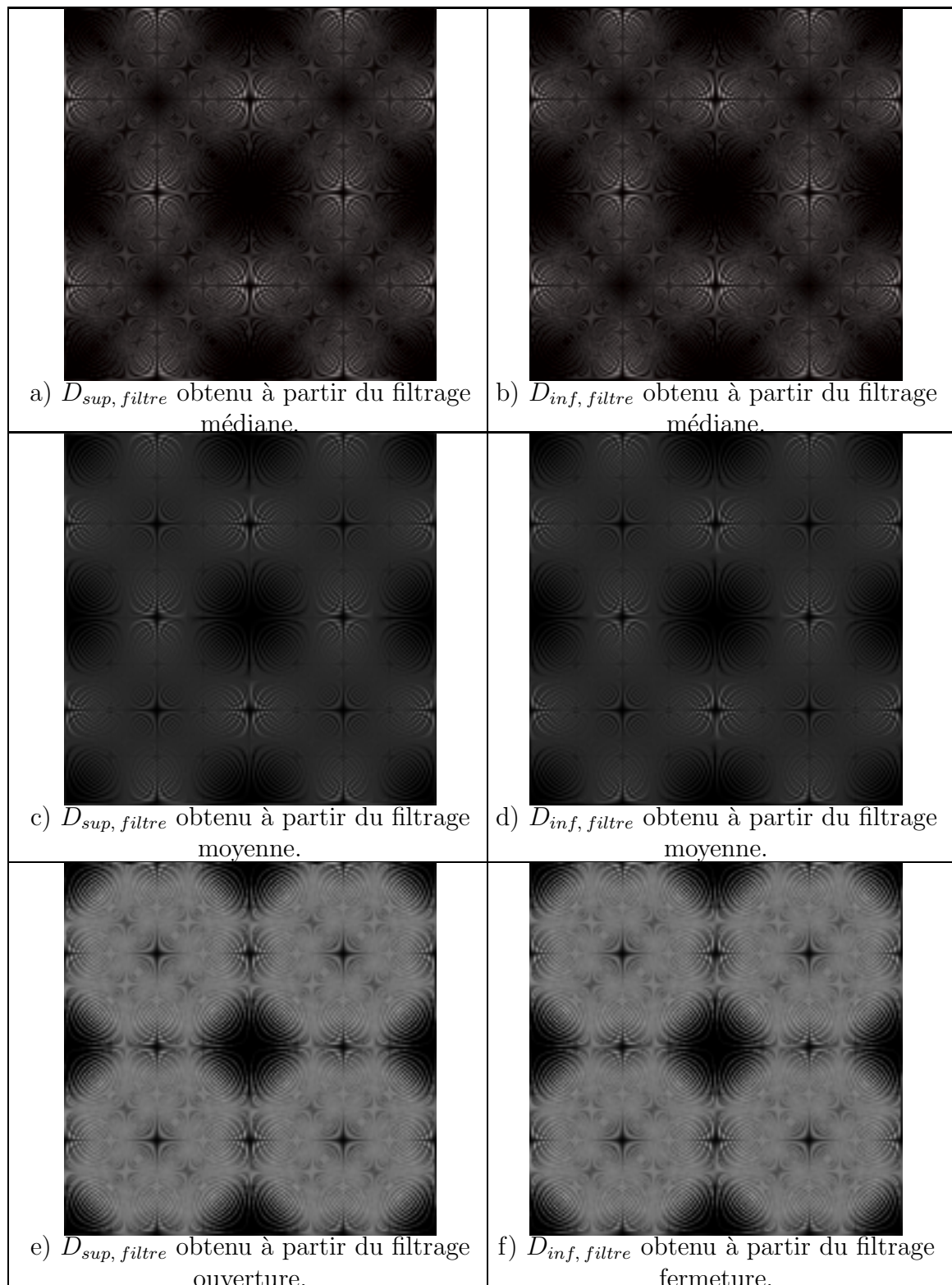


FIG. 3.10 – Images filtrées par différents filtres (médiane, moyenne, ouverture et fermeture).

FIG. 3.11 – Comparaison des détails $D_{sup, filtre}$ et $D_{inf, filtre}$ obtenus par différents filtrages.

de noyau 7×7 l'altère. L'impression de flou visible sur l'image 3.12 en témoigne.

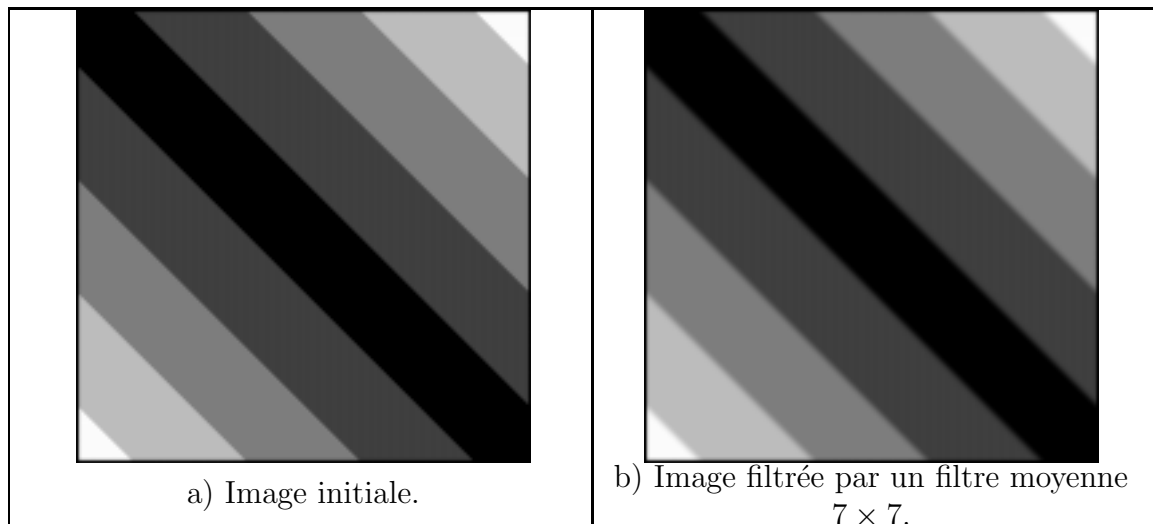


FIG. 3.12 – Exemple de la non conservation des contours dans le cas du filtre moyenne.

3.4 Synthèse : les familles de pyramides

Les caractéristiques de la pyramide morphologique mettent en évidence les différents paramètres qui en gouvernent les propriétés, il s'agit de :

- l'élément structurant : taille, forme, origine,
- le filtre,
- l'échantillonnage : pas, algorithme,
- le sur-échantillonnage : algorithme.

De façon générale, 7 paramètres sont donc à étudier, regroupés en 4 catégories distinctes. Ainsi, en fonction des paramètres, des familles de pyramides vont se dégager avec les groupes de propriétés correspondants.

– L'élément structurant

Il est défini par sa taille, sa forme, son origine.

Si l'on souhaite avoir l'isotropie spatiale du traitement, l'origine doit être au centre de l'élément structurant, et la forme doit se rapprocher au mieux d'une boule projetée dans le plan. Pour les petits éléments structurants, cela revient à prendre une forme carrée. La forme et l'origine définissent ou non l'isotropie du filtrage.

Par ailleurs, si l'on veut éliminer beaucoup de détails, on prendra un élément structurant de grande taille, ce qui entraîne que le filtrage réalisé sera plus important.

– Le filtre

C'est sa nature qui doit être définie. En fonction de sa composition d'opérations élémentaires, plus ou moins de détails sont séparés lors du filtrage.

De plus, lorsque le filtre n'est pas symétrique entre l'ouverture et la fermeture,

par exemple une ouverture suivi d'une fermeture, celui-ci favorise les détails élevés localement par rapport aux valeurs du voisinage ou vice-versa.

Le choix du filtre favorise $D_{sup, filtre}$ ou $D_{inf, filtre}$ s'il n'est pas symétrique. De plus, il filtre plus ou moins selon sa nature.

– **L'échantillonnage**

L'échantillonnage est un élément important de la décomposition. Il définit le rapport de résolution entre deux niveaux successifs de la pyramide. C'est donc lui qui caractérise la taille des objets à un niveau donné.

– **Le sur-échantillonnage**

Son importance est très variable en fonction de l'application. La façon dont on mène l'interpolation conduit d'une part à des résultats sensiblement différents et d'autre part à des temps de calculs très variables.

La combinaison de valeurs différentes pour l'échantillonnage et le sur-échantillonnage permet d'établir des rapports fractionnaires qui sont très intéressants en fusion d'images de résolutions quelconques.

3.5 Propriétés de la pyramide morphologique

3.5.1 Filtrage et dualité dans la pyramide

La dualité des filtres employés dans la pyramide morphologique est très importante : c'est son existence qui confère leur signification physique aux détails extraits dans la pyramide par les différences de fonctions ordonnées dans le treillis des images (*cf.* paragraphe 1.3.2).

Dans les paragraphes suivants, on montre ainsi les conditions qui conduisent à l'extraction des détails D_{sup} correspondant à des variations locales de luminances fortes par rapport à un environnement de luminances plus faibles (resp. pour les D_{inf} avec des luminances faibles sur des environnements plus réfléchissants, *cf.* paragraphe 3.1.1).

Soit I une image et soit deux filtres duaux F_1 et F_2 .

Filtrons morphologiquement de deux façons différentes, la première en appliquant le filtre F_1 sur l'image I , et la deuxième en appliquant le filtre F_2 sur l'image complémentaire de I , I^C .

Par définition,

$$F_1(I) = [F_2(I^C)]^C$$

Définition 11 *Soit le treillis de toutes les images. Soit m la borne supérieure sur le treillis.*

Soit I un élément du treillis. Le complémentaire est défini par symétrie par rapport au plan médian $\frac{m}{2}$.

Le complémentaire de I , I^C s'exprime par :

$$I^C = m - I$$

Les détails issus du filtre se calculent par (3.4). Pour la première méthode :

$$\begin{cases} D_{sup, F_1 \text{ sur } I} &= \sup(I, F_1(I)) - F_1(I) \\ D_{inf, F_1 \text{ sur } I} &= \sup(I, F_1(I)) - I \end{cases} \quad (3.19)$$

Pour la deuxième,

$$\begin{cases} D_{sup, F_2 \text{ sur } I^C} &= \sup(I^C, F_2(I^C)) - F_2(I^C) \\ D_{inf, F_2 \text{ sur } I^C} &= \sup(I^C, F_2(I^C)) - I^C \end{cases} \quad (3.20)$$

On a $[F_1(I)]^C = F_2(I^C)$ donc :

$$\Leftrightarrow \begin{cases} D_{sup, F_2 \text{ sur } I^C} &= \sup(I^C, [F_1(I)]^C) - [F_1(I)]^C \\ D_{inf, F_2 \text{ sur } I^C} &= \sup(I^C, [F_1(I)]^C) - I^C \end{cases} \quad (3.21)$$

$$\Leftrightarrow \begin{cases} D_{sup, F_2 \text{ sur } I^C} &= [\inf(I, F_1(I))]^C - [F_1(I)]^C \\ D_{inf, F_2 \text{ sur } I^C} &= [\inf(I, F_1(I))]^C - I^C \end{cases} \quad (3.22)$$

$$\Leftrightarrow \begin{cases} D_{sup, F_2 \text{ sur } I^C} &= [\inf(I, F_1(I)) - F_1(I)]^C \\ D_{inf, F_2 \text{ sur } I^C} &= [\inf(I, F_1(I)) - I]^C \end{cases} \quad (3.23)$$

$$(3.24)$$

Comparons maintenant les équations (3.19) et (3.24) :

Cas 1 : cas des pixels (i, j) tels que $I(i, j) > F_1(I(i, j))$

$$\begin{cases} D_{sup, F_1 \text{ sur } I} &= I(i, j) - F_1(I(i, j)) \\ D_{inf, F_1 \text{ sur } I} &= 0 \end{cases} \quad (3.25)$$

$$\begin{cases} D_{sup, F_2 \text{ sur } I^C} &= 0 \\ D_{inf, F_2 \text{ sur } I^C} &= [F_1(I(i, j)) - I(i, j)]^C \end{cases} \quad (3.26)$$

\Leftrightarrow sur le treillis des images :

$$\begin{cases} D_{sup, F_2 \text{ sur } I^C} &= 0 \\ D_{inf, F_2 \text{ sur } I^C} &= m - F_1(I(i, j)) - (m - I(i, j)) \end{cases} \quad (3.27)$$

$$\Leftrightarrow \begin{cases} D_{sup, F_2 \text{ sur } I^C} &= 0 \\ D_{inf, F_2 \text{ sur } I^C} &= -F_1(I(i, j)) + I(i, j) \end{cases} \quad (3.28)$$

$$(3.29)$$

Cas 2 : cas des pixels (i, j) tels que $I(i, j) \leq F_1(I(i, j))$

$$\begin{cases} D_{sup, F_1 \text{ sur } I} = & 0 \\ D_{inf, F_1 \text{ sur } I} = & F_1(I(i, j)) - I(i, j) \end{cases} \quad (3.30)$$

$$\begin{cases} D_{sup, F_2 \text{ sur } I^C} = & [I(i, j) - F_1(I(i, j))]^C \\ D_{inf, F_2 \text{ sur } I^C} = & 0 \end{cases} \quad (3.31)$$

\Leftrightarrow sur le treillis des images :

$$\begin{cases} D_{sup, F_2 \text{ sur } I^C} = & m - I(i, j) - (m - F_1(I(i, j))) \\ D_{inf, F_2 \text{ sur } I^C} = & 0 \end{cases} \quad (3.32)$$

et donc,

$$\Leftrightarrow \begin{cases} D_{sup, F_2 \text{ sur } I^C} = & -I(i, j) + F_1(I(i, j)) \\ D_{inf, F_2 \text{ sur } I^C} = & 0 \end{cases} \quad (3.33)$$

Donc, par équivalence entre les relations (3.25) et (3.29) d'une part et (3.30) et (3.33) d'autre part, on a sur toute l'image :

$$\boxed{\begin{cases} D_{sup, F_1 \text{ sur } I} = & D_{inf, F_2 \text{ sur } I^C} \\ D_{inf, F_1 \text{ sur } I} = & D_{sup, F_2 \text{ sur } I^C} \end{cases}} \quad (3.34)$$

Les couples de filtres suivants sont duaux si l'on utilise un même élément structurant lors de leurs applications : γ/φ , $\gamma\varphi/\varphi\gamma$, $\gamma\varphi\gamma/\varphi\gamma\varphi$.

Le filtre $\frac{1}{2}(\gamma + \varphi)$ demande une analyse particulière :

$$\begin{aligned} \frac{1}{2}(\gamma + \varphi)(I^C) &= \frac{1}{2}\gamma(I^C) + \frac{1}{2}\varphi(I^C) \\ &= \frac{1}{2}(\varphi(I))^C + \frac{1}{2}(\gamma(I))^C \\ &= \frac{1}{2}((\varphi + \gamma)(I))^C \\ &= \frac{1}{2}((\gamma + \varphi)(I))^C \end{aligned} \quad (3.35)$$

$\frac{1}{2}(\gamma + \varphi)$ est autodual. Dans ce cas,

$$\begin{cases} D_{sup, \frac{\gamma+\varphi}{2} \text{ sur } I} = & D_{inf, \frac{\gamma+\varphi}{2} \text{ sur } I^C} \\ D_{inf, \frac{\gamma+\varphi}{2} \text{ sur } I} = & D_{sup, \frac{\gamma+\varphi}{2} \text{ sur } I^C} \end{cases} \quad (3.36)$$

On notera que la loi de composition de ce filtre introduisant une combinaison linéaire, il ne s'agit plus d'un filtre morphologique au sens des propriétés de croissance et d'idempotence. Toutefois, il a été étudié dans ce travail compte tenu des résultats produits dans

différentes expérimentations numériques.

Appliquons ce résultat à la pyramide morphologique, en discret. Nous réalisons donc deux pyramides, l'une avec le filtre F_1 et l'image I_0 , l'autre avec le filtre F_2 et l'image $J_0 = I_0^C$. Au niveau 0, nous avons effectivement

$$\begin{cases} D_{sup,0,F_1 \text{ sur } I_0} = D_{inf,0,F_2 \text{ sur } J_0} \\ D_{inf,0,F_1 \text{ sur } I_0} = D_{sup,0,F_2 \text{ sur } J_0} \end{cases} \quad (3.37)$$

Ainsi, les images filtrées issues des deux pyramides s'écrivent

$$\begin{cases} IF_{0,F_1 \text{ sur } I_0} = I_0 - D_{sup,0,F_1 \text{ sur } I_0} + D_{inf,0,F_1 \text{ sur } I_0} \\ JF_{0,F_2 \text{ sur } J_0} = J_0 - D_{sup,0,F_2 \text{ sur } J_0} + D_{inf,0,F_2 \text{ sur } J_0} \end{cases} \quad (3.38)$$

Le complémentaire de $IF_{0,F_1 \text{ sur } I_0}$ s'exprime par

$$\begin{aligned} [IF_{0,F_1 \text{ sur } I_0}]^C &= [I_0 - D_{sup,0,F_1 \text{ sur } I_0} + D_{inf,0,F_1 \text{ sur } I_0}]^C \\ &= J_0 - [D_{sup,0,F_1 \text{ sur } I_0}]^C + [D_{inf,0,F_1 \text{ sur } I_0}]^C \\ &= J_0 - (m - D_{sup,0,F_1 \text{ sur } I_0}) + (m - D_{inf,0,F_1 \text{ sur } I_0}) \\ &= J_0 + D_{sup,0,F_1 \text{ sur } I_0} - D_{inf,0,F_1 \text{ sur } I_0} \\ &= J_0 - D_{sup,0,F_2 \text{ sur } J_0} + D_{inf,0,F_2 \text{ sur } J_0} \\ &= JF_{0,F_2 \text{ sur } J_0} \end{aligned} \quad (3.39)$$

L'étape suivante de la décomposition est l'échantillonnage. $IF_{0,F_1 \text{ sur } I_0}$ est transformé en $I_{1, \text{ech sur } IF_0}$ et $JF_{0,F_2 \text{ sur } J_0}$ en $J_{1, \text{ech sur } JF_0}$.

Il faut donc maintenant contrôler les conditions que doit remplir l'échantillonnage afin de conserver la propriété de complémentarité des images issues de la pyramide. Le résultat dépend évidemment du schéma d'échantillonnage retenu. Dans le cas où l'échantillonnage est une simple réduction appliquée sur les deux images, le résultat ne modifie pas les comptes numériques des données. Le résultat est donc conservé, et l'on a $I_{1, \text{ech sur } IF_0} = [J_{1, \text{ech sur } JF_0}]^C$. Dans ce cas, tout au long de la pyramide, les propriétés suivantes sont vraies

$$\begin{cases} D_{sup,i,F_i \text{ sur } I_{i-1}} = D_{inf,i, \text{ dual de } F_i \text{ sur } J_{i-1}} \\ D_{inf,i,F_i \text{ sur } I_{i-1}} = D_{sup,i, \text{ dual de } F_i \text{ sur } J_{i-1}} \end{cases} \quad (3.40)$$

$$IF_{i,F_i \text{ sur } I_i} = [JF_{i, \text{ dual de } F_i \text{ sur } J_i}]^C \quad (3.41)$$

et

$$I_{i+1, \text{ech sur } IF_i} = [J_{i+1, \text{ech sur } JF_i}]^C \quad (3.42)$$

En numérique, le cas du filtre $\frac{1}{2}(\gamma + \varphi)$ est particulier. En effet, durant le traitement algorithmique, les comptes numériques des images sont traités comme des entiers non signés. L'opération de division incluse dans l'application de ce filtre conduit à des valeurs non entières sur certains pixels. L'algorithme développé arrondit ces valeurs à l'entier par défaut le plus proche. Par définition, l'arrondi par défaut le plus proche d'un nombre non entier n'est jamais le même sur la valeur d'un pixel ou sur sa valeur complémentaire. L'emploi de ce filtre conserve donc en théorie les propriétés des images énoncées ci-dessus, mais en pratique, les arrondis numériques les rendent fausses.

L'expression du filtre moyenne est donnée par l'équation (3.9). Dans le cas de la première pyramide sur l'image I_0 , elle s'écrit

$$I_{i+1,ech\ sur\ IF_i}(x_{k_{i+1}}, y_{l_{i+1}}) = \sum_{p=0}^{(n_i-1)} \sum_{q=0}^{(m_i-1)} \frac{IF_{i, F_i\ sur\ I_i}(x_{k_{i+1}} \times n_i - p, y_{l_{i+1}} \times m_i - q)}{n_i \times m_i} \quad (3.43)$$

tandis que pour la seconde pyramide, effectuée sur J_0 , elle s'écrit

$$\begin{aligned} J_{i+1,ech\ sur\ JF_i}(x_{k_{i+1}}, y_{l_{i+1}}) &= \sum_{p=0}^{(n_i-1)} \sum_{q=0}^{(m_i-1)} \frac{JF_{i, dual\ de\ F_i\ sur\ J_i}(x_{k_{i+1}} \times n_i - p, y_{l_{i+1}} \times m_i - q)}{n_i \times m_i} \\ &= \sum_{p=0}^{(n_i-1)} \sum_{q=0}^{(m_i-1)} \frac{m - IF_{i, F_i\ sur\ I_i}(x_{k_{i+1}} \times n_i - p, y_{l_{i+1}} \times m_i - q)}{n_i \times m_i} \\ &= m - \sum_{p=0}^{(n_i-1)} \sum_{q=0}^{(m_i-1)} \frac{IF_{i, F_i\ sur\ I_i}(x_{k_{i+1}} \times n_i - p, y_{l_{i+1}} \times m_i - q)}{n_i \times m_i} \\ &= [I_{i+1,ech\ sur\ IF_i}(x_{k_{i+1}}, y_{l_{i+1}})]^C \end{aligned} \quad (3.44)$$

L'expression de l'échantillonnage médian est donnée par l'équation (3.10). Dans le cas d'un nombre de termes impairs, la médiane conserve les propriétés énoncées dans ce paragraphe. Dans le cas d'un nombre de termes pairs, l'algorithme retenu calcule la moyenne des deux termes centraux.

Ainsi, en théorie, échantillonnages moyen et médian conservent les propriétés de complémentarité entre les deux pyramides. En pratique, les opérations de division conduisent à des arrondis annulant ces propriétés.

3.5.2 Propriétés de la reconstruction

Jusqu'à présent, l'algorithme de reconstruction était le suivant :

- interpolation de l'image basse résolution i ,
- ajout des détails, pour former l'image basse résolution $i + 1$,
- interpolation de l'image basse résolution $i + 1$,
- ...

Soit Int_k la transformation « sur-échantillonnage d'une image du niveau $k + 1$ au niveau k ». Avec l'algorithme précédent, la reconstruction au niveau j s'écrit :

$$IR_j = D_j + Int_j(D_{j+1} + Int_{j+1}(D_{j+2} + \dots + Int_{N-2}(D_{N-1} + Int_{N-1}(I_N)))) \quad (3.45)$$

Si l'algorithme d'interpolation est tel que

$$Int_k(X + Y) = Int_k(X) + Int_k(Y), \quad (3.46)$$

alors l'équation précédente devient :

$$IR_j = D_j + Int_j(D_{j+1}) + Int_j(Int_{j+1}(D_{j+2})) + \dots + Int_j(\dots Int_{N-2}(D_{N-1})) + Int_j(\dots Int_{N-1}(I_N))$$

Interpoler pas à pas une image ou l'interpoler directement à la haute résolution conduit alors au même résultat, on obtient donc :

$$IR_j = D_j + \widetilde{D}_{j+1}^j + \widetilde{D}_{j+2}^j + \dots + \widetilde{D}_{N-1}^j + \widetilde{I}_N^j \quad (3.47)$$

où \widetilde{Image}_r^j correspond à l'interpolation de l'image à la résolution r jusqu'à la résolution j , c'est-à-dire

$$\widetilde{Image}_{j-x}^j = Int_j[\dots Int_{j-x+2}[Int_{j-x+1}(Image_{j-x})]]$$

L'algorithme de reconstruction défini par (3.47) est plus long car l'interpolation est réalisée séparément sur toutes les images qui sont sommées ensuite, au lieu de réaliser l'interpolation sur la somme.

Cette propriété est néanmoins intéressante et montre l'importance du schéma d'interpolation retenu sur les propriétés de la pyramide. Nous pouvons voir quels types d'interpolation satisfont le critère de linéarité énoncé dans l'équation (3.46). Comme nous l'avons déjà vu, la reconstruction est exacte quel que soit le type d'interpolation réalisé, l'addition exprimée dans l'équation (3.45) est donc réellement une addition, sans saturation possible.

Il est maintenant nécessaire d'étudier quels algorithmes d'interpolation satisfont l'équation (3.46). L'algorithme de duplication est le schéma le plus simple. Si p varie de 0 à $n_i - 1$ et q de 0 à $m_i - 1$, l'ensemble des pixels $(x_{k_i}, y_{l_i}) = (n_i \times x_{k_{i+1}} - p, m_i \times y_{l_{i+1}} - q)$ vont recevoir la valeur prise par la fonction au pixel $(x_{k_{i+1}}, y_{l_{i+1}})$. Ce schéma conserve la propriété énoncée ci-dessus (cf. figure 3.13).

Un ré-échantillonnage bilinéaire réalise des calculs de moyennes pondérées sur les valeurs des pixels les plus proches. Le principe de l'interpolation bilinéaire est indiqué sur la figure 3.14. La propriété (3.46) est également conservée; il est cependant nécessaire de noter que ce résultat n'est acquis d'une part qu'en dehors des zones de bordures de l'image et et d'autre part que si aucun arrondi n'est fait lors des divisions aux différentes étapes de calculs. En pratique, lors d'une interpolation bilinéaire, la propriété (3.46) ne sera pas conservée et l'équation (3.47) sera donc fautive, même si le résultat ne différera que d'arrondis et donc d'un nombre très faible de compte numérique.

Des algorithmes d'interpolation testés, seul, l'algorithme d'interpolation bicubique ne satisfait pas cette condition.

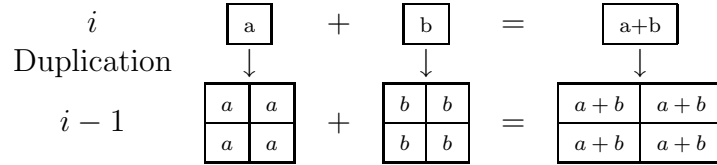


FIG. 3.13 – L’algorithme de duplication vérifie la propriété (3.46). La première ligne représente la somme des éléments au niveau i , la deuxième ligne au niveau $i - 1$. La première colonne représente le pixel courant de l’ensemble X au niveau i , puis sa duplication au niveau $i - 1$, la deuxième colonne sur l’ensemble Y . Enfin la troisième colonne montre que sommer la première ligne puis dupliquer le résultat, ou dupliquer séparément les éléments de X et Y puis sommer le résultat est équivalent.

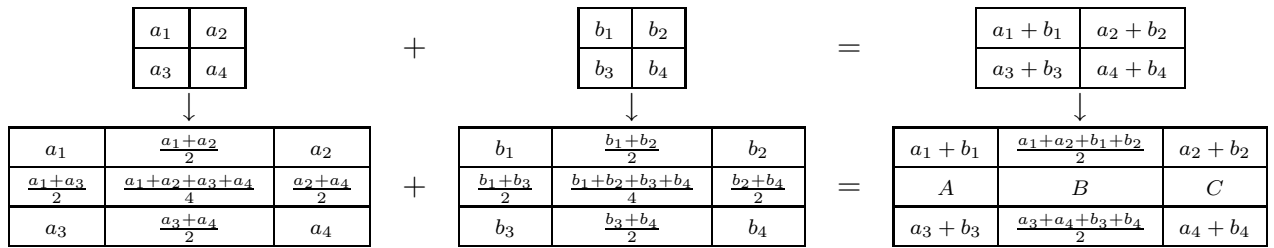


FIG. 3.14 – L’algorithme d’interpolation bilinéaire vérifie la propriété (3.46). La première ligne représente la somme des éléments au niveau i , la deuxième ligne au niveau $i - 1$. La première colonne représente le pixel courant de l’ensemble X au niveau i , puis son interpolation bilinéaire au niveau $i - 1$, la deuxième colonne sur l’ensemble Y . Enfin la troisième colonne montre que sommer la première ligne puis interpoler bilinéairement le résultat, ou interpoler bilinéairement séparément les éléments de X et Y puis sommer le résultat est équivalent. Sur ce schéma, $A = \frac{a_1+a_3+b_1+b_3}{2}$, $B = \frac{a_1+a_2+a_3+a_4+b_1+b_2+b_3+b_4}{4}$ et $C = \frac{a_2+a_4+b_2+b_4}{2}$.

3.5.3 Navigation dans les différents niveaux de la pyramide

Les propriétés de la pyramide morphologique - décomposition et transformée inverse exacte - ouvrent une perspective intéressante à la navigation dans une scène lorsque l'on veut accéder à l'information complète dans une partie donnée de l'image.

Une première localisation grossière sera évidemment facilitée par la perception de l'ensemble d'une région à un niveau de faible résolution spatiale dans la pyramide.

Ainsi, la navigation entre les échelles va permettre, à partir d'une échelle quelconque donnée de suivre la reconstruction d'un pixel ou d'un ensemble de pixels d'intérêt dans les niveaux inférieurs ou supérieurs de la pyramide. En effet, un pixel à une résolution donnée se décompose en un ensemble de pixel à la résolution immédiatement supérieure, qui eux-mêmes se transforment en un ensemble plus important de pixels, et ainsi de suite. Afin de suivre un pixel, il faut donc connaître les paramètres suivants : étage de départ dans la pyramide (a), coordonnées du pixel que l'on veut suivre (x_{k_a}, y_{l_a}), pas de décimation (n_i, m_i) aux étages i , ainsi que le nombre d'étages que l'on veut remonter (n). A partir de tous ces paramètres, on montre qu'au $n^{\text{ième}}$ étage au dessus du pixel sélectionné, celui-ci est devenu le rectangle s'étendant des coordonnées

$$\left(\prod_{i=a-1}^{a-n} n_i \right) x_{k_a} - \left(\prod_{i=a-1}^{a-n} n_i - 1 \right)$$

à

$$\left(\prod_{i=a-1}^{a-n} n_i \right) x_{k_a}$$

en ligne et

$$\left(\prod_{i=a-1}^{a-n} m_i \right) y_{l_a} - \left(\prod_{i=a-1}^{a-n} m_i - 1 \right)$$

à

$$\left(\prod_{i=a-1}^{a-n} m_i \right) y_{l_a}$$

en colonne.

Pour l'étage inférieur, le pixel sélectionné (x_{k_a}, y_{l_a}) devient le pixel de coordonnées (X, Y) où $X = \frac{x_{k_a}}{n_a}$ et $Y = \frac{y_{l_a}}{m_a}$ en arrondissant le résultat de la division à l'entier immédiatement supérieur. La sélection de zones de l'image à toutes les échelles se fait ainsi simplement de façon automatique.

Ces formules permettent alors de conserver la sélection d'une zone géographique unique aux différentes résolutions pour parvenir à son information spatiale complète.

Si on conçoit que (dans une application opérationnelle) la désignation des rectangles d'intérêt est effectuée par un utilisateur, ce principe ouvre une perspective intéressante de navigation interactive dans des pavés emboîtés.

3.6 Conclusion

Ce chapitre décrit le principe de la pyramide morphologique et en étudie les différents paramétrages possibles. Ensuite, les effets des filtres morphologiques sont comparés à ceux d'autres filtres. Enfin, la dernière partie étudie ses propriétés.

Les caractéristiques importantes de la pyramide morphologique peuvent se résumer par les points suivants :

- le choix physique *a priori* de la paramétrisation en fonction des propriétés des objets recherchés dans la décomposition,
- le respect de l'hypothèse *a priori* de non-équidistribution de la variabilité locale des signaux décomposés,
- l'existence d'une transformée inverse exacte dont on verra l'importance dans la fusion spatio-temporelle d'images.

Cela induit une paramétrisation de la pyramide qui dépend de la nature de l'information recherchée. Cette approche multi-échelle de la perception d'une surface permet donc une conduite contrôlée du processus de décomposition d'une scène en se déplaçant dans l'espace des résolutions possibles pour sa description.

Troisième partie

**ANALYSE MULTI-RESOLUTION
DES SURFACES TERRESTRES**

Chapitre 4

Analyse et extraction d'attributs par la pyramide morphologique

4.1 Introduction

De nombreuses études ont montré que les surfaces terrestres sont organisées en un ensemble d'objets emboîtés de taille et formes différentes [3]. Une approche performante de la description et de la caractérisation de scènes consiste donc à extraire et séparer ces différents éléments. Pour cela, l'organisation spatiale d'un paysage se prête bien à une décomposition pyramidale de données simulant une perception multi-échelle [75]. De plus, les propriétés de la morphologie mathématique permettent de caractériser des éléments par leur taille, leur forme et leur réflectance par rapport à l'environnement.

En conséquence, comme le montrent les propriétés de la pyramide décrite au chapitre précédent, les filtres non linéaires combinées à l'analyse multi-échelle permettent :

1. la séparation à chaque échelle des composantes du paysage représentant les détails,
2. la répartition de ces détails en fonction de la variation de réflectance locale (soit une augmentation soit une diminution de cette réflectance) sans hypothèse implicite d'équi-distribution du filtrage linéaire.

Ainsi, l'utilisation de la pyramide morphologique permet de définir un protocole unique permettant la coopération entre extraction et analyse d'attributs dans une image quelconque. Tout d'abord, le paramétrage de la pyramide, et notamment les choix du filtre, de l'élément structurant, de l'échantillonnage et du nombre d'étages, est conditionné par les éventuelles connaissances a priori sur la scène. Ensuite, le processus de décomposition est réalisé et les détails de chaque niveau sont sauvegardés. Les éléments emboîtés de la scène sont alors séparés et contenus dans les images de détails $D_{sup, filtre, i}$ et $D_{inf, filtre, i}$ et dans les images successives des différents niveaux de pyramide. Evidemment, des traitements supplémentaires sont dans certains cas nécessaires pour extraire précisément les éléments caractéristiques des attributs de la scène.

L'analyse et l'extraction d'attributs ne fait appel qu'à la partie décomposition de la pyramide morphologique. Les images intéressantes dans ce cadre sont celles aux résolutions successives (I_i) et celles des détails provenant du filtre ($D_{sup, filtre, i}$ et $D_{inf, filtre, i}$). Les premières servent à estimer les objets susceptibles d'être extraits à l'échelle i et éventuellement à affiner le choix des paramètres ; les deuxièmes permettent la caractérisation des attributs de surface des objets. Dans ce chapitre, différentes applications étayeront les intérêts de chacun des deux types d'images.

4.2 Choix des paramètres

Lorsque les caractéristiques des attributs à extraire sont définis, nous pouvons étudier l'influence théorique des paramètres de la pyramide et proposer des choix de paramétrages. La première constatation concerne les images exploitées lors de l'analyse d'un paysage. Ce sont, comme nous l'avons déjà évoqué I_i , $D_{sup, filtre, i}$ et $D_{inf, filtre, i}$ à chaque niveau. Les images de détails $D_{sup, deci, i}$ et $D_{inf, deci, i}$ ne sont ici d'aucun intérêt, pas plus que les images reconstruites aux différents niveaux. De ce fait, il n'est utile, de calculer ni $D_{sup, deci, i}$ ni $D_{inf, deci, i}$, ni même d'effectuer la partie recomposition de la pyramide morphologique. Supprimer ces deux points élimine tout sur-échantillonnage du processus, et donc rend inutile le choix d'un algorithme d'interpolation. Il ne faut donc étudier que l'influence des paramètres filtre, élément structurant et échantillonnage.

Les paramètres filtre et élément structurant définissent les caractéristiques des éléments de paysage extraits. Notamment, le filtrage peut favoriser des objets de réflectances plus ou moins élevées que le fond voisin dans la scène, tandis que les caractéristiques de l'élément structurant permettent de définir la taille, la forme et l'éventuelle orientation des objets.

Le pas d'échantillonnage définit le rapport de résolution entre deux étages de la pyramide morphologique et donc la résolution de détection des objets à extraire. Le type d'échantillonnage réalisé n'intervient pas directement dans le résultat. Cependant, l'équation (3.7) montre que l'obtention de IF_i , et donc ses caractéristiques sont conditionnées par le type d'échantillonnage. Enfin, le nombre d'étages permet, en relation avec le pas d'échantillonnage, de définir la taille maximale des objets extraits au dernier niveau de la décomposition.

Ainsi, lors de l'analyse d'une surface donnée, une connaissance à priori des objets à extraire permet de conduire le paramétrage *à priori*. Dans le cas contraire, les paramètres de la pyramide morphologique seront choisis de façon la plus neutre possible, afin de garder le plus d'éléments potentiellement significatifs possibles. Si l'on ne connaît pas à priori les résolutions d'intérêt de l'étude, le pas de décimation sera choisi le plus petit possible, c'est-à-dire dyadique. De plus, si les caractéristiques des éléments sont inconnus, le filtre sera pris de façon symétrique entre l'ouverture et la fermeture ($\frac{\gamma_{B_i}(I_i(x_{k_i}, y_{l_i})) + \varphi_{B_i}(I_i(x_{k_i}, y_{l_i}))}{2}$) à tous les étages afin de ne favoriser ni les réflectances fortes, ni les faibles et l'élément structurant sera carré, centré et de taille immédiatement supérieure à celle de la décimation (donc 3×3) afin de n'extraire que les éléments strictement liés à la résolution de l'étage.

FIG. 4.1 – Image de Lena, de taille 256×256 .

Enfin, il est utile de rappeler en conclusion que chacun de ces paramètres se définit séparément à chaque étage i de la pyramide.

4.3 Influence des paramètres

L'objectif de cette partie est de comparer les éléments extraits lors des différents paramètres. Il sera intéressant de voir entre autres l'évolution de $D_{inf, filtre, i}$ et $D_{sup, filtre, i}$ en fonction des étages.

4.3.1 Evolution des images de détails en fonction de la résolution

Dans cette partie, la pyramide choisie est conforme aux recommandations du paragraphe précédent (échantillonnage dyadique par moyennage, élément structurant 3×3 et filtre $\frac{\gamma \pm \varphi}{2}$). Cinq étages de décomposition sont calculés sur l'image de Lena (*cf.* figure 4.1).

Les figures référencées 4.2 représentent les images $D_{inf, filtre, i}$ avec $i = 0, \dots, 3$. Ces quatre images montrent l'évolution des détails lorsque l'on descend dans la pyramide morphologique. D'une part, la taille des images est de plus en plus petite et part de 256×256 pour $D_{inf, filtre, 0}$, puis 128×128 , 64×64 et enfin 32×32 pour $D_{inf, filtre, 3}$. L'effet de pixellisation est facilement discernable sur les deux derniers niveaux. D'autre part, les éléments extraits sont de plus en plus gros. L'extraction de l'image de détails de tailles grossissantes est clairement illustrée.

Les figures 4.3 comparent les résultats obtenus au premier étage par $D_{inf, filtre, 0}$ (*cf.* figure 4.3a)) et $D_{sup, filtre, 0}$ (*cf.* figure 4.3b)). Les détails extraits sur une image ne se retrouvent pas dans l'autre. Les éléments sombres de l'image 4.1 se retrouvent dans $D_{inf, filtre, 0}$, les éléments clairs dans $D_{sup, filtre, 0}$.

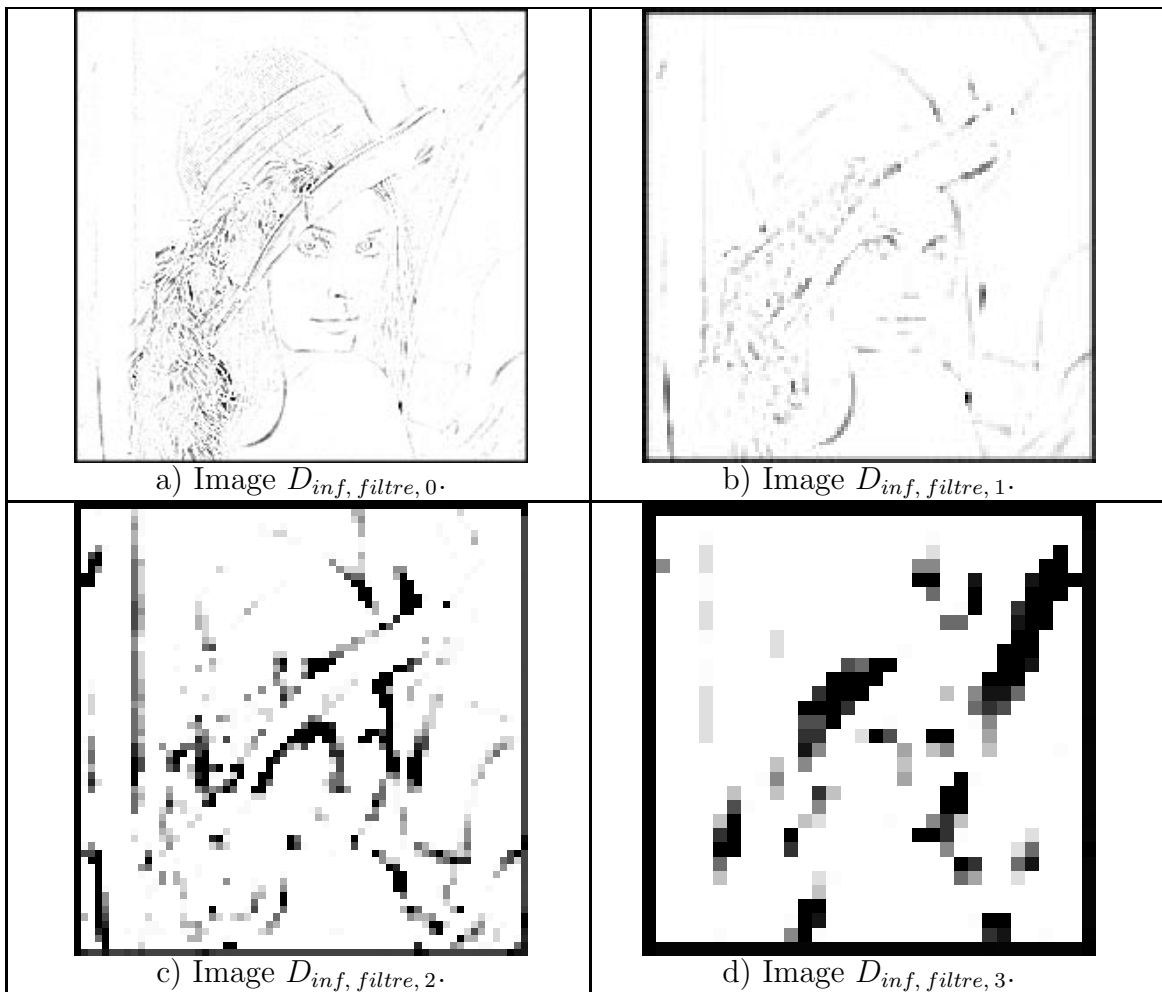


FIG. 4.2 – Comparaison des détails $D_{inf, filtre, i}$ de quatre niveaux obtenus sur l'image de Léna.



FIG. 4.3 – Comparaison des détails $D_{inf, filtre, 0}$ et $D_{sup, filtre, 0}$ obtenus à partir de l'image de Léna.

4.3.2 Influence de la taille de l'élément structurant

Comparons maintenant les effets d'un élément structurant 3×3 et un élément structurant 9×9 , au premier niveau de la pyramide. Les résultats montrés sur la figure 4.4, comparent les deux images filtrées IF_0 et les deux images de détails $D_{sup, filtre, 0}$ obtenues avec un filtre identique, $\frac{\gamma \pm \varphi}{2}$. Des différences similaires existent si l'on compare les deux images de détails $D_{inf, filtre, 0}$. L'image IF_0 obtenue avec un élément structurant 9×9 est plus floue que celle obtenue avec un élément structurant 3×3 , il est logique de retrouver plus d'informations dans l'image $D_{sup, filtre, 0}$ obtenue à partir de l'élément structurant 9×9 que celle obtenue à partir de celui 3×3 . Plus la taille des voisinages considérés pour le calcul du filtre est grand, plus celle des éléments extraits va l'être.

4.3.3 Influence du pas d'échantillonnage

Ce paragraphe a pour but de comparer les images obtenues avec deux pas d'échantillonnage, le premier d'un rapport 2, le deuxième d'un rapport quatre. Dans le premier cas, les images ont des tailles successives aux différents niveaux de 256×256 , 128×128 , puis 64×64 , tandis que dans le deuxième cas, les images ont une taille de 64×64 dès le deuxième niveau. Il est donc judicieux de comparer les éléments extraits par les deux premiers étages par échantillonnage dyadique et par le premier étage lors d'un échantillonnage de rapport 4.

– Cas d'un échantillonnage de rapport 2

L'élément structurant de taille juste supérieure est celui de rayon 1. Le changement de taille de 256×256 à 64×64 se fait en deux étages et fournit 4 images de détails.

– Cas d'un échantillonnage de rapport 4

L'élément structurant de taille juste supérieure est celui de rayon 2. Le changement

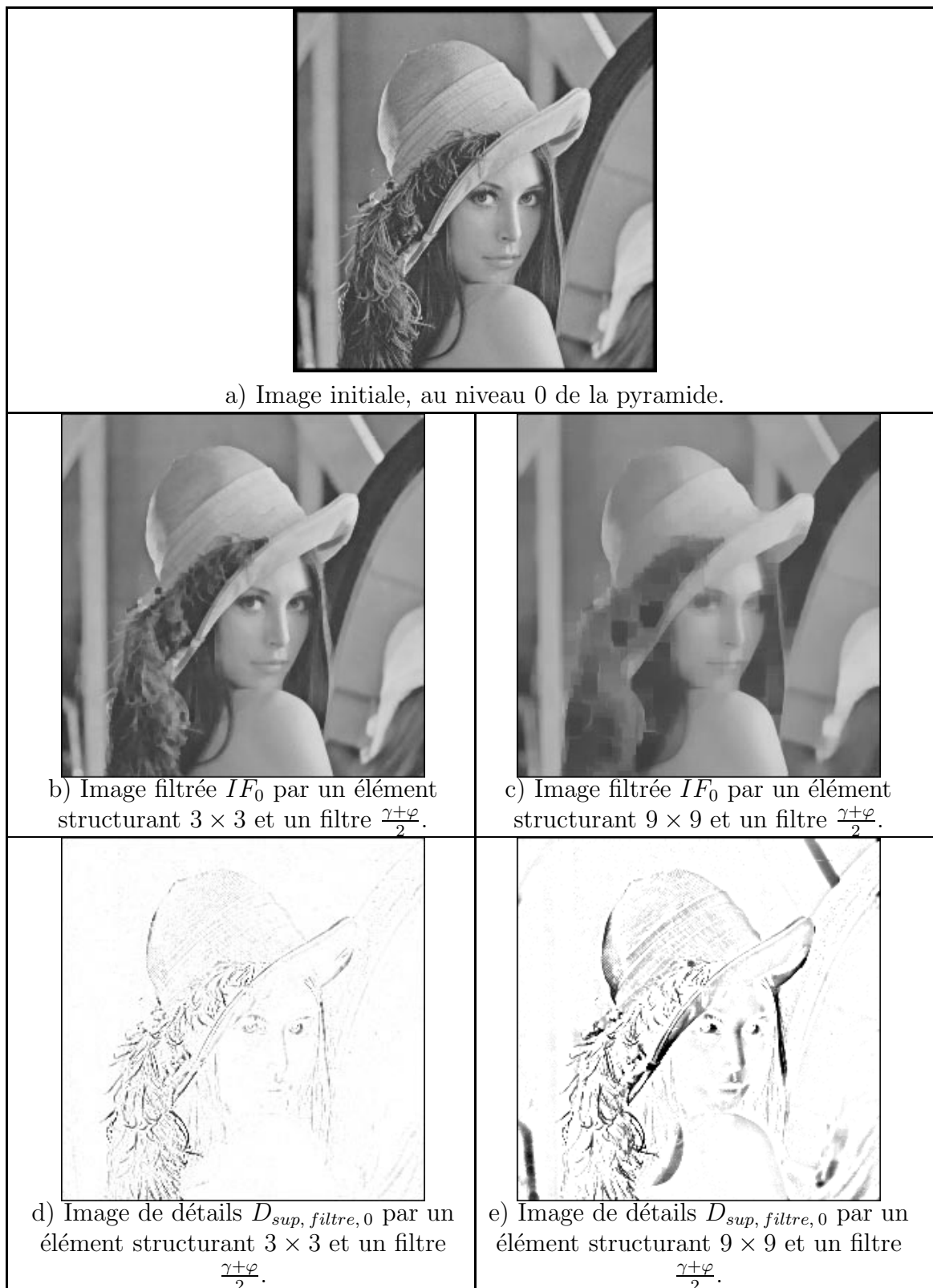


FIG. 4.4 – Comparaison du premier étage de la pyramide obtenu à partir de deux éléments structurant (3×3 à gauche et 9×9 à droite).

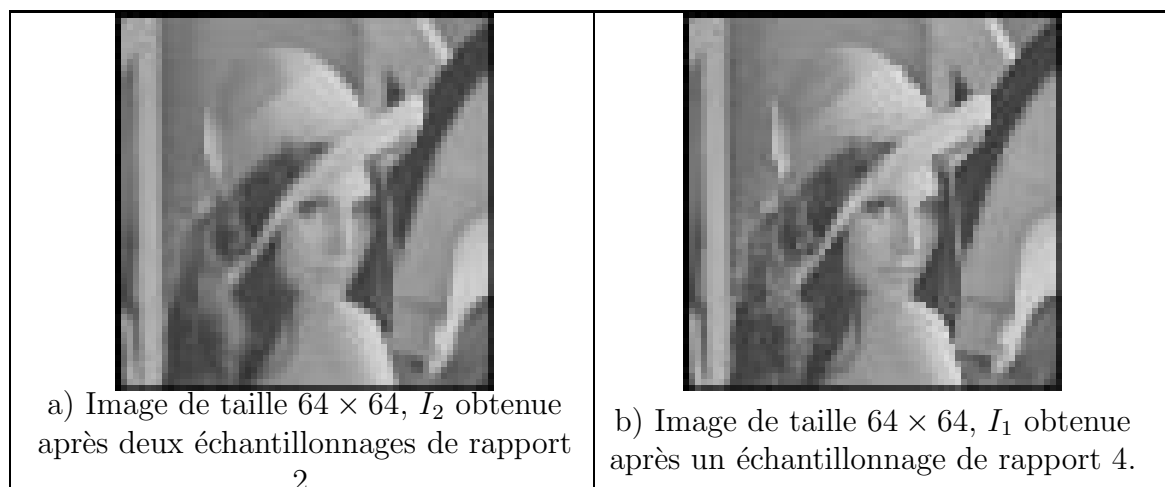


FIG. 4.5 – Comparaison des images I_1 obtenue avec un échantillonnage de rapport 4 et un élément structurant de rayon 2, et I_2 avec un échantillonnage de rapport 2 et un élément structurant de rayon 1. Les autres paramètres de la pyramide sont communs (échantillonnage par moyennage et filtre $\frac{\gamma+\varphi}{2}$).

de taille de 256×256 à 64×64 s'effectue en un seul étage et fournit 2 images de détails. L'image I_1 est légèrement décalée par rapport à I_0 , car le pas d'échantillonnage est pair, et son origine n'est donc pas centrée. Ce phénomène ne se voit pas dans le cas de deux échantillonnages dyadiques (*cf.* figure 4.5).

Par ailleurs, l'échantillonnage de rapport 2 permet obtenir quatre images de détails dans le cas d'un échantillonnage dyadique, au lieu d'en obtenir deux dans le cas d'un échantillonnage de rapport 4 (*cf.* figure 4.6). Cette séparation plus fine des éléments de la scène dans le cas d'un rapport dyadique est un avantage pour les applications nécessitant l'extraction d'objets de petites tailles.

4.3.4 Influence de l'algorithme d'échantillonnage

Un échantillonnage « simple » consistant à ne garder qu'une ligne sur n , conduit à un décalage géométrique de l'image plus marqué que dans le cas d'un échantillonnage calculée par moyennage. Dans le cas des petits pas d'échantillonnage que nous préconisons, un échantillonnage de type moyenne ou médiane n'offre pas de différence significative.

De même, sur cet exemple, l'influence des différents filtres n'est pas facile à percevoir entre deux images mises côte à côte et ne sera donc pas illustré.

4.4 Extraction d'objets en fonction de leur taille

Cette section présente les résultats publiés dans les actes du colloque IEEE/IGARSS'01 (International Geoscience and Remote Sensing Symposium) sous la référence [58]. L'exemple

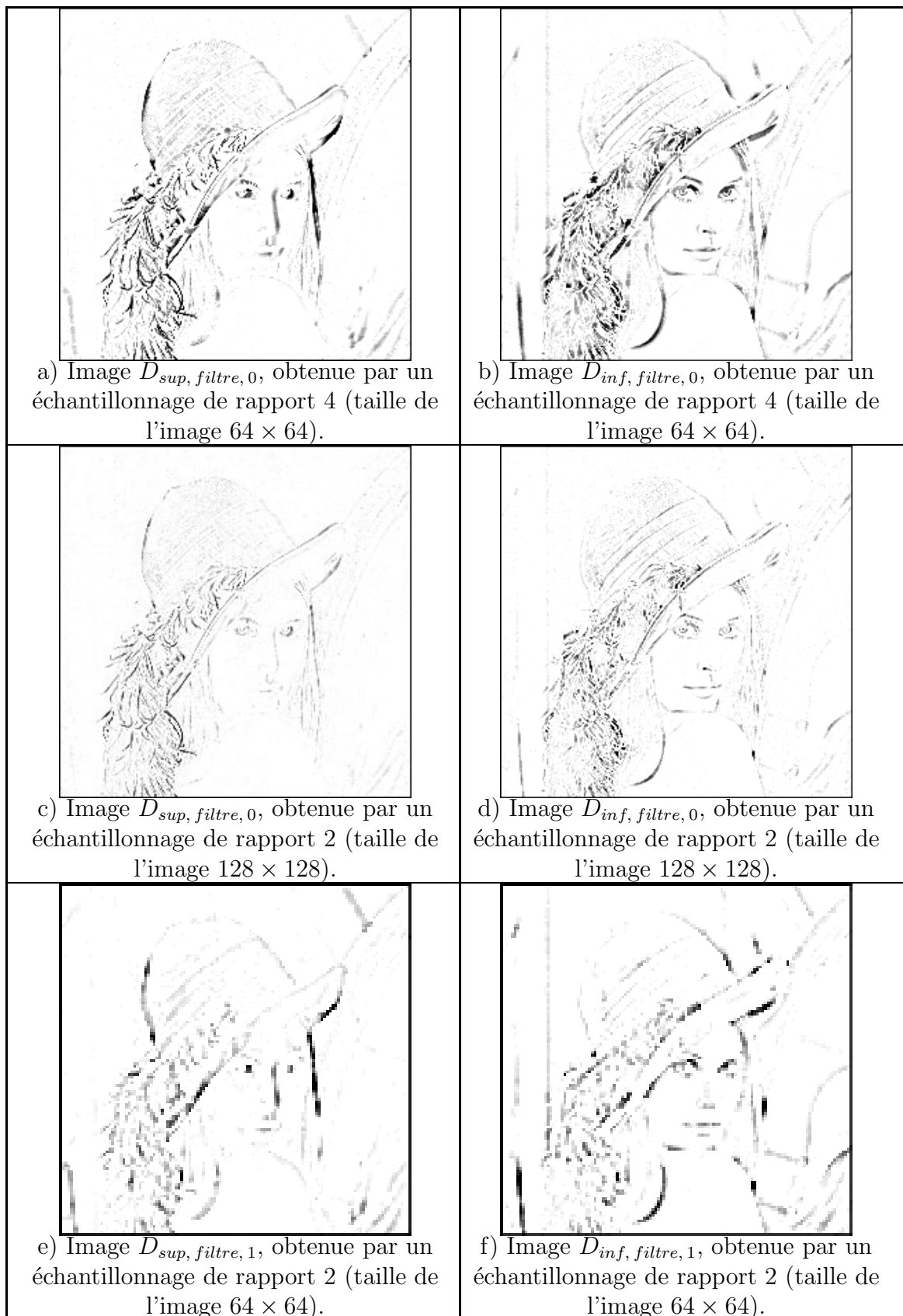


FIG. 4.6 – Comparaison des images de détails obtenues avec un échantillonnage de rapport 4 et un élément structurant de rayon 2, ou avec un échantillonnage de rapport 2 et un élément structurant de rayon 1. Les autres paramètres de la pyramide sont communs (échantillonnage par moyennage et filtre $\frac{\gamma+\varphi}{2}$).

proposé retrace l'extraction d'objets de tailles différentes mais de contraste identique par rapport au fond de la scène.

L'image analysée est une photographie aérienne du Sénégal, numérisée sur une zone proche du fleuve Niger (*cf.* figure 4.7).

Plusieurs composants de paysage sont regroupés sur cette scène. On distingue un ensemble d'éléments ponctuels sombres répartis sur toute l'image, un réseau hydrographique marqué par des traces sombres coupant l'image en deux ainsi que des éléments de forêts au Nord-Ouest de la scène. Ces éléments, que l'on souhaite caractériser ont donc tous une réflectance plus faible que le fond de l'image. Ils seront extraits dans les images $D_{inf, filtre, i}$. De plus, leurs tailles respectives sont facilement séparables. Ainsi, cette seule analyse visuelle permet de prédire que des niveaux de détection croissants seront utiles pour extraire les petits éléments de végétation, puis le réseau hydrographique, et enfin les éléments de forêt. En fait, ces trois composants représentent des couverts végétaux de natures différentes (brousses pour les petits éléments ponctuels, arbres surplombant le réseau hydrographique et forêts), ce qui explique la similitude de leur réponse spectrale. Les paramètres choisis pour cette étude sont un échantillonnage dyadique de type moyenne, un filtre $\frac{\gamma+\varphi}{2}$ et un élément structurant de rayon 1.

La visualisation des images I_i et $D_{inf, filtre, i}$ montre qu'aux deux premiers niveaux de perception sont extraites les brousses tigrées, au troisième le réseau hydrographique et au cinquième la forêt. Ces résultats sont montrés sur les figures 4.8 à 4.10.

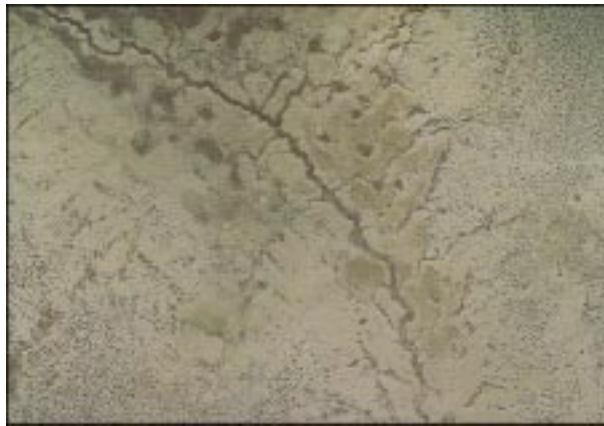


FIG. 4.7 – Image initiale sur le Sénégal.

L'union des détails $D_{inf, filtre, 0}$ et $D_{inf, filtre, 1}$ caractérise les plus petits objets extraits lors de la décomposition et décrit la répartition des brousses tigrées dans l'image. La figure 4.8 montre le résultat de cette union après binarisation. Ce seuillage est généralement indispensable pour supprimer les variations de signal très faible - de l'ordre de quelques comptes numériques - relevant plus du bruit de l'image que de l'extraction d'attributs.

La décomposition au troisième niveau et l'extraction des objets correspondant sur l'image de détails $D_{inf, filtre, 2}$ montre sur la figure 4.9 principalement les forêts de ripisylves

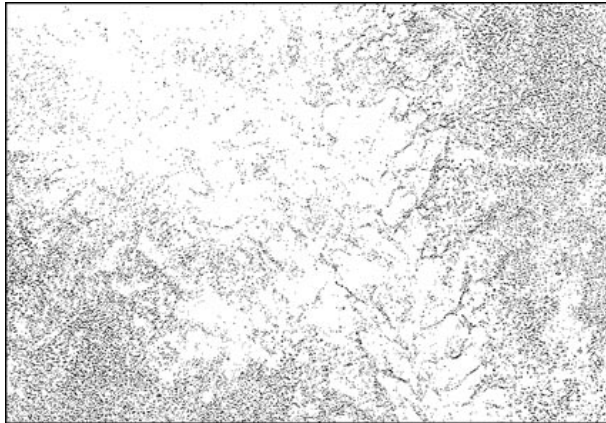


FIG. 4.8 – Union de $D_{inf, filtre, 0}$ et $D_{inf, filtre, 1}$ puis binarisation.

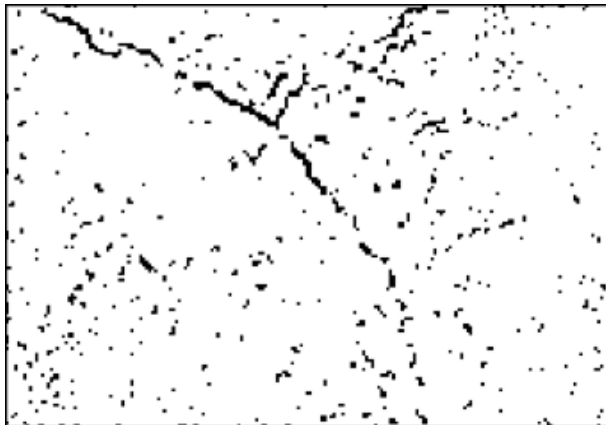


FIG. 4.9 – Binarisation de $D_{inf, filtre, 2}$.

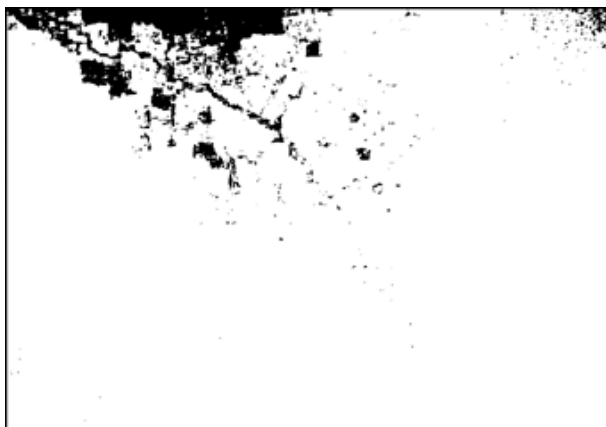


FIG. 4.10 – Union de $D_{inf, filtre, 3}$ et $D_{inf, filtre, 4}$, fermeture morphologique et binarisation du résultat.

(bords de cours d'eau) de faible largeur et ayant un agencement linéaire en bordure du réseau hydrographique.

L'union des détails aux niveaux quatre et cinq ($D_{inf, filtre, 3}$ et $D_{inf, filtre, 4}$), suivie d'une fermeture morphologique conduit à caractériser les forêts denses dans la zone Nord de la scène. Les objets correspondants ont évidemment une taille supérieure aux autres. Le bruit visible dans l'image de la forêt sur la figure 4.10 provient du seuil de binarisation choisi (des traitements simples pourraient facilement améliorer cette représentation dans le cas d'une application thématique).

Cette application illustre clairement la capacité de séparation des différents composants de paysage en fonction de leurs échelle et taille caractéristiques. Dans un cas comme celui présenté ici, où les objets sont facilement discriminables, les traitements à réaliser, après le calcul des détails, consistent en un simple seuillage, afin de binariser les images résultats.

4.5 Extraction et reconstruction d'éléments d'images

Cette section présente les résultats d'une application de la pyramide morphologique à l'analyse publiée dans la revue *Image Analysis and Stereology* [61].

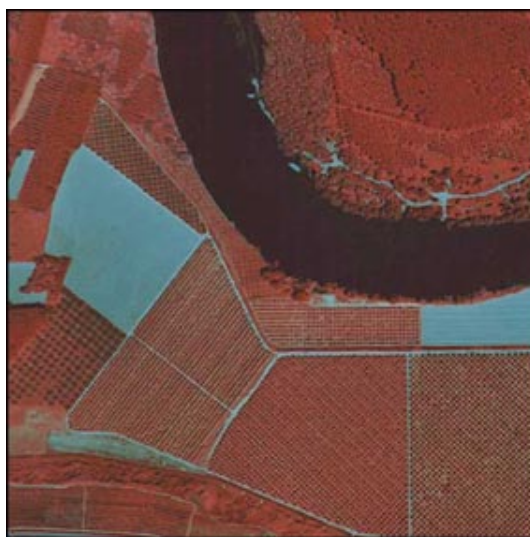


FIG. 4.11 – Image initiale sur la région de Montauban.

L'application est effectuée sur une photographie aérienne infra-rouge couleur numérisée en R, V, B sur une zone proche de Montauban, France (*cf.* figure 4.11) de taille 512×512 pixels.

Cette image contient des routes, vergers et sols nus, et est traversée par la Garonne. Les résultats sont calculés sur la bande obtenue dans le vert. Les éléments les plus fins

sont composés sur cet exemple des arbres plantés de façon périodique et de réflectances sombres, mais aussi des éléments de sol nu entre les arbres. La fréquence spatiale de plantation des arbres est équivalente à la fréquence de visibilité des sols entre ceux-ci. Seule la séparation D_{sup} / D_{inf} permet de les distinguer. Les chemins sont également des éléments fins. Finalement, les champs non cultivés et le fleuve constituent les éléments les plus gros.

En conséquence, les premiers niveaux de détails $D_{sup, filtre, 0}$ et $D_{inf, filtre, 0}$ extraient les lignes d'arbres et les frontières des champs correspondant au premier niveau de texture. Par exemple, la figure 4.12 représente l'image $D_{sup, filtre, 0}$, obtenue par un élément structurant carré de rayon 1 et d'origine centrée et par le filtre $\frac{\gamma+\varphi}{2}$ après binarisation du résultat.

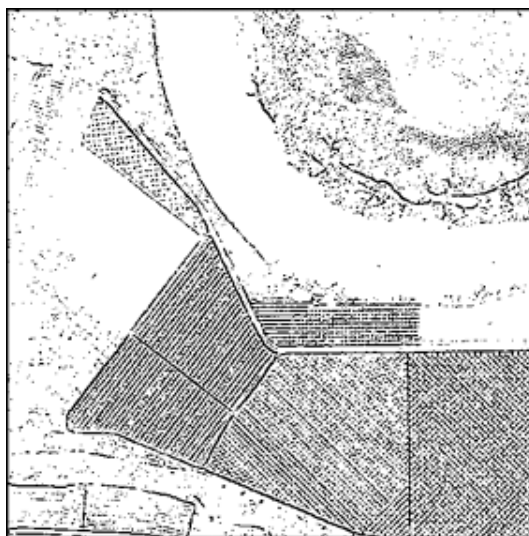


FIG. 4.12 – Image de détail $D_{sup, filtre, 0}$, obtenue par un élément structurant de rayon unité.

Il apparaît que la caractérisation précise d'un composant de paysage va nécessiter des traitements plus élaborés dans lesquels $D_{sup, filtre, i}$ et $D_{inf, filtre, i}$ servent de support à la définition de marqueurs des objets. Généralement, un marqueur est obtenu par seuillage de l'image de détails le contenant. S'il est contenu dans les images $D_{sup, filtre, i}$, la reconstruction doit être faite vers les niveaux de réflectances les plus basses. Au contraire, s'il est contenu dans les images $D_{inf, filtre, i}$, la reconstruction doit être faite vers les niveaux de réflectances élevées.

Dans le cas de l'image $D_{sup, filtre, 0}$, le seuillage a été fait de façon à obtenir un marqueur des chemins. Ensuite, on reconstruit l'image binaire obtenue dans l'image de départ. Le résultat est montré sur la figure 4.13 où les bordures de champs sont séparées des traces de sols entre les parcelles dans une représentation binaire.

Ensuite, les éléments les plus grossiers vont être extraits dans les niveaux les plus bas de la pyramide. Dans le cas de l'échantillonnage dyadique adopté dans cette application, le champ non cultivé et le fleuve apparaissent respectivement sur les images $D_{sup, filtre, 5}$

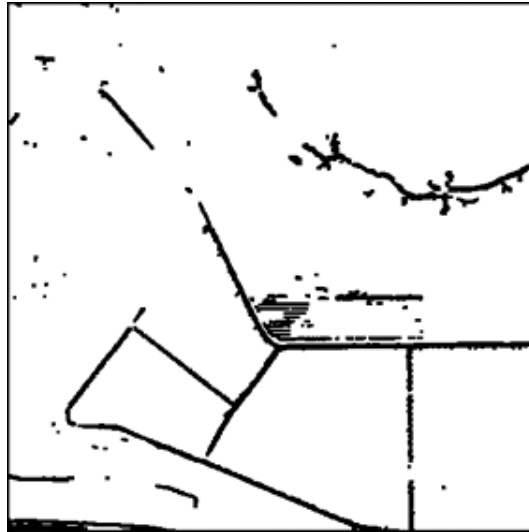


FIG. 4.13 – Image de détail $D_{sup, filtre, 0}$, obtenue par un élément structurant de rayon unité, après seuillage et reconstruction.

et $D_{inf, filtre, 5}$. La taille des images à ce niveau est $2^5 = 32$ fois inférieure à celle de l'image de départ, soit 16×16 pixels. Cette faible taille nous oblige à réaliser une reconstruction afin de reconstituer à pleine résolution les objets extraits. La figure 4.14 souligne l'effet du processus multi-résolution obtenu après le cinquième niveau et présente ces résultats : les petits encarts donnent les marqueurs, tandis que la figure globale, montre, après reconstruction, les sols nus en noir et le fleuve en gris.

La généralisation de cette approche d'extraction de composants d'une scène passe donc par l'association de traitements de reconstruction aux différents niveaux de la pyramide.

Dans les premiers niveaux, la reconstruction s'opère par dilatation conditionnelle des images binaires des détails (considérés comme des marqueurs) dans la fonction image d'origine en niveau de gris.

Dans les niveaux permettant d'extraire des composants de plus grande taille, les traitements de reconstruction s'opère selon le même principe mais, généralement, la condition est une propagation limitée à une faible variabilité spatiale du signal.

4.6 Analyse et pré-segmentation d'images

4.6.1 Principe

Les images de détails $D_{sup, filtre, i}$ et $D_{inf, filtre, i}$ contiennent les détails, ou éléments fins d'une image et correspondent à des formes caractéristiques de la scène. L'image résultat de la décomposition, I_N représente les plus gros objets, ou les éléments de grande taille de

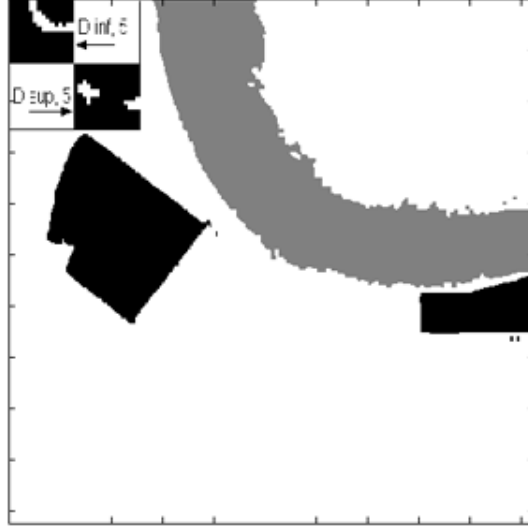


FIG. 4.14 – Image de détail au niveau 4, après reconstruction dans l'image de départ.

la scène. Les images $D_{sup, dec, i}$ et $D_{inf, dec, i}$ contiennent les éléments correspondant aux pertes d'information induites par l'échantillonnage et ne représentent donc pas d'éléments d'information propres à la scène observée, mais propres à l'échantillonnage.

On met en parallèle ces observations avec l'objectif d'une pré-segmentation d'image : il s'agit de supprimer de la scène les éléments introduisant une variabilité du signal à l'intérieur d'une région et produisant une texture rendant l'étape de segmentation délicate. Ces éléments correspondent aux petits éléments fins d'une image et sont donc à rapprocher du contenu des images $D_{sup, filtre, i}$ et $D_{inf, filtre, i}$.

C'est pourquoi, on propose ici une étape de pré-segmentation bâtie sur la pyramide morphologique de la façon suivante :

1. décomposition de la scène par la pyramide morphologique en N étages,
2. reconstruction de la scène à l'aide des images $D_{sup, dec, i}$ et $D_{inf, dec, i}$ uniquement.

Mathématiquement, l'équation (3.16) devient l'équation (4.1) :

$$\begin{aligned} IR_i &= IRF_i + D_{sup, dec, i} - D_{inf, dec, i} \\ &= IRF_i + D_{dec, i} \end{aligned} \quad (4.1)$$

L'équation (3.45) s'écrit maintenant :

$$IR_{pré-segmentation, j} = D_{dec, j} + Int_j(D_{dec, j+1} + Int_{j+1}(D_{dec, j+2} + \dots + Int_{N-2}(D_{dec, N-1} + Int_{N-1}(I_N)))) \quad (4.2)$$

La différence entre les équations (3.45) et (4.2) provient du fait que, dans le cas où l'on souhaite créer une image de pré-segmentation, seules les images de détails provenant de l'échantillonnage sont ajoutées lors du processus de reconstruction. Par la suite, nous

qualifierons de « texture » les éléments de détails provenant du filtrage car leur association correspond à l'agencement spatial des motifs de l'image.

L'image reconstruite représente alors l'image apte à être segmentée d'où les principaux éléments de texture ont été supprimés.

Par contre, ces éléments de texture peuvent être reconstruits séparément. Dans ce cas, la reconstruction commence par l'ajout des images $D_{sup, filtre, N-1}$ et $D_{inf, filtre, N-1}$ à une image « vide ». La reconstruction ne s'effectue qu'à partir des images de détails provenant du filtre. La reconstruction se synthétise de façon corollaire à l'équation (3.45) grâce à l'équation (4.3) :

$$IR_{texture, j} = D_{filtre, j} + Int_j(D_{filtre, j+1} + Int_{j+1}(D_{filtre, j+2} + \dots + Int_{N-2}(D_{filtre, N-1}))) \quad (4.3)$$

Les différences entre les deux équations proviennent du fait que

1. les images de détails insérées ne proviennent que du filtrage et non plus de l'échantillonnage et du filtrage,
2. la reconstruction n'est pas basée sur une image basse résolution I_N , mais sur une image « vide ».

En plus de la pré-segmentation ainsi décrite, il apparaît donc que l'on peut décomposer l'information de texture contenue dans la scène et extraite par les premiers niveaux d'une pyramide morphologique.

4.6.2 Exemples

L'image choisie représente une photographie aérienne couleur sur une région agricole et forestière dans le Tarn (*cf.* figure 4.15a). Les résultats sont présentés sur la composition colorée résultant de la superposition des trois canaux de numérisation rouge, vert et bleu. Chaque canal est traité de façon identique, tant lors de la décomposition que lors de la recombinaison. Les paramètres de la pyramide morphologique sont un échantillonnage dyadique par moyennage, un élément structurant de rayon 1, un filtre correspondant à la demi-somme (ouvert + fermé) et une interpolation par simple duplication.

Les figures 4.15 et 4.16 permettent de comparer, à chaque niveau de la pyramide l'image basse résolution issue de la décomposition et l'image reconstruite sans les détails provenant du filtrage. Chaque ligne de cette figure représente un niveau. Dans ce type d'application, plus le nombre d'étages réalisé est important et plus l'image résultat au niveau zéro sera filtrée. Le choix du nombre d'étages en fonction de l'image et du niveau de pré-segmentation désiré est ici un élément primordial. Les images montrées sur ces deux figures sont obtenues après saturation des comptes numériques sur 8 bits.

Les éléments contenus dans I_0 (*cf.* figure 4.15a) se retrouvent soit dans $IR_{pré-segmentation, 0}$ (*cf.* figure 4.16b), soit dans $IR_{texture, 0}$ (*cf.* figure 4.17d). De façon générale, I_i (*cf.* figures 4.15 et 4.16 colonne de gauche) se décompose en $IR_{pré-segmentation, i}$ (*cf.* figures 4.15 et 4.16

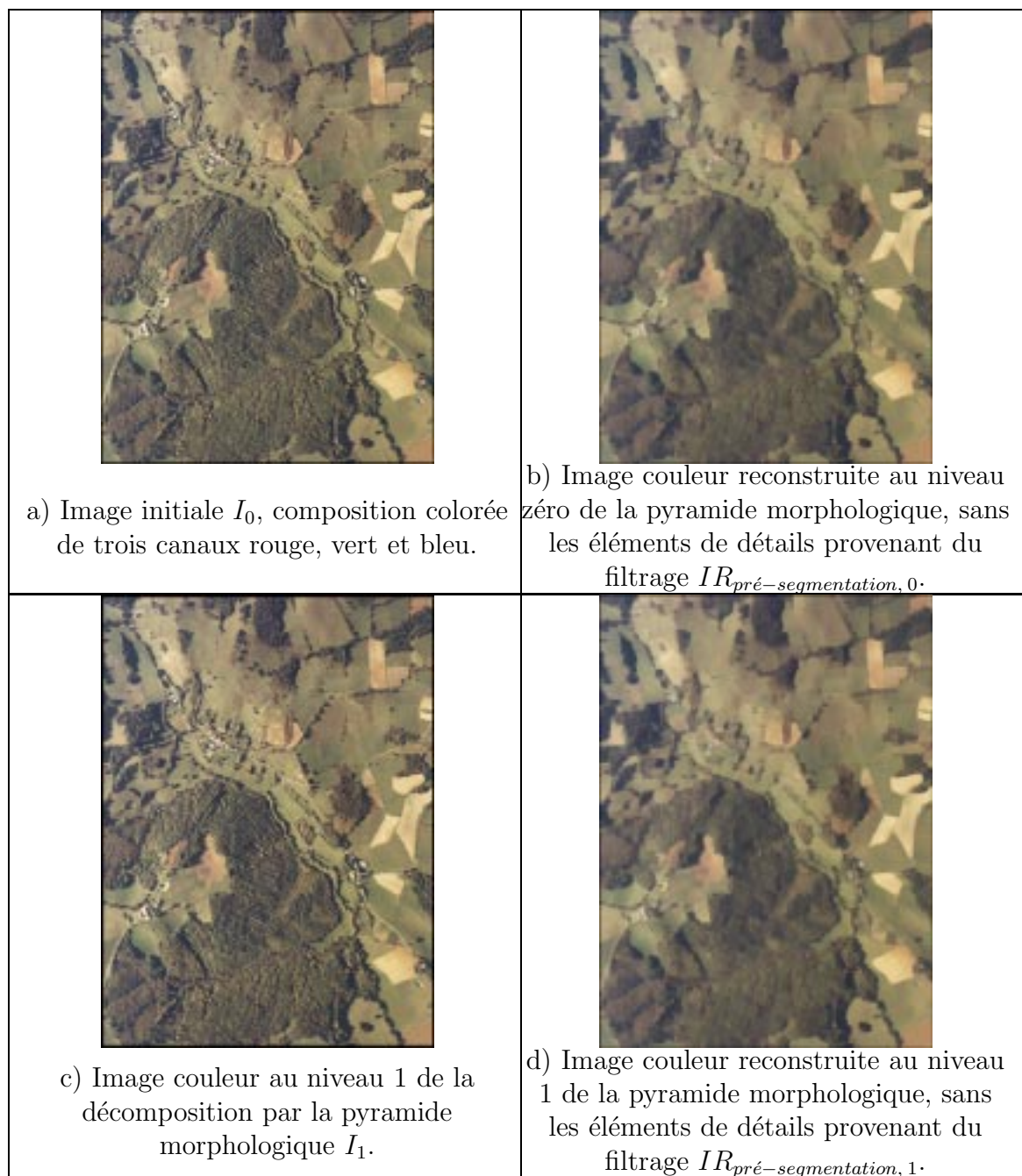


FIG. 4.15 – Comparaison entre, à gauche, les images basse résolution obtenues par la pyramide morphologique aux niveaux 0 et 1 et à droite, les images reconstruites à l'aide des seuls détails provenant de l'échantillonnage.

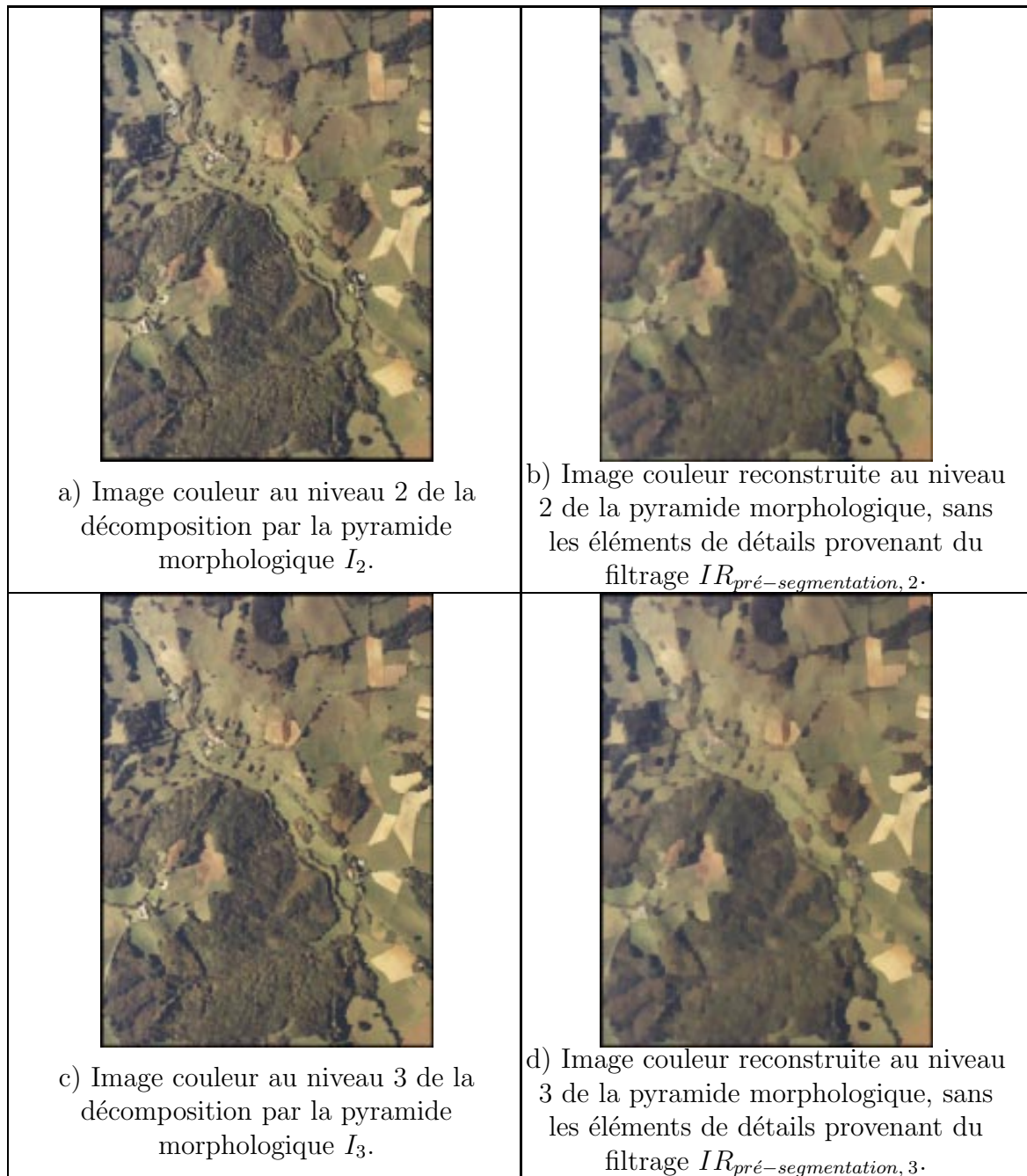


FIG. 4.16 – Comparaison entre, à gauche, les images basse résolution obtenues par la pyramide morphologique aux niveaux 2 et 3, et à droite, les images reconstruites à l'aide des seuls détails provenant de l'échantillonnage.

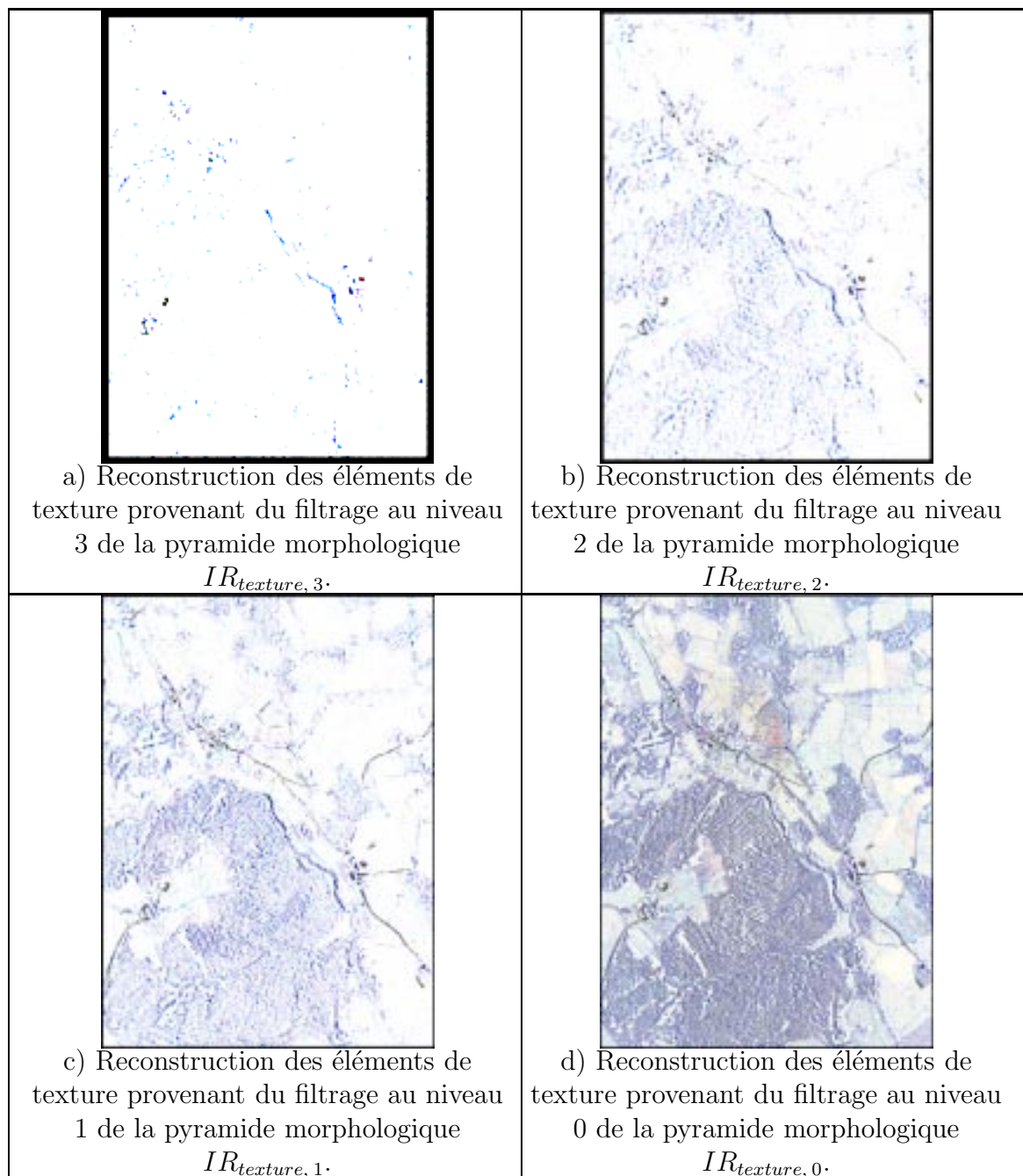


FIG. 4.17 – Reconstruction des seuls éléments de texture de l'image aux différents niveaux de la pyramide.

colonne de droite) et $I_{texture, i}$ (cf. figure 4.17). On voit donc que l'image $IR_{pré-segmentation, i}$ est une version lissée de I_i , visible à chaque niveau, si l'on compare la colonne de gauche et la colonne de droite des figures 4.15 et 4.16. La variabilité des paysages y est moins grande, ce qui se traduit par des surfaces plus homogènes. En termes statistiques, l'écart type de $IR_{pré-segmentation, i}$ est inférieur à celui de I_i . La variabilité supprimée se retrouve dans les images $I_{texture, i}$.

En conclusion, des essais de segmentation devraient être conduits sur les images $IR_{pré-segmentation, 0}$ et I_0 et leurs résultats comparés. L'influence du nombre d'étages réalisé pour reconstruire $IR_{pré-segmentation, 0}$ mérite également une étude particulière.

4.7 Analyse et pré-compression de données

4.7.1 Principe

L'idée développée ici est que les informations contenues dans les images de détails ne sont pas toutes primordiales à la restitution de l'image initiale. En effet, dans toutes les applications d'analyse vues aux sections précédentes, un seuillage préalable est réalisé avant l'affichage des images de détails. Suivant cette constatation, une image reconstruite approximant l'image de départ peut être obtenue par le processus de reconstruction dans lequel les images $D_{sup, filtre, i}$ et $D_{inf, filtre, i}$ sont injectées après seuillage. Cette image représente une version compressée de l'image de départ, pour laquelle des informations revêtant peu d'importance sont supprimées.

Dans ce cas, l'équation (3.45) devient l'équation (4.4) :

$$\begin{aligned} IR_{compression, j} = & D_{deci, j} + D_{filtre, seuil, j} \\ & + Int_j(D_{deci, j+1} + D_{filtre, seuil, j+1} Int_{j+1}(D_{deci, j+2} \\ & + D_{filtre, seuil, j+2} + \dots + Int_{N-2}(D_{deci, N-1} + D_{filtre, seuil, N-1} + Int_{N-1}(I_N)))) \end{aligned} \quad (4.4)$$

Le paramètre important de cette équation est ainsi le taux de seuillage (nombre de pixels \leq valeur du seuil) des images de détails. Si celui-ci est de 100%, on se retrouve dans le cas précédent de la pré-segmentation. Cette étape de pré-compression constitue donc une généralisation du procédé précédent. Par exemple, un seuillage supprimant les comptes numériques faibles de $D_{sup, filtre, i}$ et $D_{inf, filtre, i}$ (de valeur inférieure à environ 5) ne modifie pas l'aspect visuel de l'image.

4.7.2 Exemple

L'image utilisée dans cet exemple est une photographie aérienne infra-rouge couleur numérisée (photographie IGN, 1977, région de Château-Thierry, Aisne, France de taille 1181×1149) et représente une région agricole à une résolution métrique.

Cette image est décomposée par la pyramide morphologique avec les mêmes paramètres qu'au paragraphe précédent 4.6 et reconstruite avec des images de détails provenant du

	i=0	i=1	i=2	i=3
$D_{sup, filtre, i}$	44%	39%	34%	33%
$D_{sup, filtre, i}$	40%	28%	23%	24%

TAB. 4.1 – Pourcentage de pixels mis à zéro par le seuillage à 10 en fonction de l'étage du canal vert de l'image de la région agricole.

filtre et seuillées. La figure 4.18 montre l'image initiale en haut à gauche et les résultats de différents seuillages. La dernière image représente une reconstruction à l'aide d'un seuillage de 100% et correspond donc au cas de la section précédente.

L'image 4.18b) représente le cas où sur les images $D_{sup, filtre, i}$ et $D_{inf, filtre, i}$, les valeurs de pixels inférieures à 10 ont été mises à zéro. En dessous de ce seuil, en effet, les différences sont difficilement perceptibles à l'œil. Le pourcentage de pixels mis à zéro par ce seuillage dépend évidemment de l'image de départ et du niveau de la pyramide. Dans notre exemple, sur le canal vert, ce pourcentage est indiqué sur le tableau 4.1 et vaut entre 20 et 45%. Ce nombre ne prend pas en compte les pixels de valeur 0 déjà présents avant le seuillage et qui apportent également leur contribution à une pré-compression.

Ces constatations sont très intéressantes dans le cas où l'on s'intéresse à un schéma de compression dans un objectif de visualisation des images. Le nombre de pixels de valeur nulle est grandement augmenté par ce simple seuillage. Plusieurs schémas de compression possibles à l'aide de la pyramide morphologique sont évoqués dans la section suivante. Le pourcentage de compression qui y est annoncé peut s'améliorer par la prise en compte du pourcentage de zéros supplémentaires calculé ici.

Les images 4.18c) et d) montrent l'influence d'un seuillage de plus en plus élevé : l'image reconstruite est de plus en plus lissée et la perte d'information est ainsi de plus en plus grande.

4.7.3 Reconstruction « à la carte »

Ces applications ouvrent de nombreuses autres perspectives à la reconstruction.

- De façon encore plus générale, il peut être intéressant, en fonction de la scène observée de ne conserver que les détails *sup*, ou les détails *inf*, ou bien encore les détails *sup* de certains étages et les *inf* d'autres étages. Le choix des éléments à conserver est fonction du paysage et de la quantité d'informations contenue dans chaque image de détail.
- Au lieu de seuiller les images de détails $D_{sup, filtre, i}$ et $D_{inf, filtre, i}$, et ainsi de mettre à zéro tous les pixels de valeur inférieure au seuil, on peut choisir de supprimer une bande de valeurs comprises entre un seuil bas et un seuil haut. En effet, par définition même des détails, les pixels les plus nombreux sont ceux ayant un compte numérique faible. Les supprimer revient à supprimer de l'information sur de nombreux pixels. L'effet visuel de la suppression de peu de pixels, mais de forts comptes numériques serait intéressant à observer.

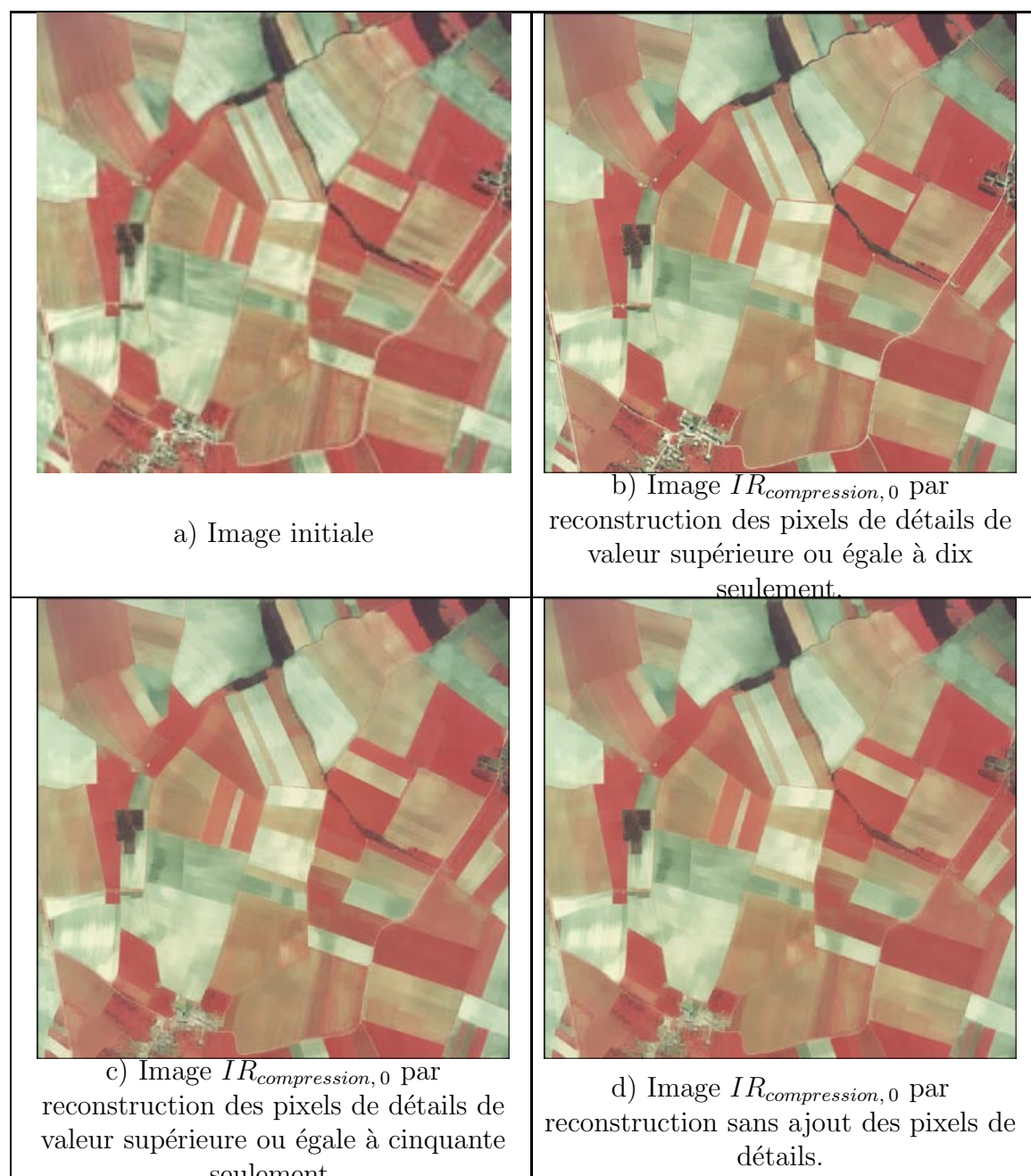


FIG. 4.18 – Comparaison des différents niveaux de seuillage des images de détails provenant du filtre sur l'image des champs. Les seuils varient de 0 où tous les détails sont conservés (cela correspond à l'image initiale) à 255 où tous les détails sont supprimés (cela correspond au cas évoqué dans la section 4.6).

- Une étude analogue à celle menée ici sur les détails $D_{sup, filtre, i}$ et $D_{inf, filtre, i}$ peut évidemment être réalisée sur les images $D_{sup, deci, i}$ et $D_{inf, deci, i}$.

4.8 Perspective pour la compression

L'intérêt de la pyramide morphologique peut également s'envisager dans un processus de compression d'images. En effet, au lieu de transmettre les images entières, codées sur 8 bits, il suffit de transmettre les images de détails et la dernière image basse résolution. La dernière image basse résolution est alors codée sur 8 bits et les images de détails sont codées sur 4 bits, ce qui peut entraîner une perte d'information.

Dans cette optique et de façon très préliminaire, quatre schémas de codage et de transmission ont été testés, pouvant conduire à des taux de compression plus ou moins importants. Ces schémas sont représentés sur la figure 4.19 où $n \times n$ représente la taille de l'image de départ et N le nombre d'étages de la pyramide morphologique réalisé avant la transmission. Le pas d'échantillonnage entre deux étages est supposé égal à deux.

Le nombre de pixels de l'image de départ est n^2 , ce qui représente donc $8 \times n^2$ bits. Le nombre de pixels de l'image I_N est $\frac{n^2}{4^N}$, ce qui correspond à transmettre $8 \times \frac{n^2}{4^N}$ bits.

- Le premier schéma, montré sur la figure 4.19a), consiste à ajouter $D_{inf, filtre, i}$ et $D_{inf, deci, i}$ à chaque étage (resp. $D_{sup, filtre, i}$ et $D_{sup, deci, i}$) et à ne transmettre comme détails que $D_{inf, i}$ et $D_{sup, i}$, codés sur 4 bits non signés. I_N est également transmise. Les images de détails correspondant à l'étage i ont une taille $\frac{n^2}{4^{(i-1)}}$ chacune, le nombre de bits à transmettre est donc $4 \times \frac{n^2}{4^{(i-1)}}$ à l'étage i . Deux images de détails étant transmises à chaque étage, le nombre total de bits en fonction de l'étage est

$$\boxed{8 \times \frac{n^2}{4^N} + 8 \times \sum_{i=1}^N \frac{n^2}{4^{(i-1)}}}. \text{ Le rapport de compression est donc :}$$

$$\left(8 \times \frac{n^2}{4^N} + 8 \times \sum_{i=1}^N \frac{n^2}{4^{(i-1)}}\right) \times \frac{1}{8 \times n^2} = \frac{1}{4^N} + 4 \sum_{i=1}^N \frac{1}{4^{(i-1)}} = \frac{1}{4^N} + 4 \times \sum_{i=1}^N \frac{1}{4^i} = \frac{1}{4^N} + 4 \times \frac{1}{4} \times \frac{1 - (\frac{1}{4})^N}{1 - \frac{1}{4}}$$

$$\text{Soit } \boxed{\frac{1}{4^N} + \frac{4}{3} \times \left(1 - \left(\frac{1}{4}\right)^N\right)}.$$

Quand on augmente le nombre d'étages, le taux limite de compression est $4/3$, soit 133%. Utilisée seule, cette méthode n'est donc pas rentable. Calcul fait, on obtient les taux de compression en fonction de l'étage indiqués sur le tableau 4.2.

- Le deuxième schéma de compression est montré sur la figure 4.19b). I_N est transmise sur 8 bits. Une seule image de détail est transmise, de taille $n \times n$; elle est la somme de toutes les images de détails intermédiaires à tous les niveaux, et est

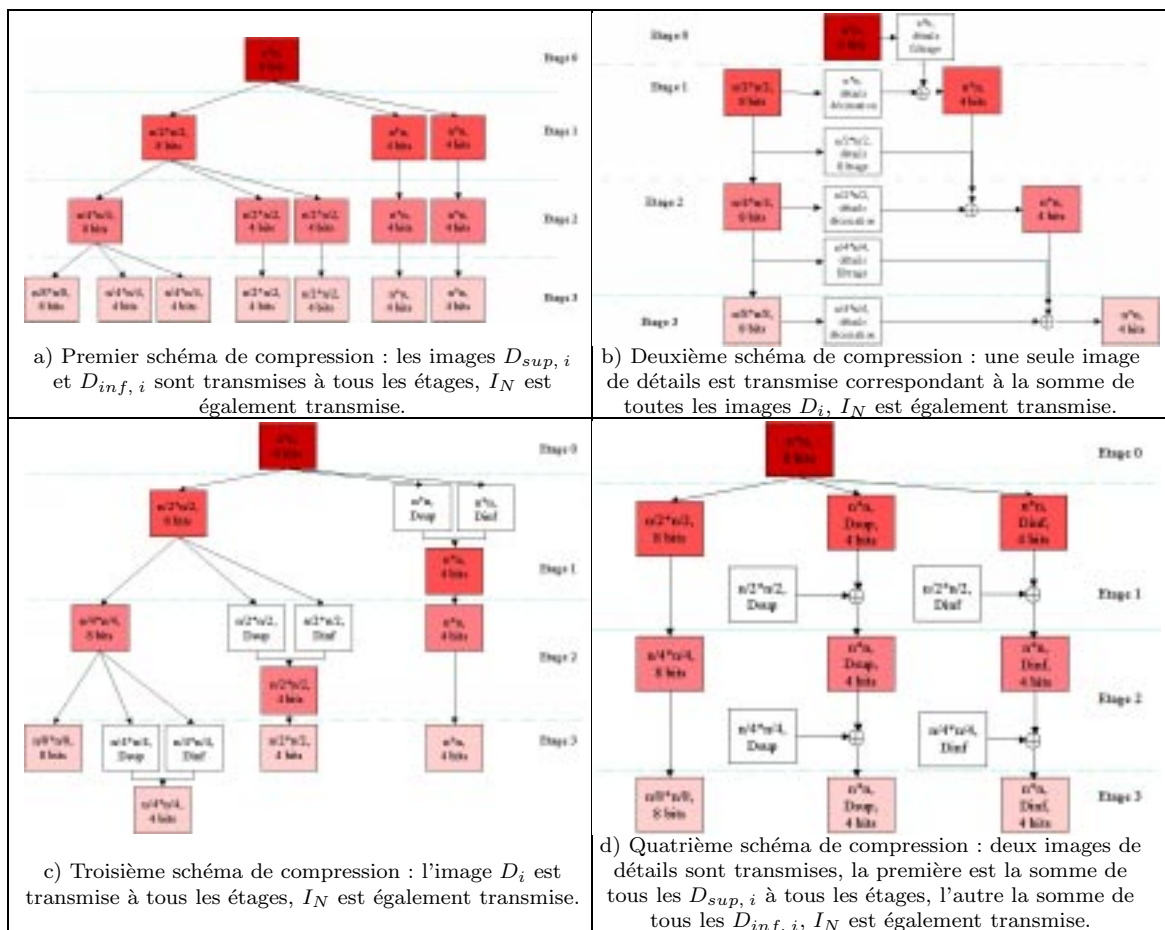


FIG. 4.19 – Comparaison des quatre schémas de transmission testés dans le cas d'une transmission par 1, 2 ou 3 étages de la pyramide. L'image I_N est toujours transmise, la différence provient de la transmission des images de détails. Les images rose clair représentent l'ensemble des images à transmettre si la pyramide comporte trois étages, celle rose moyen si la pyramide comporte deux étages, les rouges, si elle n'en comporte qu'un.

Nombre d'étages	1	2	3	4
Rapport de compression en % pour le premier schéma	125%	131,25%	132,8%	132,8%
Rapport de compression en % pour le deuxième schéma	75%	56,2%	51,6%	50,4%
Rapport de compression en % pour le troisième schéma	75%	68,7%	67,2%	66,7%
Rapport de compression en % pour le quatrième schéma	125%	106,25%	101,6%	100,4%

TAB. 4.2 – Comparaison des taux de compression théorique de chaque schéma de compression étudié.

codée sur 4 bits signés. Cette image se code à l'aide de $4 \times n^2$ bits. Au total, cette transmission coûte $8 \times \frac{n^2}{4^N} + 4 \times n^2$ bits et fournit un rapport de compression de $\frac{1}{4^N} + \frac{1}{2}$ et tend vers un rapport de compression de 50% avec le nombre d'étages. Le tableau 4.2 indique l'évolution des taux de compression en fonction du nombre d'étages réalisé.

- Le troisième schéma de compression est montré sur la figure 4.19c). I_N est transmise sur 8 bits. Une seule image de détail est transmise à chaque étage représentant D_i ; elle est codée sur 4 bits signés. Le nombre de bits utile pour transmettre ce schéma est

$$8 \times \frac{n^2}{4^N} + 4 \times \sum_{i=1}^N \frac{n^2}{4^{(i-1)}}, \text{ soit un rapport de compression de } \frac{1}{4^N} + \frac{2}{3} \times \left(1 - \left(\frac{1}{4}\right)^N\right),$$

le taux maximum de compression est $2/3$, soit 66%. Le tableau 4.2 indique l'évolution des taux de compression en fonction du nombre d'étages réalisé.

- Le quatrième schéma de compression est montré sur la figure 4.19d). I_N est transmise et codée sur 8 bits. Au total, deux images de détails de taille $n \times n$ sont transmises, l'une représente la somme de tous les $D_{sup,i}$ à chaque étage, l'autre la somme de tous les $D_{inf,i}$. Ces deux images sont codées sur 4 bits non signés. Le

nombre total de bits à transmettre est $8 \times \frac{n^2}{4^N} + 8 \times n^2$, soit un rapport de compression de $\frac{1}{4^N} + 1$ qui tend vers 100%.

Plus le taux de compression correspondant à une image est élevé, plus la perte d'information causée par le codage sur 4 bits va être importante. Ainsi, la méthode 1 va être très proche d'un codage sans perte, la méthode 2 va en être plus éloignée.

Lors de la transmission des images de détails, il n'est pas nécessaire de transmettre les valeurs nulles dans les images de détails. Bien évidemment, plus on transmet d'images de détails, plus celles-ci contiennent de zéros, et plus le pourcentage de pixels à transmettre est faible. Il reste à voir quelle est la probabilité moyenne de zéros que l'on a pour chaque

type de détails (méthodes 1 à 4), et de voir ainsi de combien cela réduit le nombre de bits à transmettre. De plus, si les détails sont trop importants, outre la diminution du nombre de bits à transmettre, vient l'accroissement du taux d'erreur induit par la compression. Lorsque l'on transmet les détails sup et les détails inf sur une seule image, les détails ne sont en fait plus codés qu'entre 0-16, puisqu'il faut transmettre un bit de signe, cela augmente bien entendu le taux d'erreur. Il faut ensuite identifier le filtre qui donne le meilleur compromis, entre d'une part la minimisation de la densité de probabilité des détails provenant du filtrage d'une part et provenant de l'échantillonnage d'autre part. Il faut donc évaluer le taux d'erreur induit par le codage sur 4 bits, ainsi que le taux de zéros qui apparaît dans l'image. A priori, les codages où l'on transmet beaucoup d'images de détails (et qui ne sont donc pas très rentables) sont celles qui doivent contenir le plus de zéros (et donc qui peuvent devenir rentable) mais également celles qui engendrent le moins d'erreur. Elles ne sont donc pas à rejeter au premier abord.

En fonction du schéma de compression et du filtre utilisé, le nombre de pixels transmis sans erreur et le taux de compression sont très variables. Dans le cas du schéma de compression 1, il suffit en moyenne de ne transmettre que 36% des pixels, les autres étant nuls. Pour les trois autres schémas de compression, le taux de pixels nul dépend fortement du nombre d'étages réalisé. Les résultats, synthétisés sur le tableau 4.3, montrent que les résultats sont maintenant beaucoup plus proches. Seule la méthode 4 garde un taux de compression faible. En ce qui concerne la perte d'information, les résultats sont très variables en fonction du filtre étudié, du nombre d'étages réalisé et du type d'images. Cependant, le schéma 1 est celui qui occasionne le moins de perte - entre 2 et 3 % seulement - et le schéma 2 celui qui en occasionne le plus - entre 15 et 30 % -. Le schéma de codage 3 occasionne des pertes entre 7% et 13%, le schéma 4 entre 6% et 17%.

Nombre d'étages	1	2	3	4
Rapport de compression en % pour le premier schéma	61%	51%	48,81%	48,20%
Rapport de compression en % pour le deuxième schéma	57%	46,75%	45,56%	45,64%
Rapport de compression en % pour le troisième schéma	57%	46,25%	43,62%	42,98%
Rapport de compression en % pour le quatrième schéma	61%	62,5%	70,6%	77,4%

TAB. 4.3 – Taux de compression comparé des quatre schémas dans le cas où les pixels de valeur nulle ne sont pas transmis.

En conclusion, cette étude préliminaire, réalisée sans faire appel aux méthodes académiques, montre que le schéma 4 n'est pas bon et que le 2 semble inférieur au 3 dans tous les domaines. En fonction de la qualité que l'on veut mettre en avant (taux de compression ou perte), les schémas 1 et 3 semblent les meilleurs. Cependant, il serait nécessaire de compléter cette étude :

- en tenant compte de la non transmission éventuelle des pixels de faible compte numérique dans les images de détails, selon le procédé décrit à la section 4.7. Dans

ce cas, le nombre de pixels moyens à transmettre dans les images de détails peut descendre jusqu'à 30 ou 40% seulement du nombre de pixels total. Ce pourcentage varie en fonction du seuil choisi,

- par un plus grand choix d'images afin de tester des variabilités de paysages plus grandes,
- par un paramétrage complet de la pyramide, notamment en testant plusieurs pas d'échantillonnage entre les étages,
- en évaluant visuellement les résultats afin de déterminer si les pertes sont gênantes ou pas,
- de quantifier les pertes, non seulement en nombre, mais aussi en écart par rapport à la valeur nominale,
- et enfin, en améliorant les résultats grâce à des algorithmes de compression déjà existants.

4.9 Conclusion

La méthode d'analyse par la pyramide morphologique permet de séparer les différents composants d'une scène en fonction de leur réflectance par rapport à leur environnement et de leur taille. On peut notamment insister sur deux points :

1. une analyse différente peut être réalisée indépendamment sur chaque niveau et sur chaque composant,
2. des connaissances *à priori* facilitent le choix des paramètres et rendent l'analyse déterministe.

En outre et selon les problèmes posés, des traitements ultérieurs sont parfois nécessaires.

Ces traitements ultérieurs sont de plusieurs natures : tout d'abord, il arrive que les éléments à extraire ne correspondent pas exactement à un niveau de résolution, mais s'échelonnent sur plusieurs niveaux. Dans ce cas, comme nous l'avons montré dans l'exemple de la section 4.4, il est nécessaire de faire l'union des détails contenant les objets d'intérêts. Ensuite, le seuillage peut ne pas être suffisant pour caractériser les objets, notamment les objets les plus grossiers, extraits sur des images de petites tailles. Dans ce cas, il est nécessaire d'extraire un marqueur des images de détails et de reconstruire ce marqueur dans l'image de départ. Cela se fait de la façon suivante :

- Calcul des marqueurs par binarisation de l'image de détail,
- Sur - échantillonnage du résultat à la taille de l'image initiale,
- Reconstruction de l'image du marqueur binaire dans l'image initiale.

La pyramide morphologique semble ainsi une bonne approximation de la perception multi-niveau d'éléments emboîtés dans un paysage donné.

En conclusion, on voit que la méthode présentée dans ce chapitre permet l'extraction d'objets dans une image. Ce type d'extraction est déterminé par les propriétés physiques

des objets et modélisé par les paramètres de la pyramide. Même si les D_{sup} ou les D_{inf} ne caractérisent pas à eux seuls les objets, un traitement supplémentaire simple le permet. En fait, dans la plupart des cas, une classe d'objets n'est pas exactement extraite à un niveau. Pour le faire, il faut généralement considérer les détails comme marqueurs d'un objet d'intérêt (*cf.* paragraphe 4.5).

Ensuite, la paramétrisation du schéma de décomposition / recombinaison de la pyramide morphologique permet de créer des images reconstruites pouvant servir de base à la segmentation et à l'amélioration d'un schéma de compression.

De même, la décomposition par la pyramide morphologique permet également une forme d'extraction passe-bande (plus précisément une extraction de classe de détails de taille donnée) décrite spatialement et explicitement dans le domaine de l'image.

Quatrième partie

**FUSION DE DONNEES
MULTI-SOURCES**

Chapitre 5

État de l'art sur la fusion de données multi-résolution

5.1 Définition de la fusion de données multi-sources

5.1.1 Introduction

Une image numérique est une représentation discrète d'une cible analogique obtenue par une chaîne d'acquisition. La fusion est un processus visant à mélanger des données d'origines et de contenus différents. De nombreuses définitions ont été proposées ; on peut citer celle de Wald [145] :

« Data fusion is a formal framework in which are expressed the means and tools for the alliance of data originating from different sources. It aims at obtaining information of greater quality ; the exact definition of 'greater quality' will depend upon the application. » (La fusion de données est un cadre formel dans lequel on exprime les outils et les moyens pour allier des données provenant de sources différentes. Son but est d'obtenir des informations de qualité accrue ; la définition de 'qualité accrue' dépendra de l'application).

Dans ce chapitre, on limitera l'étude de la fusion de données provenant de capteurs différents (« Multi-Sources ») à l'observation d'une cible commune. Les différents types de capteurs sont très nombreux : ils peuvent par exemple fournir un signal temporel 1D représentant l'évolution d'une caractéristique de la cible au cours du temps, ou au contraire un signal 2D représentant la cible en un instant fixé. Certains capteurs fournissent une série d'images montrant l'évolution dans le temps de la représentation considérée. Outre ces différences, les caractéristiques capturées peuvent s'avérer très différentes : taille de la cible, longueur d'onde, composition, ... Devant cette extraordinaire multiplicité de données pouvant être fusionnées, les méthodes et processus de fusion sont nécessairement très nombreux.

Cependant, des éléments communs apparaissent dans tous les processus de fusion :

1. L'objectif de la fusion est l'un des deux (ou les deux) points suivants (*cf.* figure 5.1)

- soit améliorer la fiabilité d'un jugement par l'apport d'une information redondante,
- soit améliorer la capacité d'interprétation par l'apport d'une information complémentaire.

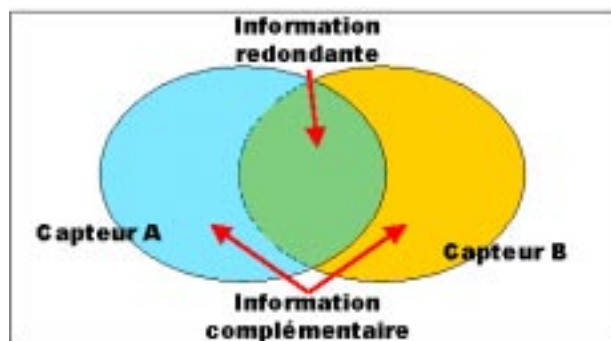


FIG. 5.1 – Objectifs généraux de la fusion.

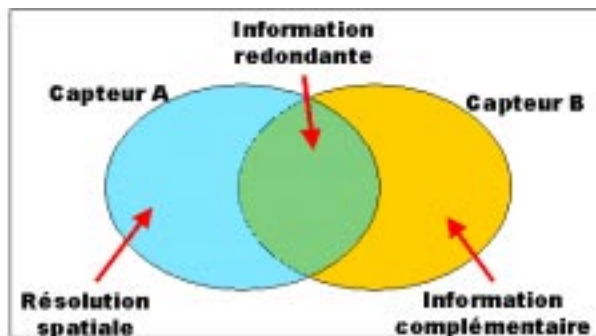


FIG. 5.2 – Restriction des objectifs dans le cadre de notre étude.

2. La fusion peut s'effectuer à plusieurs niveaux de traitement [101] :

- au niveau des données brutes, physiques,
- au niveau de l'extraction des caractéristiques,
- au niveau symbolique.

Lors d'un processus de fusion complexe, ces différents niveaux seront parfois combinés.

3. Dans tous les cas, le but est d'extraire toute l'information utile des données sources, de ne pas introduire d'artéfacts qui nuiraient soit à l'interprétation, soit aux traitements ultérieurs, et qui soit robuste et fiable vis à vis des imperfections (telle que la superposition des données).

4. Un problème majeur dans tous les problèmes de fusion de données reste le recalage préalable de celles-ci [85].

La fusion de données est indispensable dans nombreux domaines d'applications scientifiques :

- la robotique (guidage automatique d'après des capteurs visuels, tactiles, ... ; fusion de caméra stéréo ; reconnaissance automatique de cible ; ...),
- l'imagerie médicale (fusion d'images CT (computed tomography) et MR (par résonance magnétique) ; chirurgie assistée par ordinateur ; recalage spatial de surface 3D ; ...),
- les chaînes de fabrication (contrôle non destructif des produits manufacturés ; robotisation des chaînes ; ...),
- la défense militaire,
- la télédétection des surfaces terrestres.

Par la suite, nous ne considérerons que le cas de la fusion de capteurs fournissant une (ou plusieurs) image(s) de résolutions spatiales différentes d'une même cible. Dans ce cadre, une ou plusieurs image(s) a (ont) une résolution spatiale meilleure que d'autres qui procurent en complément un intérêt supplémentaire pour l'application considérée. Dans notre étude, le schéma 5.1 se réduit au schéma 5.2.

En conséquence, nous détaillerons plus particulièrement les deux domaines les plus proches de notre étude, l'imagerie médicale (fusion MR et CT) et la télédétection (fusion d'images satellites).

Nous ne détaillerons que les méthodes fusionnant les données au niveau pixel. Celles à un niveau de représentation plus élevé ne seront pas traitées, bien que certaines approches soient comparables avec ce qui se fait au niveau pixel (voir [112] pour une méthode de désagrégation par exemple).

5.1.2 Les différents types de données en télédétection

Le chapitre 1 a montré les paramètres entrant en jeu dans la composition d'une image satellite. Un processus de fusion idéal devrait pouvoir prendre en compte tous ces paramètres. Dans la réalité, cependant, il est nécessaire soit de corriger au préalable les différences entre les divers paramètres, soit de les négliger, soit encore de savoir les analyser sur le résultat de la fusion.

De façon globale, en télédétection, il nous faut distinguer deux grands types de données (*cf.* figure 5.3) : les images optiques ($\lambda = [0.4 \mu m - 2 \mu m]$ environ) et les images radar ($\lambda = [1 cm - 100 cm]$ environ). Cette distinction conduit naturellement à trois catégories de fusion ayant des applications différentes :

- fusion Image Optique / Image Optique,
- fusion Image RADAR / Image RADAR,
- fusion Image Optique / Image RADAR.

D'autres types de longueur d'onde, comme l'infra-rouge thermique (IRT) et les rayons gamma servent aussi de base à la fusion.

Trois types de capteurs optiques fournissant des images de résolutions spatiales différentes co-existent :

- Les capteurs panchromatiques (P ou PAN) sont ceux ayant la meilleure résolution spatiale,
- les capteurs multispectraux (XS) fournissent plusieurs images spectralement distinctes permettant de réaliser des compositions colorées et ont une meilleure résolution spectrale,
- les capteurs *Grand Champ* ont la moins bonne résolution spatiale, mais la meilleure résolution temporelle.

Leurs caractéristiques sont synthétisées dans le tableau 5.1.

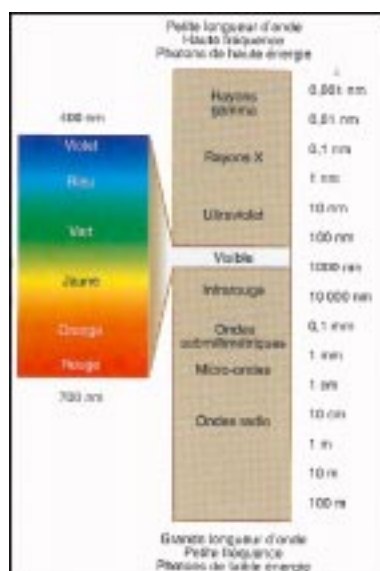


FIG. 5.3 – Spectre électromagnétique.

	Résolution spectrale	Résolution Spatiale	Résolution temporelle
Capteur PAN	Large ($\leq 0.2 \mu\text{m}$)	Bonne ($\leq 10 \text{ m}$)	Moyenne ($\sim 1 \text{ mois}$)
Capteur XS	Etroite ($< 0.1 \mu\text{m}$)	Moyenne ($\leq 30 \text{ m}$)	Moyenne ($\sim 1 \text{ mois}$)
Capteur grand champ	Etroite ($< 0.1 \mu\text{m}$)	Mauvaise ($\leq 1 \text{ km}$)	Bonne ($\sim 1 \text{ jour}$)

TAB. 5.1 – Caractéristiques des principaux types de capteurs optiques.

Le tableau 5.2, quant à lui, indique les fusions effectivement réalisées que l'on peut trouver dans la littérature.

	Capteur PAN	Capteur XS	Capteur grand champ
Capteur PAN		OUI	
Capteur XS	OUI		OUI
Capteur grand champ		OUI	

TAB. 5.2 – Types de fusion optique / optique recensés dans la littérature.

Fusion Capteur PAN / Capteur XS : Elle mélange des données de capteurs de bonne résolution spatiale, mais monospectrale (*PAN*) avec des capteurs de moins bonne résolution spatiale, mais de meilleure résolution spectrale et fournissant plusieurs images permettant une composition colorée (*XS*) (cf. figure 5.4). Le canal panchromatique est en général sensible dans les longueurs d'onde visibles, il est de ce fait peu corrélé avec les canaux multispectraux infra-rouge.

Cette fusion multi-résolution est la plus couramment employée en télédétection. Les données *PAN* et *XS* co-existent en effet depuis de nombreuses années et leur pérennité se poursuit avec les nouveaux satellites.

Le rapport de résolution spatiale entre les deux types d'images est assez faible (2 ou 3) bien que l'arrivée récente de capteurs panchromatiques très haute résolution tende à augmenter ce rapport.

Il y a en outre très peu de mélange temporel entre ces deux types d'images. En général, les images sont acquises à des dates proches ; dans le cas contraire, l'apport d'information est fait sans tenir compte des dates respectives des données.

Les applications sont de natures diverses (cartographie, croissance urbaine, environnement, agriculture, ...).

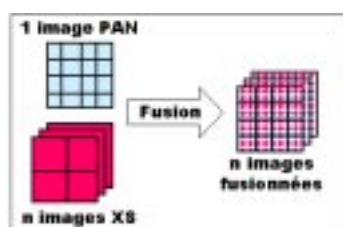


FIG. 5.4 – Objectifs de la fusion entre des données panchromatiques et des données multispectrales.

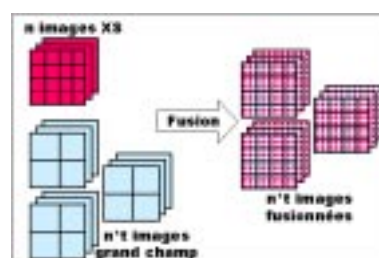


FIG. 5.5 – Objectifs de la fusion entre des données multispectrales et des données grand champ.

Fusion Capteur XS / Capteur Grand Champ : Ce type de fusion est plus marginal car des données adaptées sont moins courantes. En effet, les images fusionnées ont un

grand rapport de résolution spatiale et nécessitent la gestion des différences de résolution temporelle. De plus, si les gabarits spectraux des images *XS* et *Grand Champ* ne sont pas identiques, un problème de mélange spectral, bien que moins important que dans le cas de la fusion *P/XS*, se pose.

Actuellement, et depuis 1998, seuls les capteurs SPOT4-HRVIR (*XS*) et SPOT4-VEGETATION (*Grand Champ*) possèdent trois bandes spectrales identiques et permettent ainsi de ne gérer que les différences spatiale et temporelle.

Une méthode couramment employée pour étudier ces méthodes est la désagrégation de pixels mixtes [161, 163, 162, 164], [31], [20]. La méthode de superrésolution [100] a également été développée pour ce type de fusion.

L'application principale est le suivi des cultures agricoles.

Ce type de fusion est celui qui nous a conduit à développer notre pyramide morphologique.

Autres types de fusion :

Premièrement, des données *XS* sont parfois mélangées à des données hyperspectrales, de faible résolution spatiale, mais de haute résolution temporelle [84]. Cette étude a été faite sur des données simulées du capteur MERIS et des images du capteur LANDSAT TM [83, 84], [162].

Ensuite, des données multispectrales sont mélangées entre elles, la résolution spatiale des images n'est pas nécessairement différente, mais les techniques employées s'apparentent beaucoup aux méthodes multi-résolution commençant par un ré-échantillonnage de toutes les données à la même échelle. Le but de cette fusion est de faciliter les traitements ultérieurs (classification, segmentation) en rehaussant les informations importantes. Chavez et Kwarteng [15] mélangent les 6 bandes de Landsat TM pour en extraire une composition colorée interprétable par ACP. En 2001, Scheunders et De Backer [121] utilisent des ondelettes pour ce type de fusion.

De plus, la fusion entre une donnée panchromatique et des données *Grand Champ* peut au premier abord s'avérer très intéressante. En effet, à partir d'une image haute résolution spatiale et de données à hautes résolutions spectrale et temporelle, des images de synthèse à haute résolution spatiale, spectrale et temporelle pourraient être créées. Cela accumule cependant les difficultés et aucune méthode n'a encore été trouvée dans la littérature pour pallier les problèmes soulevés.

En outre, certains travaux réalisent la fusion de données entre des images optiques et radar. Notamment, Yésou *et al.* fusionnent des données SeaSat avec SPOT HRV [154] pour améliorer la visibilité des structures et teste deux méthodes, l'ACP et l'IHS dans lequel l'image Seasat remplace la mesure d'intensité. On peut également citer les travaux de Paradella *et al.* dans ce domaine : il fusionne des images Landsat TM et SAR [97].

L'objectif de la fusion visible / infra-rouge thermique, quand à elle, est de ramener la bande infra-rouge thermique à la résolution spatiale des bandes visibles. Par exemple, Liu et Moore [65] ramènent la bande TM6 de Landsat à 30 m de résolution. Zhukov *et*

al. [163] réalisent également ce type de fusion.

Enfin, d'autres types de fusion existent, citons notamment la fusion radar / rayon gamma, importante en géologie [97] et la fusion radar / radar [46].

La fusion, pour être performante, doit prendre en compte tous les paramètres. En général, les auteurs se focalisent sur un traitement de quelques paramètres seulement, et considère les autres comme identiques entre les deux images. La superposition («*registration*» en anglais) des images reste cependant le problème majeur de toutes les techniques de fusion [27].

Dans ces conditions, avec autant de paramètres aléatoires, l'intérêt et l'apport de la fusion par rapport à une étude séparée des données sont discutés dans [80].

5.1.3 Les différents types de données en imagerie médicale

En imagerie médicale, deux types d'images du cerveau sont utiles et montrent des détails différents. Les premières, appelées MR (Magnetic Resonance - résonance magnétique) permettent de caractériser les tissus, tandis que les secondes, appelées CT (Computed Tomographic) mettent en évidence les os.

Les images issues d'un seul processus ne suffisent pas aux demandes cliniques pour le diagnostic et le traitement. Cependant, observer séparément une série d'images multimodales d'un même objet n'est pas une meilleure solution. La fusion de ces données est donc une étape primordiale [89].

5.2 Revue des méthodes de fusion multirésolution

Cette section décrit les principales méthodes de fusion de données multi-résolution au niveau pixel. Nous détaillerons les méthodes dans leur contexte, en précisant si elles peuvent être étendues ou non à d'autre type de fusion.

Il n'existe actuellement pas de travaux comparant ces différentes méthodes de façon générale. Le travail présenté en annexe et réalisé en collaboration avec la société REALIX Technologies (Labège, France) et le CNES [10] traite de cette question. Les résultats n'en sont cependant pas encore publiés.

Dix «*méthodes*» de fusion ont été recensées dans la littérature. Elles sont basées sur les principes :

1. de la Combinaison Arithmétique de Bandes (CA),
2. du codage RVB (RVB),
3. des changements d'espaces IHS ou Lab (IHS, Lab),
4. du Filtrage Passe-haut (FPH),
5. de l'analyse en composantes principales (ACP),

6. des champs aléatoires markoviens (CAM),
7. de la régression linéaire,
8. de la désagrégation de pixels (DES),
9. de la transformée en ondelettes (TO),
10. des algorithmes pyramidaux (AP).

Elles sont décrites dans cette section et regroupées en famille en fonction des domaines physiques dont elles sont issues. Les trois familles sont les statistiques, la colorimétrie et le traitement du signal.

La section 5.3 réalise ensuite une synthèse des méthodes et des objectifs.

5.2.1 Méthodes Statistiques

La première méthode statistique décrite est la **désagrégation de pixels mixtes**. Cette méthode n'est pas à proprement parlé une méthode de fusion de données, cependant, c'est la méthode la plus utilisée pour améliorer l'interprétation de données grand champ et la caractérisation de cultures.

Le principe en est le suivant : l'utilisateur dispose de données multispectrales à haute résolution et d'une série temporelle de données *grand champ*. La problématique est la suivante : les données multispectrales sont peu nombreuses, et ne permettent donc pas de suivre l'évolution des cultures sur une saison, mais chaque pixel qui compose ces images est supposé pur, c'est-à-dire ne contenant qu'un type de végétation. Au contraire, certains pixels des images *grand champ* sont mixtes, c'est-à-dire qu'ils mélangent la réponse spectrale de plusieurs thèmes. L'hypothèse faite dans ce cas, est que la réflectance d'un pixel mixte est la moyenne des réflectances de chaque thème à la date considérée, pondérée par le pourcentage de surface que le thème occupe dans le pixel.

Deux applications existent. La première suppose que les données multispectrales permettent de cartographier les champs de la zone étudiée, mais ne permettent pas de reconnaître les cultures sous-jacentes. Dans ce cas, à chaque champ est associé une probabilité pour que la culture qu'il contient soit le thème i . Connaissant les réflectances théoriques de chaque thème à toutes les dates données, cette méthode exprime les réflectances théoriques de chaque pixel mixte à chaque date et réalise ensuite un calcul de régressions sur toutes les dates des données grand champ afin de calculer les probabilités pour que telle culture soit composée de tel ou tel thème.

La deuxième approche suppose que les données multispectrales sont suffisantes pour calculer un plan d'occupation des sols, et que les cultures puissent ainsi être identifiées. Dans ce cas, on cherche à déterminer la courbe de phénologie d'un thème d'intérêt et de la comparer avec la courbe théorique afin d'estimer la bonne croissance ou non de la culture, en fonction du pixel. Dans ce cas, les inconnues ne sont pas les thèmes, mais leur réflectance à la date t . Cette question est aussi résolue par désagrégation de pixels mixtes, par régression sur ceux-ci. Les travaux Zhukov *et al.* [161, 163, 162, 164], Minghelli *et al.* [83, 84], Faivre et Fisher [31] et Chechali *et al.* [20] sont les plus complets. Il existe également des méthodes réalisant une désagrégation sur les composantes spectrales et non

spatiales des images, c'est le cas de Van Der Meer [79, 81] et Hu *et al.* [48].

La deuxième méthode issue du domaine des statistiques est basée sur la théorie de l'**Analyse en Composantes Principales** (ACP). Les données d'entrée, quelconques, sont ré-échantillonnées à la même taille. Une ACP est réalisée sur toutes ou sur certaines données. Un nombre défini de composantes principales est ensuite gardé soit pour être le produit fusionné soit pour être combiné avec les données restantes. De nombreuses procédures existent.

Par exemple, Yésou *et al.* réalisent la fusion entre des données Seasat et SPOT HRV [154]. Les trois bandes multispectrales de SPOT - HRV et la bande Radar de Seasat servent d'entrée à l'ACP. Les composantes principales sont ensuite analysées : il apparaît que la deuxième composante exprime au mieux la différence entre les données SPOT et la bande Seasat. Une composition colorée entre les trois composantes principales peut être réalisée. Un autre exemple, encore donné par Yésou, réalise la fusion entre des données Landsat TM et SPOT PAN [155]. Les six bandes spectrales de Landsat TM sont combinées par une ACP. La composante la plus hautement corrélée avec le canal panchromatique est gardée (dans l'exemple, la composante 3). La composante principale 3 et le canal panchromatique sont convolués pour fournir l'image fusionnée.

L'interprétation de compositions colorées à partir de trois composantes principales peut être difficile, et de plus des informations intéressantes peuvent être perdues dans les autres composantes. Pour cela, certaines méthodes [15] pondèrent les images d'entrée de l'ACP.

Le troisième domaine statistique utilisé est celui de la **régression linéaire**. Deux approches existent. Le principe de la première est d'exprimer une image en fonction d'une autre par régression linéaire. Cette méthode apparaît à la fin des années 1980 [136, 103] et a été également utilisée par Pavel et Sharma [98].

Elle a été reprise et complétée notamment par Diemer et Hill [26]. Leur méthode se décompose en 3 étapes :

Etape 0 : Dégradation du canal panchromatique

Le canal panchromatique est dégradé de la haute résolution à la basse résolution, par une interpolation bilinéaire.

Etape 1 : Calcul du modèle

Sur des voisinages de taille 5×5 , il réalise une régression linéaire entre l'image panchromatique et le canal XS choisi. Cette étape fournit les coefficients de la régression (a : offset, b : gain, e : résidus). Le calcul du coefficient de corrélation (r) est également réalisé. a, b, e et r sont donc des images de la même taille que les données multispectrales.

Soit y le canal multispectral d'entrée et x le canal panchromatique dégradé. La formule de régression sur le voisinage s'écrit :

$$\begin{cases} y' = a + b \times x \\ e = y - y' \end{cases} \quad (5.1)$$

Etape 2 : Application du modèle

On suppose que cette régression, valable pour les voisinages à la basse résolution, est

également valable pour les voisinages correspondants à la haute résolution. La même équation de régression est alors appliquée. Les coefficients a , b et e sont conservés, la valeur du panchromatique prise est celle de l'image d'origine, on en déduit la valeur du canal multispectral interpolé.

La formule appliquée à l'étape 2 est donc :

$$y_{haute\ résolution} = a + b \times x_{haute\ résolution} + e \quad (5.2)$$

où $y_{haute\ résolution}$ est le canal XS fusionnée, a et b les coefficients de régression calculés, $x_{haute\ résolution}$ le canal panchromatique et e le résidu interpolé à la haute résolution.

Les zones de faible corrélation peuvent être traitées séparément.

Par extension, nous avons classé dans cette catégorie les méthodes statistiques basées sur des comparaisons locales entre les images [88].

La deuxième approche basée sur la régression linéaire consiste à définir les images issues des capteurs comme régression linéaire de la scène sous-jacente. Grâce à une approche bayésienne, il est alors possible d'estimer la scène par maximum de vraisemblance ou maximum *a posteriori* [129, 130, 98].

Enfin, le dernier type de méthode fait appel à la modélisation par **champ aléatoire markovien**. Le principe de cette fusion est le suivant : tout d'abord, la formule de Bayes permet d'exprimer la valeur de chaque pixel de l'image fusionnée en fonction de ceux d'une image d'entrée, connue. Mais cette formule ne peut être résolue ; une hypothèse de champ aléatoire markovien est alors faite et permet d'exprimer la valeur du pixel en fonction uniquement de son voisinage. Cette hypothèse permet de simplifier l'équation que l'on peut alors résoudre [66].

5.2.2 Méthodes colorimétriques

La première méthode consiste à réaliser un **codage RVB**. Les données d'entrée doivent être composées d'une seule image haute résolution spatiale et de trois images de résolution spatiale plus basse, mais de résolution spectrale plus élevée. Cette méthode s'applique donc dans le cas de la fusion de 3 bandes XS avec un canal panchromatique. Après avoir ré-échantillonnées toutes les images au même nombre de pixels, l'image à haute résolution spatiale remplace l'image de contenu spectral le plus proche. Peu efficace, cette méthode n'est plus actuellement utilisée. Toutin [137] la compare à d'autres méthodes.

Le deuxième type de fusion basé sur la colorimétrie fait appel au changement d'espace **RVB** \rightarrow **IHS (IHS)**. Les données d'entrée doivent être composées d'une seule image haute résolution spatiale et de trois images de résolution spatiale plus basse, mais de résolution spectrale plus élevée. Cette méthode s'applique donc dans le cas de la fusion de 3 bandes XS avec une donnée autre. Le principe de base de cette méthode est le suivant [147, 46, 28] :

- changement d'espace des données multispectrales de l'espace RVB (Rouge Vert Bleu) vers l'espace IHS (Intensity Hue Saturation en anglais, soit Intensité Teinte Saturation en français),
- remplacement de l'image I par l'image haute résolution,
- changement d'espace de IHS vers RVB.

L'espace IHS pouvant être défini de plusieurs façons, le passage de l'espace IHS à l'espace RVB peut s'effectuer de plusieurs façons [119], [54].

Cette méthode a subi de nombreuses améliorations, permettant notamment de fusionner plus de trois images basse résolution [6, 7] ou bien encore d'ajuster la dynamique des images d'intensité ou de saturation [9, 120]. Un ajustement peut également être réalisé avant la transformation RVB / IHS [97]. [49] réalise une transformation RGB / Lab au lieu d'une transformation RVB / IHS.

La dernière méthode détaillée dans cette catégorie est la fusion par **combinaison arithmétique de bandes spectrales**. Cette méthode est classée dans la partie colorimétrie car elle est souvent utilisée dans un but d'amélioration visuelle des données. De toutes les méthodes utilisées, c'est la plus simple, son efficacité dépend cependant des données d'entrée. Les deux types d'images sont mélangés par addition [13, 14] et / ou soustraction et / ou produit [159, 65, 64, 33], après ré-échantillonnage des données à la même taille. Cliche *et al.* [24] réalise une addition ou un produit en fonction des longueurs d'onde respectives des données haute et basse résolution. Pradines [102] réalise ce type de fusion sans ré-échantillonnage préalable.

La transformée de Brovey qui fusionne les bandes 2, 4 et 5 de Landsat TM avec le canal panchromatique de SPOT suivant la formule

$$\begin{aligned}
 ROUGE &= \frac{\text{bande 5}}{\text{bande 2} + \text{bande 4} + \text{bande 5}} * SPOT_{PAN} \\
 VERT &= \frac{\text{bande 4}}{\text{bande 2} + \text{bande 4} + \text{bande 5}} * SPOT_{PAN} \\
 BLEU &= \frac{\text{bande 2}}{\text{bande 2} + \text{bande 4} + \text{bande 5}} * SPOT_{PAN}
 \end{aligned} \tag{5.3}$$

est un exemple connu.

Parfois, les coefficients de pondération dépendent du canal panchromatique [102], ou seulement des canaux multispectraux (Transformée de Brovey) ou des deux [103].

5.2.3 Méthodes issues du traitement du signal

La première méthode de fusion issue du domaine du traitement du signal est basée sur le **filtrage passe-haut**. Le principe consiste à filtrer l'image haute résolution par un filtre passe-haut et d'injecter ensuite les hautes fréquences dans les images multispectrales. Cette méthode a été développées en premier par Schowengert [122].

Au premier abord, cette méthode semble être la plus naturelle. En effet, l'objectif premier de la fusion P / XS est d'injecter dans les données multispectrales les détails (hautes fréquences spatiales) visibles dans la donnée panchromatique. Cependant, en télédétection, le mélange spectral induit par cette méthode détériore la qualité spectrale du résultat.

C'est pourquoi, Chavez et Howell [14] propose de filtrer par un filtre passe-haut faible le canal panchromatique, puis d'ajouter le résultat aux bandes XS, afin de limiter l'apport d'information de l'image panchromatique et de mieux préserver l'information spectrale.

La deuxième méthode est basée sur la **transformée en ondelettes**. Elle permet de fusionner une image à haute résolution spatiale avec une image à basse résolution spatiale. La transformée en ondelettes d'une image transforme celle-ci en une image de plus basse résolution et trois images de détails représentant les détails horizontaux, verticaux ou diagonaux [69]. En continu, cette décomposition est exacte.

Une transformée en ondelettes est réalisée sur l'image à haute résolution spatiale. L'image basse résolution issue de la transformée est alors remplacée par celle du capteur basse résolution. Finalement, la transformée en ondelettes inverse est réalisée. Ce processus est répété pour tous les canaux basse résolution spatiale.

Il faut cependant faire attention à la corrélation entre les longueurs d'onde du canal multispectral choisi et du panchromatique. En fait, les techniques réellement employées améliorent ce système. Wald *et al.* [107, 108, 109], Nunez *et al.* [92, 91, 93, 94] et Yocky [156, 157] sont les principales équipes travaillant sur la fusion de données de télédétection par ondelettes. D'autres travaux existent [39, 160, 121, 105, 104].

Le dernier type de méthode nous intéresse plus particulièrement, il s'agit de celui basé sur les **algorithmes pyramidaux**.

Les algorithmes pyramidaux fournissent des méthodes très variées. D'une façon très générale, une image (ou deux) est (sont) décomposée(s) par un algorithme pyramidal en éléments de résolution spatiale décroissante. Les données extraites de chaque image sont ensuite combinées. Plusieurs approches existent.

La première est utilisée dans un processus de fusion P / XS :

Stefano Baronti, Luciano Alparone et Bruno Aiazzi ont développé une méthode appelée Generalized Laplacian Pyramid (GLP) ([2] et [1]). Les données d'entrée de l'algorithme sont une donnée panchromatique et une donnée multispectrale, correspondant au canal que l'on veut traiter. Cette méthode est une généralisation de la pyramide de Laplacien à des facteurs d'échelle rationnels, et notamment pas nécessairement à une puissance de 2. Le filtrage et le sous-échantillonnage sont fait en une seule étape. La fenêtre de filtrage n'est pas glissante sur tous les pixels, mais juste à ceux correspondant à un pixel basse résolution.

Deux fonctions sont utiles : la fonction « reduce » et la fonction « expand ». La fonction « reduce » réduit la taille d'une image d'un facteur q donné. La fonction « expand » quant à elle augmente la taille d'une image d'un facteur p donné.

Ainsi, pour dégrader une image d'un facteur $\frac{p}{q} > 1$ (opération $reduce_{p/q}$), on combine les fonctions « expand » par q et « reduce » par p . Le traitement commence toujours par l'expansion, afin de réduire la perte d'information de l'image dégradée. Cette image peut ensuite être interpolée à la taille de l'image de départ par le processus « expand » par p suivi de « reduce » par q (opération $expand_{p/q}$).

Le processus de fusion est le suivant :

Etape 1 : Traitement sur le canal Panchromatique, de taille $(\frac{p}{q}LIG \times \frac{p}{q}COL)$

On réalise une décomposition par la pyramide de Laplacien généralisée. Si l'on note le canal panchromatique G_0 , on calcule les images de Laplacien aux deux premiers niveaux :

$$\begin{cases} \widetilde{L}_0 = PAN - expand_{\frac{p}{q}}\{reduce_{\frac{p}{q}}(PAN)\} \\ \widetilde{L}_1 = reduce_{\frac{p}{q}}(PAN) - expand_{\frac{p}{q}}\{reduce_{\frac{p}{q}}(reduce_{\frac{p}{q}}(PAN))\} \end{cases} \quad (5.4)$$

\widetilde{L}_0 est donc de taille $(\frac{p}{q}LIG \times \frac{p}{q}COL)$ et \widetilde{L}_1 de taille (LIG, COL) .

Etape 2 : Traitement sur la donnée multispectrale, de taille (LIG, COL)

On interpole le canal multispectral XS_l à la taille du canal panchromatique par expansion par p et réduction par q en $XS_{upgrade}$.

Etape 3 : Calculs finaux

On calcule un coefficient d'ajustement w_l à partir de la bande spectrale XS_l et de \widetilde{L}_1 .

On en déduit :

$$XS_{l, fusionne} = w_l \times \widetilde{L}_0 + XS_{upgrade} \quad (5.5)$$

avec

$$w_l = \sqrt{\frac{\sum_{m=0}^{M-1} \sum_{n=0}^{N-1} var[XS_l(m, n)]}{\sum_{m=0}^{M-1} \sum_{n=0}^{N-1} var[\widetilde{L}_1(m, n)]}}$$

la variance étant calculée sur un voisinage 7×7 ($M=N=7$).

Ainsi, le schéma global peut se synthétiser par la formule :

$$XS_{l, fusionne} = \sqrt{\frac{\sum_{m=0}^{M-1} \sum_{n=0}^{N-1} var[XS_l(m, n)]}{\sum_{m=0}^{M-1} \sum_{n=0}^{N-1} var[\widetilde{L}_1(m, n)]}} \times \widetilde{L}_0 + expand_{\frac{p}{q}}(XS_l) \quad (5.6)$$

Soit :

$$\begin{aligned} XS_{l, fusionne} = & \\ & \sqrt{\frac{\sum_{m=0}^{M-1} \sum_{n=0}^{N-1} var[XS_l(m, n)]}{\sum_{m=0}^{M-1} \sum_{n=0}^{N-1} var[\widetilde{L}_1(m, n)]}} \times (PAN - expand_{\frac{p}{q}}\{reduce_{\frac{p}{q}}(PAN)\}) \\ & + expand_{\frac{p}{q}}(XS_l) \end{aligned} \quad (5.7)$$

Le filtre est un filtre passe-bande de fréquence de coupure égal au rapport de résolution entre les deux données. En général, c'est un filtre gaussien. Les filtres sont différents pour les différentes bandes spectrales.

La deuxième est utilisée pour la fusion de données *CT* et *MR* :

Cette méthode est employée par Marshall et Matsopoulos en 1994 dans laquelle les images *CT* et *MR* subissent le même traitement [77], [78].

1. Chaque image est décomposée par itération d'un processus de filtrage morphologique et d'échantillonnage. On crée ainsi deux pyramides, la pyramide *CT* et la pyramide *MR*.
2. A partir de chacune de ces pyramides, une pyramide de différence est créée de la façon suivante :
 - à chaque niveau correspond une image de différence,
 - l'image de la pyramide *CT* (resp. *MR*) du plus bas niveau est interpolée à la taille de l'image du niveau immédiatement supérieur,
 - le résultat est lissé,
 - la différence entre le résultat lissé et l'image de la pyramide *CT* (resp. *MR*) du niveau immédiatement supérieur est calculée.
 L'ensemble de ces différences constitue deux nouvelles pyramides : la pyramide de différence *CT* et la pyramide de différence *MR*.
3. A chaque étage, on compare les deux pyramides de différences. L'image résultat est le maximum pixel à pixel à chaque niveau. La pyramide ainsi créée synthétise les hautes fréquences des deux images et est appelée pyramide différence intermédiaire.
4. La reconstruction se fait par itération du processus {interpolation, lissage de l'image de la pyramide différence intermédiaire} et ajout de toutes les images résultats.

Le filtre morphologique est par exemple une combinaison ouverture-fermeture par un élément structurant 3×3 , avec un échantillonnage dyadique.

La troisième est utilisée pour la fusion de données *CT* et *MR* :

Nous décrivons ici la méthode employée par Mukhopadhyay et Chanda en 2001. Les images de départ, *CT* et *MR* ont un nombre de pixels identiques. Les deux images subissent donc le même traitement [89].

1. Chaque image est filtrée par des ouvertures et des fermetures d'éléments structurants grossissants. Pour chaque image, à chaque niveau d'éléments structurants on obtient donc deux images ; la première est le résultat de l'ouverture, la deuxième celui de la fermeture.

2. Les différences entre les niveaux sont calculées dans chaque cas. A chaque niveau correspond quatre images différences. Pour chacune des deux images obtenues précédemment, une image différence provient de la différence des ouvertures et l'autre de la différence des fermetures.
3. A chaque niveau, on calcule le maximum pixel à pixel entre les différences issues des ouvertures (resp. des fermetures) des deux images. On obtient donc deux images à chaque niveau, une synthétisant les différences des ouvertures, l'autre les différences des fermetures.
4. On somme les différences des ouvertures (S^{OP}) et les différences des fermetures (S^{CL}) à tous les niveaux. (OP est mis pour OPENING (*ouverture*) et CL pour CLOSING (*fermeture*)).
5. M_X est calculé comme étant le maximum pixel à pixel entre l'image CT fermée par le plus gros élément structurant et l'image MR fermée par le plus gros élément structurant.
 M_N est calculé comme étant le minimum pixel à pixel entre l'image CT ouverte par le plus gros élément structurant et l'image MR ouverte par le plus gros élément structurant.
 On calcule ensuite $A = moyenne(M_X, M_N)$. A représente l'image basse fréquence synthétisant les images.
6. Le produit fusionné est calculé par

$$\text{Produit Fusionné} = A + \frac{1}{2}S^{OP} - \frac{1}{2}S^{CL}$$

5.3 Synthèse des processus de fusion d'images

Dans cette partie, les méthodes seront référencées par le sigle entre parenthèses donné au paragraphe 5.2.

Les méthodes de fusion sont principalement basées sur deux principes de fonctionnement (*cf.* tableau 5.3) : l'injection d'éléments à haute résolution dans les éléments à basse résolution ou le mélange spectral et / ou colorimétrique.

Les méthodes basées sur l'injection de structures mettent en valeur les caractéristiques spatiales différenciant les images haute et basse résolution et mettent donc l'accent sur les éléments spatiaux de haute résolution. Au contraire les méthodes basées sur les mélanges spectraux ou colorimétriques mettent en avant l'aspect visuel des images multispectrales.

Le tableau 5.4 recense les types d'images fusionnées en fonction de la méthode utilisée. Les références multiples au sein d'un même crochet regroupent les travaux d'une même équipe. Ce tableau ne se prétend pas exhaustif, mais présente néanmoins un large

Injection d'éléments haute résolution	Mélange spectral ou colorimétrique	Autres
FPH	ACP	CA
TO	IHS	CAM
AP	Lab	Désaggrégation
RL	RVB	

TAB. 5.3 – Fonctionnement des différentes méthodes de fusion multi-résolution recensées.

éventail de méthodes. Certaines méthodes combinent les différents types de processus et se situent ainsi plusieurs fois dans une même colonne, alors qu'elles ne sont pas citées dans les références types d'un schéma de fusion. C'est notamment le cas de Chitwong et Cheevasuvit [22] dont la méthode utilise la transformation RVB / IHS, le filtrage passe-haut et l'Analyse en Composantes Principales. D'autres articles se situent également sur plusieurs lignes, ce sont ceux comparant plusieurs méthodes. Leurs résultats seront détaillés dans la suite.

On peut également rajouter les travaux de Sharma et Pavel [129] qui fusionnent des images provenant de caméra TV et de capteur IR et de Filiberti *et al.* [33] sur la fusion de données panchromatique et hyperspectrale par combinaisons arithmétiques de bandes. Les éléments suivants ressortent de ce tableau :

- Les méthodes basées sur les transformations RVB / IHS sont très utilisées.
- Les méthodes basées sur la transformée en ondelettes connaissent maintenant un fort développement dans le cadre de la fusion de données panchromatique et multispectrales.
- Les méthodes basées sur la combinaison arithmétique de bandes et sur les filtrages passe-haut sont très utilisées, mais uniquement dans le cadre de la fusion de données panchromatique et multispectrales.
- Les méthodes basées sur les champs aléatoires markoviens sont marginales.
- La méthode de fusion à base d'algorithme pyramidal que nous proposons dans ce mémoire est originale et novatrice : elle est la seule se situant à l'intersection avec la fusion de données multispectrales et grand champ. Cette remarque témoigne de la difficulté que nous aurons pour évaluer son efficacité. Il est cependant important de citer à cet endroit les travaux de Poglio *et al.* [100] pour fusionner le même type de données avec une méthode baptisée superrésolution.

Quelques auteurs ont essayé de comparer les différentes méthodes et leurs effets sur les images fusionnées. Le tableau 5.5 décrit les méthodes comparées par différents auteurs. On constate qu'ils comparent de 2 à 4 méthodes, pas nécessairement les mêmes. Les méthodes de fusion par algorithme pyramidal et par champ aléatoire markovien ne sont jamais comparées avec les autres. Au contraire, la méthode par changement d'espace RVB / IHS sert de référence à toutes les autres et est systématiquement étudiée. Dans ces conditions, il est difficile d'évaluer l'efficacité des méthodes par algorithme pyramidal et par champ aléatoire markovien.

	P / XS	XS / Grand champ	XS / Hyper- spec- tral	XS / XS	RADAR / RA- DAR	RADAR / VIR	VIR / IRT	RADAR / Rayon Gamma	XS / Gamma	MR / CT
Désagreg.	[161, 162]	[31], [20]	[163], [84, 83]	[163, 164], [79, 81], [48]			[163, 164]			
CA	[24], [102], [13, 14], [153], [33], [159], [64], [143], [140]									
RVB	[137]									
IHS	[147], [9], [16], [6, 7], [137], [111], [143], [92], [140], [22], [120], [28], [49]				[46]	[154], [97], [46], [152]	[97], [46]	[120]		
FPH	[122], [14, 16], [155], [137], [111], [22], [140]									
ACP	[16], [155], [137], [111], [22], [140]			[15]		[154], [152]				
CAM	[66]									
Régression linéaire	[136], [103], [26]			[88], [136], [129], [130]	[98]	[98]	[136]			
TO	[104], [107, 108, 109], [92, 91, 93, 94], [156, 157], [39], [111], [160], [140], [22], [121]			[121]						
AP	[2, 1]	[36, 56, 59, 60]								[77, 78], [89]

TAB. 5.4 – Classement des différentes méthodes de fusion en fonction des types d'images sur lesquels elles sont utilisées.

	[16]	[33]	[137]	[157]	[111]	[143]	[160]	[159]	[140]
Désag.									
CA		×				×	×	×	×
RVB			×						
IHS	×		×	×	×	×	×	×	×
FPH	×		×		×				
ACP	×		×		×		×	×	×
CAM									
TO				×	×		×		×
AP									

TAB. 5.5 – Les méthodes comparées par différents auteurs.

De façon générale, il ressort cependant que les méthodes ACP et FPH distordent les caractéristiques spectrales et que la transformée en ondelettes préserve au mieux la qualité spectrale et l'énergie du signal.

Pour Filiberti *et al.* [33], les méthodes [24], [147], [13], [16], [9] ont une bonne interprétation visuelle tandis que [102] et [103] ont une meilleure précision radiométrique.

Le tableau 5.6 sépare les méthodes effectuant un traitement local sur un voisinage glissant sur toute l'image traitée de celles effectuant un traitement global. De façon générale, les méthodes réalisant un traitement global sont plus rapides, mais moins adaptables et moins efficaces que celles réalisant un traitement global.

Méthode globale	Méthode locale
ACP	FPH
IHS	TO
RVB	AP
CA	
Désagrégation	
CAM	

TAB. 5.6 – Classement des différents traitements en fonction du calcul local ou global qu'ils réalisent.

Enfin, les méthodes issues du domaine de la colorimétrie sont généralement axées sur la qualité visuelle des images fusionnées, celles issues du domaine du traitement du signal sur la qualité spectrale, tandis que celles basées sur les statistiques réalisent des hypothèses reliant les images basse et haute résolution.

5.4 Conclusion

Les méthodes de fusion sont principalement basées sur trois types de théorie provenant du traitement du signal, de la colorimétrie et des statistiques. Nous avons ainsi pu identifier 10 types majeurs de méthodes. Parfois, les méthodes mélangent ces divers processus. De ce fait, il ressort de cette étude que les méthodes de fusion sont nombreuses en télédétection et n'ont jamais été comparées de façon systématique entre elle. Les études déjà réalisées montrent cependant qu'aucune méthode ne serait meilleure qu'une autre : tout dépend de l'objectif de la fusion (qualité visuelle / qualité spectrale) ainsi que des types de surfaces observées.

Chapitre 6

La fusion de données par pyramide morphologique

La nomenclature des termes utilisés dans ce chapitre est décrite dans le tableau 2.

6.1 Principe

Le principe de fusion de données développé dans ce chapitre est basé sur le concept de la pyramide morphologique selon les caractéristiques suivantes.

Les données d'entrée du processus de fusion comprennent au minimum deux images, une à haute résolution et l'autre à basse résolution, sur une même zone géographique et si possible acquises dans les mêmes longueurs d'onde.

La partie décomposition de la pyramide est réalisée sur l'image haute résolution jusqu'à ce que la dernière image filtrée soit de même taille que l'image basse résolution provenant du deuxième capteur [56, 36].

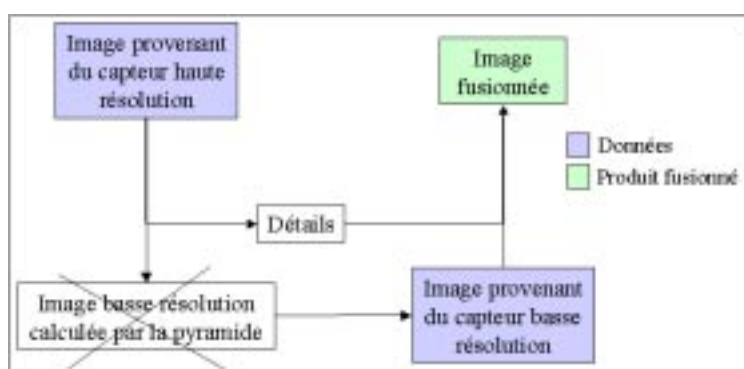


FIG. 6.1 – Schéma de la fusion de données par pyramide morphologique.

Le nombre d'étages et le pas d'échantillonnage sont fixés par le rapport de résolution entre les deux images. De plus, l'échantillonnage doit être défini par un carré (de taille

$n_i \times n_i$ en fonction de l'étage) de façon à respecter la proportionnalité entre les lignes et les colonnes haute et basse résolutions.

Supposons par exemple que l'image haute résolution ait une résolution de 1 m, et l'image basse résolution de 8 m, une solution envisageable consiste à réaliser une pyramide de trois étages, avec un pas de décimation dyadique. Au contraire, si la résolution spatiale de l'image basse résolution est de 9 m, une pyramide de deux étages avec un pas de décimation 3×3 est plus adaptée. S'il n'y a pas de rapport de résolution simple entre les deux, il faut soit dégrader une image pour avoir des rapports de résolution convenables, soit introduire des rapports de résolutions entre étages non constants, voire non entiers.

Ensuite, les images « dernière image filtrée » et « image basse résolution provenant du capteur » sont interverties. La reconstruction est faite en prenant comme base l'image du capteur basse résolution (cf. figure 6.1). L'image reconstruite est alors le produit fusionné entre une image basse et une image haute résolution. Les choix du filtre, de l'élément structurant, de l'algorithme d'échantillonnage et de ré-échantillonnage restent à faire.

6.2 Mise en oeuvre

6.2.1 Calcul théorique de l'image fusionnée

Dans le cas général, l'équation de fusion se synthétise de façon corollaire à celle de la reconstruction (cf. équation (3.45)) par

$$IFusion_j = D_j + Int_j(D_{j+1} + Int_{j+1}(D_{j+2} + \dots + Int_{N-2}(D_{N-1} + Int_{N-1}(IBR)))) \quad (6.1)$$

où $IFusion_j$ représente l'image fusionnée au niveau j de la pyramide. Cependant, si l'algorithme d'interpolation utilisé satisfait la condition $Int_k(X + Y) = Int_k(X) + Int_k(Y)$, l'image fusionnée peut alors s'exprimer de façon corollaire à l'algorithme décrit par l'équation (3.47),

$$IFusion_0 = D_0 + \widetilde{D}_1^0 + \widetilde{D}_2^0 + \dots + \widetilde{D}_{N-1}^0 + \widetilde{IBR}^0 \quad (6.2)$$

où IBR est l'image provenant du capteur basse résolution et D_i les détails du niveau i , obtenus par décomposition de l'image du capteur haute résolution.

Au niveau 0 sans fusion, l'équation (3.47) s'écrit :

$$IR_0 = D_0 + \widetilde{D}_1^0 + \widetilde{D}_2^0 + \dots + \widetilde{D}_{N-1}^0 + \widetilde{I}_N^0 \quad (6.3)$$

Ainsi, en se souvenant que $IR_0 = I_0$ (cf. équation (3.18)), le calcul de la différence

entre l'image haute résolution initiale et l'image haute résolution fusionnée donne

$$\begin{aligned}
I_0 - IFusion_0 &= IR_0 - IFusion_0 \\
&= D_0 + \widetilde{D}_1^0 + \widetilde{D}_2^0 + \dots + \widetilde{D}_{N-1}^0 + \widetilde{I}_N^0 \\
&\quad - (D_0 + \widetilde{D}_1^0 + \widetilde{D}_2^0 + \dots + \widetilde{D}_{N-1}^0 + \widetilde{IBR}^0) \\
&= \widetilde{I}_N^0 - \widetilde{IBR}^0
\end{aligned} \tag{6.4}$$

De ce fait, en théorie, dans le cas où l'interpolation choisie satisfait l'hypothèse de départ - ce qui est le cas pour les algorithmes de duplication et d'interpolation bilinéaire - la différence entre l'image initiale et l'image fusionnée ne dépend que des deux images basse résolution, celle calculée par la pyramide et celle provenant du capteur et du schéma d'interpolation, mais pas des détails injectés.

Cela a plusieurs conséquences :

- a) Si l'interpolation réalisée est une simple duplication de pixels, les images \widetilde{I}_N^0 et \widetilde{IBR}^0 sont constituées de paliers de côté $\prod_{i=0}^{N-1} n_i$ où n_i est le pas de décimation au niveau i et N le nombre d'étages de la pyramide. La différence $I_0 - IFusion_0$ est donc elle aussi uniquement constituée de paliers de la même taille dépendant de la différence entre les deux images basse résolution I_N et IBR .
- b) Pour que la différence $I_0 - IFusion_0$ soit petite, il faut minimiser la différence entre I_N et IBR .

6.2.2 Codage des images sur 8 bits

Les images employées sont codées sur 8 bits, donc entre 0 et 255. Localement, l'image du capteur basse résolution et la dernière image filtrée peuvent être significativement différentes. Cela se produit notamment dans les zones où le processus de décomposition a filtré un objet présent sur l'image du capteur basse résolution, c'est-à-dire dans les zones de fort gradient. De même, cette situation se rencontre lorsque les images n'ont pas fait l'objets de corrections photométriques par la même chaîne de prétraitement (ce qui est souvent le cas en observation de la Terre car les données ne sont pas toujours fournies en réflectance). De ce fait, lors de la reconstruction, il peut arriver que l'ajout des détails D_{sup} fournisse localement un compte numérique supérieur à 255 et que la soustraction des détails D_{inf} crée des comptes numériques négatifs.

Ce problème peut être pris en compte de trois façons.

- On ne tient pas compte du codage des images sur 8 bits. L'image fusionnée est alors appelée IHR_{SS} (Image Haute Résolution Sans Saturation). On a

$$IFusion_0 = IHR_{SS}$$

Dans le cas d'une interpolation par simple duplication, la différence est effectivement sous forme de paliers $\prod_{i=0}^{N-1} n_i \times \prod_{i=0}^{N-1} n_i$.

- On tient compte du codage sur 8 bits en saturant toutes les valeurs de $IFusion_0$ sortant du cadre $[0 - 255]$. Toutes les valeurs négatives sont mises à 0 et toutes les valeurs supérieures à 255 sont mises à 255. L'équation (6.2) n'est plus valable dans les zones où il y a saturation. Dans ces zones là, la différence n'est plus sous forme de paliers, même dans le cas d'une interpolation par duplication. L'image obtenue est appelée IHR_{AS} (Image Haute Résolution Avec Saturation).
- On tient compte du codage sur 8 bits en effectuant une translation puis une contraction du compte numérique. Si la plus petite valeur de l'image fusionnée est $-MIN$, avec $MIN > 0$, alors on transforme l'image fusionnée en $I_{translatée} = IFusion_0 + MIN$. Ensuite, si MAX est le maximum de $I_{translatée}$, alors l'image codée sur 8 bits est calculée par homothétie par $I_{tr+ho} = I_{translatée} \times \frac{MAX}{255}$. Ainsi, $I_{tr+ho} = \frac{MAX}{255} \times IFusion_0 + MIN \times \frac{MAX}{255}$. La différence s'exprime par

$$IR_0 - I_{tr+ho} = \underbrace{(IR_0 - IFusion_0)}_{\text{Constante par palier}} - \underbrace{MIN \times \frac{MAX}{255}}_{\text{Constante}} + \underbrace{\left(1 - \frac{MAX}{255}\right) \times IFusion_0}_{\text{Non Constant}}$$

La contraction perd totalement la propriété de paliers de l'image différence. L'histogramme de cette image est distordu.

6.3 Choix des paramètres de la pyramide

L'entrée de l'algorithme de fusion est composée de deux images, l'une à haute résolution et l'autre à basse résolution. La décomposition de la pyramide va être réalisée de sorte que I_N ait la même résolution que l'image issue du capteur basse résolution, IBR . Cette contrainte va limiter le choix à la fois du pas d'échantillonnage et du nombre d'étages de décomposition à faire.

Les paramètres restant à évaluer sont le filtre, l'élément structurant, les algorithmes d'interpolation et de décimation. Pour évaluer l'impact de ces paramètres, supposons que le jeu de données d'entrée soit composé d'une image haute résolution acquise à une date connue et une image basse résolution, acquise à la même date, sur la même scène. L'algorithme de fusion idéal fournirait en sortie une image fusionnée identique à l'image d'entrée haute résolution. Sur un tel jeu de données, les paramètres optimaux sont ceux minimisant la différence entre l'image initiale et l'image fusionnée. Comme le montre l'équation (6.4), cette différence dépend de la différence entre I_N et IBR et de l'algorithme d'interpolation uniquement. Cette constatation témoigne de la différence entre le choix des paramètres pour l'analyse et pour la fusion de données. Notamment, le schéma d'interpolation revêt ici une importance capitale, alors qu'il n'a aucune influence dans le processus d'extraction d'éléments lors de l'analyse.

IBR étant fixé, minimiser la différence $I_N - IBR$ revient à calculer les paramètres de la décomposition rendant I_N le plus proche possible de IBR . Le choix du filtre et de

l'élément structurant doivent être adaptés à cette contrainte.

Tout d'abord, un capteur basse résolution ne va favoriser ni les éléments plus clairs ni ceux plus foncés que l'environnement. Un filtre adapté à cette situation est $\frac{\varphi+\gamma}{2}$. Les tests des différents paramétrages de la pyramide morphologique détaillés à la section 6.4 confirme ce choix. De même, le capteur basse résolution va traiter de façon isotrope la zone géographique correspondant au pixel basse résolution.

Ainsi, l'élément structurant doit être symétrique et d'origine centrée. Il faut définir un rayon d'élément structurant qui filtre de façon équivalente au filtrage correspondant au capteur basse résolution. Les différents tests réalisés montrent que, pour un échantillonnage dyadique, un élément structurant de rayon minimal 1 filtre déjà légèrement trop. En conclusion, le filtre retenu sera $\frac{\varphi+\gamma}{2}$ avec un élément structurant de rayon juste supérieur au pas de décimation à l'étage considéré.

Le choix de l'algorithme d'échantillonnage sera fait grâce aux résultats des tests de la section 6.4. Une fois la différence entre I_N et IBR minimisée, seul l'algorithme de ré-échantillonnage va modifier le résultat. Plus l'algorithme calculera les valeurs manquantes sur de gros voisinages, plus la différence entre les deux images sera lissée. Le rendu visuel sera donc très différent d'un algorithme à l'autre.

6.4 Evaluation de la qualité de la fusion

Ainsi que nous l'avons déjà spécifié dans le chapitre 5 sur l'état de l'art des méthodes de fusion multi-résolution, aucune méthode ne répond aux mêmes spécificités que celle de notre étude. Nous avons donc basé notre évaluation sur deux critères :

- influence du choix des paramètres et du codage réalisé sur les résultats et validation des choix théoriques,
- comparaison avec d'autres techniques de fusion par algorithmes pyramidaux implémentés pour la circonstance.

6.4.1 Evaluation des résultats

Dans la littérature, dans le cadre de la fusion de données optiques, peu d'indicateurs ont été mis au point pour tester la qualité visuelle et spectrale des images résultats ([71] et [110]). Les indicateurs utilisés sont les suivants :

- erreur relative sur la moyenne,
- erreur relative sur la variance ou l'écart-type,
- coefficient de corrélation,
- écart-type de la différence,
- probabilité d'avoir en un pixel une erreur relative inférieure à un seuil donné.

Ils sont calculés de façon globale sur toute l'image.

La comparaison de deux images n'est pas une chose aisée car de nombreux paramètres peuvent être étudiés. Cette étude est séparée en deux : la première partie étudie les paramètres algorithmes de décimation, filtre, élément structurant et nombre d'étages, tandis que la deuxième étudie l'influence de l'algorithme d'interpolation. Cette séparation des paramètres est possible, car on a vu que l'algorithme d'interpolation ne modifie pas les images issues de la décomposition, et que plus les images IBR et I_N étaient proches, plus l'image fusionnée le serait de l'image de référence.

6.4.1.1 Les paramètres algorithmes d'échantillonnage, filtre, élément structurant et nombre d'étages

L'algorithme d'interpolation est systématiquement une duplication de pixels. Dans cette étude, nous avons défini 13 indicateurs globaux puis 7 indicateurs locaux. Les premiers fournissent une indication de la ressemblance entre les images initiale et fusionnée, tandis que les deuxièmes fournissent une localisation des zones où les différences sont les plus importantes. Nous supposons que nous avons toujours une image de référence, nommée f , que la deuxième image, g , doit approcher au mieux. Les images sont codées sur 8 bits, les valeurs des pixels (x, y) sont donc entières, positives ou nulles.

Les indicateurs globaux Les indicateurs globaux retenus sont indiqués dans le tableau 6.1 de la page 119. Les indicateurs IG_5 , IG_{11} et IG_{13} n'ont pas de sens si l'image de référence est nulle en tout point $f(x, y)$. Cela n'est pas pénalisant car comparer une image avec une image nulle ne sera jamais fait.

Dans un premier temps, les propriétés théoriques de ces indicateurs sont étudiées. Leur propriété de symétrie est montrée sur le tableau 6.2. Deux indicateurs sont dits symétriques si $Indicateur(f, g) = Indicateur(g, f)$. Les indicateurs symétriques traitent indifféremment l'image fusionnée et la référence ; les autres prennent en compte le fait que l'image de référence contient les « vraies » valeurs. En général, les indicateurs non symétriques proviennent du domaine de la physique, les symétriques des mathématiques.

Le tableau 6.3 montre l'influence d'une translation et d'une homothétie sur l'image g . f est considérée comme référence et donc comme constante. Les résultats s'interprètent comme suit : a et b sont respectivement des constantes de translation et d'homothétie. Le symbole \ signifie qu'il n'existe pas de lien simple entre $Indicateur(f, g)$ et $Indicateur(f, ag)$, $Indicateur(f, g+b)$ et $Indicateur(f, ag+b)$. *Inchangé* signifie $Indicateur(f, g) = Indicateur(f, ag)$... Une valeur signifie que le résultat est modifié par ajout ou multiplication par le terme indiqué.

Les indicateurs sont ensuite classés suivant leur nature.

1. Les distances (IG_1 , IG_3 et IG_{12})

Définition 12 Une distance est définie sur un ensemble E , si, étant donnée une application d de $E \times E$ dans l'ensemble des réels ≥ 0 telle que, quel que soit x, y, z dans E :

Distance d_1 $IG_1 = \frac{1}{255 \times n \times m} \sum_{x=1}^{lig} \sum_{y=1}^{col} f(x, y) - g(x, y) $	Différence du carré $IG_2 = \frac{1}{255} \sqrt{\frac{\sum_{x=1}^{lig} \sum_{y=1}^{col} (f(x, y) - g(x, y))^2}{n \times m}}$
Différence symétrique 1 $IG_3 = \frac{1}{n \times m} \sum_{x=1}^{lig} \sum_{y=1, f(x, y)g(x, y) \neq 0}^{col} \frac{ f(x, y) - g(x, y) }{\sup(f(x, y), g(x, y))}$	Différence symétrique 2 $IG_4 = \frac{1}{n \times m} \sum_{x=1}^{lig} \sum_{y=1, f(x, y)g(x, y) \neq 0}^{col} \frac{(f(x, y) - g(x, y))^2}{\sup(f(x, y), g(x, y))^2}$
Erreur relative sur la moyenne $IG_5 = \begin{cases} \frac{\text{moyenne}(f) - \text{moyenne}(g)}{\text{moyenne}(f)} & \text{si } f \text{ non nul} \\ \infty & \text{si } f \text{ est identiquement nul} \end{cases}$	
Erreur relative sur l'écart type $IG_6 = \begin{cases} \frac{\text{écart-type}(f) - \text{écart-type}(g)}{\text{écart-type}(f)} & \text{si } \sigma(f) \neq 0 \\ \infty & \text{si } \sigma(f) = 0 \end{cases}$	
Moyenne de l'image différence $IG_7 = \frac{1}{255} \text{moyenne}(f - g)$	Ecart-type de la différence $IG_8 = \frac{1}{255} \text{écart-type}(f - g)$
Valeur maximale de la différence $IG_9 = \frac{\max f - g }{255}$	Indice de corrélation entre les deux images $IG_{10} = \rho(f, g) = 1 - \frac{\text{cov}(f, g)}{\sigma(f)\sigma(g)}$
Corrélation entre les deux images $IG_{11} = \begin{cases} 1 - \frac{\sum f(x, y)g(x, y)}{\sum f(x, y)^2} & \text{si } f \text{ non nul} \\ \infty & \text{si } f \text{ identiquement nul} \end{cases}$	Contraste $IG_{12} = \frac{1}{n \times m} \sum_{x=1}^{lig} \sum_{y=1, f(x, y)g(x, y) \neq 0}^{col} \frac{ f(x, y) - g(x, y) }{f(x, y) + g(x, y)}$
$IG_{13} = \text{Pourcentage}(\frac{ f-g }{f} < \text{seuil})$ pour f non identiquement nul, indique le pourcentage de pixels dont l'erreur relative est inférieure à un seuil donné.	

TAB. 6.1 – Liste des indicateurs globaux utilisés : f est l'image de référence, g l'image à comparer. f et g sont de taille (lig, col) et le pixel courant est (x, y) .

- 1) $d(x, y) = 0$ est équivalent à $x = y$,
- 2) $d(x, y) = d(y, x)$,
- 3) $d(x, z) \leq d(x, y) + d(y, z)$.

	Symétrique	Non Symétrique		Symétrique	Non Symétrique
IG_1	✓		IG_8	✓	
IG_2	✓		IG_9	✓	
IG_3	✓		IG_{10}	✓	
IG_4	✓		IG_{11}		✓
IG_5		✓	IG_{12}	✓	
IG_6		✓	IG_{13}		✓
IG_7		$I_7(f, g) = -I_7(g, f)$			

TAB. 6.2 – Etude de la propriété de symétrie des différents indicateurs.

	$a \times g$	$g + b$	$a \times g + b$		$a \times g$	$g + b$	$a \times g + b$
IG_1	\	\	\	IG_8	\	<i>inchangé</i>	\
IG_2	\	\	\	IG_9	\	\	\
IG_3	\	\	\	IG_{10}	<i>inchangé</i>	<i>inchangé</i>	<i>inchangé</i>
IG_4	\	\	\	IG_{11}	$\times a$	$+\frac{b+\sum f_i}{\sum f_i^2}$	$\times a + \frac{b+\sum f_i}{\sum f_i^2}$
IG_5	\	$-\frac{b}{moy(f)}$	\	IG_{12}	\	\	\
IG_6	\	<i>inchangé</i>	\	IG_{13}	\	\	\
IG_7	\	$-b$	\				

TAB. 6.3 – Etude de l'influence d'une translation et d'une homothétie sur l'image g sur les différents indicateurs.

Soit f , g , et h trois images et $f(x, y)$, $g(x, y)$, $h(x, y)$ la valeur du pixel (x, y) de f , g , h respectivement.

• IG_1 est une distance :

Le calcul d' IG_1 se réalise de la manière suivante : on calcule tout d'abord la distance d_1 entre tous les pixels de f et de g , puis on calcule la moyenne de toutes les distances. Les deux premières propriétés des distances ne sont pas affectées par la moyenne. Il est également aisé de voir que cela n'affecte pas non plus la propriété d'inégalité triangulaire.

• IG_3

$$\frac{|f(x,y)-g(x,y)|}{\sup(f(x,y),g(x,y))} \leq 1 \text{ pour } f(x,y)g(x,y) \neq 0.$$

Pour montrer que c'est une distance, on procède comme précédemment, en démontrant que le calcul fait sur chaque pixel ($d(f(x, y), g(x, y)) = \frac{|f(x,y)-g(x,y)|}{\sup(f(x,y),g(x,y))}$) est une distance.

◇ $d(f(x, y), f(x, y)) = 0$

◇ $d(f(x, y), g(x, y)) = d(g(x, y), f(x, y))$

Les deux premières propriétés sont triviales.

◇ $d(f(x, y), h(x, y)) \leq d(f(x, y), g(x, y)) + d(g(x, y), h(x, y))$

Pour démontrer cela, nous écrivons l'inégalité triangulaire sur trois pixels quel-

conques de valeurs x , y et z strictement positives et montrons qu'elle est vraie (cf. tableau 6.4).

$$\frac{|x - z|}{\text{sup}(x, z)} \leq \frac{|x - y|}{\text{sup}(x, y)} + \frac{|y - z|}{\text{sup}(y, z)}$$

Si une des valeurs est nulle, la distance à ce point est égale à un. On a donc :

▷ Si x ou z est nul, l'inégalité triangulaire devient $1 \leq 1 + (d > 0)$. Elle est donc vraie.

▷ Si y est nul, elle devient $d(x, z) \leq 2$ ce qui est toujours vrai.

Si deux ou trois valeurs sont nulles, l'inégalité triangulaire est toujours respectée.

★ $x \geq y \geq z$	★ $x \geq z \geq y$
$\frac{x-y}{x} + \frac{y-z}{y} \stackrel{?}{\geq} \frac{x-z}{x}$	$\frac{x-y}{x} + \frac{z-y}{z} \stackrel{?}{\geq} \frac{x-z}{x}$
$\Leftrightarrow x - y + \frac{x}{y}(y - z) \stackrel{?}{\geq} x - z$	$\Leftrightarrow x - y + \frac{x}{z}(z - y) \stackrel{?}{\geq} x - z$
$\Leftrightarrow (\frac{x}{y} - 1) \times (y - z) \stackrel{?}{\geq} 0$ vrai	$\Leftrightarrow (\frac{x}{z} + 1) \times (z - y) \stackrel{?}{\geq} 0$ vrai
★ $y \geq x \geq z$	★ $y \geq z \geq x$
$\frac{y-x}{y} + \frac{y-z}{y} \stackrel{?}{\geq} \frac{x-z}{x}$	$\frac{y-x}{y} + \frac{y-z}{y} \stackrel{?}{\geq} \frac{z-x}{z}$
$\Leftrightarrow y - x + y - z \stackrel{?}{\geq} y - y\frac{z}{x}$	$\Leftrightarrow y - x + y - z \stackrel{?}{\geq} y - y\frac{x}{z}$
$\Leftrightarrow -(x + z) + y(1 + \frac{z}{x}) \stackrel{?}{\geq} 0$	$\Leftrightarrow -(x + z) + y(1 + \frac{x}{z}) \stackrel{?}{\geq} 0$
$\Leftrightarrow (\frac{y}{x} - 1) \times (x + z) \stackrel{?}{\geq} 0$ vrai	$\Leftrightarrow (\frac{y}{z} - 1) \times (x + z) \stackrel{?}{\geq} 0$ vrai
★ $z \geq x \geq y$	★ $z \geq y \geq x$
$\frac{x-y}{x} + \frac{z-y}{z} \stackrel{?}{\geq} \frac{z-x}{z}$	$\frac{y-x}{y} + \frac{z-y}{z} \stackrel{?}{\geq} \frac{z-x}{z}$
$\Leftrightarrow \frac{z}{x}(x - y) + z - y \stackrel{?}{\geq} z - x$	$\Leftrightarrow \frac{z}{y}(y - x) + z - y \stackrel{?}{\geq} z - x$
$\Leftrightarrow (\frac{z}{x} + 1) \times (x - y) \stackrel{?}{\geq} 0$ vrai	$\Leftrightarrow (\frac{z}{y} - 1) \times (y - x) \stackrel{?}{\geq} 0$ vrai

TAB. 6.4 – Vérification de l'inégalité triangulaire dans le cas de l'indicateur IG_3 .

• IG_{12} est une distance :

$|f(x, y) - g(x, y)| \leq f(x, y) + g(x, y) \forall (x, y)$ donc $\frac{|f(x, y) - g(x, y)|}{f(x, y) + g(x, y)} \leq 1$ Comme pour l'indicateur 1, l'indicateur 12 consiste à calculer sur tous les pixels le contraste, puis à moyenner le résultat. La définition du contraste est la suivante :

Définition 13 Soit $f(x, y)$ et $g(x, y)$ les niveaux de gris respectifs du pixel (x, y) des images f et g respectivement. Le contraste $C(f(x, y), g(x, y))$ entre ces deux pixels s'exprime par :

$$\begin{cases} C(f(x, y), g(x, y)) = \frac{|f(x, y) - g(x, y)|}{f(x, y) + g(x, y)} & \text{si } (f(x, y), g(x, y)) \neq (0, 0) \\ C(f(x, y), g(x, y)) = 0 & \text{si } (f(x, y), g(x, y)) = (0, 0) \end{cases}$$

C'est une distance [158].

2. Les calculs de moyennes (IG_5, IG_7)
Ces indicateurs sont sensibles aux translations effectuées sur une image qui décalent la moyenne. Ce sont des moments d'ordre 1.
3. Les calculs d'écart-type (indicateurs IG_6, IG_8)
Ces indicateurs permettent de comparer l'étalement des histogrammes des deux images. Ils sont insensibles aux translations faites sur une image et réalisent des calculs de moments d'ordre 2.
4. Les calculs de corrélation (IG_{10} et IG_{11})
 - L'indice de corrélation IG_{10}
Deux images f et g sont totalement corrélées si elle sont liées par une relation linéaire du type $g = af + b$. L'indice de corrélation entre f et g est alors de 1.
Ainsi, un décalage dans la moyenne d'une image n'affecte pas son indice de corrélation.
 - La corrélation IG_{11}
Sans posséder les mêmes propriétés d'invariance à la translation et à l'homothétie et de symétrie que l'indice de corrélation IG_{10} , IG_{11} s'interprète cependant de façon identique.
5. Autres (indicateurs IG_2, IG_4, IG_9 et IG_{13})
 - Mathématiquement, le calcul des indicateurs IG_2 et IG_4 est sensiblement équivalent à celui des indicateurs IG_1 et IG_3 . On peut donc s'attendre à des propriétés identiques.
 - L'indicateur IG_9 met en exergue le pire pixel de l'image, il fournit donc des valeurs élevées et est très sensible aux valeurs aberrantes.
 - L'indicateur IG_{13} est relativement différent des autres, puisqu'il donne une valeur représentant un nombre de pixels satisfaisant à un seuil donné. Il est donc assez proche d'un indicateur local et nécessite le choix du seuil.

Pour tester les indicateurs, nous avons comparé plusieurs images, de façon à dégager de manière empirique les grandes tendances comportementales. Nous avons ensuite établi les comportements similaires entre indicateurs. Pour chaque image, 38 paramétrages ont été définis (*cf.* tableau 6.5). Ensuite, la fusion est réalisée entre plusieurs types d'images, plusieurs nombres d'étages et plusieurs codages des résultats; pour chaque cas de figure, un graphique du type de celui montré figure 6.2 est calculé. Il représente les résultats de tous les indicateurs, sur les 38 filtres considérés.

Les indicateurs $IG_1, IG_2, IG_4, IG_6, IG_8, IG_{10}, IG_{12}$ et IG_{13} ont des comportements identiques. Les indicateurs IG_5 et IG_7 ont le même comportement, cela se comprend aisément d'après leurs définitions. Ils sont cependant difficiles à interpréter. L'indicateur IG_{11} diffère légèrement des autres filtres. L'indicateur 9 ne travaillant que sur une valeur, la valeur maximale, est très sensible au bruit; nous ne le conserverons pas comme indicateur physique interprétable.

Physiquement, les notions de contraste, de distance d_1 et de corrélation entre deux images sont bien connues. La probabilité d'erreur par rapport à un seuil donné représente le pourcentage de pixels de l'image fusionnée qui s'éloigne trop de l'image de référence. La tolérance à l'erreur désirée pour le produit fusionné fixe le seuil. Par exemple, si l'on

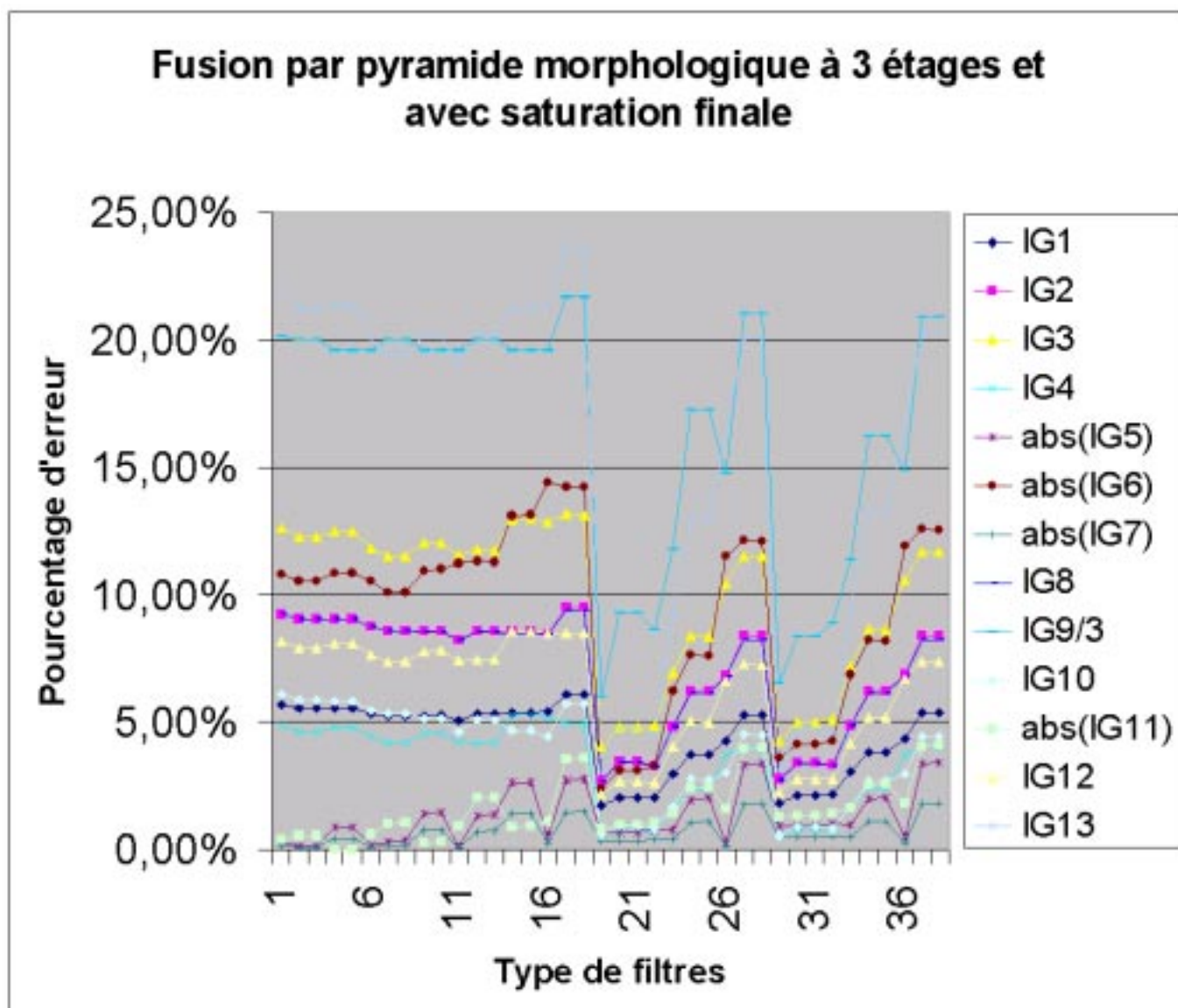


FIG. 6.2 – Exemple de résultats donnés par les indicateurs globaux, pour 38 filtres différents de la pyramide morphologique. La fusion est réalisée sur le site de Lamasquère entre deux images de résolution 1 m et 8 m ; la décomposition est réalisée par trois étages et un échantillonnage dyadique, le codage de l'image résultat est obtenu par saturation de $IFusion_0$.

souhaite que le produit fusionné ait moins de 10% de pixels qui diffèrent de plus de 2% entre l'image initiale et l'image fusionnée, le seuil est fixé à 2%, et l'indicateur doit donner un résultat inférieur à 10% pour que la fusion soit jugée acceptable.

Evaluer les paramètres de la fusion peut se réaliser de deux façons différentes :

1. De manière empirique, on teste sur de nombreuses images et de nombreux paysages les valeurs que prennent les indicateurs. Mais la diversité des images à comparer et des paysages rend très difficile le choix des seuils. Le travail mené en collaboration avec le CNES et la société Réalix Technologies et détaillé en annexe propose une méthode d'évaluation basée sur un nombre statistiquement significatif de paysages et d'images, et concerne la comparaison de différentes méthodes de fusion P/XS .
2. De manière théorique, chaque utilisateur fixe un seuil de ressemblance en deçà duquel la fusion n'a plus d'intérêt pour lui. Cette méthode possède l'avantage d'être directement liée aux besoins mais l'inconvénient d'être indépendante des indicateurs. En effet, ceux-ci, même normalisés entre 0 et 100%, ont des échelles de variations différentes. En fait, cela est plus particulièrement adapté à l'indicateur IG_{13} qui « compte ». Pour les trois autres, il est plus fiable d'avoir un jugement relatif :

« *Telle fusion est meilleure que telle autre car l'indicateur donne un résultat plus faible* »

 qu'un jugement absolu :

« *Un indicateur fournit un résultat de 2%, un autre de 7%, la fusion est-elle bonne ?* »

 Il faut alors que les utilisateurs fixent un seuil, en tenant compte non seulement du taux d'erreur qu'il veulent pour la fusion, mais aussi des propriétés physiques des indicateurs. Cela n'est pas facile à faire.

Le principe de la pyramide et les considérations de codage sur 8 bits imposent la comparaison de plusieurs jeux d'images :

- l'image basse résolution provenant du capteur avec celle calculée par la pyramide (IBR et I_N),
- l'image haute résolution provenant du capteur avec celle obtenue par la pyramide, sans saturation (I_0 et IHR_{SS}),
- l'image haute résolution provenant du capteur avec celle obtenue par la pyramide, avec saturation finale (I_0 et IHR_{AS}),
- l'image haute résolution provenant du capteur avec celle obtenue par la pyramide, avec saturation par translation et homothétie (I_0 et IHR_{tr+ho}).

Les indicateurs globaux se calculent effectivement sur ces 4 jeux de données. Afin de tester la fusion par pyramide morphologique, l'influence de ses différents paramètres et sa robustesse, la fusion a été réalisée à partir d'une image basse résolution calculée et non pas obtenue par un capteur. Dans notre cas, l'image haute résolution (1 m) est une photographie aérienne numérisée (cf. figure 6.3); les images basse résolution respectives (2 m, 4 m, 8 m et 16 m) sont obtenues par moyennage de l'image haute résolution. La superposition des différentes images est ainsi garantie. Les indicateurs globaux indiquent quel(s) est (sont) le(s) meilleur(s) paramétrage(s) de la pyramide pour réaliser la fusion

n°	Forme ES	Taille ES	Filtre	Type Déci
1	linéaire	1×2	$\frac{\gamma+\varphi}{2}$	simple
2	linéaire	1×2	γ puis φ	simple
3	linéaire	1×2	γ puis φ puis γ	simple
4	linéaire	1×2	φ puis γ	simple
5	linéaire	1×2	φ puis γ puis φ	simple
6	carré	2×2	$\frac{\gamma+\varphi}{2}$	simple
7	carré	2×2	γ puis φ	simple
8	carré	2×2	γ puis φ puis γ	simple
9	carré	2×2	φ puis γ	simple
10	carré	2×2	φ puis γ puis φ	simple
11	carré	3×3	$\frac{\gamma+\varphi}{2}$	simple
12	carré	3×3	γ puis φ	simple
13	carré	3×3	γ puis φ puis γ	simple
14	carré	3×3	φ puis γ	simple
15	carré	3×3	φ puis γ puis φ	simple
16	carré	5×5	$\frac{\gamma+\varphi}{2}$	simple
17	carré	5×5	γ puis φ	simple
18	carré	5×5	γ puis φ puis γ	simple
19	linéaire	1×2	$\frac{\gamma+\varphi}{2}$	médiane
20	linéaire	1×2	γ puis φ	médiane
21	linéaire	1×2	γ puis φ puis γ	médiane
22	carré	2×2	$\frac{\gamma+\varphi}{2}$	médiane
23	carré	3×3	$\frac{\gamma+\varphi}{2}$	médiane
24	carré	3×3	γ puis φ	médiane
25	carré	3×3	γ puis φ puis γ	médiane
26	carré	5×5	$\frac{\gamma+\varphi}{2}$	médiane
27	carré	5×5	γ puis φ	médiane
28	carré	5×5	γ puis φ puis γ	médiane
29	linéaire	1×2	$\frac{\gamma+\varphi}{2}$	moyenne
30	linéaire	1×2	γ puis φ	moyenne
31	linéaire	1×2	γ puis φ puis γ	moyenne
32	carré	2×2	$\frac{\gamma+\varphi}{2}$	moyenne
33	carré	3×3	$\frac{\gamma+\varphi}{2}$	moyenne
34	carré	3×3	γ puis φ	moyenne
35	carré	3×3	γ puis φ puis γ	moyenne
36	carré	5×5	$\frac{\gamma+\varphi}{2}$	moyenne
37	carré	5×5	γ puis φ	moyenne
38	carré	5×5	γ puis φ puis γ	moyenne

TAB. 6.5 – Liste des 38 filtres testés, l'échantillonnage est dyadique et le sur-échantillonnage est une simple duplication. les éléments structurants linéaires sont horizontaux, leur origine est le pixel le plus à gauche; l'origine des éléments structurants 2×2 est le pixel en haut à gauche.

et quel type de saturation est le plus adéquat.

Pour les 38 types de filtres définis (*cf.* tableau 6.5) les effets du rapport de résolution entre les images haute et basse résolution et donc du nombre d'étages de la pyramide et du type de saturation ont été évalués. Un exemple de résultat est montré sur la figure 6.2 page 123 avec une pyramide à trois étages, et à saturation finale. Les images utilisées correspondent au canal rouge aux résolutions 1 m et 8 m. Cette figure représente les valeurs de chaque indicateur, en pourcentage, en fonction des 38 filtres sélectionnés pour faire la fusion. Quelle que soit la configuration, les résultats sont quasiment identiques : les meilleurs filtres sont les numéros 19 et 29. Cependant, les résultats diffèrent assez sensiblement lorsque l'on teste l'image fusionnée avec saturation par translation et homothétie.

– Remarques générales sur les indicateurs

La différence théorique entre l'image initiale et l'image fusionnée n'est que le reflet de celle entre les images basse résolution calculée et l'image provenant du capteur. Ainsi, les indicateurs travaillant uniquement sur la différence entre les images four-

nissent des résultats identiques, que l'on travaille sur les images basse résolution ou sur les images haute résolution avec codage de l'image fusionnée sans saturation (IHR_{SS}). Il s'agit des indicateurs globaux 1, 2, 7, 8 et 9.

– **Influence du nombre d'étages**

Plus le nombre d'étages diminue, et plus les indicateurs donnent de bons résultats, mais la forme de la courbe reste inchangée. Le choix d'un filtre ne dépend donc pas du nombre d'étages que l'on fait. Il y a une exception cependant pour l'indicateur 9 dont on a déjà souligné le manque de fiabilité. En fait, dans le cas de la fusion avec des données réelles, le nombre d'étages n'est pas un paramètre à proprement parlé et est fixé par le rapport de résolution entre les images des capteurs haute et basse résolutions.

– **Influence des autres paramètres de la pyramide**

En comparant les indicateurs sur l'image basse résolution, on s'aperçoit qu'il y a trois comportements différents :

$$G_1 = \{IG_1, IG_2, IG_3, IG_4, IG_6, IG_8, IG_{10}, IG_{12}, IG_{13}\}$$

$$G_2 = \{IG_5, IG_7\}$$

$$G_3 = IG_9 \text{ et } IG_{11}$$

Globalement, pour tous les filtres, plus l'élément structurant est gros, moins bons sont les résultats, et cela est encore plus marqué lorsque la décimation est moyenne ou médiane.

Le groupe G_1 donne de plus mauvais résultat que les autres les filtres 1 – 18, ayant une décimation simple, cette tendance est beaucoup moins marquée pour G_2 .

G_2 donne les filtres 16, 26 et 36 nettement meilleurs que leurs voisins, tandis que G_1 les donne similaires au précédent, voire moins bons pour IG_6 . Ces filtres ont des gros éléments structurants 5×5 , et un filtre $\frac{\gamma+\varphi}{2}$.

G_2 semble plus sensible à la qualité du filtre qu'à celle de l'élément structurant, c'est le contraire pour G_1 . Même si l'allure des courbes est identique, on observe des comportements différents pour certains filtres avec les indicateurs IG_{11} et IG_9 .

– **Influence du type de codage**

Elle dépend des indicateurs.

Les indicateurs $\{IG_1, IG_2, IG_3, IG_5, IG_6, IG_7, IG_9, IG_{11}\}$ donnent des résultats très similaires pour les images IBR , IHR_{SS} et IHR_{AS} . Bien que les résultats pour IHR_{tr+ho} soient plus mauvais, la forme des deux types de courbes reste identique. L'indicateur 4 n'indique IHR_{tr+ho} que légèrement plus mauvais.

L'indicateur 10 donne de moins bons résultats avec IBR qu'avec toutes les autres images. Les résultats $IG_{10}(IHR_{SS})$ et $IG_{10}(IHR_{tr+ho})$ sont identiques, puisque la transformation $ag + b$ n'affecte pas cet indicateur.

L'indicateur 8, enfin, donne des résultats identiques entre IBR et IHR_{SS} , mais les résultats changent dès qu'intervient une saturation. IHR_{AS} donne des résultats

meilleurs, tandis que IHR_{tr+ho} en donne de légèrement moins bons.

L'indicateur 6 correspond à l'erreur relative entre deux écarts types. L'image IHR_{AS} est obtenue par saturation de IHR_{SS} , la dynamique de l'image diminue et donc son écart type aussi. De ce fait, en calculant

$$IG_6(IHR_{SS}) - IG_6(IHR_{AS}) = \frac{\sigma(IHR_{SS}) - \sigma(IHR_{AS})}{\sigma(I_0)}, \text{ on trouve}$$

$$IG_6(IHR_{SS}) - IG_6(IHR_{AS}) > 0$$

Cela est vérifié en pratique.

Indépendamment du nombre d'étages et du type de codage, le meilleur élément structurant est le plus petit, le *linéaire* 1×2 et le meilleur filtre est $\frac{\gamma+\varphi}{2}$.

En fonction du codage, la meilleure décimation est soit la médiane, pour IHR_{AS} soit la moyenne pour IHR_{tr+ho} . Mais les résultats sont plus partagés que pour l'élément structurant ou le filtre, la décimation simple représente quand même 23% des meilleurs résultats. En fait, certains indicateurs sont sensibles à la décimation et considèrent mauvaise la décimation simple, d'autres au contraire ne font pas de différence.

La saturation « translation+homothétie » changeant énormément la statistique de l'image, elle sera abandonnée au profil de la saturation finale, même si l'interpolation pourrait jouer un rôle à ce niveau là.

En conclusion, **le filtre optimal est $\frac{\gamma+\varphi}{2}$, et la décimation est à choisir entre une décimation moyenne ou médiane.** L'élément structurant optimal est le plus petit possible, plus il grossit, moins les résultats sont bons. Ainsi, l'élément 1×2 linéaire est celui qui donne les meilleurs résultats. Cependant, de façon à ne pas privilégier une direction par rapport à une autre on préférera utiliser un élément structurant carré 2×2 ou même 3×3 de façon à avoir isotropie du filtrage.

Les indicateurs locaux Introduisons maintenant les indicateurs locaux. Au lieu, comme précédemment, de travailler sur l'image entière et d'en déduire une valeur par indicateur, l'image est maintenant découpée en petites parties, nommées « *voisinages* » pour lesquelles on déduit une valeur par indicateur. Ainsi, une transformée locale est définie par l'origine de la transformation, la forme du voisinage, la taille du voisinage, et la nature de la transformation. De façon générale, on notera I le nombre de pixels contenu dans un voisinage.

La nature intrinsèque des indicateurs locaux ne diffèrent pas de celle des indicateurs globaux. Seul l'espace sur lequel est appliqué ces indicateurs change. Les indicateurs locaux ne se calculent que sur les jeux haute résolution sur lesquels un voisinage significatif peut être défini. Nous l'avons choisi correspondant à la taille d'un pixel basse résolution interpolé à la haute résolution. En effet le pixel basse résolution constitue l'entité la plus grossière générée par la pyramide, c'est sur elle que viennent se greffer les détails. Prendre un voisinage plus gros conduirait à comparer des zones de nature différente, prendre des voisinages plus petits conduiraient à ne pas tenir compte de liens inter-voisinages.

Chaque voisinage est donc de taille $\prod_{i=0}^{N-1} n_i \times \prod_{i=0}^{N-1} n_i$; il y en a autant sur une image haute résolution que de pixels dans l'image *IBR*.

On note $DIFF = f - g$ la différence entre l'image de référence et l'image à comparer. Chaque pixel d'une image peut être représenté par le voisinage auquel il appartient, et le numéro du pixel à l'intérieur du voisinage. $f_{p,q}$ (resp. $g_{p,q}$ et $DIFF_{p,q}$) est un pixel de f (resp. g et $DIFF$) où p représente l'indice de voisinage et q la numérotation du pixel à l'intérieur du voisinage.

Lorsque le voisinage est défini, p est une constante, et les calculs portent donc sur l'ensemble des pixels inclus dans ce voisinage. La variable est alors q comprise entre 1 et I . Sur un voisinage donné, nous effectuons les calculs suivants :

Moyenne de la différence $IL1_p = \text{moyenne}_q(DIFF_{p,q}) = \frac{1}{I} \sum_{q=1}^I DIFF_{p,q}$	Ecart-type de la différence $IL2_p = \sigma_q(DIFF_{p,q})$
Erreur relative sur la moyenne $IL3_p = \begin{cases} \frac{\text{moyenne}_q(f_{p,q}) - \text{moyenne}_q(g_{p,q})}{\text{moyenne}_q(f_{p,q})} = \frac{\sum_{q=1}^I f_{p,q} - \sum_{q=1}^I g_{p,q}}{\sum_{q=1}^I f_{p,q}} & \text{si } f_p \text{ est non nul} \\ \infty & \text{si } f_p \text{ est identiquement nul} \end{cases}$	
Erreur relative sur l'écart-type $IL4_p = \begin{cases} \frac{\sigma_q(f_{p,q}) - \sigma_q(g_{p,q})}{\sigma_q(f_{p,q})} & \text{si } \sigma_q(f_{p,q}) \neq 0 \\ \infty & \text{si } \sigma_q(f_{p,q}) = 0 \end{cases}$	
Distance d_1 $IL5_p = \text{moyenne}_q DIFF_{p,q} = \frac{1}{I} \sum_{q=1}^I DIFF_{p,q} $	Contraste $IL6_p = \frac{1}{I} \sum_{q=1, f_{p,q}, g_{p,q} \neq 0}^I \frac{ f_{p,q} - g_{p,q} }{f_{p,q} + g_{p,q}}$
Indice de corrélation $IL7_p = \frac{\text{cov}_q(f_{p,q}, g_{p,q})}{\sigma_q(f_{p,q}) \times \sigma_q(g_{p,q})}$	

TAB. 6.6 – Liste des indicateurs locaux étudiés : f est l'image de référence, g l'image à comparer ; p est l'indice de voisinage et q celui des pixels à l'intérieur d'un voisinage.

Ainsi, pour une image de taille (lig, col) , chaque indicateur local fournit une matrice de $\frac{lig \times col}{I}$ pixels, nommée « *matrice de voisinages* » par la suite.

Les propriétés de symétrie et de comportement à la translation et l'homothétie des indicateurs locaux sont indiquées dans les tableaux 6.7 et 6.8.

Dans le tableau 6.8, $\text{moy}(f_p)$ représente la moyenne des valeurs des pixels du voisinage p de f .

Les indicateurs IL_5 et IL_6 sont des distances, IL_1 et IL_3 des calculs de moyenne, IL_2

	Symétrique	Non Symétrique		Symétrique	Non Symétrique
IL 1		$IL_1(f, g) = -IL_1(g, f)$	IL 5	✓	
IL 2	✓		IL 6	✓	
IL 3		✓	IL 7	✓	
IL 4		✓			

TAB. 6.7 – Etude de la symétrie des indicateurs locaux.

	$a \times g$	$g + b$	$a \times g + b$		$a \times g$	$g + b$	$a \times g + b$
IL_1	\	-b	\	IL_5	\	\	\
IL_2	\	Inchangé	\	IL_6	\	\	\
IL_3	\	$\frac{-b}{moy(f_p)}$	\	IL_7	Inchangé	Inchangé	Inchangé
IL_4	\	Inchangé	\				

TAB. 6.8 – Influence d'une translation et d'une homotéthise sur les indicateurs locaux.

et IL_4 des calculs d'écart-type, et IL_7 est un calcul de corrélation.

Les matrices de voisinage qui résultent de l'application des indicateurs locaux permettent

1. d'obtenir un nouvel indicateur global par calcul de moyenne ou d'écart-type sur les matrices de voisinages.

En fait, après des tests préliminaires, le premier critère fournit des résultats similaires aux indicateurs globaux, il est donc jugé inutile et abandonné.

2. de localiser et quantifier les zones de forte erreur pour chaque indicateur.

Le but de la localisation et de la quantification est de situer sur l'image fusionnée les plus fortes erreurs. Plusieurs possibilités sont envisageables.

- Les valeurs des pixels les plus élevées sont cartographiées et superposées sur l'image de référence. Les pixels de valeurs les plus élevées peuvent être sélectionnés par les pixels dont la valeur est supérieure à la moyenne, ou par les pixels dont la valeur est supérieure à la valeur maximale divisée par 2, ou encore par la sélection d'un nombre fixe de pixels, par exemple les 10% dont les valeurs sont les plus élevées (*cf.* figure 6.3).

Les deux premières solutions dépendent très fortement de la distribution des matrices de voisinage. Ainsi, en fonction de la ressemblance entre les deux images et de l'indicateur, il peut y avoir soit trop de pixels représentés pour la localisation, auquel cas l'information est noyée, soit très peu voire pas de pixels.

La dernière ne tient pas compte du tout de la distribution et l'on peut craindre soit de visualiser trop de pixels, et donc des pixels ayant de bonnes valeurs, soit trop peu et d'oublier des zones à problèmes. La solution retenue est d'estimer au cas par cas un pourcentage de pixels à garder.

- Les matrices de voisinages sont traitées comme des images sur lesquelles on estime si l'indicateur prend des valeurs importantes et où l'on détecte l'éventuelle présence d'artéfacts. La deuxième condition se met en évidence par le calcul de

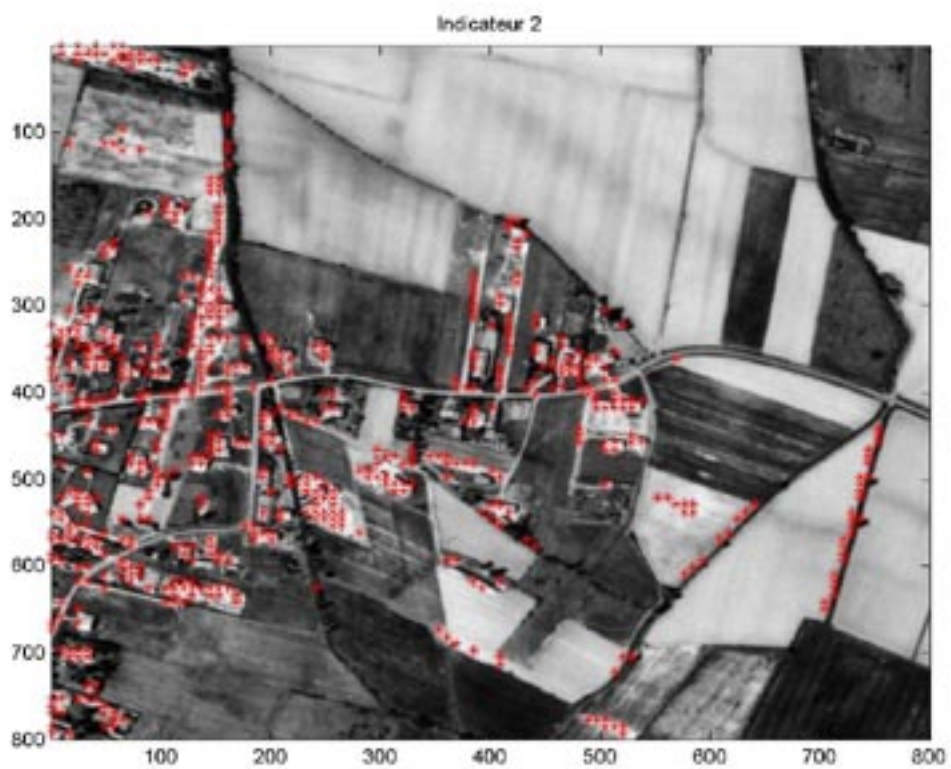


FIG. 6.3 – Exemple de localisation de 10% des valeurs les plus élevées de l'indicateur local 2 sur une photographie aérienne numérisée de la zone de Lamasquère, Haute-Garonne, France.

gradient×*compte numérique* et fournit une nouvelle image où sont mis en exergue les pixels de valeurs élevées et isolées. De façon à éviter l'effet directionnel du gradient classique, nous utiliserons un gradient morphologique. Il est défini par *Dilatation - Erosion*, pour un élément structurant donné.

3. de recouper les résultats des différents indicateurs.

Il est ensuite nécessaire de recouper les différents résultats pour voir si ceux-ci concordent ou au contraire fournissent des zones d'artéfacts différentes. On synthétise là les différents résultats.

4. de décider si les deux images sont ressemblantes ou non.

Cette étape fournit l'appréciation finale quant à la ressemblance ou non des deux images.

Les voisinages choisis étant de taille $\prod_{i=0}^{N-1} n_i \times \prod_{i=0}^{N-1} n_i$, sur l'image différence codée sans saturation, le voisinage est un palier sur lequel toutes les valeurs sont identiques. Nous indiquerons par le terme *voisinage* le voisinage sur une image quelconque mais par *palier* le voisinage de l'image différence sans saturation. Soit C_p la valeur d'un palier p de l'image différence $DIFF = I_0 - IHR_{SS}$ et soit M le nombre de paliers dans une image.

Quel que soit q , $DIFF_{p,q} = C_p$. Or $DIFF_{p,q} = I_{0,p,q} - IHR_{SS,p,q} = C_p$, donc $I_{0,p,q} = IHR_{SS,p,q} + C_p$.

Sur un voisinage donné, I_0 et IHR_{SS} ne diffèrent que par un décalage C_p de leur compte numérique. Les propriétés suivantes apparaissent alors dans les calculs intermédiaires sur un palier :

- la moyenne du palier p est égale à C_p . Ainsi, $IL1_p = C_p$,
- l'écart-type du palier p est égal à 0. Ainsi, $IL2_p = 0$,
- la moyenne de l'image initiale haute résolution est égale à celle de l'image fusionnée plus C_p sur un voisinage. Ainsi, $IL3_p = \frac{C_p}{Moyq(I_{0,p,q})}$,
- l'écart-type de l'image haute résolution initiale et celui de l'image fusionnée sont égaux sur un voisinage. Ainsi, $IL4_p = 0$,
- la distance d_1 entre l'image initiale haute résolution et l'image fusionnée correspond à la valeur absolue du palier $|C_p|$. Ainsi, $IL5_p = |C_p|$,
- $IL6 = \frac{|C_p|}{Moyq(I_{0,p,q} + IHR_{SS,p,q})}$,
- le coefficient de corrélation entre l'image haute résolution initiale et l'image fusionnée est égal à 1. Ainsi, $IL7_p = 1$.

Dans ce cas, les indicateurs ont de nombreuses propriétés. Il est inutile de calculer IL_2 , IL_4 et IL_7 qui sont constants. IL_1 et IL_3 sont proportionnels, ainsi que IL_5 et IL_6 . IL_1 et IL_5 ne diffèrent de plus que par une valeur absolue. En fait, la seule information est la quantification de la valeur des paliers. Cela change lorsque intervient la saturation, où il faut quantifier de combien et où se fait la différence.

Comme prévu, la saturation se fait dans les zones de gradients (*cf.* figure 6.3 page 130). La fusion réalisée dans cet exemple est la numéro 35, entre des images à 1 et 8 m de résolution, sur le canal rouge. Les conclusions de l'exploitation des indicateurs concordent :

- les meilleurs filtres donnés par les indicateurs correspondent aux filtres créant le moins de zones non plates,
- moins le nombre d'étages réalisés est grand, plus le nombre de zones non plates est faible.

6.4.1.2 Le paramètre algorithme d'interpolation

Il faut donc étudier maintenant l'effet de l'algorithme d'interpolation. Pour cela, la pyramide retenue sera une pyramide de trois étages, avec un pas de décimation 2×2 par moyennage, un filtre $\frac{\gamma+\varphi}{2}$ et un élément structurant 3×3 . Les algorithmes d'interpolation testés sont la duplication de pixels, l'interpolation bilinéaire et l'interpolation bicubique. Le test a été réalisé sur les images de Lamasquère sur le canal de longueur d'onde rouge, avec des résolutions de 1 m pour la haute résolution et 8 m pour la basse résolution. Les images fusionnées sont montrées sur la figure 6.4. Les images initiales et fusionnées à l'aide d'un algorithme d'interpolation bilinéaire et bicubique sont proches visuellement, par contre l'image fusionnée produite à l'aide d'une duplication présente un carroyage visuellement désagréable. Les résultats des indicateurs globaux pour les trois algorithmes sont calculés sans saturation de l'image fusionnée (*cf.* tableau 6.9). Cet exemple, et tous les autres tests réalisés dans ce contexte, montrent que l'interpolation bilinéaire est la plus performante. L'interpolation bicubique fournit des résultats proches mais tout de même moins bon, sûrement à cause d'un filtrage excessif ; elle conduit de plus à des temps de calculs plus élevés. La duplication ne donne de bon résultat ni visuellement ni quantitativement. L'algorithme de sur-échantillonnage retenu pour la fusion de données sera donc celui faisant appel à l'interpolation bilinéaire.

	IG_1	IG_2	IG_3	IG_4	IG_5	IG_6	IG_7	IG_8	IG_9	IG_{10}	IG_{11}	IG_{12}	IG_{13}
Duplication	0,0314	0,0512	0,0732	0,0185	-0,0113	-0,0770	-0,0061	0,0508	0,3412	0,0179	-0,0191	0,0444	0,0970
Bilinéaire	0,0213	0,0338	0,0489	0,0073	-0,0107	-0,0322	-0,0058	0,0333	0,3373	0,0087	-0,0128	0,0267	0,04060
Bicubique	0,0253	0,0399	0,0584	0,0105	-0,0104	-0,0403	-0,0056	0,0395	0,3373	0,0121	-0,0133	0,0325	0,0650

TAB. 6.9 – Résultat des indicateurs globaux sur les trois types de ré-échantillonnage duplication, interpolation bilinéaire et bicubique.

6.4.2 Choix de la répartition pas de décimation / nombre d'étages

Les tests réalisés précédemment ont été conduits avec un pas de décimation dyadique, le nombre d'étages découle naturellement. Si $ligHR$ (respectivement $libBR$) est le nombre de ligne de l'image haute résolution (respectivement basse résolution), alors $\frac{ligHR}{libBR}$ donne le rapport de résolution entre la basse et la haute résolution. Si le pas de décimation est dyadique à tous les étages et que le rapport peut s'écrire 2^N , il faut réaliser N étages. Mais le rapport peut s'écrire de façons différentes et l'on peut alors choisir un autre pas de décimation et un autre nombre d'étages également adaptés. Par exemple, si $rapport = 2^{2M}$, il est également possible soit de réaliser un échantillonnage de pas 2

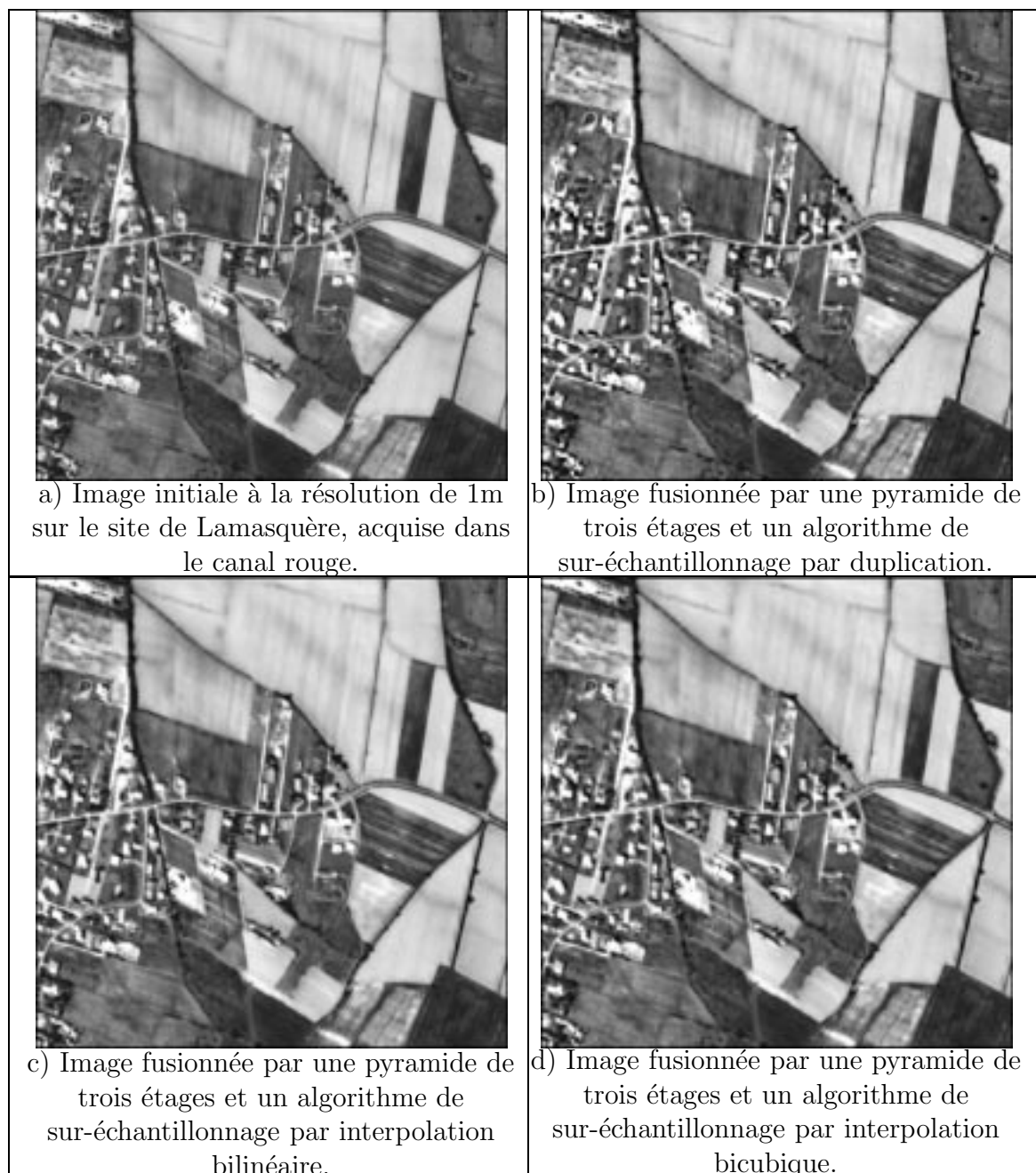


FIG. 6.4 – Comparaison entre le choix d'un algorithme de fusion avec un ré-échantillonnage par duplication (image b), par interpolation bilinéaire (image c), et par interpolation bicubique (image d). Les images sont recadrées sur 8 bits par saturation des valeurs. L'exemple montré représente une scène au dessus du village de Lamasquère, Haute-Garonne, France.

avec $2M$ étages, soit un échantillonnage de pas 4 avec M étages. Dans un cas encore plus général, si les pas de décimation ne sont pas constants d'un étage à l'autre, voire s'ils ne sont pas entiers, le choix de ces deux paramètres devient vite délicat.

Cependant, quelques constatations simples viennent aider notre choix. Tout d'abord, lorsque cela est possible, le choix du pas de décimation est choisi entier afin d'éviter des calculs d'interpolation supplémentaires qui vont insérer dans la chaîne de traitement des effets de filtrage supplémentaires qu'il sera difficile de quantifier et qui filtreront davantage encore I_N par rapport à IBR .

Ensuite, un pas de décimation entier traite de façon équivalente tous les niveaux de résolution ; il n'est pas justifié, lorsque l'on peut l'éviter de changer le pas de décimation en fonction de l'étage.

Finalement, de trop gros pas de décimation vont entraîner des difficultés dans le calcul de l'interpolation correspondante. En effet, comme exprimé au paragraphe précédent, le schéma de duplication n'est pas efficace et les schémas d'interpolation bilinéaire et bicubique deviennent imprécis lorsque le nombre de pixels à calculer est très supérieur à celui des pixels connus.

Afin de tester ces choix, deux fusions ont été réalisées sur le même jeu de données. Les images d'entrées ont des résolutions de 1 m et 16 m sur le site de Lamasquère, dans le canal de longueur d'onde rouge. Les paramètres communs sont le filtre, $\frac{\gamma+\varphi}{2}$, l'interpolation bilinéaire, et la décimation moyenne. La première fusion est réalisée par 4 étages, une décimation dyadique et un élément structurant 3×3 ; la deuxième par 2 étages, un pas de décimation 4×4 et un élément structurant 5×5 . Les résultats sont montrés sur la figure 6.5. Visuellement, la méthode à quatre étages est plus proche de l'image haute résolution de référence que la méthode à deux étages. Les résultats des indicateurs globaux sur les produits fusionnés sans saturation sont montrés dans le tableau 6.10. Les valeurs maximales sont respectivement de 293 et 279 et les valeurs minimales de -46 et -8 pour les résultats 2 étages et 4 étages. Le dépassement de dynamique est plus élevé dans le cas 2 étages. Les indicateurs 5, 7, 9 et 11 donnent de meilleurs résultats pour la fusion à deux étages, les autres pour la fusion pour la fusion à quatre étages. L'indicateur IG_{12} fournit un résultat infini pour les résultats de la fusion 2 étages, cela signifie qu'il existe au moins un pixel pour lequel $I_0(x_{k_0}, y_{l_0}) = -IFusion_0(x_{k_0}, y_{l_0})$.

	IG_1	IG_2	IG_3	IG_4	IG_5	IG_6	IG_7	IG_8	IG_9	IG_{10}	IG_{11}	IG_{12}	IG_{13}
4 étages	0,0274	0,0404	0,0615	0,0108	-0,0118	-0,0342	-0,0063	0,0399	0,3451	0,0127	-0,0133	0,0342	0,0645
2 étages	0,0540	0,0767	0,1156	0,0375	0,0093	-0,0553	0,0050	0,0765	0,3333	0,0462	0,0067	Inf	0,1656

TAB. 6.10 – Résultat des indicateurs globaux sur les deux types de fusion à 4 et 2 étages.

En conclusion, on privilégiera, lorsque cela est possible, un grand nombre d'étages avec un pas de décimation faible.

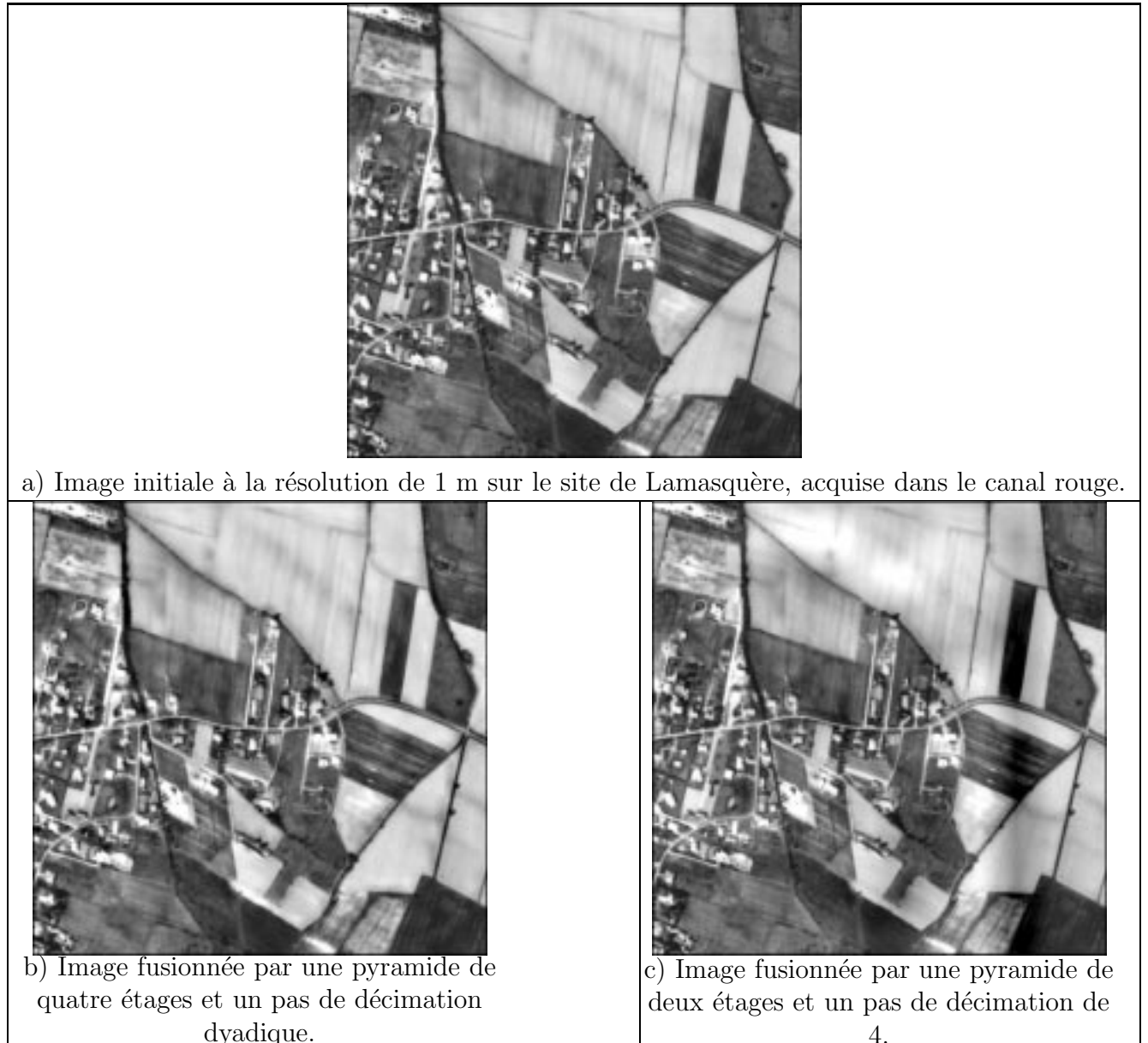


FIG. 6.5 – Comparaison entre le choix d’une pyramide a) de quatre étages et un pas de décimation 2 et b) de deux étages et un pas de décimation 4. Les images sont recadrées sur 8 bits par saturation des valeurs.

6.4.3 Comparaison de la pyramide morphologique avec d'autres algorithmes pyramidaux

Dans cette section, une comparaison est faite entre le schéma de pyramide morphologique retenu et d'autres algorithmes pyramidaux dans le but de faire de la fusion de données multi-résolution. Les paramètres retenus dans la pyramide morphologique sont ceux préconisés : filtre $\frac{\gamma+\varphi}{2}$, échantillonnage dyadique par moyennage, interpolation bilinéaire, élément structurant 3×3 et 3 étages de décomposition. Les images d'entrée sont les images sur le site de Lamasquère, acquises dans la bande spectrale rouge et avec un rapport de résolution $1 \text{ m} \rightarrow 8 \text{ m}$.

Pour mener à bien cet objectif, nous avons implémenté l'algorithme pyramidal décrit par Tanimoto et Pavlidis [134] et détaillé dans la section 2.3. Comme aucun algorithme de reconstruction n'est prévu dans leur processus, nous avons retenu un ré-échantillonnage par interpolation bilinéaire, identique à celui réalisé dans la pyramide morphologique. Rappelons que l'algorithme de Tanimoto et Pavlidis ne fonctionne que pour des rapports de résolution de 2, et que trois étages sont donc nécessaires. A l'exception de l'indicateur IG_6 (erreur relative sur l'écart type), la pyramide morphologique obtient de meilleurs résultats que la pyramide proposée par Tanimoto et Pavlidis [134].

	IG_1	IG_2	IG_3	IG_4	IG_5	IG_6	IG_7	IG_8	IG_9	IG_{10}	IG_{11}	IG_{12}	IG_{13}
Pyr. morpho.	0,0213	0,0338	0,0489	0,0073	-0,0107	-0,0322	-0,0058	0,0333	0,3373	0,0087	-0,0128	0,0267	0,04060
[134]	0,0389	0,0599	0,757	0,0122	-0,0705	-0,0121	-0,0379	0,0465	0,4706	0,0183	-0,0576	0,0413	0,1143
[8]	0,0341	0,0517	0,0827	0,0234	-0,0018	-0,1150	-0,0010	0,0517	0,2938	0,0147	-0,0180	0,0504	0,1025

TAB. 6.11 – Résultat des indicateurs globaux pour la comparaison de la pyramide morphologique et d'autres algorithmes pyramidaux dans le cadre de la fusion de données.

La pyramide de Burt et Adelson [8] a également été programmée avec le noyau de filtrage recommandé par les auteurs. La recombinaison proposée par ceux-ci se fait par ajout successif de la pyramide de laplacien à la dernière image basse résolution. Afin d'adapter l'algorithme à la fusion de données, l'ajout successif des éléments de la pyramide de laplacien s'effectue sur l'image du capteur basse résolution. Comme pour l'algorithme de Tanimoto et Pavlidis, la pyramide de Burt et Adelson ne fonctionne qu'avec des pas d'échantillonnage dyadique. Les résultats sont meilleurs pour la pyramide de Burt et Adelson pour les indicateurs IG_5 et IG_7 (indicateurs de moyenne) et pour l'indicateur de valeur maximale IG_9 . Visuellement, le résultat est en faveur de la pyramide morphologique.

6.5 Recalage de données

Lorsque l'on souhaite fusionner deux images, il est nécessaire que celles-ci soient au préalable recalées afin d'être géométriquement superposables. Cette section explique comment la pyramide morphologique pourrait être un outil d'aide à cette superposition.

Les différences entre les données sont de plusieurs types :

- résolution spatiale,
- date de prise de vue,
- système de projection des données,
- angle de prise de vue,
- direction de balayage,
- longueur d’onde de sensibilité du capteur,
- ...

Ainsi, de façon générale, deux données ne sont pas superposables, et donc non fusionnables. Une fois que les deux données ont été remises dans le même système de projection, celles-ci sont superposables spatialement au mieux par une transformation composée d’une rotation (pour compenser les directions de balayage différentes), d’une translation et d’une homothétie. Dans ce qui suit, on suppose que les données sont dans le même système de projection et que les déformations dues à la trajectoire du satellite ou de l’avion sont corrigées.

Sous ces hypothèses, les seules transformations qui différencient deux images sont une translation, une rotation et/ou un facteur d’échelle. Celui-ci intervient si les données sont à des résolutions différentes. Dans l’étude prospective menée ici, nous avons utilisé le coefficient de corrélation afin de tester la qualité de superposition de deux images. A cet effet, l’image haute résolution est décomposée par la pyramide morphologique, jusqu’à la taille de l’image provenant du capteur basse résolution. Le jeu de données pris pour cette étude est un jeu test pour lequel aucune rotation n’est nécessaire. L’image simulée du capteur basse résolution couvre un zone géographique plus grande que l’image haute résolution. Ainsi, l’algorithme de recalage doit détecter la position relative des pixels des deux images.

Pour cela, on calcule la corrélation pour toutes les positions de décalage possibles, et on cherche celle qui donne la corrélation maximale. Ensuite, une fois cette position trouvée, la corrélation est calculée pour différents angles de rotation entre les deux images jusqu’à l’obtention d’une corrélation maximale. Dans le cas étudié, aucune rotation n’était nécessaire entre les images et on cherchait à retrouver cette valeur nulle.

Les résultats de cette étude préliminaire sont présentés ici. Pour peu que les paramètres de la pyramide soient convenables, la translation est compensée dans tous les cas de figure. Par contre, l’étape de rotation pose problème, quel que soit l’algorithme d’interpolation réalisé. Une rotation par méthode plus proche voisin n’est pas capable de reproduire une rotation sur des faibles angles de rotation ; une interpolation bilinéaire voire bicubique crée sur l’image après rotation un effet de filtrage passe-bas qui fausse les résultats : en effet, comme cela a été constaté dans le chapitre 6, l’image issue de la pyramide est plus filtrée que l’image issue du capteur basse résolution. La rotation se faisant sur l’image de plus grande taille pour éviter l’introduction de bordure noire, le filtrage induit tend à créer une corrélation plus grande entre l’image ayant subi une rotation de $\pm 1^\circ$ ou 2° et l’image issue de la pyramide qu’entre l’image sans rotation et l’image issue de la pyramide.

Afin de compléter cette étude et d'évaluer si la pyramide morphologique est un outil de recalage d'images performant, il serait nécessaire :

- d'étudier le principe proposé à un plus grand nombre d'images,
- de chercher des algorithmes de rotation plus performants,
- d'évaluer les résultats sur un jeu de données réel,
- de comparer les résultats obtenus à d'autres méthodes de recalage.

D'une manière générale, on voit que la pyramide morphologique pourrait être exploitée comme les autres approches multi-échelle dans le recalage d'images. Elle apporterait sa capacité à réintroduire l'information de détail décrite explicitement lors de la décomposition.

6.6 Conclusion

Le principe de la fusion de données par pyramide morphologique permet de mélanger des données de résolutions différentes.

Les paramètres à utiliser pour avoir une fusion optimale sont un filtre $\frac{\gamma+\varphi}{2}$, un échantillonnage par moyennage et un sur-échantillonnage par interpolation bilinéaire. Le pas de décimation et le nombre d'étages sont à choisir en fonction du rapport de résolution entre les deux images avec la contrainte de prendre un pas de décimation le plus petit possible. L'élément structurant est pris de taille juste supérieure au pas d'échantillonnage.

Le principe de la fusion par pyramide morphologique fournit de plus des résultats visuels et statistiques satisfaisants.

Chapitre 7

La fusion avec des données multi-temporelles

7.1 La problématique

Ce chapitre décrit l'application du processus de fusion par pyramide morphologique à la fusion de données haute résolution spatiale / basse fréquence temporelle et de données *grand champ* basse résolution spatiale / haute fréquence temporelle. L'objectif de fusion temporelle est de mettre à jour l'image haute résolution spatiale par l'apport de l'information contenue dans les images grand champ à haute fréquence temporelle.

Cet objectif est donc de créer, à la résolution spatiale des images haute résolution, autant d'images de synthèse que l'on a d'images *grand champ*. Ces images de synthèse doivent permettre de mettre à jour les informations basse résolution dans les images haute résolution aux dates des images *grand champ*, tout en conservant les éléments de meilleures résolutions spatiales, supposés constants dans l'intervalle de temps (route, contour de champ, ...).

Ce processus est possible car il n'y a pas de contrainte de date entrant en jeu dans l'algorithme de fusion.

Soit HRS_k , $k = 1 \dots p$ (pour High Resolution Sensor), l'ensemble des images du capteur haute résolution spatiale, ayant une résolution HR (pour High Resolution). Soit CRS_l , $l = 1 \dots q$, (pour Coarse Resolution Sensor) l'ensemble des images du capteur grand champ, de résolution CR (pour Coarse Resolution). La résolution temporelle du capteur *grand champ* est supérieure à celle du capteur haute résolution, en conséquence $q > p$. Le processus de fusion est le suivant :

1. chaque image HRS_k est décomposée jusqu'à la résolution CR . Ainsi, chaque HRS_k fournit des détails D_{sup} et D_{inf} , appelés $D_{sup, filtre, i, k}$, $D_{inf, filtre, i, k}$, $D_{sup, deci, i, k}$ et $D_{inf, deci, i, k}$ où i est l'indice de niveau dans la pyramide et k l'indice correspondant à la date des images HRS_k .

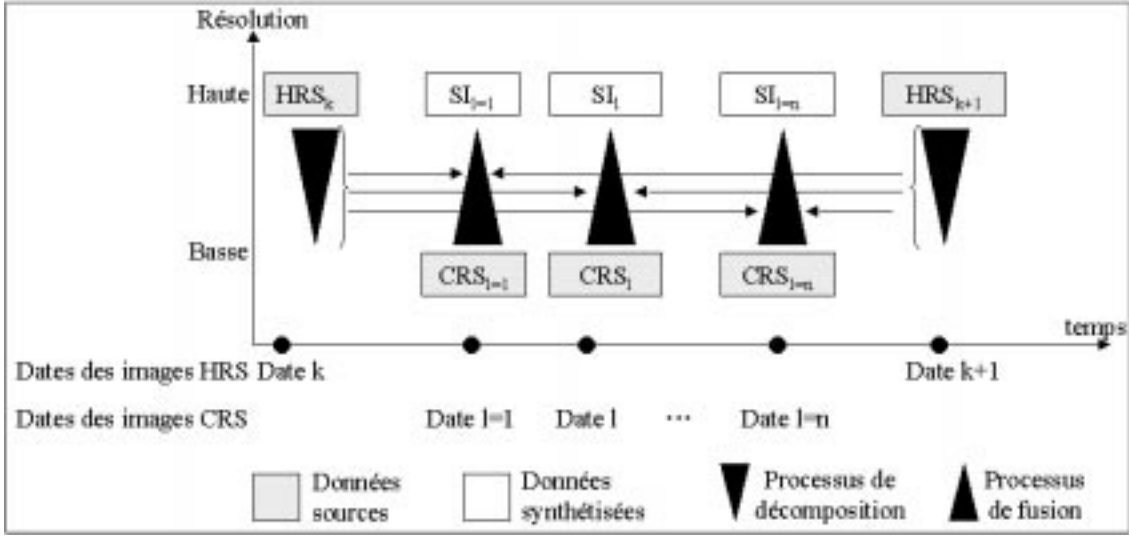


FIG. 7.1 – Schéma de fusion multitemporelle entre les images haute résolution HRS et basse résolution CRS .

2. Le processus de fusion tel que vu au chapitre 6 est répété séparément sur chaque image CRS_l . Il est donc itéré q fois.

Deux cas sont possibles :

- Les images CRS sont encadrées temporellement par deux images, HRS_k et HRS_{k+1} (cf. figure 7.1). En conséquence, différents composants de détails vont être injectés dans l'image CRS_l en fonction de la proximité de la date l avec les dates k et $k+1$.

Supposons qu'il y ait n images basse résolution, allant de la date 1 à la date n (CRS_1 à CRS_n) intercalées entre HRS_k et HRS_{k+1} . Supposons également que les images CRS soient acquises à des intervalles de temps réguliers (tous les jours ou tous les 10 jours par exemple).

Dans ce cas, n images de synthèse $SI_l, l = 1, \dots, n$ sont créées à partir de HRS_k, HRS_{k+1} et de CRS_1 à CRS_n .

Soit $\{D_k\}$ (resp. $\{D_{k+1}\}$) l'ensemble des détails obtenus à partir de l'image HRS_k (resp. de HRS_{k+1}).

Soit $\{D_l\}$ l'ensemble des détails qui seront insérés dans CRS_l pendant le processus de fusion.

Alors, on aura, comme le montre la figure 7.1,

$$\{D_l\} = \frac{l}{n+1}\{D_k\} + \frac{n+1-l}{n+1}\{D_{k+1}\} \quad (7.1)$$

L'image SI_l est calculée à partir des images CRS_l et $\{D_l\}$ suivant le processus de reconstruction décrit dans le chapitre 6.

- Les images CRS_l ne sont encadrées que par une seule image, HRS_k . Dans ce cas, les détails $\{D_k\}$ sont entièrement injectés lors du processus de fusion avec l'image CRS_l . L'image SI_l ainsi obtenue est calculée à partir de CRS_l et $\{D_k\}$.

7.2 Validation de l'algorithme de fusion temporelle

7.2.1 Jeu de données

Pour estimer son efficacité, le processus de fusion temporelle doit être testé sur un jeu de données satisfaisant les conditions suivantes :

- le jeu de données doit être composé au minimum de trois images haute résolution, HRS_1 , HRS_2 et HRS_3 et d'une image basse résolution spatiale acquise à la même date que HRS_2 , CRS_2 . Une image de synthèse à la date 2, SI_2 est créée par le processus de fusion et mélange les informations contenues dans les images HRS_1 , HRS_3 et CRS_2 . L'image HRS_2 sert de référence.
- Les gabarits spectraux des images HRS et CRS doivent être identiques. Dans ce cas, le processus de fusion optimal est obtenu lorsque SI_2 est proche de HRS_2 .
- Les différentes images doivent être parfaitement recalées entre elles.

Ce schéma permet de valider ou d'invalider le procédé de fusion temporelle par pyramide morphologique. La deuxième condition nous pousse à utiliser les images acquises par le capteur haute résolution SPOT 4 - HRVIR et le capteur grand champ SPOT 4 - VEGETATION, ces deux capteurs ayant trois longueurs d'onde communes.

Par contre, n'ayant pu obtenir un jeu comprenant trois images haute résolution, nous avons été obligé de simuler HRS_1 et HRS_2 à partir de HRS_3 . Un exemple est montré sur des images du Sénégal, dans une zone située à l'ouest de la ville de Saint-Louis. Cet exemple a été publié dans [60].

HRS_3 a été acquise le 21 Septembre, dans la bande XS2 (longueur d'onde rouge), comme le montre la figure 7.4. Afin de simuler la progression dans la saison des pluies, l'image HRS_1 est calculée par la pondération de l'image HRS_3 dans les zones où les surfaces en eau prédominent [59]. Le traitement a consisté à réduire la surface occupée par les zones d'eau en modifiant la signature spectrale de HRS_3 , à l'aide notamment du canal de longueur d'onde bleu. L'image HRS_1 simule ainsi une image acquise plus tôt dans la saison des pluies. L'image HRS_2 est simulée par la moyenne entre les images HRS_1 et HRS_3 ,

$$HRS_2 = \frac{HRS_1 + HRS_3}{2} \quad (7.2)$$

D'après nos simulations, l'image HRS_1 est acquise le 11 août, HRS_3 le 21 septembre et donc, HRS_2 simule une image acquise le 1^{er} septembre. Ce processus de simulation est évidemment limité, mais assure une superposition exacte des images, ainsi qu'une radiométrie cohérente.

Les figures 7.2a) et 7.2b) montrent les différences entre d'une part HRS_1 et HRS_2 et d'autre part entre HRS_2 et HRS_3 . La LUT des images a été inversée, les différences principales apparaissent ainsi en noir sur fond blanc. Les zones variant le plus durant la saison des pluies sont les zones d'eau. De plus, la similitude entre ces deux images montrent une évolution comparable entre le 11 août et le 1^{er} septembre d'une part et entre le 1^{er} septembre et le 21 septembre d'autre part.

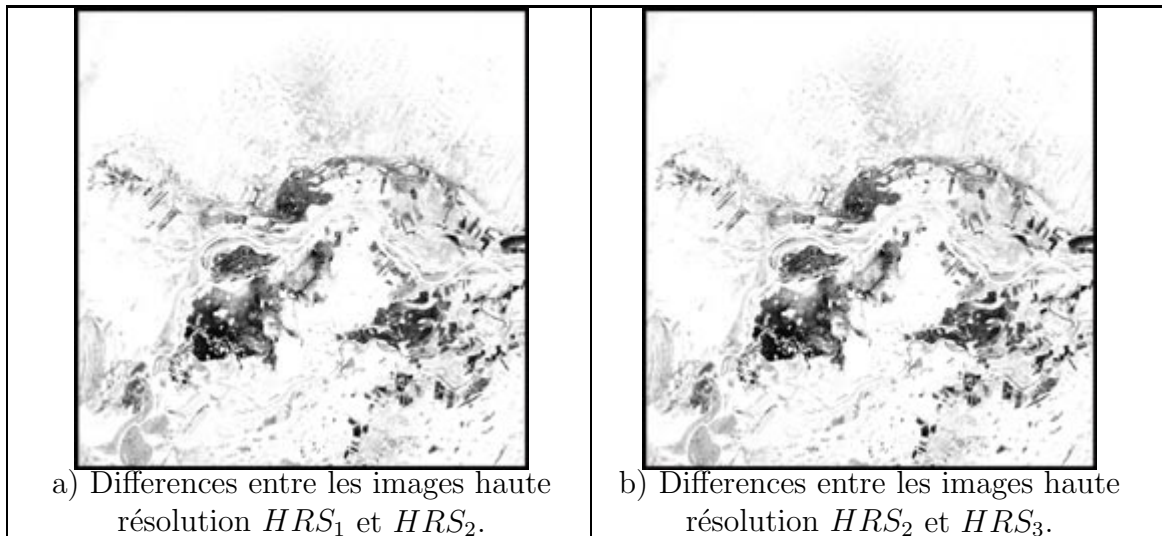


FIG. 7.2 – Differences entre les images haute résolution a) HRS_1 et HRS_2 et b) HRS_2 et HRS_3 .

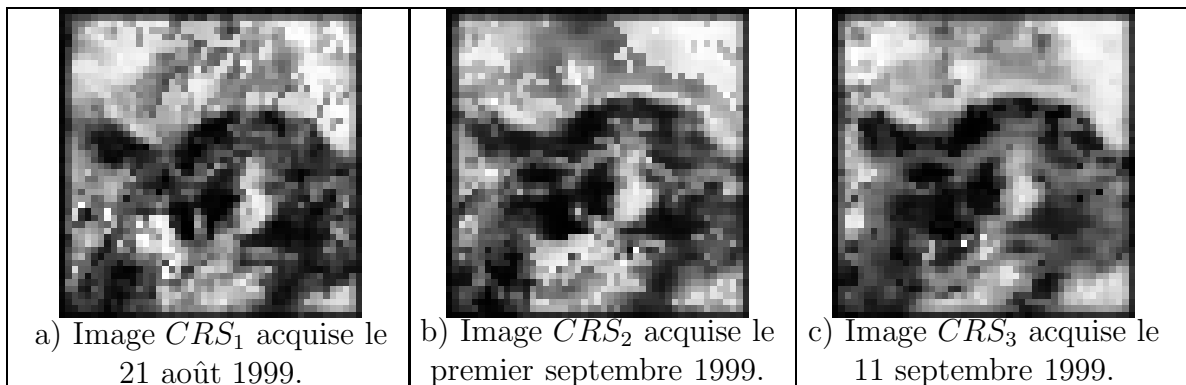


FIG. 7.3 – Images basse résolution issues du capteur VEGETATION acquises le a) 21 août 1999 (CRS_1), b) premier septembre 1999 (CRS_2), c) 11 septembre 1999 (CRS_3).

Les images du capteur *grand champ* SPOT4-VEGETATION sont disponibles tous les 10 jours. Les images du 21 août, CRS_1 , du premier septembre, CRS_2 et du 11 septembre, CRS_3 sont montrées sur les figures 7.3a) à 7.3c). Contrairement aux images simulées HRS , des différences importantes existent, témoignant notamment d'une forte évolution de la surface couverte par les zones d'eau.

Après la fusion, les images de synthèse créées sont SI_1 , SI_2 et SI_3 , comme le montre la figure 7.5. L'image SI_2 est alors comparée à l'image servant de référence HRS_2 . Le tableau 7.1 synthétise les dates des différentes images.

7.2.2 Processus de fusion

La résolution spatiale du capteur VEGETATION est de 1 km, celle du capteur HRVIR de 20 m. En tenant compte de cette contrainte, les paramètres suivants ont été choisis

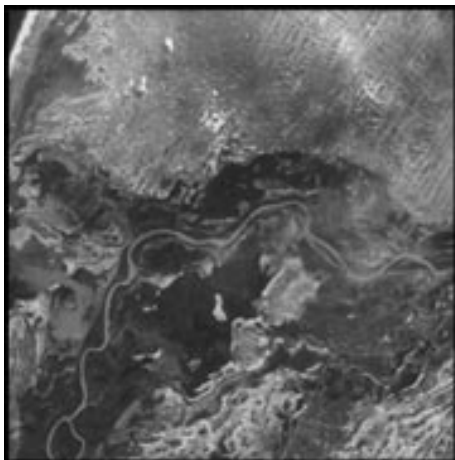


FIG. 7.4 – Image du capteur HRVIR dans la bande spectrale rouge, HRS_3 .

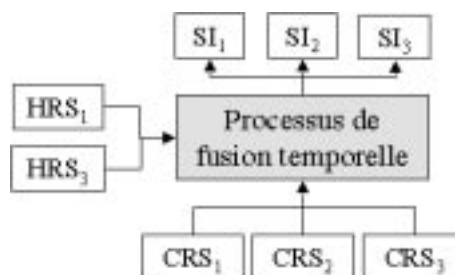


FIG. 7.5 – Schéma des entrées et sorties du processus de fusion multitemporelle dans le cas du site du Saint-Louis, Sénégal.

11 août	21 août	1 ^{er} septembre	11 septembre	21 septembre
HRS_1		HRS_2		HRS_3
	CRS_1	CRS_2	CRS_3	
	SI_1	SI_2	SI_3	

TAB. 7.1 – Tableau de synthèse des dates des images mises en jeu dans le processus de validation de la fusion de données temporelles sur le site de Saint-Louis, Sénégal : $HRS_{1,2,3}$, $CRS_{1,2,3}$ et $SI_{1,2,3}$.

pour réaliser la décomposition des images HRS :

- les images HRS sont dégradées à une résolution de 31,25 m,
- l'échantillonnage est dyadique,
- la décomposition est réalisée en 5 niveaux.

Le sur-échantillonnage est réalisé par un processus d'interpolation bilinéaire, le filtre pris est $\frac{\gamma+\varphi}{2}$ et l'élément structurant est de rayon 1, centré et symétrique, conformément aux recommandations émises pour le choix des paramètres dans la section 6.3.

n étant égal à trois, les détails insérés dans SI_l selon le principe énoncé dans l'équation (7.1) sont

$$\{D_l\} = \frac{l}{4}\{D_{HRS_1}\} + \frac{4-l}{4}\{D_{HRS_2}\} \quad (7.3)$$

7.2.3 Résultats

Dans le cas parfait, SI_2 devrait être identique à HRS_2 . Visuellement, HRS_2 et SI_2 sont proches (*cf.* figures 7.6a) et 7.6b)). Les différences sont réparties sur toute l'image. Le tableau 7.2 compare les moyennes et écarts-types des images HRS_2 et SI_2 . Le coefficient de corrélation entre ces deux images est de 0,83. Ces statistiques ne sont cependant pas suffisantes pour évaluer la qualité du processus de fusion. L'écart-type de l'image

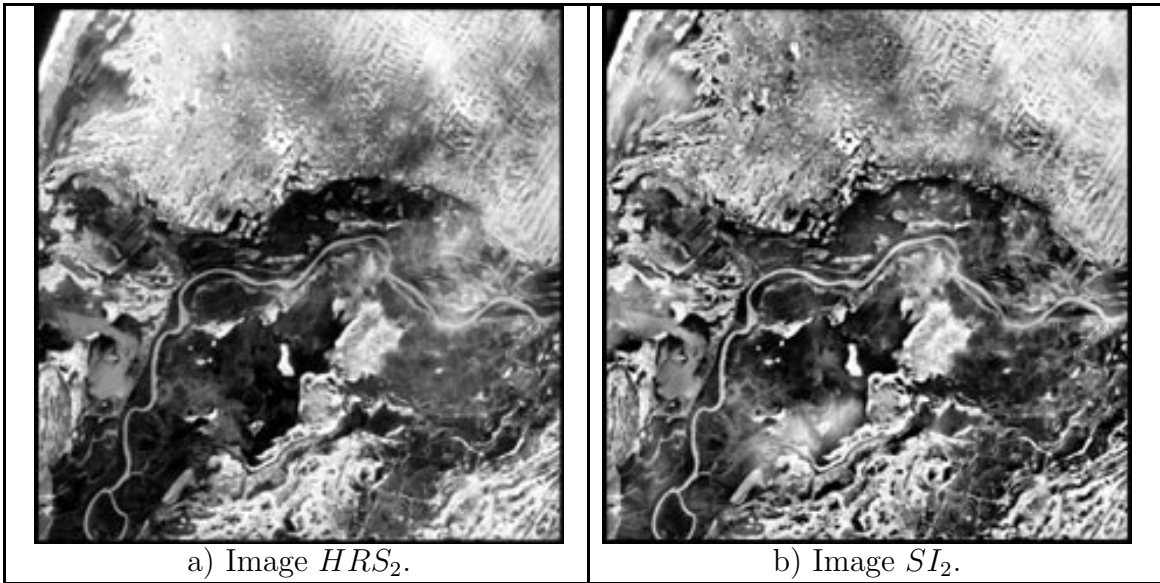


FIG. 7.6 – Comparaison des images haute résolution initiale et fusionnée a) HRS_2 et b) SI_2 .

de différence entre SI_2 et HRS_2 est supérieur à 19. Cette valeur, très importante, est probablement la conséquence de deux phénomènes : premièrement, le processus de simulation de l'image HRS_2 n'est sûrement pas optimisé et deuxièmement les corrections radiométriques effectuées sur les images des capteurs HRVIR et VEGETATION ne sont pas identiques.

	Moyenne	Ecart type
HRS_2	101.80	35.81
SI_2	101.64	27.93

TAB. 7.2 – Comparaison des statistiques entre les images HRS_2 et SI_2 .

Un point mérite également d'être souligné : aucun artefact n'apparaît sur les images synthétisées. Après cette étude, bien que les résultats soient calculés sur un jeu de données ne correspondant pas à celui préconisé, l'approche de fusion de données temporelles par pyramide morphologique sera considérée comme robuste. Elle l'est cependant visuellement.

7.3 Evolution temporelle d'un paysage terrestre

Cet exemple propose de mélanger la bande spectrale rouge d'une photographie aérienne numérisée de résolution 1,25 m à la bande spectrale XS2 d'une image acquise par le capteur SPOT 4 - HRVIR à la résolution spatiale de 20 m sur la même zone. Ces deux

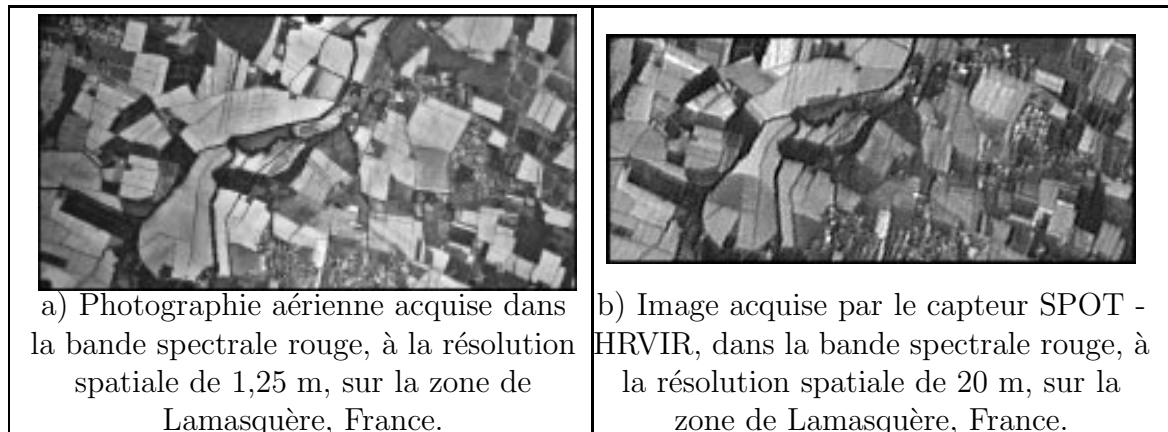


FIG. 7.7 – Comparaison des images initiales a) haute résolution et b) basse résolution acquises sur la zone de Lamasquère, Haute-Garonne, France.

images sont montrées sur les figures 7.7a) et 7.7b). Les images ont été au préalable recalées géométriquement.

Elles ont été acquises autour du village de Lamasquère situé au sud-ouest de la France, près de Toulouse. Le paysage est composé de parcelles agricoles, de maisons isolées et de quelques routes. Le rapport de résolution entre les deux images est de 16. Nous avons vu (*cf.* section 6.3) que les meilleurs résultats sont obtenus avec un taux d'échantillonnage faible entre chaque étage, et un élément structurant de taille juste supérieure. En conséquence, nous avons choisi dans cette application de réaliser la décomposition sur l'image aérienne numérisée par une pyramide de quatre étages, avec un échantillonnage dyadique. L'élément structurant est de rayon 1, centré et symétrique. L'image acquise par le capteur HRVIR sert de base à la reconstruction.

Les deux images n'ont pas été acquises à la même date, en conséquence, des variations de réflectances apparaissent sur les champs et les sols. D'après le processus de fusion, l'image de synthèse créée doit intégrer les informations basse résolution de l'image du capteur HRVIR et les informations haute résolution issues de la photographie aérienne.

Afin d'améliorer la visualisation des effets de la fusion, deux zones ont été sélectionnées, elles sont montrées sur les figures 7.8a) à 7.8c) et 7.9a) à 7.9c). Sur le premier extrait, deux zones sont particulièrement intéressantes, elles sont numérotées 1 et 2 sur les images 7.8a) et 7.8c). Sur l'image 7.8a), la première partie correspond à des champs agricoles non encore couverts par des cultures et ont donc un aspect similaire à celui de sols nus ; la deuxième partie représente une zone agricole après irrigation. Sur l'image 7.8b), acquise ultérieurement, les cultures sur la zone 1 ont commencé à recouvrir le sol, tandis que sur la zone 2, l'eau s'est infiltrée dans le sol dont la réflectance diminue. Le résultat de la fusion sur cet extrait, montré sur la figure 7.8c), démontre la mise à jour de l'information de culture grâce à l'image à 20 m de résolution, mais garde intactes les informations spatiales haute résolution (frontière des champs, routes, ...). Il est notamment important de

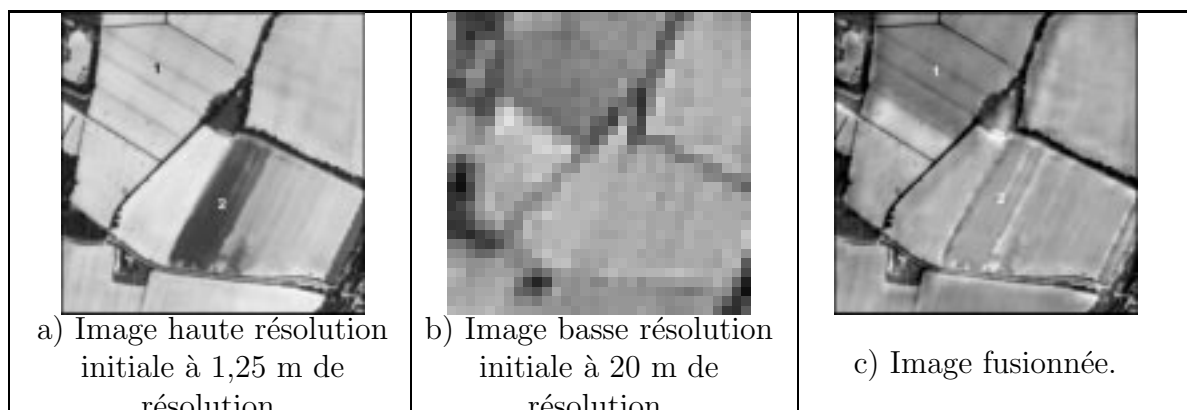


FIG. 7.8 – Premier extrait sur la zone de Lamasquère, France : a) Extrait issu de la photographie aérienne numérisée à la résolution de 1,25 m ; b) Extrait issu de l'image acquise par le capteur HRVIR à la résolution de 20 m ; c) Extrait issu de l'image fusionnée.

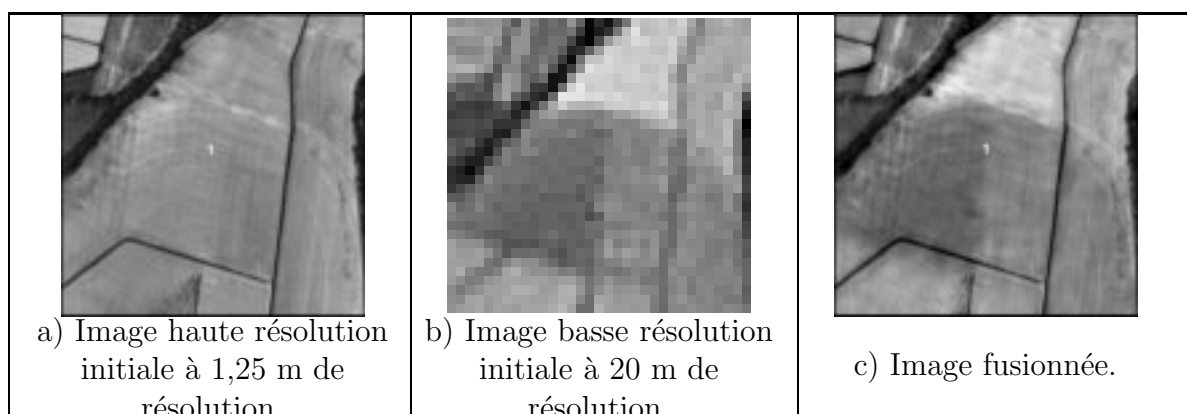


FIG. 7.9 – Deuxième extrait sur la zone de Lamasquère, France : a) Extrait issu de la photographie aérienne numérisée à la résolution de 1,25 m ; b) Extrait issu de l'image acquise par le capteur HRVIR à la résolution de 20 m ; c) Extrait issu de l'image fusionnée.

constater que la frontière du champ numéroté 2 a été gardée, même si celle-ci, ayant une réflectance identique à celle de son environnement à la date d'acquisition par le capteur satellite, n'est plus visible sur l'image 7.8b).

Sur le deuxième extrait, l'effet d'un arrosage circulaire sur une culture est facilement visible. Parce que l'arrosage n'était pas actif lors de l'acquisition de l'image aérienne (*cf.* figure 7.9a)), il existe une grande différence sur cette zone entre les deux images d'entrée du processus de fusion. L'image acquise par HRVIR (*cf.* figure 7.9b)) tient compte des propriétés spectrales de l'eau, qui diminue la réflectance du sol. La fusion de ces deux images, dont le résultat est montré sur la figure 7.9c), met en avant l'effet de l'arrosage circulaire sur les cultures entre ces deux dates.

Cet exemple illustre bien la capacité de la fusion à mettre à jour des informations basse résolution (par exemple l'évolution des cultures) soumises à une évolution rapide tout en gardant les informations à haute résolution spatiale.

7.4 Suivi temporel à partir d'images multispectrales et grand champ

Comme nous l'avons vu dans la section précédente, le principe de fusion est très efficace pour mettre à jour des informations basse résolution dans des images haute résolution. Ce point est très important, notamment pour mettre à jour des informations de capteurs satellites haute résolution à l'aide de capteurs satellites haute fréquence temporelle. En fait, ces deux types de données sont nécessaires pour réaliser du suivi de cultures : les images haute résolution spatiale sont indispensables pour discriminer les champs et les données à haute fréquence temporelle sont indispensables pour surveiller la croissance et la sénescence des plantes. Cette partie cherche à démontrer l'avantage d'utiliser un produit fusionné combinant des données haute résolution spatiale et temporelle plutôt que d'utiliser séparément ces deux types d'images.

Cette partie montre un exemple fusionnant une image du capteur SPOT 4 - HRVIR dégradée à 31,25 m et une image du capteur SPOT 4 - VEGETATION à la résolution kilométrique, toutes deux acquises dans les bandes spectrales rouge, proche infra-rouge et moyen infra-rouge. Le rapport de résolution entre les deux images est 32. La zone étudiée ici couvre une surface géographique de $68 \times 76 \text{ km}^2$ sur le site d'Ekrafane à 300 km au nord-est de Niamey au Niger. Les paramètres de la pyramide sont, classiquement, un échantillonnage dyadique, le filtre $\frac{\gamma+\varphi}{2}$ couplé à un élément structurant de rayon 1, centré et symétrique, une décomposition réalisée en 5 étages et une interpolation bilinéaire.

Dans une première expérience, la date d'acquisition est la même pour les deux images, le 17 août 1998. L'image initiale du capteur HRVIR et la partie extraite du capteur VEGETATION sur la même zone sont montrées sur les figures 7.10a) et 7.10b). Leurs histogrammes sont montrés sur les figures 7.10c) et 7.10d) respectivement. Les calculs seront effectués sur la seule bande spectrale rouge.

Afin de pouvoir comparer ces deux images, toutes les statistiques ont été réalisées sur la partie centrale de l'image acquise par le capteur HRVIR, composée de valeurs non nulles et sur la partie correspondante de l'image VEGETATION. La moyenne et l'écart type sont de 103,35 et 12,89 pour l'image acquise par HRVIR, 105,51 et 9,60 pour celle acquise par le capteur VEGETATION et 107,67 et 12,95 pour le produit fusionné.

Si aucune fusion n'est faite, il a été prouvé dans le chapitre 3 que l'image initiale et l'image reconstruite sont identiques. Dans le cas de la fusion, les différences proviennent essentiellement de celles entre l'image acquise par le capteur VEGETATION et l'image

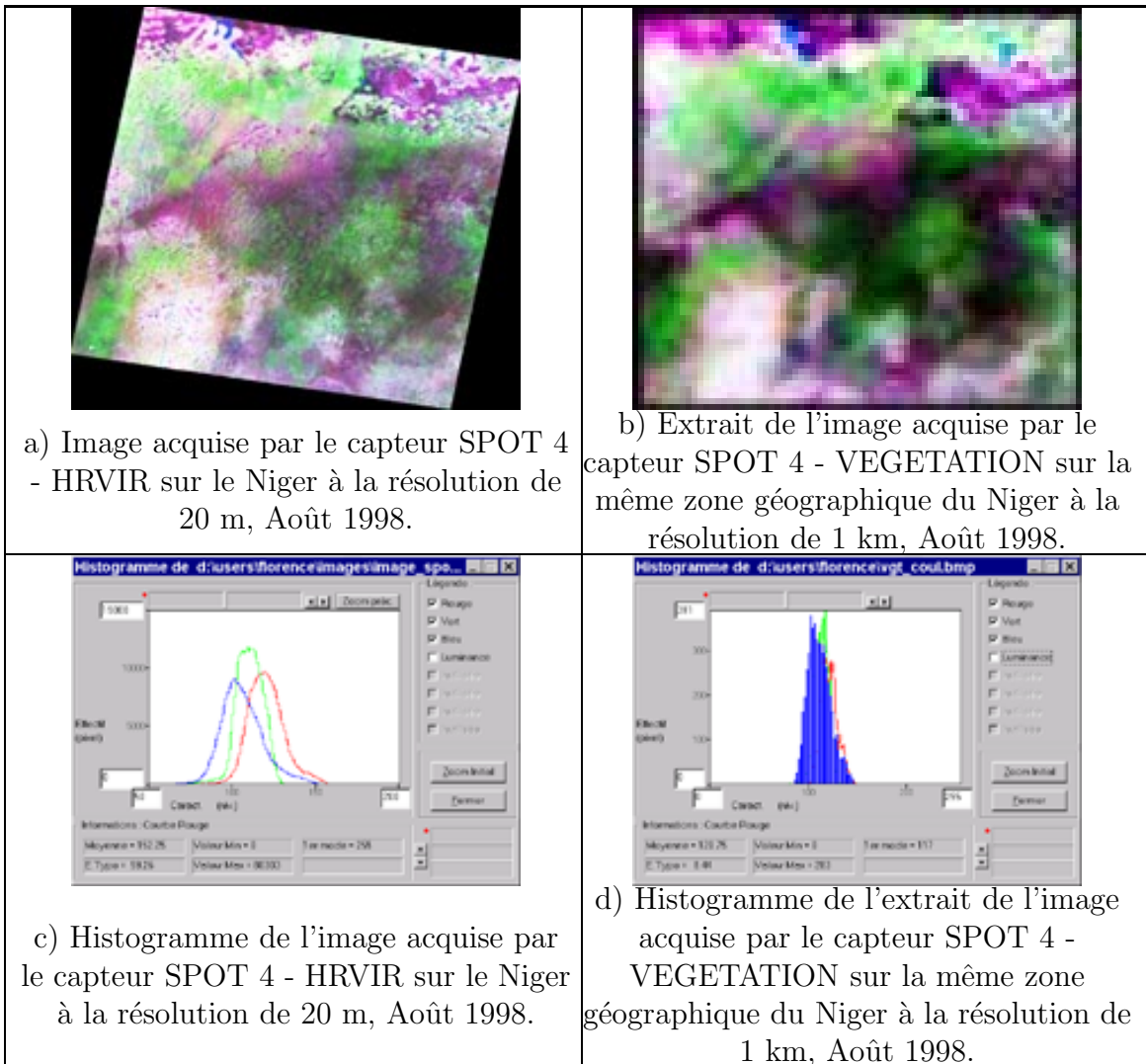


FIG. 7.10 – Comparaison des images initiales a) acquise par SPOT 4 - HRVIR et b) acquise par SPOT 4 - VEGETATION et de leurs histogrammes (c) et d)).

basse résolution issue de la pyramide, I_5 (cf. paragraphe 6.2.1).

L'information contenue dans l'image fusionnée provient d'une part des images de détails et d'autre part de l'image acquise par le capteur VEGETATION. Par nature, les images de détails, représentant les éléments fins, fournissent une variabilité de paysage plus grande que celle apportée par l'image acquise par le capteur VEGETATION. Comme dans cet exemple l'écart type de l'image acquise par le capteur VEGETATION est plus basse que celui de l'image acquise par le capteur HRVIR, la variabilité va essentiellement provenir des images de détails représentant l'image haute résolution et donc l'écart-type de l'image fusionnée est plus proche de celui de l'image haute résolution.

De plus, si l'interpolation satisfait la condition $Int_k(X + Y) = Int_k(X) + Int_k(Y)$, et dans le cas d'une fusion avec éventuellement dépassement de la dynamique, l'équation

(6.4) s'écrit

$$I_0 - IFusion_0 = \widetilde{I}_N^0 - \widetilde{IBR}^0 \quad (7.4)$$

Ainsi,

$$moyenne(I_0 - IFusion_0) = moyenne(\widetilde{I}_N^0 - \widetilde{IBR}^0) \quad (7.5)$$

où $moyenne(X)$ représente la valeur moyenne de l'image X et donc

$$moyenne(IFusion_0) = moyenne(I_0) - moyenne(\widetilde{I}_N^0) + moyenne(\widetilde{IBR}^0) \quad (7.6)$$

Le schéma d'interpolation peut changer la valeur moyenne de l'image de différence $I_N - IBR$; cependant, la nature même de l'interpolation permet de supposer que les variations seront faibles. En conséquence, la moyenne du produit fusionné va être étroitement liée à celle de l'image haute résolution à laquelle on soustrait la moyenne de l'image I_N et on ajoute celle de l'image IBR . Comme la décomposition filtre généralement plus une image que ne le fait un capteur basse résolution, cette différence est positive et la moyenne du produit fusionné est supérieure à celle de l'image haute résolution. Dans l'exemple proposé, $moyenne(I_0) - moyenne(I_5) + moyenne(IBR) = 103,34 - 101,09 + 105,50 = 107,76$. Cette valeur est effectivement proche de la moyenne de l'image fusionnée, 107,67, l'écart entre les deux est inférieur à 0,1%.

En conséquence, à une date donnée, l'image fusionnée est assez similaire à l'image haute résolution initiale, ce qui est satisfaisant, car aucune information d'évolution de paysage n'est ajoutée. La différence entre les deux images provient d'une part des caractéristiques différentes des deux capteurs en jeu et d'autre part des différences de corrections radiométriques appliquées entre les deux images d'entrée.

A la suite de cette première expérience, une deuxième fusion est réalisée sur le même principe. L'image acquise par le capteur HRVIR est fusionnée avec une image acquise sur la même zone par le capteur VEGETATION 15 jours plus tard, le premier septembre 1998 exactement.

La figure 7.11 montre un extrait sur la partie Nord-Ouest de la scène, représentant l'image initiale acquise par le capteur HRVIR le 17 août 1998 (image a), les deux images acquises par le capteur VEGETATION le 17 août (image b) et le 1^{er} septembre (image c), et les deux images fusionnées résultats (images d et e). Cette figure est intéressante pour plusieurs raisons. Tout d'abord, elle permet une comparaison visuelle sur un extrait significatif entre l'image acquise par le capteur HRVIR et l'image fusionnée obtenue à la même date (images a et d). D'autre part, elle permet de caractériser l'évolution du paysage en l'espace de deux semaines seulement au travers des images acquises par le capteur VEGETATION (images b et c) seulement ou bien par comparaison des deux images fusionnées (images d et e). Globalement, la comparaison des extraits des images acquises par le capteur VEGETATION montre que la réflectance du fond décroît dans la bande

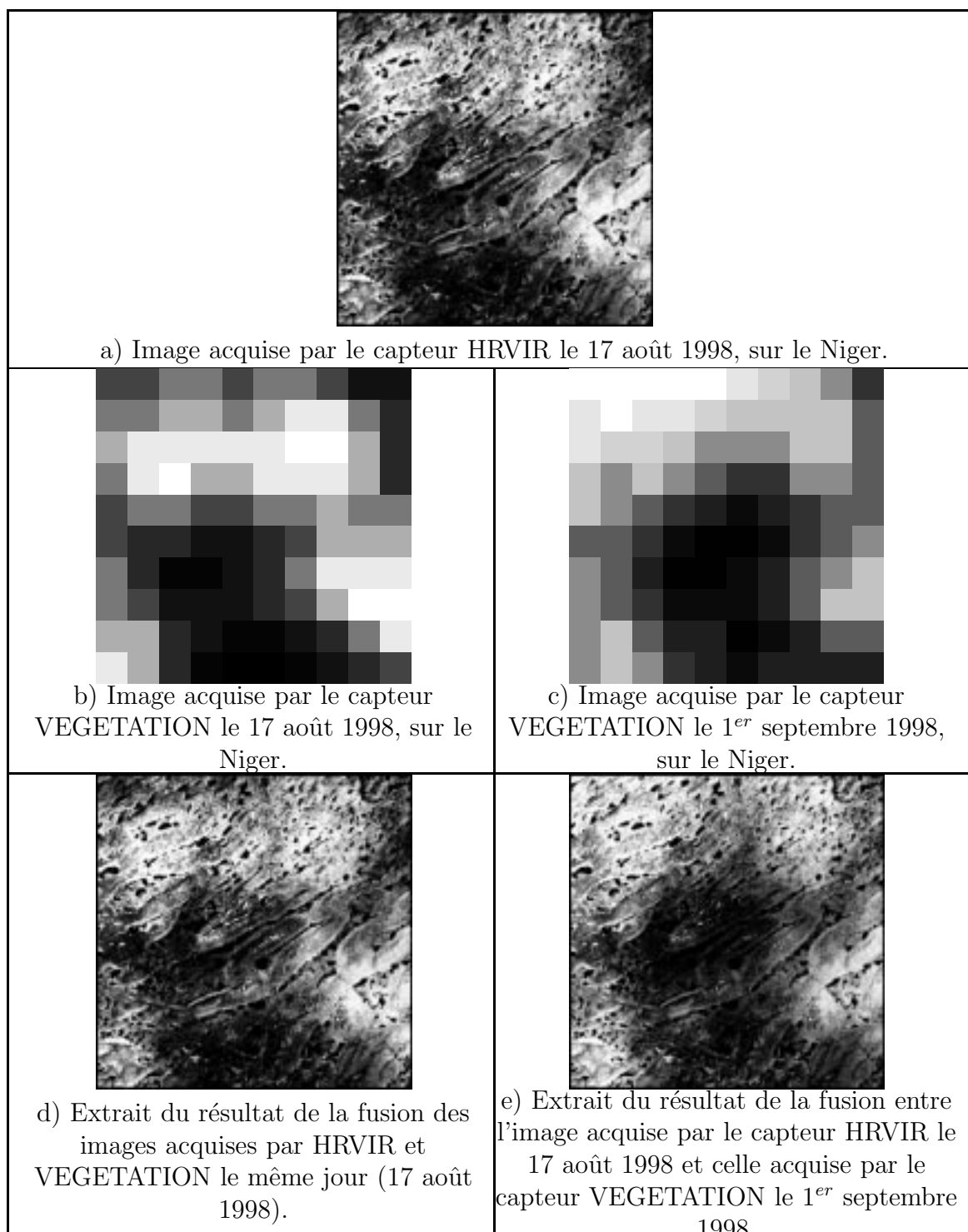


FIG. 7.11 – Résultat de la fusion une date et deux dates sur une même scène du Niger. Les images présentées sur les figures a à e correspondent à un extrait du nord ouest de la scène globale.

spectrale rouge dans la partie centrale de l'image - les pixels sont plus sombres - tandis qu'elle augmente en haut à gauche de cette image - les pixels sont plus clairs. Physiquement, cela signifie que le couvert herbacé se densifie entre le 17 août et le 1^{er} septembre au centre de l'image. Au contraire, les motifs dessinés par les arbres et contenus dans les détails ne changent pas pendant la même période comme en témoigne l'image acquise par HRVIR. Enfin, les images fusionnées permettent une localisation précise des évolutions, plus facilement interprétables que sur les images acquises par le capteur VEGETATION.

Le résultat de cette fusion est très intéressant car il permet de quantifier précisément, dans ces régions sub-sahariennes l'évolution de la surfaces en deux semaines au cours de la saison des pluies.

7.5 Conclusion

Les résultats obtenus sur la fusion spatio-temporelle d'images de télédétection impliquent un rappel et entraînent deux enseignements intéressants.

Tout d'abord, il faut rappeler que l'intérêt de ces résultats est une conséquence des propriétés de la pyramide morphologique présentée dans ce mémoire. En effet, l'existence d'une transformée inverse exacte de la décomposition des images permet une fusion complète car elle introduit la totalité de l'information spatiale extraite.

Ensuite, le premier enseignement est d'ordre essentiellement applicatif. Il concerne évidemment la capacité de cette fusion à construire de nouvelles possibilités de surveillance quand il est nécessaire d'avoir à la fois des données à haute résolution spatiale et à haute fréquence temporelle dans un même jeu de données. Une étape ultérieure de l'étude consistera à effectuer cette surveillance sur un jeu complet de données, comprenant des images à haute fréquence temporelle sur tout un cycle annuel et quelques images haute résolution spatiale dans la même période de temps, afin de tester les algorithmes proposés sur une durée plus importante.

Enfin, le second enseignement est d'ordre purement méthodologique. Les résultats de fusion d'images provenant de capteurs différents sont une approche nouvelle de la désagrégation de l'hétérogénéité qui est intégrée et occultée dans les pixels mixtes des instruments à faible résolution angulaire (comme VEGETATION, résolution de 1 km²).

Or, la décomposition de ces pixels grossiers est un enjeu important pour améliorer le suivi des surfaces et c'est pour cette raison que de nombreux travaux concernent la restitution de profils temporels de réflectance. Ainsi, des résultats significatifs ont été obtenus en recherchant l'inversion du modèle additif de composition de réflectance. Par exemple, les résultats obtenus dans le couplage des données à haute résolution spatiale et à haute fréquence temporelle [21, 19, 30, 96, 106, 144] ont montré qu'un nombre significatif de composants de la surface (8 à 12 par exemple) peut être caractérisé en réflectance avec

une bonne résolution temporelle. Cependant, cette capacité pose le problème d'une bonne représentation de la variabilité des réponses spectrales. En outre, elle s'appuie sur le besoin d'une bonne connaissance des champs spatiaux d'isovaleurs de réflectance.

Par rapport à ces travaux, la fusion spatio-temporelle avec la pyramide morphologique est donc une orientation nouvelle intéressante car cette condition préalable n'existe plus. En outre, les éléments spatiaux de paysage masqués dans les pixels grossiers sont réintroduits par la pyramide car ils sont connus explicitement grâce aux fonctions images des détails obtenus dans la décomposition.

Cinquième partie
CONCLUSION

Chapitre 8

Conclusion

Ce mémoire présente le principe de la pyramide morphologique et décrit son application à l'analyse d'images et à la fusion de données, notamment spatio-temporelle. Le choix de cette approche et de l'étude de ses propriétés découle des caractéristiques d'hétérogénéité que l'on rencontre généralement dans les images d'observation spatiale des surfaces terrestres. Précisément, la composition complexe d'une scène en objets munis d'attributs spatiaux et temporels très variables implique une analyse multi-niveaux afin de séparer les composants emboîtés et imbriqués à différentes échelles d'espace et de temps.

Dans cette perspective, la disponibilité de capteurs à très haute résolution spatiale ou à haute fréquence temporelle permet et permettra de plus en plus la perception d'une large gamme de caractéristiques de surfaces. C'est dans ce contexte actuel qu'il s'est trouvé motivant de développer cette approche multi-résolution d'analyse et de fusion par pyramide morphologique.

Les résultats montrent que les propriétés de la décomposition et sa transformée inverse respectent les hypothèses génériques de : 1) non équidistribution de la variabilité des signaux filtrés dans l'analyse ; 2) meilleur compromis possible de combinaison spatio-temporelle d'images dans la fusion ; 3) invariance par rapport aux transformées affines.

L'étude des principales grandes familles de paramétrisations possibles de la pyramide morphologique a montré également sa capacité d'approche déterministe d'extraction et de synthèse d'information grâce aux propriétés des opérateurs morphologiques.

Un dernier point très important à souligner est la nouvelle approche de décomposition de l'hétérogénéité que représente l'analyse multi-résolution non linéaire et la fusion spatio-temporelle correspondante. L'extraction de profils temporels de comportement à une échelle fine représente en effet une direction intéressante à approfondir.

Il apparaît donc clairement que la pyramide morphologique permet de nombreuses applications. Ce mémoire, en s'appuyant sur les propriétés de la pyramide morphologique, démontre son intérêt en analyse d'images et extraction d'attributs et en fusion de données,

éventuellement spatio-temporelle, notamment dans le domaine de l'observation de la terre. Il ouvre également des perspectives prometteuses dans plusieurs domaines du traitement d'images, tels que la compression de données et leur recalage géométrique.

Ainsi, d'autres perspectives sont apparues au cours de ce travail. Certaines, comme la navigation dans la pyramide (*cf.* section 3.5.3), la pré-segmentation (*cf.* section 4.6), la pré-compression (*cf.* section 4.7) et la compression de données (*cf.* section 4.8) ont déjà fait l'objet d'une petite pré-étude dans ce mémoire ; d'autres, comme le recalage des données multi-résolution (*cf.* section 6.5) y ont juste été évoquées. Toutes constituent des perspectives prometteuses amenées par cette thèse.

Premièrement, la navigation entre les échelles consiste à repérer, à une échelle quelconque donnée, un pixel ou un ensemble de pixels d'intérêt et de suivre son (leur) évolution dans les niveaux inférieurs ou supérieurs de la pyramide.

Ensuite, la pyramide morphologique peut contribuer à un processus de compression. En effet, l'apport de la pré-compression évoquée à la section 4.7 devrait permettre de réaliser un schéma de compression multi-résolution performant puisqu'elle devient un facteur multiplicatif de la compression de données proprement dite. De plus, elle peut faire bénéficier de ses propres performances car les calculs s'appliquent au plus proches voisins et reposent sur des comparateurs.

Finalement, une autre perspective intéressante est le recalage entre deux images. Les données satellites que l'on souhaite coupler proviennent en général de capteurs et plate-formes différents. C'est pourquoi, après la chaîne d'acquisition d'une image, il est nécessaire de rendre les données superposables. Là encore, la pyramide morphologique offre des perspectives intéressantes.

En conclusion, dans le cadre de l'observation de la Terre par satellite, les diverses applications de la pyramide morphologique ont un intérêt majeur. Par exemple, les applications thématiques comme la surveillance de couverts végétaux ou les études épidémiologiques nécessitent d'avoir des mises à jour de données régulières, couplées à une résolution spatiale suffisante. La fusion de données s'avère alors un outil indispensable.

De plus, le développement actuel de la technologie des capteurs embarqués à bord des satellites confirme la tendance en terme de données multi-résolution. D'une part, les capteurs multi-spectraux ont une résolution de plus en plus fine ; les méthodes d'analyse d'images doivent s'adapter à ces nouveaux besoins. D'autre part, ces nouveaux types de capteurs ne font qu'accroître les rapports de résolution entre les données satellites et rendent indispensable une méthode de fusion capable de les gérer.

Sixième partie

ANNEXE

La fusion de données Panchromatique / Multispectral

Ces travaux s'insèrent dans l'étude du Développement Exploratoire « Chaîne Image THR » (cf. figure 1). Dans le cadre des études de définition du produit $P+XS$ du futur système imageur PLEIADES - HR (suite du système SPOT), conduites par le CNES avec le support du CESBIO (expertise F. Laporterie) et de la société industrielle REALIX, j'ai été amenée à travailler sur l'évaluation de méthodes de fusion de données haute résolution optique panchromatique (P) et multispectrales (XS).

L'étude s'est menée de 1999 à 2002 en parallèle avec mes travaux de thèse.

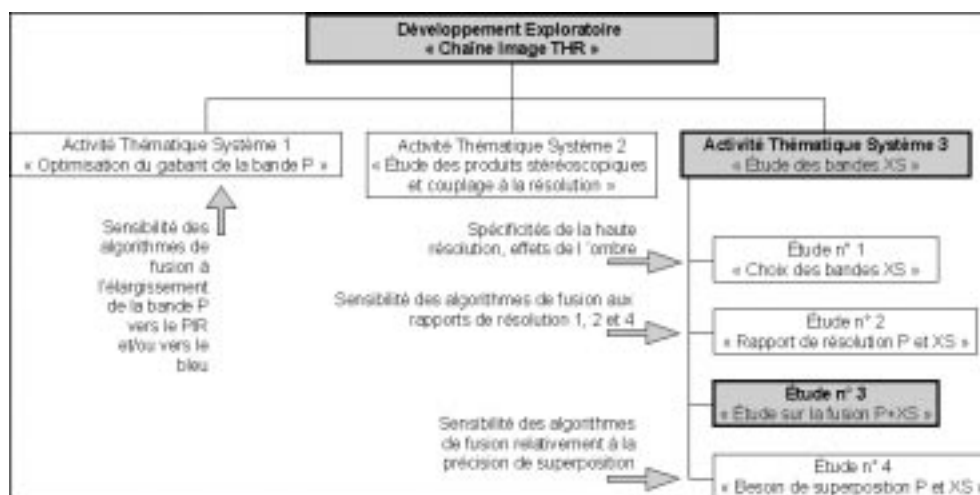


FIG. 1 – Développement Exploratoire « Chaîne Image THR ».

Ces travaux ont été structurés en deux phases principales.

Phase 1 :

- étude bibliographique et choix des laboratoires les plus performants en Europe - Responsabilité UPS-CESBIO et REALIX,

- réalisation des simulations de PLEIADES HR sur divers sites - Responsabilité CNES,
- visite des laboratoires afin d'évaluer les capacités opérationnelles des logiciels de fusion - Responsabilité REALIX et UPS-CESBIO,
- mise au point du processus permettant de comparer les cinq méthodes retenues et mise en place du référentiel - Responsabilité REALIX et UPS-CESBIO.

Phase 2 :

- traitement de la fusion de l'ensemble des données par les cinq laboratoires - Responsabilité Realix,
- évaluation qualitative auprès d'experts thématiques - Responsabilité CNES,
- analyse quantitative statistique - Responsabilité UPS-CESBIO,
- synthèse - Responsabilité REALIX et UPS-CESBIO.

La première phase concernant l'étude bibliographique sur les principaux algorithmes de fusion panchromatique / multispectraux s'est étendue d'octobre 1999 à décembre 2000. Les principaux résultats de cette phase ont été retranscrits dans le chapitre d'état de l'art de méthodes de fusion. Les quatre laboratoires européens sélectionnés sont les laboratoires :

- CENERG (Centre d'Energétique) de l'Ecole des Mines de Paris, basé à Nice Sophia Antipolis ;
- IROE (Istituto di Ricerca sulle Onde Elettromagnetiche - Consiglio Nazionale Delle Ricerche) de Florence, Italie ;
- du laboratoire SURFACES de l'université de Liège, Belgique ;
- du laboratoire RSD (Remote Sensing Department) de l'université de Trèves, Allemagne.
- A ces quatre méthodes de fusion proposées se rajoute la méthode proposée par le CNES.

La méthode du laboratoire CENERG est basée sur des transformations en ondelettes, celle de l'IROE sur un algorithme pyramidal, celle du laboratoire SURFACES sur des changements d'espace, celle de l'université de Trèves sur une régression linéaire et enfin celle du CNES sur un changement d'espace $RVB \rightarrow IHS$.

Des données simulées infra-métrique du futur satellite Pleiades ont été fournies à ces laboratoires sous la forme suivante : donnée panchromatique, correspondant à la haute résolution, et données multispectrales, jouant le rôle de la basse résolution, dans un rapport de résolution 2 ou 4. Les données multispectrales sont composées de 4 canaux, bleu, vert, rouge et proche infra rouge. Ces données ont été simulées à partir d'images prises par des capteurs aéroportés au dessus de plusieurs sites en France. De plus, un jeu de données de référence correspondant à des données multispectrales à la même résolution que les données panchromatiques est disponible. Ce dernier jeu ne sert pas pour la fusion, mais uniquement pour l'évaluation des différents algorithmes proposés par les laboratoires.

	Colmar	Moissac	Montpellier	Strasbourg	Toulouse	Vannes	Port de Bouc	Salon de Provence
Nombre de ligne	17320	11340	12820	13088	19700	13260	5504	6112
Nombre de colonne	2480	1800	1600	1408	2160	1640	1440	1408

TAB. 1 – Taille des images simulées Pleiades.

Les données traitées sont très volumineuses, la taille de chaque image est d'environ 15000 lignes par 2000 colonnes et chaque pixel est codé sur 16 bits. Des images des sites de Colmar, Moissac, Montpellier, Strasbourg, Toulouse, Vannes, Port de Bouc et Salon de Provence ont ainsi été utilisées. La taille des images haute résolution est indiquée dans le tableau 1.

Les données panchromatiques ont une résolution spatiale de 0,8 m pour les images des sites de Colmar, Moissac, Montpellier, Toulouse et Vannes, et une résolution spatiale de 0,6 m pour les images des sites de Strasbourg, Port de Bouc et Salon de Provence. Le rapport de résolution entre les données panchromatique et multispectrales est de 4 pour tous les sites, les sites de Toulouse et Montpellier possèdent en outre des données multispectrales de rapport de résolution 2 par rapport au panchromatique. Il existe également deux types d'images sur le site de Strasbourg, des données « enregistrées » et des données « non enregistrées » par rapport à l'image panchromatique. Ce jeu d'images est conçu pour tester l'effet d'un décalage géométrique entre les images panchromatique et multispectrales et évaluer la robustesse des méthodes de fusion.

De plus, des images provenant du satellite IKONOS (une image panchromatique et quatre images multispectrales dans les longueurs d'onde bleu, vert, rouge et proche infra-rouge) sur la ville de Toulouse ont également été proposées à la fusion.

Le très grand jeu de données assure une grande variabilité de paysages (*cf.* figure 2).



FIG. 2 – Exemples de la variabilité des paysages des images. De gauche à droite, sillons de champ sur le site de Colmar, terrain de sport sur le site de perpignan et zone résidentielle sur le site de Toulouse.

La suite de la première phase a donc consisté à mettre au point un processus précis décrivant la méthodologie commune à suivre par les laboratoires sélectionnés, afin de s'as-

surser que les fusions seraient faites dans des conditions identiques. C'est dans ce but que le CESBIO et REALIX se sont rendus dans les laboratoires concernés (*cf.* figure 3).

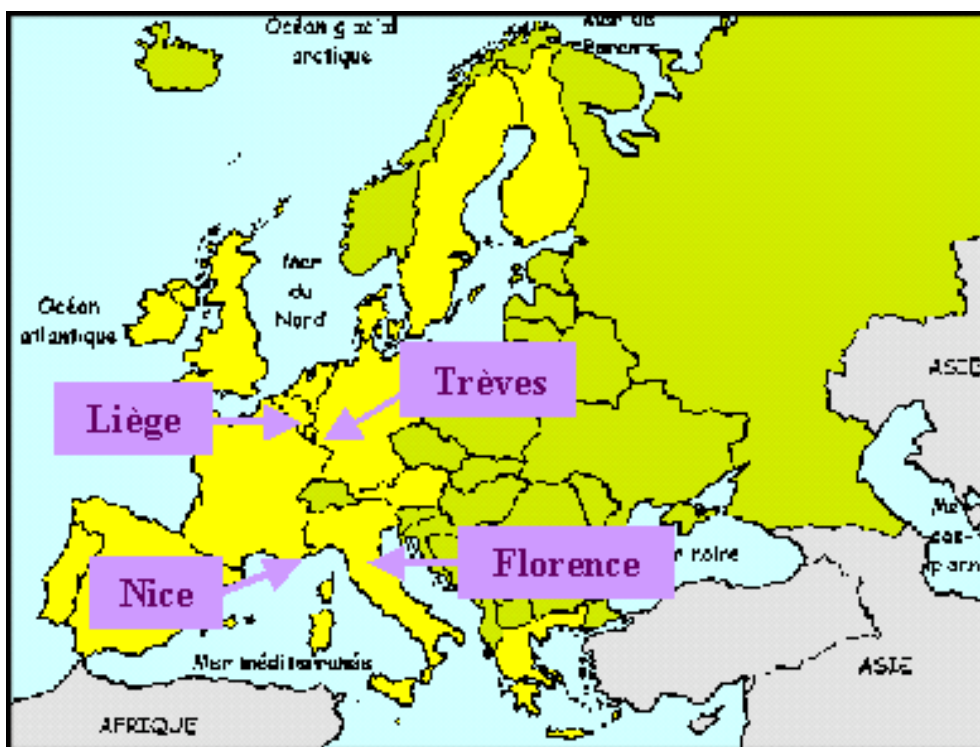


FIG. 3 – Localisation des quatre laboratoires européens sélectionnés.

La deuxième phase a donc consisté à évaluer les méthodes de fusion. Deux méthodologies ont été menées en parallèle. La première, gérée par le CNES a consisté à faire évaluer les résultats des différentes méthodes par des photo-interprètes. Ceux-ci ont en charge de désigner quelle méthode fournit visuellement les meilleurs résultats. La deuxième, à la charge de l'UPS-CESBIO, a consisté à réaliser une évaluation quantitative statistique des méthodes de fusion. Ces deux approches ont été développées en étroites relations, afin de pouvoir mettre en correspondance les deux types de résultats. Le but de l'étude quantitative est d'estimer, si des indicateurs statistiques - et lesquels - fournissent des résultats similaires aux photointerpètes.

Afin de réaliser l'étude quantitative, des imagerie représentant un objet précis ont été extraites de chaque site. Trois exemples en sont montrés sur la figure 2. Au total, 232 imagerie ont été extraites sur chaque bande spectrale de référence. Comme le montre le tableau 2 elles sont réparties en thématique urbaine, rurale, maritime et divers. Ainsi, le nombre d'imagerie de référence est $232 \times 5 = 1160$ images (4 bandes spectrales + 1 canal panchromatique). Les laboratoires CENERG, IROE, RSD et SURFACES fournissent chacun quatre images fusionnées par site correspondant aux quatre longueurs d'onde des

images multispectrales. La méthode de fusion CNES crée six images, 3 pour une composition colorée en couleurs naturelles et 3 pour une composition colorée en fausses couleurs. De plus, le site de Toulouse est étudié avec un rapport de résolution 2 et 4, et celui de Strasbourg dans le cas « normal » et le cas « déregistré ». Pour ces deux sites, le nombre d'images fusionnées obtenues est donc double. Ainsi, le nombre d'images fusionnées extraites est égal à $(26 + 22 + 24 + 33 + 31 \times 2 + 24 \times 2 + 25 + 20 + 27) \times (4 \times 4 + 1 \times 6) = 6314$.

Ensuite, deux types d'indicateurs statistiques sont calculés : les indicateurs bande à bande, qui comparent une image fusionnée avec son image de référence et les indicateurs couleurs, qui comparent trois images fusionnées formant soit une composition en couleurs naturelles soit en fausses couleurs avec les trois images de références correspondantes. Les indicateurs bande à bande sont le coefficient de corrélation, l'erreur relative sur la moyenne, l'erreur relative sur l'écart-type, l'erreur relative sur l'étendue, la moyenne de l'image de différence, l'écart-type de l'image de différence, la moyenne de l'image de contraste et l'écart-type de l'image de contraste. Les indicateurs couleurs sont l'erreur dans l'espace RVB, l'erreur dans l'espace IHS, la corrélation des images de teinte et la corrélation des images de saturation.

Les résultats obtenus ont été exploités par une analyse de variance afin d'évaluer les résultats globaux obtenus sur toutes les images, mais également les résultats par thème et ceux par bande spectrale. Les effets du rapport de résolution 2 par rapport à 4 sur le site de Toulouse et de la non registration sur le site de Strasbourg sont également étudiés. Ensuite, les résultats donnés par les différents indicateurs sont également étudiés.

Une fois les résultats de l'analyse quantitative et de l'évaluation par photointerprètes obtenus, l'ultime étape consiste à comparer les résultats obtenus et d'estimer si un ou plusieurs indicateurs quantitatifs s'approchent du résultat obtenu par les photo-interprètes.

	Colmar	Moissac	Montpellier	Perpignan	Strasbourg	Toulouse	Vannes	Salon de P.	Port de B.	
Urbain										
Aéroport					3			4		7
Bâtiments			1	5	7	4	5	3	4	29
Centre Commercial				1		1				2
Centre Ville						1				1
Cour								1		1
Cuves									11	11
Maison		1								1
Marché						1				1
Parking	1			1	4	1		2		9
Ombre						1				1
Piscines		1	2	3		1			1	8
Rond-Point				1		1		1		3
Route	5	1		3		4	1	1		15
Terrains Sport			2	2	4	1	4	1	1	15
Tuyaux									2	2
Voie Ferrée						1				1
Zones résidentielles	3		2	5	6	3	2	2	1	24
Rural										
Arbres	1			1	2	1		1		6
Champs	7	9	5	2		1	2	4		30
Chemins		1	4				2			7
Jardins				1	3	1				5
Forêts	3	1	6							10
Serres				1						1
Sillons	4	4	1	1						10
Vergers		3								3
Maritime										
Bateaux				2	2		3		3	10
Canal						1				1
Eau	1								2	3
Fleuve	1	1								2
Mer				2			3			5
Parc Ostréicole							1			1
Ponton									1	1
Sable				1			1			2
Divers										
Chantier							1			1
Travaux				1						1
Terre			1						1	2
	26	22	24	33	31	24	25	20	27	232

TAB. 2 – Description des imagerie extraites en fonction du site et de la thématique.

BIBLIOGRAPHIE

Bibliographie

- [1] Aiazzi Bruno, Alparone Luciano, Argenti Fabrizio et Baronti Stefano. – Wavelet and pyramid techniques for multisensor data fusion : a performance comparison varying with scale ratios. *In : Proceedings SPIE Europto, Image and signal processing for remote sensing V, 1999*, pp. 251–262.
- [2] Aiazzi Bruno, Alparone Luciano, Baronti Stefano et Carla Roberto. – An assessment of pyramid-based multisensor image data fusion. *In : Proceedings SPIE Europto, Image and signal processing for remote sensing, 1998*, pp. 237–247.
- [3] Auger Pierre. – *Hiérarchies et échelles en écologie*. – Naturalia Publication, 1992.
- [4] Bleau André, Guise Jacques De et LeBlanc A.-Robert. – A new set of fast algorithms for mathematical morphology. *CVGIP : Image understanding*, vol. 56, n° 2, septembre 1992, pp. 210–229.
- [5] Boucher P. et Goldberg Morris. – Transform image coding by vector quantization. *In : 9^{ème} symposium signal processing appli., Nice, France, 1983*, pp. 629–633.
- [6] Béthune Stanislas De, Muller F. et Binard M. – Adaptative intensity matching filters : a new tool for multiresolution data fusion. *In : AGARD SPP Symposium on multisensor systems and data fusion for telecommunications, remotes sensing and radar, 1997*, pp. 28/1–28/15.
- [7] Béthune Stanislas De, Muller Fabrice et Donnay Jean-Paul. – Fusion of multispectral and panchromatic images by local mean and variance matching filtering techniques. *In : Proceedings Fusion of Earth Data, Sophia Antipolis, 1998*, pp. 31–36.
- [8] Burt Peter J. et Adelson Edward H. – The laplacian pyramid as a compact image coder. *IEEE Transactions on Communications*, vol. 31, avril 1983, pp. 482–540.
- [9] Carper W. Joseph, Lillesand Thomas et Kiefer Ralph. – The use of Intensity-Hue-Saturation transformations for merging SPOT panchromatic and multispectral image data. *Photogrammetric Engineering and Remote Sensing*, vol. 56, n° 4, 1990, pp. 459–467.
- [10] Caumont Hervé, Nicolas Gaël, Laporterie Florence, Madier Jean-Pierre et Flouzat Guy. – *Fusion P et XS dans le domaine HR optique, rapport d'étude bibliographique*. – Rapport technique, REALIX, DE CNES ATS3-E3, 1999.
- [11] Chanda Bhabatosh, Kundu Malay K. et Padmaja Y. Vani. – A multi-scale morphologic edge detector. *Pattern Recognition*, vol. 31, n° 10, 1998, pp. 1469–1478.

- [12] Chassery Jean-Marc et Montanvert Annick. – *Géométrie discrète en analyse d'images*. – Traité des nouvelles technologies, série images, Hermes, 1991.
- [13] Chavez Pat. – Digital merging of landsat TM and digitized NHAP data for 1 :24,000 scale image mapping. *Photogrammetric Engineering and Remote Sensing*, vol. 52, n° 10, octobre 1986, pp. 1637–1646.
- [14] Chavez Pat et Bowell Jo Ann. – Comparison of the spectral information content of Landsat Thematic Mapper and SPOT for three different sites in the Phoenix, Arizona region. *Photogrammetric Engineering and Remote Sensing*, vol. 54, n° 12, décembre 1988, pp. 1699–1708.
- [15] Chavez Pat et Kwarteng Andrew Yaw. – Extracting spectral contrast in Landsat Thematic Mapper image data using selective Principal Component Analysis. *Photogrammetric Engineering and Remote Sensing*, vol. 55, n° 3, mars 1989, pp. 339–348.
- [16] Chavez Pat, Sides Stuart et Anderson Jeffrey. – Comparison of three different methods to merge multiresolution and multispectral data : Landsat TM and SPOT panchromatic. *Photogrammetric Engineering and Remote Sensing*, vol. 57, n° 3, mai 1991, pp. 295–303.
- [17] Chehikian A. – Algorithmes optimaux pour la génération d'images passe-bas et laplaciennes. *Traitement du signal*, vol. 9, n° 4, 1992, pp. 297–307.
- [18] Chen Ming-Hua et Yan P.-F. – A multiscaling approach based on morphological filtering. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 11, n° 7, juillet 1989, pp. 694–700.
- [19] Cherchali Selma. – *Contribution à la résolution du problème du pixel mixte en vue de l'amélioration de l'estimation de la productivité primaire nette en zone sahélienne (couplage haute et basse résolution spatiale)*. – Thèse de Doctorat, Université Paul-Sabatier, janvier 1995.
- [20] Cherchali Selma, Amram Olivier et Flouzat Guy. – Retrieval of temporal profiles of reflectances from simulated and real NOAA-AVHRR data over heterogeneous landscapes. *International Journal of Remote Sensing*, vol. 21, n° 4, 2000, pp. 753–775.
- [21] Cherchali Selma et Flouzat Guy. – The robustness of linear mixture modelling using low spatial resolution data. In : *Proceedings SPIE The European Symposium on Satellite Remote Sensing (Earth Surface Sensing), Rome, 1994*, pp. 26–30.
- [22] Chitwong Sakreya et Cheevasuvit Fusak. – Fusion of principal component of multispectral bands with PAN using HIS and wavelet transform. In : *Proceedings SPIE conference on Remote Sensing for environmental monitoring, GIS, applications, and geology, 2001*.
- [23] Cichosz Jacek et Meyer Fernand. – Morphological multiscale image segmentation. In : *Proceedings WIAMIS'97, Workshop on image analysis for multimedia interactive services, Louvain-la-Neuve (Belgium), 1997*, pp. 161–166.

- [24] Cliche G., Bonn F. et Teillet P. – Integration of the SPOT panchromatic channel into its multispectral mode for image sharpness enhancement. *Photogrammetric Engineering and Remote Sensing*, vol. 51, n° 3, mai 1985, pp. 311–316.
- [25] Crowley J.L. et Stern R.M. – Fast computation of the difference of low pass transform. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 6, 1984, pp. 212–222.
- [26] Diemer Claudius et Hill Joachim. – A local correlation approach for the fusion of remote sensing data with different spatial resolution. In : *Proceedings Fusion of Earth data, Sophia Antipolis, France, 2000*, pp. 91–98.
- [27] Djamdji Jean-Pierre, Bijaoui Albert et Manière Roger. – Geometrical registration of images - the multiresolution approach. *Photogrammetric Engineering and Remote Sensing*, vol. 59, 1993, pp. 645–653.
- [28] Dwivedi R. S., Ramana K. V., Thammappa S. S. et Singh A. N. – The utility of IRS-1C LISS-III and PAN-merged data for mapping salt-affected soils. *Photogrammetric Engineering and Remote Sensing*, vol. 67, n° 10, octobre 2001, pp. 1167–1175.
- [29] Eichmann George, Lu Chao, Zhu Jianxin et Li Yao. – Pyramidal image processing using morphology. In : *Proceedings SPIE Applications of digital image processing XI, San Diego, California, 1988*, pp. 30–37.
- [30] Faivre Robert et Fisher Alberte. – Predicting crops reflectance using satellite data observing mixed pixels. In : *Proceedings 17 th International Biometric Conference (IBC), Hamilton, Ontario, 1994*.
- [31] Faivre Robert et Fisher Alberte. – Predicting crop reflectances using data observing mixed pixels. *Journal of Agricultural, Biological and Environmental Statistics*, vol. 2, n° 1, 1997, pp. 87–107.
- [32] Fan Kuo-Chin, Chen Den-Fong et Wen Ming-Gang. – Skeletonization of binary images with nonuniformwidth via block decomposition and contour vector matching. *Pattern Recognition*, vol. 31, n° 7, 1998, pp. 823–838.
- [33] Filiberti Daniel P., Marsh Stuart E. et Schowengerdt Robert A. – Synthesis of imagery with high spatial and spectral resolution from multiple image sources. *Optical Engineering*, vol. 33, n° 8, août 1994, pp. 2520–2528.
- [34] Florencio Dinei A. F. et Schafer Ronald W. – A non-expansive pyramidal morphological image coder. In : *Proceedings IEEE International Conference Image Processing, Austin, Texas, 1994*, pp. 331–335.
- [35] Flouzat Guy, Amram Olivier, Laporterie Florence et Cherchali Selma. – Multi-resolution analysis and reconstruction by a morphological pyramid in the remote sensing of terrestrial surfaces. *Signal Processing*, vol. 81, n° 10, octobre 2001, pp. 2171–2185.
- [36] Flouzat Guy, Laporterie Florence et Amram Olivier. – Fusion d'images satellitaires à haute et basse résolution dans une pyramide morphologique. In : *23ème journée ISS, ENSMP, Paris, 2000*.

- [37] Flouzat Guy et Moueddene K. – Computer-aided interpretation of complex geological patterns in remote sensing. *In : Proceedings of IGARSS Zürich, 1986*, pp. 783–786.
- [38] Froment Jacques et Morel Jean-Michel. – Analyse multiéchelle, vision stéréo et ondelettes. *In : Lectures Notes in Mathematics, 1438, les ondelettes en 1989*, éd. par Lemarié P.G., pp. 51–80. – Springer-Verlag, 1989.
- [39] Garguet-Duport Bruno, Girel J., Chassery J.M. et Pautou G. – The use of multiresolution analysis wavelets transform for merging SPOT panchromatic and multispectral data. *Photogrammetric Engineering and Remote Sensing*, vol. 62, n° 9, septembre 1996, pp. 1057–1066.
- [40] Goldberg Morris et Wang Limmin. – Comparative performance of pyramid data structures for progressive image transmission. *IEEE transactions on Communications*, vol. 39, n° 4, avril 1991, pp. 540–548.
- [41] Goutsias John et Heijmans Henk. – Non linear multiresolution signal decomposition schemes - part I : Morphological pyramids. *IEEE Transactions on Image Processing*, vol. 9, n° 11, novembre 2000, pp. 1862–1876.
- [42] Grossman A. et Morlet J. – Decomposition of hardy functions into square integrable wavelets of constant shape. *SIAM J. Math. Anal.*, 1984, pp. 723–736.
- [43] Haralick Robert, Lin Charlotte, Lee James et Zhuang Xinhua. – Multiresolution morphology. *In : First International Conference on computer vision, 1987*. pp. 3–11. – The Computer Society of the IEEE.
- [44] Haralick Robert, Zhuang Xinhua, Lin Charlotte et Lee James. – The digital morphological sampling theorem. *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 37, n° 12, décembre 1989, pp. 2067–2090.
- [45] Haralick Robert M., Sternberg Stanley R. et Zhuang Xinhua. – Image analysis using mathematical morphology. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 9, n° 4, juillet 1987, pp. 532–550.
- [46] Harris Jeff R., Murray Richard et Hirose Tom. – IHS transform for the integration of radar imagery with other remotely sensed data. *Photogrammetric Engineering and Remote Sensing*, vol. 56, n° 12, décembre 1990, pp. 1631–1641.
- [47] Ho Yo-Sung et Gersho Allen. – A pyramidal image coder with contour-based interpolative vector quantization. *In : Proceedings SPIE Visual Communication and Image Processing IV, 1989*, pp. 733–740.
- [48] Hu Y. H., Lee H. B. et Scarpace F. L. – Optimal linear spectral unmixing. *IEEE Transactions on Geoscience and Remote Sensing*, vol. 39, n° 7, juillet 2001, pp. 1533–1536.
- [49] Imai Francisco H. et Berns Roy S. – High-resolution multi-spectral image archives - a hybrid approach. *In : Proceedings of IS&T/SID Sixth Color Imaging Conference, 1998*, pp. 224–227.
- [50] Jackway Paul T. – Gradient watersheds in morphological scale-space. *IEEE Transactions on Image Processing*, vol. 5, n° 6, juin 1996, pp. 913–921.

- [51] Jolion Jean-Michel. – Analyse d'image : le modèle pyramidal. *Traitement du signal*, vol. 7, n° 1, 1990, pp. 5–17.
- [52] Jolion Jean-Michel. – Analyse multirésolution du contraste dans les images numériques. *Traitement du signal*, vol. 11, n° 3, 1994, pp. 245–255.
- [53] Jolion Jean-Michel et Rosenfeld Azriel. – *A pyramid framework for early vision*. – Kluwer Academic Publishers, 1994.
- [54] King R. W. et Waite W. P. – Digital color space transformations. *In : Proceedings of IGARSS'84 Symposium, Strasbourg, France, 1984*, pp. 649–654.
- [55] Laferté Jean-Marc, Heitz Fabrice, Pérez Patrick et Fabre Eric. – Fusion d'images multirésolutions par modèles markoviens hiérarchiques. *In : 15ème colloque GRETSI, Juan-Les-Pins, 1995*.
- [56] Laporterie Florence, Amram Olivier, Flouzat Guy, Pilich Ewa et Gay Michel. – Data fusion thanks to an improved morphological pyramid approach : comparison loop on simulated images and application to SPOT 4 data. *In : IGARSS, Honolulu, Hawaii, 2000*.
- [57] Laporterie Florence, Flouzat Guy et Amram Olivier. – Morphological pyramid : a method for very high resolution satellite image analysis. *In : 50ème IAF, Amsterdam, The Netherlands, 1999*.
- [58] Laporterie Florence, Flouzat Guy et Amram Olivier. – Mathematical morphology multi-level analysis of trees patterns in savannas. *In : IGARSS, Sydney, Australia, 2001*.
- [59] Laporterie Florence, Flouzat Guy et Amram Olivier. – Multiresolution characterization of the water surfaces and the vegetation cover by non linear pyramid methods applied to satellites images. *In : 52ème IAF, Toulouse, France, 2001*.
- [60] Laporterie Florence, Flouzat Guy et Amram Olivier. – Spatial and temporal data fusion with morphological pyramid for terrestrial surfaces survey. *In : Proceedings SPIE 8th remote sensing symposium, Proceedings SPIE Image and Signal Processing for Remote Sensing VII, 2001*, pp. 199–210.
- [61] Laporterie Florence, Flouzat Guy et Amram Olivier. – The morphological pyramid and its applications to remote sensing : multiresolution data analysis and features extraction. *Image Analysis and Stereology*, mars 2002.
- [62] Liang Su, Ahmadi M. et Shridhar M. – Segmentation of handwritten interference marks using multiple directional stroke planes and reformalizes morphological approach. *IEEE Transactions on Image Processing*, vol. 6, n° 8, août 1997, pp. 1195–1202.
- [63] Lindeberg Tony. – Scale-space theory : a basic tool for analyzing structures at different scales. *In : Advances in applied statistics, statistics and images : 2. A supplement to Journal of Applied Statistics, 1994*, éd. par Mardia K. V., chap. 11, pp. 225–270. – Carfax Publishing Company, 1994.
- [64] Liu Jian Guo. – Smoothing filter-based intensity modulation : a spectral preserve image fusion technique for improving spatial details. *International Journal of Remote Sensing*, vol. 21, n° 18, décembre 2000, pp. 3461–3472.

- [65] Liu Jian Guo et Moore J. McM. – Pixel block intensity modulation : adding spatial detail to tm band 6 thermal imagery. *International Journal of Remote Sensing*, vol. 19, n° 13, septembre 1998, pp. 2477–2491.
- [66] Il Bjarne Ersbø, Conradsen Knut, Larsen Rasmus, Nielsen Allan et Nielsen Thomas. – Fusion of SPOT HRV XS and orthophoto data using a markov random field. *In : Proceedings Fusion of Earth Data, Sophia Antipolis, 1998*, pp. 25–30.
- [67] Mallat Stéphane. – Multifrequency channel decomposition of images and wavelet models. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 37, n° 12, décembre 1989, pp. 2091–2110.
- [68] Mallat Stéphane. – *A theory for multiresolution signal decomposition : the wavelet representation*. – Rapport technique, University of Pennsylvania, 1989.
- [69] Mallat Stéphane. – A theory for multiresolution signal decomposition : the wavelet representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 11, n° 7, juillet 1989, pp. 674–693.
- [70] Mallat Stéphane. – *A wavelet tour of signal processing*. – Academic Press, 1998.
- [71] Mangolini Marc, Ranchin Thierry et Wald Lucien. – Evaluation de la qualité des images multispectrales à haute résolution spatiale dérivées de SPOT. *Bulletin de la Société Française de Photogrammétrie et Télédétection*, vol. 137, 1998, pp. 24–29.
- [72] Manzanera Antoine et Jolion Jean-Michel. – Pyramide irrégulière : une représentation pour la vision exploratoire. *Traitement du signal*, vol. 12, n° 2, 1995, pp. 169–176.
- [73] Maragos Petros. – Morphology-based symbolic image modeling, multi-scale nonlinear smoothing and pattern spectrum. *In : IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Ann Arbor, Mi., 1988*, pp. 766–773.
- [74] Marr David. – Visual information processing : the structure and creation of visual representation. *In : Phil. Trans. R. Soc. Lond., B, London, 1980*, pp. 199–218.
- [75] Marr David. – *Vision, a computational investigation into the human representation and processing of visual representation*. – Freeman, 1982.
- [76] Marr David et Hildreth E. – Theory of edge detection. *In : Proc. R. Soc. Lond., B, 1980*, pp. 187–217.
- [77] Marshall S. et Matsopoulos G.K. – Fusion of MR and CT images of the human brain using multiresolution morphology. *In : Mathematical Morphology and Its Applications to Image Processing*, éd. par Serra J. et Soille P., pp. 317–324. – Kluwer Academic Publishers, 1994.
- [78] Matsopoulos G. K., Marshall S. et Brunt J. N. H. – Multiresolution morphological fusion of MR and CT images of the human brain. *In : IEEE Proceedings Vision, image and signal processing, 1994*, pp. 137–142.
- [79] Meer F. Van Der. – Spectral unmixing of Landsat Thematic Mapper data. *International Journal of Remote Sensing*, vol. 16, n° 16, novembre 1995, pp. 3189–3194.

- [80] Meer F. Van Der. – What does multisensor image fusion adds in terms of information content for visual interpretation. *International Journal of Remote Sensing*, vol. 18, n° 2, 1997, pp. 445–452.
- [81] Meer F. Van Der. – Iterative spectral unmixing. *International Journal of Remote Sensing*, vol. 20, n° 17, novembre 1999, pp. 3431–3436.
- [82] Meyer Yves. – Principe d’incertitude, bases hilbertiennes et algèbres d’opérateurs. *In : Séminaire Bourbaki, 1986.*
- [83] Minghelli Audrey, Polidori Laurent et Mangolini Marc. – Image fusion for the simulation of hyperspectral satellite images. *In : Proceedings Fusion of Earth Data, Sophia Antipolis, 1998*, pp. 97–102.
- [84] Minghelli-Roman Audrey, Mangolini Marc, Petit Michel et Polidori Laurent. – Spatial resolution improvement of MeRIS images by fusion with TM images. *IEEE Transactions on Geoscience and Remote Sensing*, vol. 39, n° 7, juillet 2001, pp. 1533–1536.
- [85] Mitiche Amar et Aggarwal J. K. – Multiple sensor integration / fusion through image processing : a review. *Optical Engineering*, vol. 25, n° 3, mars 1986, pp. 380–386.
- [86] Müller Frank, Brigger Patrick, Illgner Klaus et Unser Michael. – Multiresolution approximation using shifted splines. *IEEE Transactions on Signal Processing*, vol. 46, n° 9, septembre 1998, pp. 2555–2558.
- [87] Montiel Eugenia, Aguado Alberto, Garza-Jinich Maria et Alarcón J. – Image manipulation using m-filter in a pyramidal computer model. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 17, n° 11, novembre 1995, pp. 1110–1115.
- [88] Moran M. Susan. – A window-base technique for combining Landsat Thematic Mapper thermal data with higher-resolutin multispectral data over agricultural lands. *Photogrammetric Engineering and Remote Sensing*, vol. 56, n° 3, mars 1990, pp. 337–342.
- [89] Mukhopadhyay Susanta et Chanda Bhabatosh. – Fusion of 2d grayscale images using multiscale morphology. *Pattern Recognition*, vol. 34, n° 10, octobre 2001, pp. 1939–1949.
- [90] Mullon C. – Le transfert d’échelle. *In : Seminfo 4*, éd. par Editions ORSTOM Paris. – 1991.
- [91] Núñez Jorge, Otazu Xavier, Fors Octavi et Prades Albert. – Fusion and reconstruction of Landsat and SPOT images using wavelets. *In : Proceedings Fusion of Earth Data, Sophia Antipolis, 1998*, pp. 103–108.
- [92] Núñez Jorge, Otazu Xavier, Fors Octavi, Prades Albert, Palà Vincenç et Arbiol Román. – Data fusion of SPOT and LANDSAT images using additive multiresolution wavelet decomposition. *In : Proceedings EUROPTO conference on Image and Signal Processing for Remote Sensing, 1998*, pp. 202–212.

- [93] Núñez Jorge, Otazu Xavier, Fors Octavi, Prades Albert, Pala Vincenç et Arbiol Román. – Multiresolution-based image fusion with additive wavelet decomposition. *IEEE Transactions on Geoscience and Remote Sensing*, vol. 37, n° 3, mai 1999, pp. 1204–1211.
- [94] Núñez Jorge, Otazu Xavier, Fors Octavi, Prades Albert, Pala Vincenç et Arbiol Román. – Image fusion with additive multiresolution wavelet decomposition. Application to SPOT+LANDSAT images. *Journal of the Optical Society of America A*, vol. 16, n° 3, mars 1999, pp. 467–474.
- [95] O’Neill R. V. – Hierarchy theory and global change. In : *Scales and Global Change*, éd. par T. Rosswall R. G. Woodmansee et eds P. G. Risser, pp. 29–45. – SCOPE 35, John Wiley and Sons Ltd publ., 1988.
- [96] Ouaidrari H. – *Utilisation de données NOAA-AVHRR pour modéliser la production du mil dans l’écosystème sahélien (Niger)*. – Thèse de Doctorat, Université Paris VII, octobre 1994.
- [97] Paradella Waldir Renato, Bignelli Pedro Alberto, Veneziani Paulo, Pietsch R. et Toutin T. – Airborne and spaceborne synthetic aperture radar SAR integration with Landsat TM and gamma ray spectrometry for geological mapping in a tropical rainforest environment, the Carajás Mineral Province, Brazil. *International Journal of Remote Sensing*, vol. 18, n° 7, mai 1997, pp. 1483–1501.
- [98] Pavel Misha et Sharma Ravi K. – Fusion of Radar images rectification without the flat Earth assumption. In : *Proceedings of SPIE, 1996*, pp. 108–118.
- [99] Pichler Olaf, Teuner Andreas et Hosticka Bedrich. – A comparison of texture feature extraction using adaptative gabor filtering, pyramidal and tree structures wavelet transform. *Pattern Recognition*, vol. 29, n° 5, 1996, pp. 733–742.
- [100] Poglio Thierry, Blanc Philippe et Aloisi Roberto. – Applicability of superresolution to the VEGETATION sensor. In : *Proceedings Fusion of Earth data, Sophia-Antipolis, France, 2000*, pp. 73–81.
- [101] Pohl Christine et Genderen Van. – Multisensor image fusion in remote sensing : concepts, methods and applications. *International Journal of Remote Sensing*, vol. 19, n° 5, 1998, pp. 823–854.
- [102] Pradines D. – Improving SPOT images size and multispectral resolution. In : *Proceedings SPIE Earth Remote Sensing Using Landsat Thematic Mapper and SPOT sensor systems, 1986*, pp. 98–102.
- [103] Price John C. – Combining panchromatic and multispectral imagery from dual resolution satellite instruments. *Remote sensing of environment*, vol. 21, 1987, pp. 119–128.
- [104] Proença Maria. – *La transformée en ondelettes appliquée au traitement d’images de télédétection*. – Thèse en traitement d’images, 1992.
- [105] Proenca M.C. et Flouzat G. – Orthogonal wavelet analysis : a multiresolution tool applied to remote sensing images. In : *Proceedings IGARSS’91, Helsinki, 1991*, pp. 1463–1465.

- [106] Puyou-Lascassies P. – *Surveillance des comportements en paysages agricoles hétérogène avec des données satellitaires à haute et à basse résolution spatiale.* – Thèse de Doctorat, Université Paul-Sabatier, décembre 1994.
- [107] Ranchin Thierry. – *Application de la transformée en ondelettes et de l'analyse multirésolution au traitement de images de télédétection.* – Thèse en sciences de l'ingénieur, 1993, 110p.
- [108] Ranchin Thierry. – Improving spatial resolution of spaceborne images using airborne images. *In : Proceedings Fusion of Earth Data, Sophia Antipolis, 1998*, pp. 37–44.
- [109] Ranchin Thierry et Wald Lucien. – Fusion of high spatial and spectral resolution images : the ARSIS concept and its implementation. *Photogrammetric Engineering and Remote Sensing*, vol. 66, n° 1, janvier 2000, pp. 49–61.
- [110] Ranchin Thierry, Wald Lucien, Mangollini Marc et Penicand Claude. – On the assessment of merging processes for the improvement of the spatial resolution of multispectral SPOT XS images. *In : Proceedings Fusion of Earth Data, Cannes, France, 1996*, pp. 59–67.
- [111] Raptis Vassilis, Vaughan Robin, Ranchin Thierry et Wald Lucien. – An assessment of different data fusion methods for the classification of an urban environment. *In : Proceedings Fusion of Earth Data, Sophia Antipolis, 1998*, pp. 167–182.
- [112] Reginster Isabelle. – Fusion of remotely sensed data and census data for a better understanding and management of the city environment. *In : Proceedings Fusion of Earth Data, Sophia Antipolis, 1998*, pp. 205–210.
- [113] Rosenfeld Azriel. – Some useful properties of pyramids. *In : Multiresolution Symposium, 1982*, pp. 2–5.
- [114] Rosenfeld Azriel et Thurston M. – Edge and curve detection for visual scene analysis. *IEEE Transactions on Computer*, vol. TC-20, 1971, pp. 512–569.
- [115] Roswall T., Woodmandsee R. G. et Risser P. G. – *Scales and Global change.* – SCOPE 35, John Wiley and Sons Ltd publ., 1988.
- [116] Rowe J. S. – The level-of-integration concept and ecology. *Ecology*, vol. 42, 1961, pp. 420–427.
- [117] Salembier Philippe, Brigr Patrick, Casas Josep R. et Pardàs Montse. – Morphological operators for image and video compression. *IEEE Transactions on Image Processing*, vol. 5, n° 6, juin 1996, pp. 881–898.
- [118] Salembier Philippe et Serra Jean. – Morphological multiscale image segmentation. *In : Proceedings SPIE Visual Communications and image Processing, 1992*, pp. 620–630.
- [119] Schetselaar Ernst. – Fusion by ihs transform : Should we use cylindrical or spherical coordinates? *International Journal of Remote Sensing*, vol. 19, n° 4, mars 1998, pp. 759–765.
- [120] Schetselaar Ernst. – On preserving spectral balance and its advantages for geological image interpretation. *Photogrammetric Engineering and Remote Sensing*, vol. 67, n° 8, août 2001, pp. 925–934.

- [121] Scheunders Paul et Backer Steve De. – Fusion and merging of multispectral images with use of multiscale fundamental forms. *Journal of the Optical Society of America A*, vol. 18, n° 10, octobre 2001, pp. 2469–2477.
- [122] Schowengerdt Robert. – Reconstruction of multispatial, multispectral image data using spatial frequency content. *Photogrammetric Engineering and Remote Sensing*, vol. 46, n° 10, 1980, pp. 1325–1334.
- [123] Serra Jean. – *Image analysis and mathematical morphology*. – Academic Press, 1982.
- [124] Serra Jean. – Structures syntaxiques en morphologie mathématique. In : *Premier colloque image : Traitement, Synthèse, Technologie et Applications, 1984*.
- [125] Serra Jean. – *Éléments de théorie pour l'optique morphologique*. – These Université de Paris VI - Paris, janvier 1986.
- [126] Serra Jean. – *Image analysis and mathematical morphology volume 2 : Theoretical advances*. – Academic Press, 1988.
- [127] Serra Jean. – Morphological filtering : an overview. *Signal processing*, vol. 38, 1994, pp. 3–11.
- [128] Serra Jean et Salembier Philippe. – Connected operators and pyramids. In : *Proceedings SPIE, 1993*, pp. 65–76.
- [129] Sharma Ravi et Pavel Misha. – Adaptative and statistical image fusion. *Society for Information Display*, vol. XXVII, mai 1996, pp. 969–972.
- [130] Sharma Ravi K., Leen Todd K. et Pavel Misha. – Probabilistic image sensor fusion. *Advances in Neural Information Processing systems, M.S. Kearns, S. A. Solla and D.A. Cohn eds, the MIT press, Cambridge, MA*, vol. 11, 1999, pp. 824–830.
- [131] Shen J. et Castan S. – An optimal linear operator for edge detection. In : *Proceedings Conference on computer vision and pattern recognition, Miami, Florida, USA, 1986*, pp. 109–114.
- [132] Sun Fang-Kuo et Maragos Petros. – Experiments on image compression using morphological pyramids. In : *Proceedings SPIE Visual Communications and Image Processing IV, 1989*, pp. 1303–1312.
- [133] Tanimoto Steven. – Image transmission with gross information first. *Computer Graphics and Image Processing*, vol. 9, janvier 1979, pp. 72–76.
- [134] Tanimoto Steven et Pavlidis Theodosios. – A hierarchical data structure for picture processing. *Computer Graphics and Image Processing*, vol. 4, n° 2, juin 1975, pp. 104–119.
- [135] Toet A. – A morphological pyramidal image decomposition. *Pattern Recognition Letters*, vol. 9, mai 1989, pp. 255–261.
- [136] Tom V. T. – A synergistic approach for multispectral image restoration using reference imagery. In : *Proceedings IGARSS, 1986*, pp. 559–564.
- [137] Toutin Th. – Intégration de données multisources : comparaison de méthodes géométriques et radiométriques. *International Journal of Remote Sensing*, vol. 16, n° 15, 1995, pp. 2795–2811.

- [138] Trangmar B. B., Yost R. S. et Uehara G. – Application of geostatistics to spatial studies of soil properties. *Advances in Agronomy*, vol. 38, 1985, pp. 45–94.
- [139] Tripathi Nitin Kumar et Gokhale K. V. G. K. – Directional morphological image transforms for lineament extraction from remotely sensed images. *International Journal of Remote Sensing*, vol. 21, n° 17, novembre 2000, pp. 3281–3292.
- [140] Tu Te-Ming, Su Shun-Chi, Shyu Hsuen-Chyun et Huang Ping. – A new look at ihs-like image fusion methods. *Information fusion*, vol. 2, n° 3, septembre 2001, pp. 177–186.
- [141] Urban D. L., O’Neill R. V. et Shugart H. H. Jr. – Landscape ecology. a hierarchical perspective can help scientists understand spatial patterns. *BioScience*, vol. 37, 1987, pp. 119–127.
- [142] Vachier Corinne, Meyer Fernand, Gratin C. et Talbot H. – Filtrage par décomposition morphologique : application à l’extraction de structures rectilignes. In : *9ème congrès reconnaissance des formes et intelligence artificielle, Paris, 1994*, pp. 255–263.
- [143] Vaughan Robin et Oune Omar. – The fusion of TM, SPOT, IRS and KFA data for urban change detection in Tripoli. In : *Proceedings Fusion of Earth Data, Sophia Antipolis, 1998*, pp. 191–203.
- [144] Vignolles C. – *Modélisation de la production agricole à moyenne échelle. Apport du couplage des mesures par satellite à haute résolution spatiale et à haute fréquence temporelle.* – Thèse de Doctorat, Université Paul-Sabatier, décembre 1997.
- [145] Wald Lucien. – Some terms of reference in data fusion. *IEEE Transactions on Geoscience and Remote Sensing*, vol. 37, n° 3, mai 1999, pp. 1190–1193.
- [146] Wang Limmin et Goldberg Morris. – Reduced-difference pyramid : a data structure for progressive image transmission. *Journal of Optical Engineering*, vol. 28, n° 7, juillet 1989, pp. 708–716.
- [147] Welch R. et Ehlers Manfred. – Merging multiresolution SPOT HRV and Landsat TM data. *Photogrammetric Engineering and Remote Sensing*, vol. 53, n° 3, mars 1987, pp. 301–303.
- [148] Wendler Thomas et Meyer-Ebrecht Dietrich. – Proposed standard for variable format picture processing and a codec approach to match diverse imaging devices. In : *Proceedings SPIE Picture archiving and communications systems (PACS) for medical applications, 1982*, pp. 298–305.
- [149] Wilson H. R. et Bergen. – A four mechanism model for threshold spatial vision. *Vision Res.*, 1979, pp. 19–32.
- [150] Wilson H. R. et Giese S. C. – Threshold visibility of frequency gradient pattern. *Vision Res.*, vol. 17, 1977, pp. 1177–1190.
- [151] Woodcock Curtis et Strahler Alan. – The factor of scale in remote sensing. *Remote Sensing of Environment*, vol. 21, 1987, pp. 311–332.
- [152] Worawattanamateekul J., Canisius X. J. F. et Samarakoon L. – Multisource data fusion results of fused optical and SAR data for irrigated rice areas identification.

- [153] Yesou Hervé, Besnus Yves et Rolet Joël. – Extraction of spectral information from landsat TM data and merge with SPOT panchromatic imagery. *ISPR Journal of Photogrammetry and Remote Sensing*, vol. 48, n° 5, 1993, pp. 23–26.
- [154] Yesou Hervé, Besnus Yves et Rolet Joël. – Merging seasat and spot imagery for the study of geological structures in a temperate agricultural region. *Remote Sensing of Environment*, vol. 43, n° 3, mars 1993, pp. 265–279.
- [155] Yesou Hervé, Besnus Yves et Rolet Joël. – Perception of a geological body using multiple source remotely-sensed data - relative influence of the spectral content and the data resolution. *International Journal of Remote Sensing*, vol. 15, n° 12, août 1994, pp. 2495–2510.
- [156] Yocky David A. – Image merging and data fusion by means of the discrete two-dimensional wavelet transform. *Journal of the Optical Society of America A*, vol. 12, n° 9, septembre 1995, pp. 1834–1841.
- [157] Yocky David A. – Multiresolution wavelet decomposition image merger of Landsat Thematic Mapper and SPOT Panchromatic data. *Photogrammetric Engineering and Remote Sensing*, vol. 62, n° 9, septembre 1996, pp. 1067–1074.
- [158] Zéboudj Rachid. – *Filtrage, seuillage, automatique, contraste et contour : Du pré-traitement à l'analyse d'image*. – Thèse, Université de Saint-Etienne, 1988.
- [159] Zhang Y. – A new merging method and its spectral and spatial effects. *International Journal of Remote Sensing*, vol. 20, n° 10, 1999, pp. 2003–2014.
- [160] Zhou Jiang, Civco Daniel et Silander John A. – A wavelet transform method to merge Landsat TM and SPOT panchromatic data. *International Journal of Remote Sensing*, vol. 19, n° 4, 1998, pp. 743–757.
- [161] Zhukov Boris, Berger Michael, Lanzl Frantz et Kaufmann Hermann. – A new techniques for merging multispectral and panchromatic images revealing sub-pixel spectral variation. In : *Proceedings MOMS-02 Symp., Köln, Allemagne, Paris, EARSeL, 1995*, pp. 163–168.
- [162] Zhukov Boris, Oertel Dieter, Berger Michael, Kaufmann Hermann et Müller Andreas. – Unmixing and fusion of multi-sensor multi-resolution data acquired over the Makhtesh Ramon Neveg desert. In : *Proceedings Fusion of Earth Data, Sophia Antipolis, 1998*, pp. 135–140.
- [163] Zhukov Boris, Oertel Dieter et Lanzl Frantz. – Unmixing and fusion of multi-resolution images with applications to TM/landsat and simulated ASTER data. In : *Proceedings Fusion of Earth Data, Sophia Antipolis, 1998*, pp. 91–96.
- [164] Zhukov Boris, Oertel Dieter, Lanzl Frantz et Reinhäkel Götz. – Unmixing-based multisensor multiresolution image fusion. *IEEE Transactions on Geoscience and Remote Sensing*, vol. 37, n° 3, mai 1999, pp. 1212–1226.

Le mémoire présente le développement d'une représentation multi-échelle d'images par une famille de pyramides morphologiques et ses applications à l'analyse et à la fusion d'images en télédétection. La représentation hiérarchique proposée est basée sur une approche pyramidale utilisant les filtres non linéaires de la morphologie mathématique.

La partie II propose d'abord un état de l'art des transformations pyramidales puis décrit le principe de la pyramide morphologique. Ses propriétés sont étudiées au travers de différents paramétrages et des familles qui en découlent. La pyramide morphologique permet d'une part, de séparer à chaque niveau de résolution les éléments de détails par leur taille et leur réflectance par rapport à l'environnement, et, d'autre part, de représenter les images aux niveaux de perception inférieurs.

La partie III du mémoire est consacrée aux applications des pyramides morphologiques à l'analyse des surfaces observées. La décomposition des éléments imbriqués dans les scènes en signaux séparables à différentes résolutions démontre la capacité de caractérisation multi-échelle. On montre également comment des traitements de reconstruction appliqués aux détails contribuent à cette analyse. Différents exemples de paysages analysés illustrent la méthodologie développée.

La partie IV consiste à fusionner des données de résolutions différentes, notamment celles issues de capteurs à haute résolution et à grand champ de vue. L'approche de fusion par pyramide morphologique crée ainsi des données de synthèse à haute résolution spatiale et à haute fréquence temporelle qui permettent une approche nouvelle du suivi des surfaces terrestres. Les résultats de ce principe de fusion sont présentés en fonction de différents jeux de dates d'acquisition des images.

La conclusion souligne trois perspectives très prometteuses. Premièrement, la pyramide morphologique peut être utilisée comme un navigateur exploitant les différents niveaux de résolution spatiale permettant l'accès à une information plus ou moins détaillée. Ensuite, la pyramide morphologique ouvre des opportunités intéressantes au recalage d'images de résolutions différentes. Finalement, la pyramide morphologique est un cadre intéressant pour la compression de données par le choix des différents paramètres autorisés.

Mots-clés : Algorithmes Pyramidaux, Morphologie Mathématique, Télédétection, Multi-échelle, Analyse d'images, Fusion spatio-temporelle

This report presents the development of a multi-scale image representation approach performed with a family of morphological pyramids and focuses on its applications to image analysis and fusion in remote sensing. This hierarchical representation is based on a pyramidal approach using morphological nonlinear filters.

Part II deals with a state-of-the-art of pyramidal transformation and then describes the principle of the morphological pyramid. Its properties are studied under several parameterisations and under the associated families. The morphological pyramid enables, on the one hand, to split at each resolution level detail elements according to their sizes and their reflectances with regards to the neighbourhood, and, on the other hand, to represent images at lower perception levels.

Part III is devoted to the applications of the morphological pyramids to surfaces analysis. The decomposition of nested objects of scenes into separable elements clearly shows the multi-scale characterisation ability. Further processings over those elements contribute towards this analysis. Examples of several landscapes illustrate the methodology.

Part IV consists in merging multi-resolution data, more precisely those acquired by high resolution sensors and wide field of view sensors. The morphological pyramid approach for fusion creates synthesis data having both high spatial resolution and high temporal frequency, which enables a new approach for terrestrial surfaces monitoring. Fusion principle results are presented according to several images data sets.

The conclusion underlines three promising outlooks. Firstly, the morphological pyramid can be used as a navigator exploiting all the spatial resolution levels and gives access to more or less accurate information. Then, the morphological pyramid offers interesting applications to multi-resolution images registration. Finally, with the choice of parameters, the morphological pyramid is a promising frame for data compression.

Keywords : Pyramid Algorithms, Mathematical Morphology, Remote Sensing, Multiscale, Image analysis, Spatio-temporal fusion