



HAL
open science

Fusion de données : prise en compte des caractéristiques liées à l'imageur lors de la synthèse d'images multispectrales à haute résolution spatiale

Pierre Massip

► **To cite this version:**

Pierre Massip. Fusion de données : prise en compte des caractéristiques liées à l'imageur lors de la synthèse d'images multispectrales à haute résolution spatiale. Traitement du signal et de l'image [eess.SP]. École Nationale Supérieure des Mines de Paris, 2010. Français. NNT : 2010ENMP0038 . pastel-00564814

HAL Id: pastel-00564814

<https://pastel.hal.science/pastel-00564814>

Submitted on 10 Feb 2011

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

École doctorale n°84 : STIC

Doctorat ParisTech

T H È S E

pour obtenir le grade de docteur délivré par

l'École nationale supérieure des mines de Paris

Spécialité “ Informatique temps réel, robotique et automatique ”

présentée et soutenue publiquement par

Pierre MASSIP

le 19 Novembre 2010

**Fusion de données : prise en compte des caractéristiques liées à
l'imageur lors de la synthèse d'images multispectrales à haute
résolution spatiale**

Directeur de thèse : **Lucien WALD**

Maître de thèse : **Philippe BLANC**

Jury

M. Stéphane JACQUEMOUD, Professeur, IPGP, Université Paris Diderot

M. Jocelyn CHANUSSOT, Professeur, Gipsa-lab, Grenoble-INP

M. Philippe BOLON, Professeur, LISTIC, Polytech Annecy-Chambéry

M. Matthieu CORD, Professeur, LIP6lab, Université Pierre et Marie CURIE

Mme. Audrey ROMAN - MINGHELLI, Docteur, LSEET, Université du Sud Toulon-Var

M. Lucien WALD, Professeur, CEP, MINES ParisTech

M. Philippe BLANC, Docteur, CEP, MINES ParisTech

Président

Rapporteur

Rapporteur

Examineur

Examineur

Examineur

Examineur

**T
H
È
S
E**

MINES ParisTech

Centre Energétique et Procédés

1, rue Claude Daunesse, F-06904 Sophia Antipolis Cedex

Remerciements

Cette thèse s'est déroulée au sein de l'équipe OMD (Observation, Modélisation, Décision) du Centre Énergétique et Procédés - MINES ParisTech.

Je tiens à te remercier tout particulièrement Philippe. Tu as su faire du stagiaire que j'étais un vrai thésard. Avec Lucien, vous vous êtes attelés à la lourde tâche de transformer la chenille en papillon... Je vous remercie tout les deux pour votre gentillesse, votre générosité et votre sens de l'humour, mais surtout pour le savoir et savoir faire que vous m'avez transmis.

Ces trois années de thèse se sont déroulées à merveille dans la famille OMD (merci à : Lionel, Isabelle Thierry, Roseline). Le cadre de vie et de travail fut très agréable tant et si bien qu'il devient difficile de le quitter... (Existe-t-il des patches ??)

J'ai beaucoup apprécié la vie au CEP et suis ravi d'avoir rencontré : Armel, Nils, Didier, Alexandre, Alexis, Raphaël, Bassam, Christophe, Luis, Pierryves, Bella, Damien, Benoit, Arthur, Franck, Mathilde, Johann, Demi-périmètre (Pierre I)...et d'avoir participé à tous ces repas de tribus animés de discussions, de blagues (de tous niveaux) et de débats...

Mais aussi : Florent G

Lolo, c'est grâce à toi que je suis venu dans cette belle région et tu as su me supporter pendant ces trois ans...Tes encouragements ton soutien surtout, et t'avoir à mes coté fut une grande douceur... D'ailleurs grâce à toi mon fils a pu assister à ma soutenance bien au chaud dans ton petit ventre...

Maman, Papa, merci pour votre soutien dans cette aventure, Doud, Clairon, Lisou, Sonia merci pour vos encouragements.

Je remercie l'ensemble du jury de son analyse critique de mes travaux ainsi que de sa lecture chirurgicale. J'ai apprécié leurs retours et les échanges que nous avons eus ainsi que leur regard nouveau sur mes travaux.

Sommaire

1	INTRODUCTION.....	7
1.1	CONTEXTE.....	7
1.1.1	La fusion d'images.....	9
1.1.2	Quelques définitions : résolution spatiale, résolution spectrale, géométrie.....	10
1.2	OBJECTIF DE LA THESE.....	12
1.3	DEMARCHE.....	14
2	FTM, REPLIEMENT DE SPECTRE, SNR.....	17
2.1	FONCTION DE TRANSFERT DE MODULATION (FTM).....	17
2.1.1	Que représente la FTM ?.....	18
2.1.2	PSF, FTM et résolution spatiale ?.....	19
2.1.3	Quel effet a la FTM sur les images ?.....	20
2.1.4	Modélisation de la FTM.....	22
2.1.5	FTM du satellite PLÉIADES.....	24
2.1.6	Modification de la FTM d'une image.....	24
2.2	REPLIEMENT DE SPECTRE.....	27
2.2.1	Le processus d'échantillonnage.....	27
2.2.2	Phénomènes et artéfacts liés à l'échantillonnage.....	32
2.2.3	Le repliement de spectre dans les images MS et la fusion.....	33
2.3	RAPPORT SIGNAL A BRUIT.....	37
3	JEUX D'IMAGES, CONCEPT DE FUSION, EVALUATION DE QUALITE,	
OUTILS.....		39
3.1	JEUX D'IMAGES UTILISEES.....	39
3.1.1	Simulation PLÉIADES.....	40
3.1.2	Simulation SPOT-5.....	42
3.2	NOTATIONS.....	44
3.3	LE CONCEPT ARSIS.....	46
3.4	LA METHODE DE FUSION DE REFERENCE : ATWT-M3.....	48
3.4.1	ATWT.....	48
3.4.2	IMM et HRIMM, Modèle : M3.....	49
3.5	FILTRES DE DECOMPOSITION ET RECONSTRUCTION DES ONDELETTES A TROUS.....	50
3.6	OUTILS.....	52
3.6.1	Protocole d'évaluation de la qualité.....	52
3.6.2	Outils de mesure du repliement de spectre et de la netteté.....	58
3.6.3	Distance MTF_{dev}	61
4	ADAPTATION DE FTM DANS LA FUSION.....	65
4.1	INFLUENCE DE LA FTM.....	65
4.1.1	Objectif.....	65
4.1.2	La transformation de FTM dans le processus de fusion.....	66
4.1.3	Adaptation de FTM sans déconvolution optique.....	71
4.1.4	Effet de l'adaptation de FTM selon le rapport de résolution : 2 ou 4.....	72
4.2	ATWT-M3-MTFADAPTED.....	73
4.2.1	Description de la méthode.....	73

4.2.2	Evaluation des performances de la méthode	74
4.2.3	Les résultats	74
4.3	CONCLUSION.....	88
5	COMPENSATION DU RAPPORT DE FTM (MTF_{DEV})	91
5.1	ATWT-M3-MTFEQUALIZED	91
5.1.1	Principe	91
5.1.2	Description de la méthode.....	92
5.1.3	Les résultats	93
5.2	CONCLUSION.....	101
6	INFLUENCE DU REPLIEMENT DE SPECTRE ET DU SNR	103
6.1	REPLIEMENT DE SPECTRE.....	103
6.1.1	Images de simulation et repliement de spectre.....	103
6.1.2	Suppression du repliement de spectre et fusion	104
6.1.3	Amélioration de la qualité par suppression de repliement de spectre	105
6.1.4	Variation de la FTM en fonction de la suppression de repliement	108
6.1.5	Ré-échantillonnage et artefact de repliement de spectre.....	109
6.1.6	Influence de la méthode d'interpolation sur la FTM de l'image synthétisée	110
6.2	SNR ET NOUVELLES METHODES DE FUSION	113
7	CONCLUSION ET PERSPECTIVES	115
	ANNEXE A	121
	ANNEXE B :	125
	BIBLIOGRAPHIE	131

1 Introduction

1.1 Contexte

De nombreux satellites observent la Terre afin de mieux connaître notre environnement. Leurs missions sont multiples allant de la météorologie à la cartographie. Un satellite porte généralement plusieurs capteurs, dont certains sont des imageurs. Nous nous intéresserons à ces derniers, et particulièrement à ceux dont les bandes spectrales sont situées dans le visible et le proche infrarouge.

Les images fournies par la majorité des satellites commerciaux à haute résolution spatiale (SPOT-5, Ikonos, Quickbird) peuvent être classées en deux groupes : d'une part, les images à haute résolution spatiale et à basse résolution spectrale, et d'autre part les images à basse résolution spatiale, mais à haute résolution spectrale.

Les images ayant une haute résolution spatiale sont issues de l'acquisition d'une large bande en longueur d'onde qui correspond le plus souvent à une grande partie du spectre visible et pouvant inclure une partie de la plage infrarouge. Ces images sont appelées images panchromatiques (PAN).

Les images multimodales (MS) observent chacune une partie du spectre visible correspondant le plus souvent aux longueurs d'onde du bleu, du vert, du rouge et du proche infrarouge (PIR). Ces images ont une meilleure résolution spectrale, mais celle-ci est obtenue au détriment de la résolution spatiale comme nous le verrons plus tard.

La figure 1.1 donne des exemples d'images, panchromatique à haute résolution spatiale (a), et multimodales à basse résolution spatiale composées le plus souvent de quatre canaux (b). La chose la plus frappante dans la figure 1.1 est certainement la petite taille des images MS par rapport à l'image PAN : il y a 16 fois plus de pixels dans l'image PAN. Les détails géométriques y apparaissent nettement. Par contre, il n'est pas toujours facile de distinguer des objets de texture proche alors que les images MS peuvent le permettre. Par exemple, des courts de tennis peuvent avoir une taille similaire à celle d'une piscine, et les deux types d'objet ont des textures assez uniformes : ils peuvent ne pas être discriminés dans l'image PAN. Cependant, la dominante rouge (ou verte) des terrains de tennis et la dominante bleue des piscines permettent la discrimination dans les images MS

Le tableau 1.1 recense les caractéristiques de ces images pour plusieurs satellites. On observe qu'il y a souvent un rapport 4 entre les résolutions spatiales de PAN et MS.



Figure 1.1 : Illustration des images intervenant dans la fusion. (a) Image PAN, résolution 0.7 m. (b) Images MS résolution 2.8 m. Copyright CNES 2000.

Tableau 1.1 : Exemple de caractéristiques des images des satellites actuels

Caractéristiques des images satellites						
Satellite	Modalités XS				Modalité PAN	
	Nom de la bande	Couleur	Bande passante spectrale (μm)	Résolution au sol (m)	Bande passante spectrale (μm)	Résolution au sol (m)
Ikonos	B1	bleu	0,45 - 0,53	4	0,45 - 0,90	1
	B2	vert	0,53 - 0,61	4		
	B3	rouge	0,64 - 0,72	4		
	B4	PIR	0,77 - 0,88	4		
Quickbird	B1	bleu	0,45 - 0,52	2,8	0,45-0,9	0,7
	B2	vert	0,52 - 0,6	2,8		
	B3	rouge	0,63 - 0,69	2,8		
	B4	PIR	0,76 - 0,9	2,8		
SPOT-5	B1	vert	0,50 - 0,59	10	0,48 - 0,71	2,5 ou 5
	B2	jaune	0,61 - 0,68	10		
	B3	PIR	0,78 - 0,89	10		
	B4	MIR	1,58 - 1,75	10		
Pléiades (simulation)	B0	bleu	0,43 - 0,55	2,8	0,48 - 0,82	0,7
	B1	vert	0,49 - 0,61	2,8		
	B2	rouge	0,60 - 0,72	2,8		
	B3	PIR	0,75 - 0,95	2,8		

Les deux propriétés de résolution spatiale et spectrale sont techniquement liées et difficiles à concilier. Plus la bande passante est large, plus le flux lumineux qui arrive sur le capteur est grand. Dans le cas du multispectral, la largeur de la bande spectrale de chaque modalité est bien plus faible que dans le cas du PAN. Le capteur reçoit donc moins d'énergie par unité de temps. Comme les capteurs des bandes PAN et MS sont sur le même satellite et utilisent la plupart du temps le même télescope, la vitesse de défilement est la même pour toutes les bandes. Un temps d'intégration identique entrainerait un rapport signal à bruit et une dynamique trop différents entre les images PAN et MS. Pour pallier cette situation, il faut augmenter le temps d'intégration des canaux MS. Comme nous venons de le dire, la vitesse de défilement est identique pour toutes les bandes. Il est donc impossible de jouer sur

la vitesse pour modifier le temps d'intégration. La seule solution pour allonger le temps d'intégration est alors d'augmenter la taille des photosites. L'augmentation de surface des photosites et de l'espace qui les sépare entraîne alors une augmentation de la taille des pixels et une baisse de la résolution spatiale. Par ailleurs, l'observation à haute résolution des bandes MS et PAN entraînerait un stockage et une transmission de quatre fois plus de données. La compression à bord et la transmission d'un flux de données quatre fois plus important peuvent notamment se révéler être une contrainte forte sur le système. Ainsi, la résolution spectrale et spatiale des images délivrées par les satellites actuels est le résultat de compromis lié à des raisons techniques et physiques.

Les informations recherchées dans les images satellites peuvent être différentes selon les domaines pour lesquels elles sont utilisées. Dans certains cas, l'image PAN à haute résolution spatiale peut être suffisante comme par exemple la détection des axes routiers (Péteri et Ranchin 2004). Au contraire, dans le cas de la détection de parcelles l'information spectrale est primordiale (Girard et Girard 1999) Les informations spectrales sont contenues dans l'image multimodale alors que l'image panchromatique contient des détails géométriques. Les informations contenues dans ces deux types d'images (PAN et MS) sont complémentaires.

1.1.1 La fusion d'images

Disposer de l'ensemble des données avec la résolution spatiale de PAN et spectrale de MS offre des gains notables dans l'analyse des images, leur exploitation et la cartographie. Parmi de nombreux auteurs Ranchin et Thomas, citent des besoins très variés tels que la cartographie, l'étude du développement des zones urbaines, la détection de cibles, la prévention des risques ou la photo-interprétation, la détection de changement ou la prévention des risques (Ranchin *et al.* 2003; Thomas 2006). Les algorithmes de fusion d'images permettent d'allier les caractéristiques/avantages des deux types d'images.

Wald définit la fusion comme le cadre formel dans lequel s'expriment les moyens et techniques permettant l'alliance de données provenant de sources diverses (Wald 1999). Une définition plus restrictive du cas de fusion qui nous intéresse, la fusion d'image, est proposée par (Piella 2003) comme le procédé permettant la combinaison de plusieurs images en un nombre plus petit d'images, le plus souvent unique et contenant les informations « importantes » des images d'entrées.

Nous nous intéressons aux algorithmes visant à augmenter la résolution spatiale des images MS à l'aide de la haute résolution spatiale de l'image PAN. Ces algorithmes sont souvent regroupés sous le terme de *pansharpening*. Les objectifs des méthodes de *pansharpening* sont assez divers (Pohl et Genderen 1998). Wald (Wald 1999) souligne que la qualité du résultat de fusion est relative à l'application et à l'utilisateur. Il convient donc de préciser l'objectif de l'algorithme de fusion.

Le cadre de notre thèse est donc la synthèse d'images MS à haute résolution spatiale et respectant la propriété de synthèse définie par Wald (Wald *et al.* 1997) et Thomas (Thomas 2006). Cette propriété stipule que : « **l'image synthétisée doit ressembler à celle**

qu'aurait acquis le capteur MS correspondant si celui-ci possédait une meilleure résolution spatiale ».

La fusion d'image constitue un outil très appréciable pour les photo-interprètes si les algorithmes de fusion sont suffisamment performants/fiables afin que ces derniers n'aient pas à se référer aux images PAN ou MS pour confirmer une observation (Thomas 2006). La fusion peut dans le cas de la détection de parcelles, apporter une meilleure précision de mesure grâce à une meilleure résolution spatiale de l'image sans augmenter les erreurs de détection dans la mesure où les algorithmes de synthèse présentent une distorsion spectrale faible. De manière générale, il y a un réel avantage à exploiter des images à la haute résolution spatiale et haute résolution spectrale pour les algorithmes de détection, de segmentation ou encore de comptage. Par exemple, la fusion peut par l'apport d'information spectrale améliorer la discrimination entre les dégâts et les *changements naturels*, lors de catastrophes naturelles (Chesnel 2008).

Enfin, l'accès aux images issues de l'imagerie spatiale ou aéroportée a été facilité pour le public ces dernières années en particulier grâce à Google Earth (et à Géoportail). Le nombre d'utilisateurs et d'application liés à ces données a d'ailleurs beaucoup augmenté récemment. Une grande proportion des images fournies par Google Earth est issue d'un processus de pansharpening. De même, les fournisseurs d'images satellites proposent aussi de plus en plus d'images issues de tel processus (Gamba et Chanussot 2008). Cette augmentation de l'offre d'images synthétisées ainsi que l'accroissement des résolutions spatiales des nouveaux satellites (Quickbird) ou à venir (Pléiades), justifient une remise en question des algorithmes les plus performants pour les faire évoluer.

L'objectif global de la thèse est de contribuer à l'amélioration des méthodes de fusion entre les images PAN et MS qui respectent la propriété de synthèse. Les méthodes de fusion sont considérées actuellement comme plutôt performantes (Alparone *et al.* 2007; Ranchin *et al.* 2003; Thomas *et al.* 2008) et MINES-ParisTech a beaucoup œuvré dans ce sens. Nous ne proposerons pas de révolution technologique, mais une amélioration de certains des derniers défauts. De façon plus précise, nous chercherons plutôt à améliorer la représentation de la géométrie des images synthétisées en perfectionnant certaines étapes du processus de fusion. Pour bien comprendre les objectifs de notre thèse, nous commençons par définir certains termes.

1.1.2 Quelques définitions : résolution spatiale, résolution spectrale, géométrie

La résolution spatiale correspond à la capacité d'un système d'observation à identifier ou reconnaître un élément de l'image grâce à ses caractéristiques géométriques. D'après (Girard et Girard 1999), la résolution spatiale est définie comme la taille du plus petit élément qui peut être discerné dans l'image. La figure 1.2 montre bien que certains éléments ne peuvent pas être distingués à basse résolution. Il est difficile d'identifier les véhicules dans l'image à basse résolution (280 cm), alors que cette tâche ne présente aucune difficulté dans l'image à

haute résolution (70 cm). La résolution spatiale est donc une qualité essentielle pour identifier ou localiser certains éléments dans une image (Dubois *et al.* 1997).



Figure 1.2 : Illustration de la haute résolution spatiale versus basse résolution. (a) Image PAN à résolution 70 cm. (b) même image à résolution 280 cm mais ré-échantillonnée à résolution 280 cm). Copyright CNES 2000.

La résolution spectrale est l'aptitude du capteur à distinguer des rayonnements de longueurs d'onde différentes et permet ainsi de différencier les éléments d'un paysage grâce à leur signature spectrale (Girard et Girard 1999). Par exemple, la végétation a une forte signature dans le proche infrarouge alors que la mer a une signature prépondérante dans le bleu. Cette différence de signature est un moyen pour distinguer les objets. Nous avons vu précédemment dans les exemples des courts de tennis et des piscines.

La figure 1.3 présente le cas où la transition entre deux parcelles de champ est observée dans le proche infrarouge (figure 1.3b) alors qu'elle ne l'est pas dans le canal panchromatique (figure 1.3a). Ainsi dans certains cas, une bonne résolution spectrale peut être plus importante que la résolution spatiale.



Figure 1.3 : SPOT-5, Toulouse, France. Illustration d'une occultation d'objet dans la modalité PAN. (a) PAN, 2.5 m sous-échantillonnée à 10 m, (b) image dans le proche infrarouge, 10 m. Copyright Digital Globe 2002. (*In* Thomas 2008)

Dans le domaine de la qualité image des systèmes d'observation satellitaire, la notion de géométrie englobe plusieurs aspects comme la géolocalisation, l'altération des longueurs, la cohérence spatiale ou la fonction de transfert de modulation (CNES/ONERA/IGN 2008). La géométrie est aussi définie comme la capacité à délimiter les surfaces (Laporterie-Déjean *et al.* 2005). La bonne représentation de la géométrie passe par le respect de la linéarité des

structures, des contours de bâtiments nets et l'absence d'agrégation d'objets (Thomas 2006). La diversité des définitions et des publications scientifiques montre que la géométrie est une notion complexe, difficile à appréhender et à décrire.

Dans cette thèse, nous nous attachons particulièrement à la restitution des contours et des contrastes permettant une bonne distinction des structures fines – de petites échelles – dans l'image. Les aspects de géolocalisation ou d'altération des longueurs ne sont pas considérés. Nous suivrons donc Le Pennec qui définit la géométrie comme l'organisation de la radiométrie dans l'image (Pennec 2002). Nous proposons donc une expression de la propriété de synthèse en ce qui concerne la représentation de la géométrie qui pourrait être : **« la restitution de l'organisation radiométrique du paysage observé par l'image synthétisée doit ressembler à celle qui serait faite par le capteur MS correspondant si celui-ci possédait une meilleure résolution spatiale »**.

La restitution de la géométrie dans l'image dépend de caractéristiques du système imageur. En effet, la restitution fidèle d'un contour ou encore le pouvoir de distinction d'objets de petite taille d'un paysage dépend en grande partie de l'échantillonnage et de la réponse impulsionnelle du système imageur, encore appelée tache image. Le module de la transformée de Fourier de cette tache image, appelée fonction de transfert de modulation, est un moyen de caractériser la capacité du système à bien restituer la géométrie en fonction des échelles spatiales considérées.

1.2 Objectif de la thèse

Le récent concours IEEE sur la fusion d'images (Alparone *et al.* 2007) a montré que des méthodes performantes étaient maintenant disponibles. La synthèse spectrale des images fusionnées est souvent de bonne qualité ainsi que la géométrie. Cependant, nous observons encore quelques artefacts comme des halos ou des bavés de couleurs qui empêchent notamment une analyse visuelle ou automatique de manière fiable, robuste et juste (Thomas *et al.* 2008).

Mes travaux de thèse visent à faire décroître l'intensité et la fréquence d'apparition de tels artefacts, voire de les faire totalement disparaître.

Plusieurs possibilités s'offraient à nous. L'une d'entre elles consistait à développer une nouvelle méthode, combinant, par exemple, les divers éléments des méthodes primées dans le concours IEEE et reposant sur des concepts différents. Le fait de combiner deux méthodes peut dans un premier temps, combler les défauts des deux méthodes, mais il n'est pas exclu que certains défauts ne soient pas eux aussi rehaussés. Le grand nombre de publications récentes sur des méthodes hybrides n'a pas apporté d'améliorations notables. De plus, nous avons pris conscience de l'intérêt de l'accumulation d'expérience : les défauts de chacun des concepts et leurs causes sont de mieux en mieux connus (Thomas *et al.* 2008) et (Alparone *et al.* 2007). Aussi, il nous a semblé plus judicieux de se baser principalement sur un seul concept, bien connu et éprouvé, et de chercher comment perfectionner les méthodes les plus performantes dans ce concept afin d'atteindre une meilleure qualité géométrique. Nous

pensons que la prise en compte de certains paramètres des images, notamment ceux liés à l'imageur dont elles proviennent, constitue une piste de recherche suffisamment riche pour améliorer les performances des méthodes de fusion.

Wald (Wald 2002) a identifié trois concepts de base dans les méthodes de fusion :

- ARSIS (Amélioration de la Résolution par Injection de Structures) ;
- Projection substitution ;
- Contribution spectrale relative.

Après analyse de ces trois concepts, abondamment utilisés par de nombreuses méthodes publiées, il nous a semblé que le concept ARSIS (Ranchin et Wald 2000) était le plus approprié à notre but. Il est bâti sur une analyse multirésolution, propice à la modélisation de la géométrie. De plus, c'est un cadre très formalisé où les différentes étapes sont bien décrites de même que leurs relations, tout en restant ouvert à des évolutions : aucun outil mathématique n'est fixé. Enfin, il est à noter que les méthodes les plus performantes relèvent de ce concept (Alparone *et al.* 2007).

L'objectif de la thèse est donc de contribuer à l'accroissement de la qualité des images MS synthétisées à haute résolution spatiale par la prise en compte de caractéristiques des capteurs, comme la fonction de transfert de modulation (FTM) et l'échantillonnage

1.3 Démarche

Notre travail de thèse a tout d'abord consisté en l'étude des éléments intervenants (analyse multi-résolution, méthodes d'interpolations, ...) lors de la fusion d'image, notamment ceux sur lesquels nous estimons pouvoir faire des avancées. Nous avons commencé par l'étude des principaux concepts de *pansharpening*. Wald (2002) propose trois groupes : projection-substitution, amélioration de la résolution spatiale par injection de structures (ARSIS) et contribution spectrale relative. Nous avons examiné les propriétés de chacun des concepts. Le principe du concept ARSIS est basé sur une analyse multiéchelle qui permet d'injecter les détails manquants à partir des hautes fréquences de l'image PAN (Mangolini 1994, Ranchin *et al.* 2003, Thomas 2006). Cette analyse multiéchelle fait que ce concept est celui qui nous a semblé offrir le plus d'opportunités pour améliorer la représentation de la géométrie. De plus, il s'avère que les méthodes les plus avancées utilisant ce concept font partie des méthodes les plus performantes (Alparone *et al.* 2007). Ce dernier point consolide notre choix pour développer nos travaux.

Nous cherchons à créer une image qui proviendrait d'un capteur multimodal possédant la haute résolution spatiale. Pour s'approcher au mieux de ce que seraient ces images, il est nécessaire d'avoir une bonne connaissance du processus d'acquisition de ces dernières. Il sera alors possible de tenir compte des caractéristiques des images disponibles (basse résolution pour les images MS) et de donner aux images synthétisées les caractéristiques qu'elles auraient si elles provenaient d'un capteur possédant la haute résolution spatiale.

Très exactement, il nous a semblé que l'amélioration attendue pourrait être atteinte par la prise en compte des caractéristiques des capteurs susceptibles de changer entre la basse et la haute résolution spatiale. Par exemple, le gabarit spectral restera inchangé et la FTM changera. Parmi ces caractéristiques, nous en avons identifié deux dont la prise en compte pourrait mener à une amélioration notable de la qualité géométrique de l'image synthétisée : la FTM et l'échantillonnage.

La réponse du système d'observation à une impulsion lumineuse parfaite, est appelée réponse impulsionnelle du système dans le cas de filtres en traitement du signal ou tache image ou Point Spread Function (PSF) dans notre cas. La PSF est intimement liée aux performances de qualité image du système d'observation. La FTM est le module de la transformée de Fourier de la PSF, elle n'est pas la même pour le capteur PAN et les capteurs MS.

Les méthodes de type ARSIS exploitent des outils d'analyse multi-résolution (AMR). Si la bande passante de l'AMR est très dissemblable de la FTM, alors l'image synthétisée présentera un défaut de contraste (manque d'information/énergie) pour les fréquences qui correspondent aux hautes fréquences de l'image MS originale. L'approche de (Aiazzi *et al.* 2006) vise à limiter ce défaut lors de la synthèse, grâce à un filtre proche de la FTM. Ces auteurs proposent une analyse multi-résolution (AMR) dont les bancs de filtres dépendent de la modalité et dont la bande passante est proche de celle de la FTM. Cela est rendu possible grâce à la connaissance de la valeur de la FTM à la fréquence de Nyquist ($f_e/2$) et en

supposant que la FTM puisse être approché par une gaussienne. Cette approche est prometteuse, car on trouve de plus en plus d'information sur la FTM à la fréquence de coupure (Ryan *et al.* 2003) (Gaudin-Delerieu *et al.* 2008)). En l'absence d'information sur la FTM de l'instrument, il est possible d'estimer la FTM.

Nous avons choisi d'explorer une autre voie en exploitant la différence de FTM entre les images PAN et MS. La prise en compte de la différence de PSF/FTM entre les images a pour but de donner à l'image synthétisée à haute résolution la PSF qu'elle aurait si elle provenait d'un capteur possédant la haute résolution spatiale. Pour parvenir à ce résultat, nous avons identifié et exploré deux voies d'amélioration. La première modifie la FTM des images MS avant l'AMR, la deuxième effectue une correction a posteriori de l'image synthétisée.

Ces deux approches sont pratiquement implantées dans une méthode déjà reconnue comme performante (Thomas 2006) en améliorant certaines étapes de la méthode. La prise en compte de la différence de FTM entre les deux images va permettre une plus grande cohérence dans la représentation des différentes fréquences présentes dans l'image synthétisée. Cette contribution sera évaluée en comparant les différents résultats obtenus au moyen du protocole d'évaluation de qualité proposé par Thomas (Thomas 2006).

L'évaluation de la qualité aura pour objectif de définir si la meilleure prise en compte de la FTM dans le processus de fusion permet d'améliorer l'ensemble des indicateurs de qualité, si la qualité visuelle est aussi améliorée et si les artefacts sont moins présents.

L'échantillonnage est une étape qui permet de faire passer le signal analogique en signal numérique. La résolution des images est liée, de manière inversement proportionnelle à la fréquence d'échantillonnage. La résolution des images MS est donc faible comparativement aux images PAN. Dans le cas d'images satellites, nous observons la Terre qui présente beaucoup de hautes fréquences, qui sont mieux échantillonnées avec la modalité PAN qu'avec la modalité MS.

Pour les images MS, la fréquence d'échantillonnage est trop faible et une perte d'information peut se produire. Cette dernière est due au repliement de spectre induit par un échantillonnage ne respectant pas les conditions du théorème de l'échantillonnage de Shannon. Le fait d'échantillonner un signal périodise son spectre de Fourier suivant une période égale à la fréquence d'échantillonnage. Si cette dernière n'est pas suffisamment élevée alors, les spectres se superposeront au niveau des extrémités. Cette superposition est à l'origine de distorsions et de pertes d'information, qui se produisent essentiellement dans les hautes fréquences c'est-à-dire autour de la fréquence de Nyquist (la demi-fréquence d'échantillonnage) pour l'échantillonnage considéré. On peut trouver dans les images MS des zones où l'information n'est pas correcte du fait que les fréquences de certaines zones sont trop élevées. Or ces informations sont dans une certaine mesure contenues dans l'image PAN à haute résolution.

Par conséquent, nous pouvons imaginer exploiter ces informations de l'image PAN pour corriger le repliement de spectre dans les modalités MS. Cette correction, a priori ou a

posteriori, devrait apporter un gain dans la synthèse des hautes fréquences. C'est pourquoi nous avons étudié ce repliement en tentant de répondre aux questions suivantes :

- La suppression du repliement de spectre dans les images MS à basse résolution permet-elle une amélioration de la qualité de l'image synthétisée ?
- Peut-on mesurer la quantité de repliement de spectre dans une image ?

2 FTM, repliement de spectre, SNR

Dans ce chapitre, nous introduisons les caractéristiques des images qui sont importantes dans le domaine de l'imagerie spatiale : la FTM, le repliement de spectre (qui est lié à l'échantillonnage des images), et le bruit radiométrique.

La FTM caractérise la capacité de l'instrument à représenter le contraste en fonction des fréquences spatiales. Cette caractéristique de l'instrument est différente entre l'acquisition des images PAN et celle des images MS. La FTM a l'allure d'un filtre passe-bas (sauf cas particulier de certains instruments optiques interférométriques), avec des hautes fréquences représentées avec un contraste plus faible que les basses fréquences. La FTM globale du système imageur est liée à plusieurs sous-systèmes : l'optique (diffraction, qualité de réalisation et d'alignement), la partie détection numérique (zone photosensible, temps d'intégration, compensation des effets de filé), ou encore la plate-forme satellite (attitudes et stabilité).

Le repliement de spectre est un phénomène lié à l'échantillonnage de l'image. Il crée notamment des artefacts visuels dans l'image. L'image PAN a une résolution quatre fois supérieure à l'image MS, son échantillonnage est donc plus fin. L'image PAN est donc moins sujette au repliement de spectre que ne l'est l'image MS.

Enfin, le bruit radiométrique lié au bruit photonique et aux bruits électroniques de la chaîne de détection, est caractérisé par le rapport signal sur bruit.

2.1 Fonction de Transfert de Modulation (FTM)

La représentation spatiale de l'information est liée à la capacité du capteur à restituer les informations qui arrivent au capteur à travers l'instrument optique. Décrire et évaluer cette capacité est une façon de quantifier les performances de l'instrument. Celles-ci ne sont pas quantifiées à l'aide d'un seul chiffre/indicateur/paramètre, mais plutôt par une multitude. (Luxen et Förstner 2002), par exemple, donnent une liste de caractéristiques, par forcément indépendantes les unes des autres, qui quantifient les performances de l'instrument : le contraste, la luminance, la variance du bruit, la netteté, la résolution radiométrique, la granularité, la PSF (point spread function), la FTM (fonction de transfert de modulation), le pouvoir de résolution.

Cependant, de nombreux auteurs/publications s'accordent sur l'importance de la FTM pour la qualité image en ce qui concerne la géométrie dans le sens défini précédemment (Leger *et al.* 1994) et (Grois *et al.* 2006). Une représentation précise de la FTM lors du

processus de fusion devrait donc contribuer à une restitution plus précise de l'organisation radiométrique du paysage dans les images MS synthétisées.

2.1.1 Que représente la FTM ?

La FTM peut être définie de plusieurs façons :

- la capacité du système à représenter le contraste à une fréquence (spatiale) donnée (Patra *et al.* 2002) (Estribeau et Magnan 2004).
- une fonction qui exprime le contraste avec lequel pourra être représentée une sinusoïde spatiale de fréquence donnée (Patra *et al.* 2002).
- le rapport entre le contraste de la scène (le contraste obtenu avec un instrument idéal/parfait, mais qui n'existe pas) et le contraste obtenu/présent dans l'image acquise par l'instrument. On peut aussi voir la FTM comme le rapport du contraste d'une scène avant et après l'instrument.
- la réponse en fréquence (spatiale) du système d'observation. La FTM correspond au contraste de perception du système vis-à-vis des composantes fréquentielles du paysage observé.

Toutes ces définitions sont valides, bien sûr. Elles diffèrent par la perception qu'elles donnent de la FTM.

La FTM est un rapport de contraste, ses valeurs possibles seront comprises dans l'intervalle $[0, 1]$. La FTM décroît avec la fréquence, avec une allure assez proche de celle d'un sinus cardinal (CNES/ONERA/IGN 2008).

Bien que la FTM soit fonction de la fréquence, il est fréquent de la voir définie/quantifiée à l'aide de ses seules valeurs pour les fréquences de Nyquist en lignes et en colonnes. C'est d'ailleurs à travers une valeur minimum à la fréquence de Nyquist que la FTM apparaît le plus souvent dans les cahiers des charges des satellites d'observation. La partie de la FTM qui nous intéresse est contenue dans la bande de fréquence comprise entre 0 et la fréquence de Nyquist. Dans les figures représentant la FTM, la fréquence sur l'axe des abscisses sera normalisée par rapport à la fréquence d'échantillonnage de l'image PAN. La figure 2.1 suivante donne un exemple d'une FTM typique de la voie PAN d'un satellite d'observation.

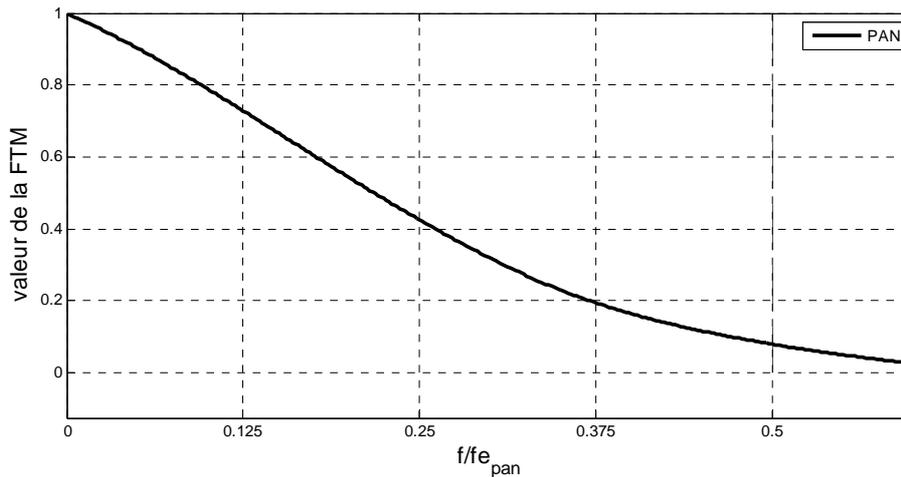


Figure 2.1 : FTM typique de la voie PAN d'un satellite d'observation.

La PSF/FTM dépend de différentes caractéristiques et dimensions des éléments optiques et numériques de l'instrument d'observation. La partie optique d'un satellite d'observation est constituée principalement d'un miroir primaire et de nombreux miroirs en entrée du télescope secondaire. Ainsi le dimensionnement, la qualité de réalisation et la stabilité en vol des formes et alignements de tous les éléments optiques influencent les performances en FTM de l'instrument, mais c'est le dimensionnement du miroir primaire qui a le rôle le plus critique de la partie optique (Lamard *et al.* 2004).

La partie numérique de l'instrument a, elle aussi, un rôle non négligeable dans l'allure de la FTM globale de l'instrument. Un détecteur idéal donnerait un contraste identique, quelle que soit la fréquence, tant que celle-ci est inférieure à la fréquence de Nyquist. Cependant, la technologie et la taille de surface d'intégration influencent l'allure de la FTM de l'instrument (Shcherback et Yadid-Pecht 2001) (Grais *et al.* 2006) (Yadid-Pecht 2000) (Estribeau et Magnan 2004). L'image acquise par l'instrument est affectée de la FTM globale de l'instrument qui est le résultat notamment de la combinaison de la FTM liée à la partie optique et à la partie numérique. À cette FTM de niveau instrument viennent s'ajouter des termes de dégradation de FTM de niveau plate-forme (stabilité et précision du pointage durant la prise de vue) ainsi que la FTM liée aux effets de l'atmosphère.

2.1.2 PSF, FTM et résolution spatiale ?

De nombreux systèmes de traitement de signal et plus particulièrement d'acquisition, sont caractérisés par leur réponse à une impulsion (Dirac) appelée réponse impulsionnelle. La représentation de ce Dirac par l'instrument est appelée Point Spread Function (PSF), qui peut se traduire par tache image (ou fonction d'étalement du point). La tache image observée par l'instrument pour mesurer la PSF étant considérée « parfaite », les défauts de représentation de cette tache sont attribués à l'instrument d'observation.

Ce paramètre est particulièrement important pour les instruments de télédétection, d'astronomie et même de microscopie. La PSF représente l'étalement du point, c'est-à-dire l'influence qu'aura un Dirac en entrée du système optique sur les pixels voisins. La FTM est la fonction duale de la PSF puisqu'elle est le module de la transformée de Fourier de la PSF.

Plus le système d'observation est performant, et plus l'étalement dû à la PSF sera réduit. Pour la FTM, plus le système est performant, plus la FTM est élevée sur le domaine de Nyquist. L'espace de Nyquist est la zone comprise entre $-f_e/2$ et $f_e/2$. Bien que la FTM soit déduite de la PSF, il est plus facile de raisonner ici dans le domaine des fréquences; c'est pour cette raison que nous utiliserons principalement la FTM plutôt que la PSF.

La figure 2.2 présente les allures des PSF et FTM pour des images à haute (HR) et basse (BR) résolution spatiale. Dans la figure de gauche, l'abscisse représente la distance au centre de la tache. Plus cette distance est grande, plus la tache est large. Dans cet exemple, on voit très bien qu'une PSF large correspond à la basse résolution spatiale. L'image de droite représente la FTM. On note que la FTM-HR permet de détecter des fréquences spatiales jusqu'à 0.75 environ, alors qu'on ne détectera pas les fréquences supérieures à 0.25 avec une FTM-BR.

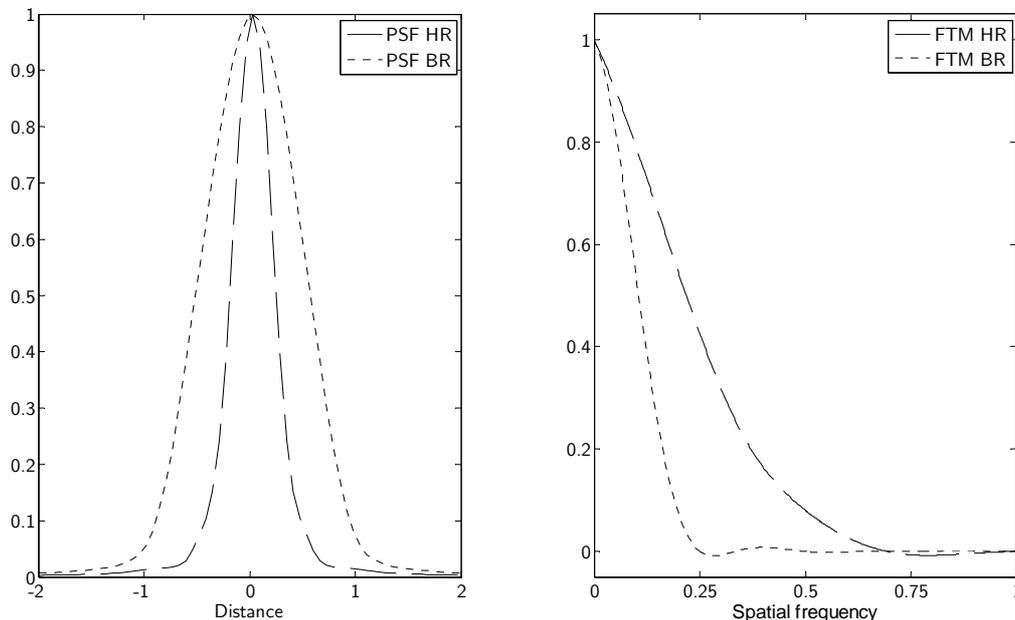


Figure 2.2 : PSF et FTM. HR: haute résolution; BR: basse résolution

2.1.3 Quel effet a la FTM sur les images ?

Les paragraphes précédents ont présenté les éléments qui modifient la FTM. La FTM est une fonction qui décroît avec la fréquence, ce qui signifie que plus les fréquences spatiales constituant une structure sont élevées et plus le contraste sera faible et tendra vers 0 (Thomas et Wald 2006).

Le fait que la FTM diminue avec la fréquence apporte du flou dans l'image acquise par l'instrument. La figure 2.3 illustre l'effet de la FTM sur une image en appliquant des FTM de niveaux très décroissants à une image de synthèse créée à des fins d'illustration. L'image de la figure 2.3-a) a une FTM idéale (égale à 1) sur le domaine de Nyquist; dans ce cas les contours sont nets et le contraste est important. Pour les images des figure 2.3-b)-c), on constate que les contours deviennent flous et que le contraste diminue à mesure que le niveau de FTM décroît. La combinaison de ces deux phénomènes sur les contours et le contraste

transforme les stries noires et blanches situées sur la droite de l'image en stries grises de plus en plus difficiles à distinguer.

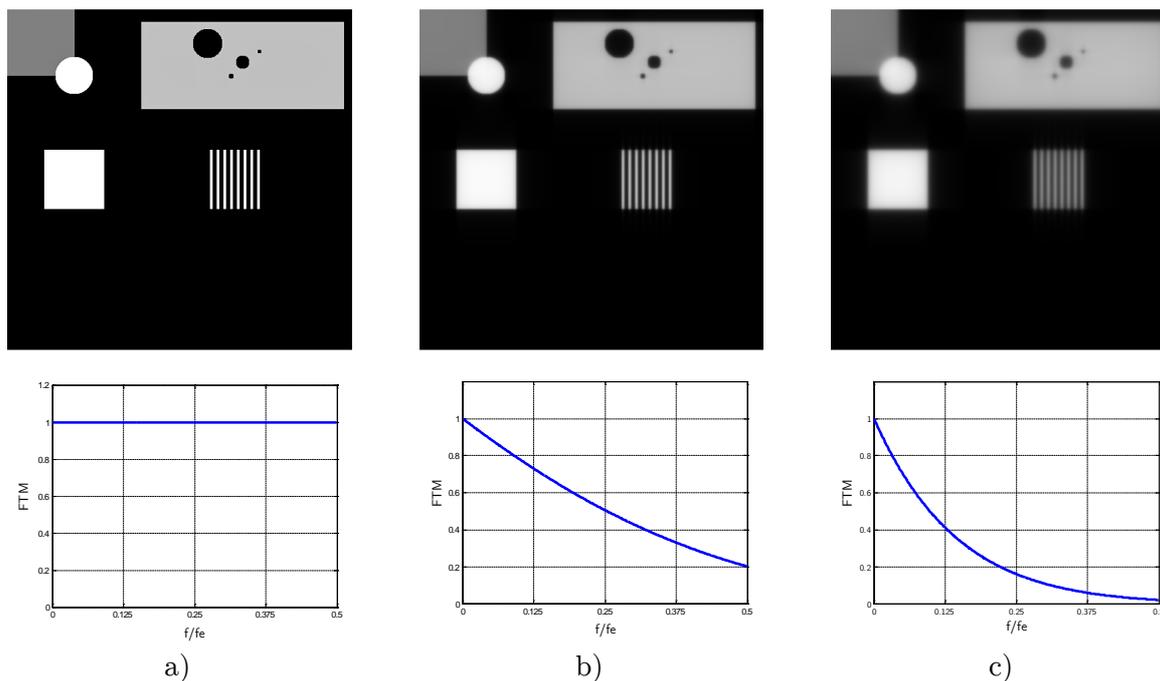


Figure 2.3 : Effet sur la mire de différentes allures de FTM. Chaque mire se trouve au dessus de la FTM qui lui a été appliquée.

Pour mieux comprendre le phénomène, nous présentons dans la figure 2.4 le profil horizontal extrait des images de la figure 2.3 au niveau des stries noires et blanches. On peut observer sur cette figure, qu'une FTM faible entraîne des fronts montants/descendants moins marqués. La courbe en pointillés représentant le cas où le niveau de FTM est le plus faible fournit un contraste nettement plus faible que dans le cas idéal, représenté en trait plein. La FTM est équivalente à un filtre passe-bas. Celui-ci conduit la suppression ou l'atténuation des hautes fréquences. Cette suppression/atténuation a pour effet de rendre les transitions moins « rapides », et la faible rapidité de croissance ne permet pas d'atteindre la valeur atteinte dans le cas idéal. Les transitions moins rapides du signal font que les contours des structures présentes dans l'image s'étendent sur un plus grand nombre de pixels et rendent les contours moins précis, et moins bien définis. La FTM a aussi pour effet de faire diminuer le contraste. La diminution du contraste associé à des contours moins bien définis provoque un effet de flou.

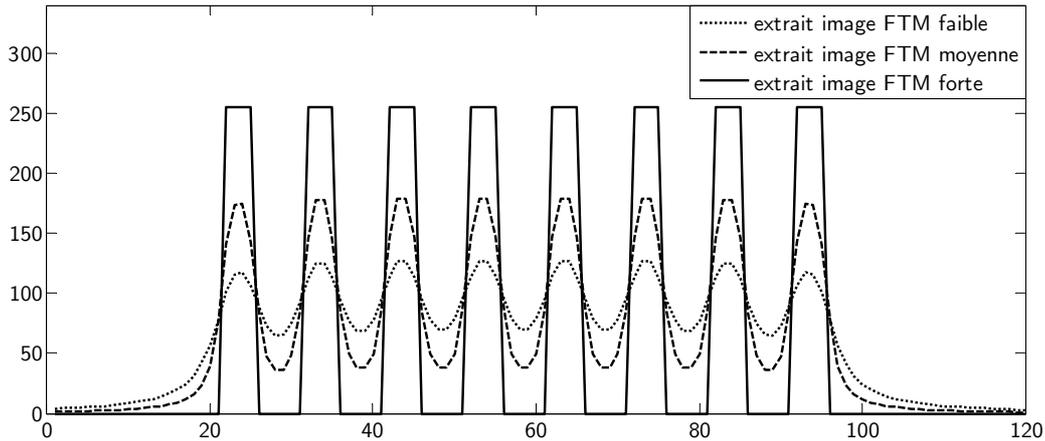


Figure 2.4 : Profil horizontal des images de la figure 2.3 au niveau du carré strié au milieu et à droite

Le pouvoir de résolution est une propriété obtenue par le croisement de la FTM, qui caractérise le contraste dans l'image et de la CSF (Contrast Sensitivity Function) (Luxen et Förstner 2002). La CSF caractérise le contraste le plus faible qu'il est possible de distinguer. Le pouvoir de résolution quantifie la netteté des images provenant de l'instrument. La netteté est un critère important pour la qualité géométrique, car celle-ci permet de bien faire ressortir la géométrie présente dans l'image.

2.1.4 Modélisation de la FTM

La FTM est liée aux transformations du signal entre le rayonnement lumineux réfléchi par la Terre (analogique, support infini) et le signal numérique (discret, à support fini). Ces phénomènes ont lieu dans la partie optique du télescope ainsi que dans la partie électronique de la détection. La stabilité de la plate-forme et l'atmosphère ont aussi une contribution.

Des dégradations du signal ont lieu à travers la chaîne d'acquisition de l'image comme : les turbulences atmosphériques, la dispersion et l'atténuation par les aérosols, les diffractions et aberrations optiques, l'intégration spatiale et temporelle, ainsi que les effets liés au détecteur et aux mouvements de l'instrument (Chen *et al.* 2009)(Delvit *et al.* 2003). La FTM totale peut être exprimée comme suit :

$$FTM = FTM_{opt} \times FTM_{atm} \times FTM_{PF} \times FTM_{det} \quad 2.1$$

La FTM_{PF} représente la FTM liée à la plate-forme. Elle peut se modéliser par la FTM de filé et de désynchronisation le long de la trace du satellite dû aux mouvements du détecteur pendant l'intégration temporelle.

$$FTM_{PIF}(f_x, f_y) = sinc(\alpha_{filé} f_y p) \quad 2.2$$

où :

- f correspond aux coordonnées dans l'espace des fréquences spatiales projetées au sol ($f \in \mathbb{R}^2$) ;
- p correspond à la projection de la dimension du détecteur au sol.

Le paramètre $\alpha_{\text{filé}}$ permet de prendre en compte les détecteurs TDI (Time Delay Integration) multi-phase, et la possibilité d'un temps d'intégration plus court que le temps d'échantillonnage.

La FTM_{det} est la FTM liée à la partie numérique de l'instrument et modélise la FTM associée à la réponse du détecteur. Le modèle suivant est le modèle classique correspondant à un détecteur de sensibilité de type trapézoïdale :

$$FTM_{\text{det}}(f) = \text{sinc}(fp) \text{sinc}(\alpha * fp) \quad 2.3$$

où :

- a est un paramètre lié au détecteur, établi en fonction de la valeur FTM de détection mesurée à Nyquist.

La FTM optique modélise la diffraction optique spectrale, notée $FTM_{\text{opt}}^{\lambda}$, directement liée au module de l'autocorrélation $AC[P](q)$ de la transmission complexe de l'instrument optique P (Marion 97) :

$$FTM_{\text{opt}}^{\lambda}(f) = AC[P](\lambda H f) \quad 2.4$$

Avec :

$$AC[P](q) = \iint P(p - q)P^*(p) dp \quad 2.5$$

où :

- H correspond à l'altitude de la plate-forme.

La FTM optique est liée à la diffraction optique. Tous les autres termes de dégradations de FTM viennent se multiplier à cette FTM. La FTM optique constitue la « borne supérieure » de la FTM de l'instrument.

C'est avec cette modélisation complète de la FTM que nous avons réalisé les simulations du satellite PLÉIADES.

Il est souvent difficile d'obtenir des informations complètes sur la FTM des instruments du fait que cette information est « critique » pour l'industriel à l'origine de l'instrument. Le plus souvent, l'information à laquelle il est possible d'avoir accès est la valeur de la FTM instrument pour la fréquence de Nyquist. En l'absence d'information sur l'allure de la FTM, celle-ci est souvent modélisée par une gaussienne dans le domaine du *pansharpening* (Nunez *et al.* 1999)(Aiazzi *et al.* 2003).

Dans nos travaux nous connaissons la FTM des images. Cependant, dans la pratique, il arrive de devoir estimer les paramètres d'un modèle. Il ne faut donc pas prendre un modèle comportant trop de paramètres, car l'ensemble des paramètres ne pourra pas être estimé et ne serait pas cohérent (Delvit *et al.* 2003).

Il existe plusieurs types d'algorithmes pour estimer la FTM d'une image. Certains déduisent la PSF à partir de la LSF/ESF (line spread function, edge spread function) qui

peut être observée à travers des cibles (Viallefont-Robinet et Léger 2010) où des structures naturellement présentes dans l'image (Thomas et Wald 2006) (Kohm 2004).

2.1.5 FTM du satellite PLÉIADES

Dans ce manuscrit, les illustrations sont souvent faites dans le cas particulier du satellite PLÉIADES, qui doit être lancé par le CNES en fin 2010. L'allure des FTM est assez similaire aux autres satellites d'observation de la Terre dans le domaine du visible. La figure 2.5 présente l'allure de la FTM pour les différentes modalités.

La décroissance de la FTM des différentes modalités MS est plus rapide que celle de l'image PAN. Cette différence de décroissance est due à la différence de taille des détecteurs, elle illustre la différence de résolution spatiale entre PAN et MS. On observe aussi un léger écart entre les FTM des bandes spectrales de l'image MS. Ceci est dû à la variation de longueur d'onde qui entraîne des diffractions différentes dans la partie optique de l'instrument. La modalité B3 dans le proche infra-rouge, présente le plus faible contraste pour toutes les fréquences.

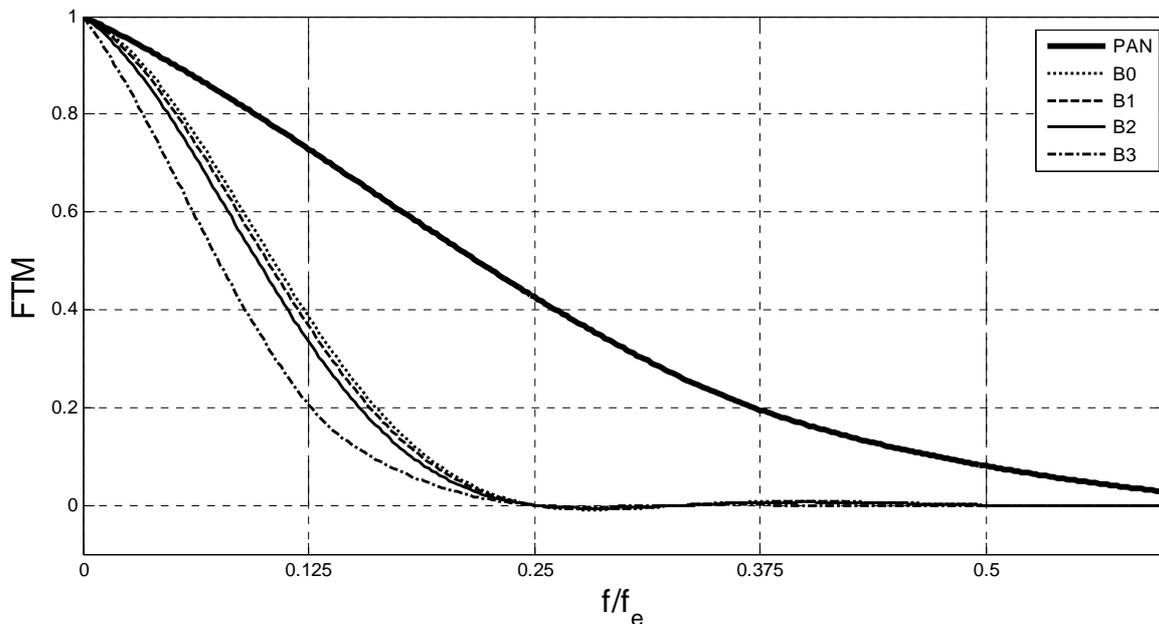


Figure 2.5 : FTM des images MS et PAN. f_e représente la fréquence d'échantillonnage pour PAN

2.1.6 Modification de la FTM d'une image

La FTM (et la PSF) que nous présentons dans ce chapitre modélise certaines atténuations que subit le signal lors de son acquisition. Nous qualifierons l'image sans atténuations par la FTM, d'image réelle. L'image ayant subi les perturbations modélisées par la FTM sera « l'image acquise ».

Il est possible d'appliquer à une image une FTM par convolution si l'on fait l'hypothèse d'un système d'observation considéré comme linéaire et spatialement invariant

(CNES/ONERA/IGN 2008). Cette hypothèse est une approximation faite communément, car assez proche de la réalité.

La convolution de l'image réelle (o) avec la Point Spread Function (PSF, noté h) donne l'image acquise (u) comme présentée dans l'équation ci-dessous. L'opération de convolution sera notée $*$.

$$u(x) = h(x) * o(x) \quad 2.6$$

À l'inverse, pour obtenir l'image réelle (o) à partir de l'image acquise (u), il faut effectuer une déconvolution. Cette opération consiste à donner à l'image la restitution des contrastes que l'image aurait eue si l'instrument avait une FTM unitaire, et ce, quelle que soit la fréquence. De cette façon, les contrastes de l'image seront représentés de façon fidèle, c'est-à-dire que la représentation du contraste sera indépendante de la fréquence spatiale.

L'image déconvoluée n'est qu'une estimation de l'image réelle, car l'opération de déconvolution est un problème mal posé au sens d'Hadamard. Pour ce type de problème, on ne sait pas s'il existe une solution, si cette solution est unique et si elle est stable vis-à-vis des bruits de mesure.

Le but de la déconvolution est d'obtenir une estimation de l'image réelle en appliquant la transformée inverse de celle subie par l'image. Deux types de déconvolution existent selon que la FTM est connue ou non. Dans le cas où la FTM est inconnue alors ce type de déconvolution est appelé déconvolution aveugle (traduction de blind deconvolution). Il est assez répandu, mais ne permet pas vraiment de « maîtriser » la FTM de l'image déconvoluée. Si la FTM est connue, alors c'est une « simple » déconvolution. Plusieurs techniques existent pour effectuer la déconvolution : le filtrage inverse, le filtrage de Wiener, le filtrage de Kalman récursif (Kundur et Hatzinakos 1996).

Le passage dans le domaine des fréquences par la transformée de Fourier transforme l'opération de convolution (dans le domaine spatial) en une simple multiplication. De même, la déconvolution se transforme en division dans le domaine des fréquences. L'ajout et la suppression de la FTM se font ainsi respectivement grâce à la multiplication et à la division dans le domaine des fréquences. Les opérations de multiplication et division sont moins coûteuses en ressource et le passage dans l'espace de Fourier est rapide. Nous noterons en majuscules les images ou noyaux de convolution dans le domaine des fréquences. Ainsi, la convolution devient :

$$U(f) = O(f) \cdot H(f) \quad 2.7$$

De même pour la déconvolution qui devient alors :

$$O(f) = I(f) / H(f) \quad 2.8$$

La déconvolution devenant une division dans le domaine de Fourier, il faut prendre des précautions afin de ne pas effectuer de division par zéro. Pour pallier ce problème, il est fréquent de faire une régularisation simple de l'inverse de la FTM.

$$FTM_{\varepsilon}^{-1} = \frac{FTM^*}{(\max(|FTM|^2, \varepsilon^2))} \quad 2.9$$

Le cas que l'on vient de présenter est théorique, car dans la pratique les images contiennent du bruit. Le bruit peut être caractérisé par le rapport signal sur bruit (SNR). Le niveau de SNR des images est différent selon les imageurs. Certains imageurs ont des niveaux de SNR élevés qui rendent l'étape de déconvolution plus ou moins critique. Dans le cas du satellite Quickbird, le niveau de SNR ne permet pas de faire de restauration de FTM sur l'image MS, mais celle-ci est possible sur les images PAN (Aiazzi *et al.* 2006). Le satellite PLÉIADES a été conçu pour que les images fournies aient un SNR élevé et ainsi permettre de faire de la déconvolution ou de la restauration de FTM. Les premières mesures au sol de l'instrument donnent des SNR supérieurs à 130 pour les bandes MS et supérieurs à 147 dB pour la bande PAN (Gaudin-Delerieu *et al.* 2008). En considérant le bruit de mesure (σ), les équations 2.5 et 2.6 deviennent :

$$U(f) = O(f) \cdot H(f) + \sigma(f) \quad 2.10$$

$$O(f) = \frac{U(f) - \sigma}{H(f)} = \frac{U(f)}{H(f)} - \frac{\sigma(f)}{H(f)} \quad 2.11$$

Le terme de bruit σ/H peut poser problème, car la déconvolution peut amplifier le bruit. Le rôle de la déconvolution est de rehausser certaines composantes fréquentielles. Ce rehaussement est effectué indifféremment sur les composantes fréquentielles du bruit et du signal. La FTM réduit le contraste pour les hautes fréquences alors que la densité spectrale du bruit est constante. Le contraste étant déjà fort pour les basses fréquences, il ne sera pas rehaussé alors qu'il le sera pour le bruit présent dans ces zones homogènes (Latry *et al.* 2001), ce qui rend la déconvolution sensible au bruit (Sroubek et Flusser 2003). Ce processus peut amplifier le bruit présent dans l'image (Banham et Katsaggelos 1997) (Aiazzi *et al.* 2006) (Latry *et al.* 2001).

Les fabricants et les fournisseurs d'images ont une bonne connaissance des caractéristiques des satellites qu'ils fabriquent ou des images qu'ils produisent. Mais la FTM typique de l'imageur n'est pas une information toujours connue ou accessible à tous les utilisateurs. De plus, on ne connaît pas toujours les traitements déjà appliqués aux images dont on dispose.

De telles lacunes constituent un problème important pour mettre en place de nouvelles méthodes de fusions performantes. Il est toutefois possible d'avoir recours à des images dites de simulation qui permettent de pallier ce problème. Ces images simulent celles fournies par le satellite à partir d'images aéroportées. Ce type d'image permet d'éviter les problèmes de méconnaissance des traitements appliqués aux images, dans le cadre de recherche et d'analyse de performance d'algorithmes de *pansharpening*.

2.2 Repliement de spectre

L'acquisition d'une image/signal numérique peut être divisée en deux parties : une analogique/optique, et une autre numérique quand le signal est numérisé. La numérisation du signal est composée de deux opérations : l'échantillonnage et la quantification du signal analogique.

Le signal analogique est de support infini. Il est impossible de le représenter numériquement. Contrairement au signal analogique, le signal numérique a un support fini. L'étape d'échantillonnage a pour but de réduire le support du signal analogique de façon à pouvoir le représenter numériquement. Le célèbre théorème sur l'échantillonnage de Shannon donne les conditions (suffisantes) permettant d'échantillonner un signal de façon réversible. L'échantillonnage est une étape importante dans l'acquisition du signal. Si le dimensionnement de ce dernier n'est pas convenable alors le signal restitué par l'échantillonnage n'assure pas forcément une bonne représentation du signal analogique. Un échantillonnage dont la fréquence est insuffisante peut entraîner une différence entre la structure et la représentation obtenue avec les échantillons.

La quantification affecte une valeur numérique à l'amplitude d'un échantillon du signal analogique. La dynamique du signal dépend du nombre de bits alloués pour coder la valeur de l'échantillon. Par exemple, si la valeur de l'échantillon est codée sur 10 bits alors $2^{10}=1024$ niveaux seront possibles. La partie de quantification est assimilable à un bruit uniforme d'écart type $q/\sqrt{12}$, où q est le pas de quantification et dont la valeur correspond à la dynamique divisée par le nombre de niveaux possible.

Dans cette partie sera présentée l'opération d'échantillonnage ainsi que le phénomène d'aliasing/repliement de spectre qui peut apparaître si les conditions de réalisation de cette opération ne sont pas optimales.

2.2.1 Le processus d'échantillonnage

L'échantillonnage est la première étape qui permet de passer d'un signal analogique à un signal numérique. Échantillonner un signal, c'est mesurer la valeur du signal analogique à intervalle régulier. L'intervalle est selon le cas un temps ou une distance. Dans le domaine des fréquences, l'échantillonnage revient à la convolution du spectre du signal par un peigne de Dirac de période égale à la fréquence d'échantillonnage.

Soit un peigne de Dirac $p(x)$:

$$p(x) = \sum_{n=-\infty}^{+\infty} \delta(t - nT) \quad 2.12$$

L'échantillonnage $f_e(x)$ peut être exprimé à l'aide de la formule de l'échantillonnage :

$$f_e(x) = f(x)p(x) = \sum_{n=-\infty}^{+\infty} f(x)\delta(t - nT) = \sum_{n=-\infty}^{+\infty} f(nT)\delta(t - nT) \quad 2.13$$

Et dans Fourier :

$$F_e(f) = \frac{1}{T} \sum_{k=-\infty}^{+\infty} F(f - \frac{k}{T}) \quad 2.14$$

Cette opération est caractérisée par un paramètre : le pas d'échantillonnage qui doit être défini relativement aux propriétés du signal à traiter. Du point de vue de l'image numérique, le pas d'échantillonnage est souvent considéré comme la résolution.

Le théorème de l'échantillonnage de Shannon donne une condition suffisante, mais non nécessaire, pour définir le pas d'échantillonnage afin d'obtenir une restitution fidèle (c'est-à-dire réversible) du signal échantillonné :

Un signal à support spectral compact est échantillonné sans perte d'information, si la fréquence d'échantillonnage est au moins égale au double de la plus haute fréquence du spectre.

Dans le domaine spatial, cette condition peut se traduire par le fait que les détails les plus fins de l'image avant échantillonnage doivent être au moins deux fois plus grands que la résolution spatiale du capteur. Si l'image est échantillonnée en respectant le théorème de l'échantillonnage alors les échantillons seront une bonne représentation de l'image ou du signal observé. Dans ces conditions, l'échantillonnage est qualifié de réversible, c'est-à-dire que le signal reconstruit à partir des échantillons sera identique en théorie au signal analogique original.

Par la suite, nous allons présenter les conséquences engendrées par le non-respect de ces conditions. Nous décrirons le phénomène de repliement de spectre aussi appelé aliasing, dans le domaine temporel/spatial puis dans le domaine fréquentiel.

2.2.1.1 Repliement de spectre dans le domaine spatial

Dans la figure 2.6, le signal analogique est représenté en trait plein, les échantillons sont symbolisés par des petits carrés situés sur le signal analogique à intervalles réguliers et le signal en pointillés est le signal que l'on peut reconstruire à l'aide des échantillons. On observe bien dans cet exemple que le signal analogique et le signal reconstruit à partir des échantillons n'ont pas la même fréquence. Or la fréquence est une des informations les plus importantes que le signal transporte, cet échantillonnage ne respectant pas les conditions de Shannon, il déforme le signal et détruit l'information de manière irréversible.

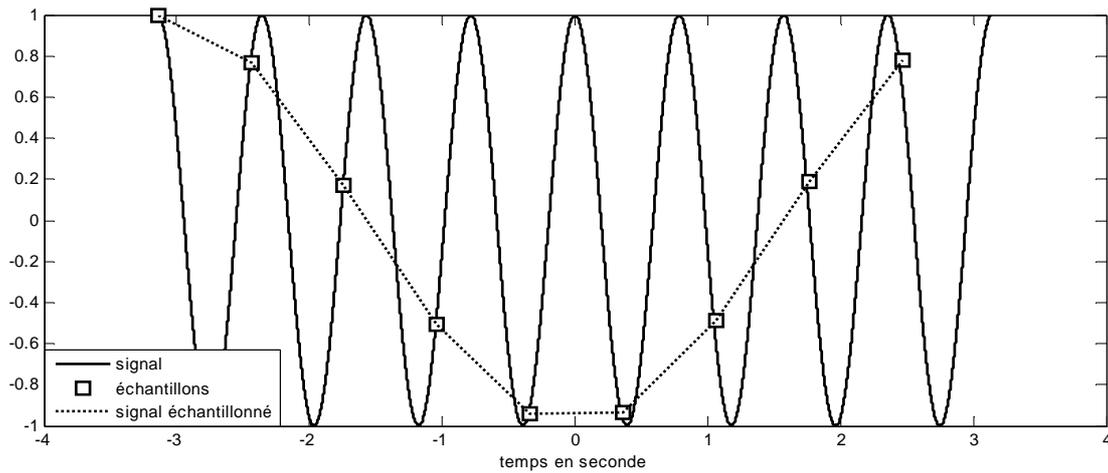


Figure 2.6 : Conséquence du non-respect des conditions du théorème de l'échantillonnage de Shannon dans le domaine spatial/temporel

La figure 2.6 a été réalisée dans le cas d'un signal à une dimension et pour des conditions bien au-delà de la condition limite dans un but d'illustration.

Pour une image, le principe est le même, mais en deux dimensions. Dans la figure 2.7, on peut distinguer au niveau des zones encadrées en blanc le changement de représentation d'une même structure du fait d'un échantillonnage trop lâche, c'est-à-dire éloigné des conditions de Shannon. L'exemple le plus visible est celui du cercle de droite (C_A) : les stries/lignes perpendiculaires aux plus longs côtés du bâtiment se sont transformées en larges stries obliques à la suite du repliement de spectre

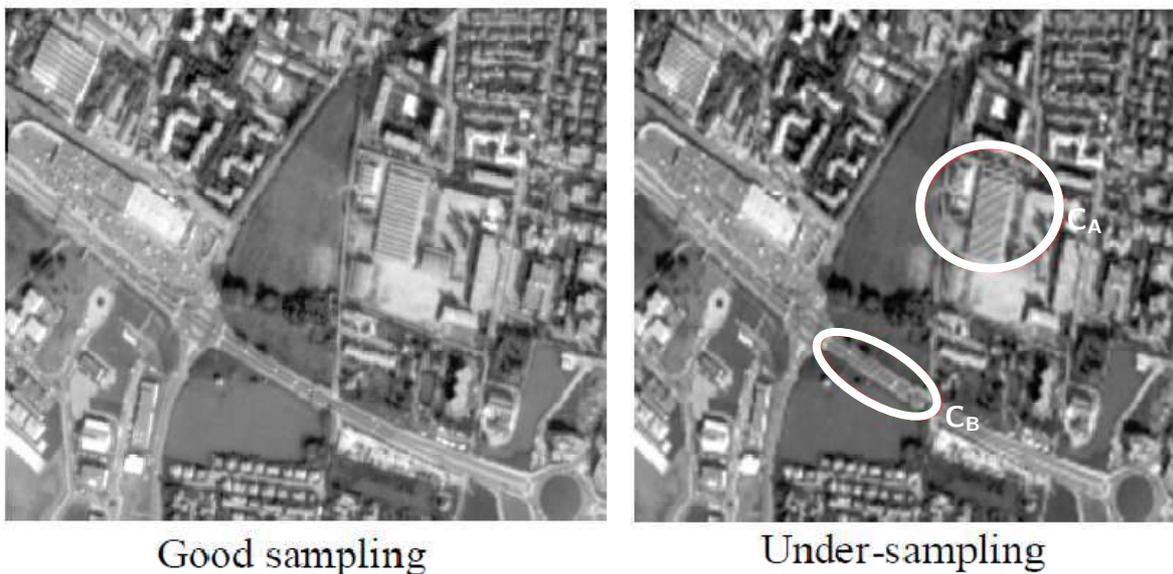


Figure 2.7 : illustration de mauvais échantillonnage donnant lieu à du repliement de spectre. Extraite d'une présentation de Lebegue CNES, 2009

2.2.1.2 Repliement de spectre dans le domaine fréquentiel

Dans cette partie, nous abordons le phénomène de repliement de spectre dans le domaine fréquentiel. C'est d'ailleurs par ce point de vue que l'on comprend son nom « repliement de

spectre ». Le fait d'échantillonner un signal va entraîner la périodisation du spectre du signal à échantillonner, et ce, à chaque multiple/module de la fréquence d'échantillonnage. Dans la figure 2.8, on peut voir de manière schématique en 1D, le spectre du signal original représenté en trait épais ainsi que ses copies (en traits fins), lesquelles sont centrées sur des multiples de la fréquence d'échantillonnage.

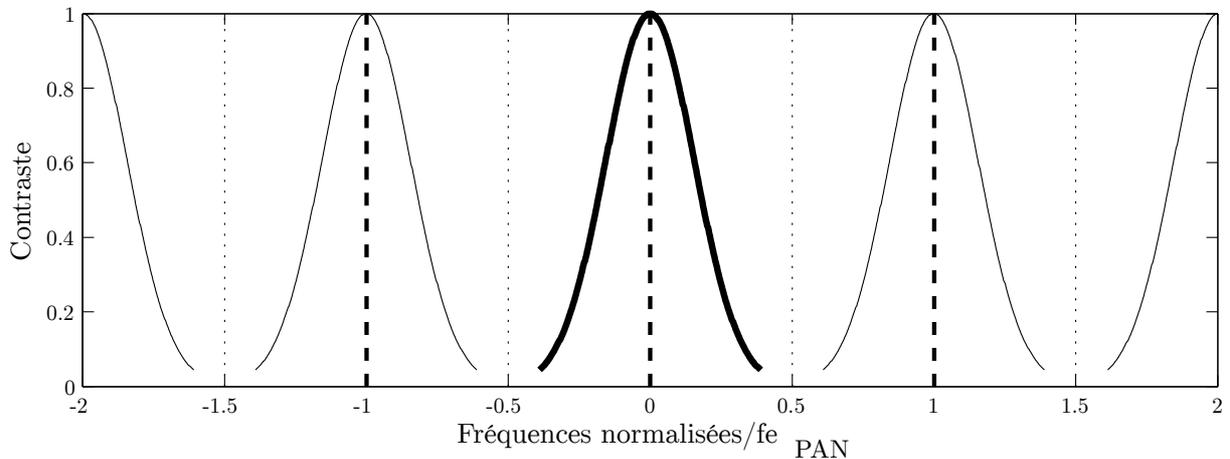


Figure 2.8 : échantillonnage et réplication du spectre

Dans la figure 2.8, la fréquence maximale du spectre du signal original (en trait épais) est inférieure à la demi-fréquence d'échantillonnage aussi appelé fréquence de Nyquist. L'échantillonnage présenté ici respecte les conditions du théorème d'échantillonnage de Shannon. Pour revenir au signal analogique, il faut effectuer le processus inverse en supprimant les copies/répliques du spectre central comme illustré dans la figure 2.9, en théorie, par l'application d'un filtre passe-bas parfait en sinus cardinal.

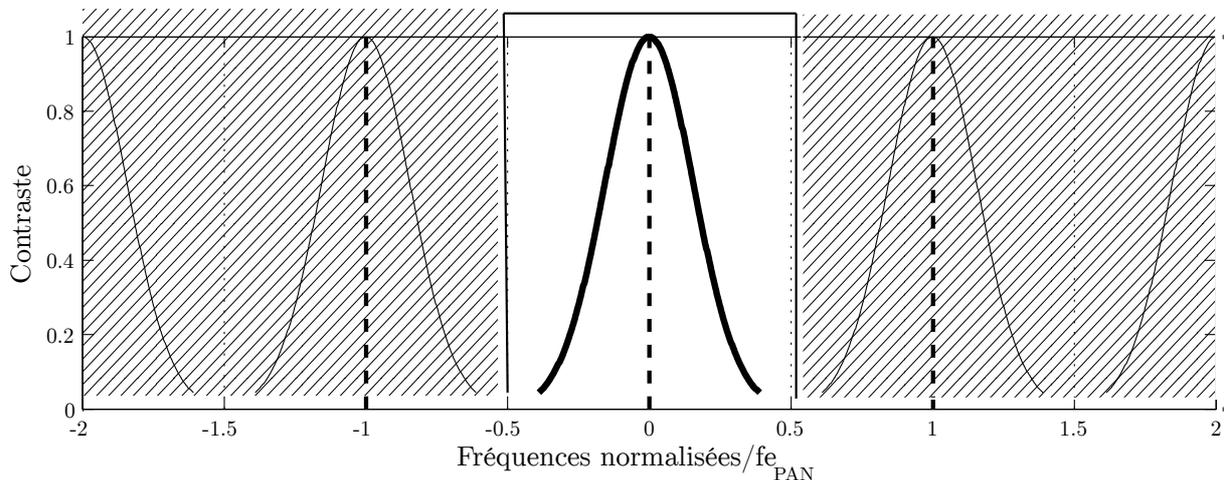


Figure 2.9 : Passage du signal échantillonné au signal analogique

Pour que l'opération d'échantillonnage soit réversible, il faut que le spectre du signal original et le spectre du signal reconstruit à partir des échantillons soient identiques. On observe bien dans la figure 2.9 que le spectre central n'est pas modifié au cours du procédé d'échantillonnage ainsi que par la reconstruction du signal. L'échantillonnage est alors un procédé réversible puisqu'il respecte les conditions du théorème de Shannon.

Si la fréquence maximale contenue dans le signal est supérieure à la demi-fréquence d'échantillonnage alors, le spectre du signal original sera présent au-delà de la fréquence de Nyquist. Or, la duplication du spectre qui se produit lors de l'échantillonnage se fait tous les multiples de la fréquence d'échantillonnage. La partie du spectre qui se situe au-delà de cette fréquence va se replier/superposer avec ses copies. Ce phénomène est illustré dans la figure 2.10. La superposition de ces spectres engendre une distorsion d'information puisque des composantes au-delà de Nyquist non nulles viennent se superposer à une autre fréquence des composantes utiles. Les zones qui présentent ce risque sont hachurées dans la figure 2.10.

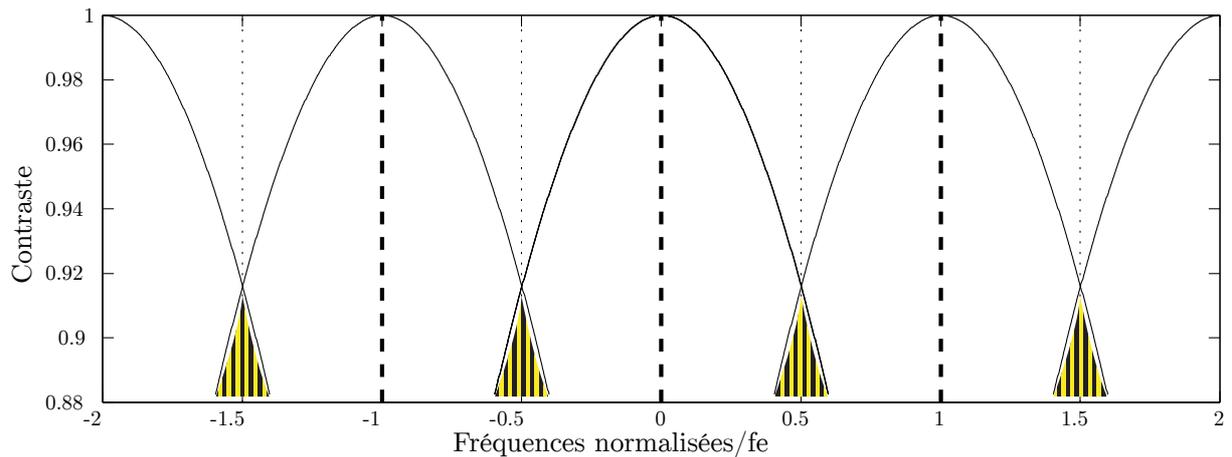


Figure 2.10 : Repliement spectral.

Dans ces zones hachurées, les spectres s'additionnent. L'allure du spectre s'en trouve modifiée. La figure 2.11 illustre la déformation induite par l'addition des spectres pour les fréquences où ils se superposent. Dans la figure 2.11 est représenté le spectre du signal original qui aurait été filtré par un filtre anti-repliement parfait qui, en pratique, n'existe pas car d'extension infinie. La courbe en trait plein est le spectre que l'on obtient du fait du repliement. On constate un rehaussement du spectre dans les hautes fréquences. Cette modification de l'allure entraîne la non-réversibilité du processus dans ces conditions. Le signal reconstruit à l'aide des échantillons n'assure pas une représentation fiable du signal échantillonné.

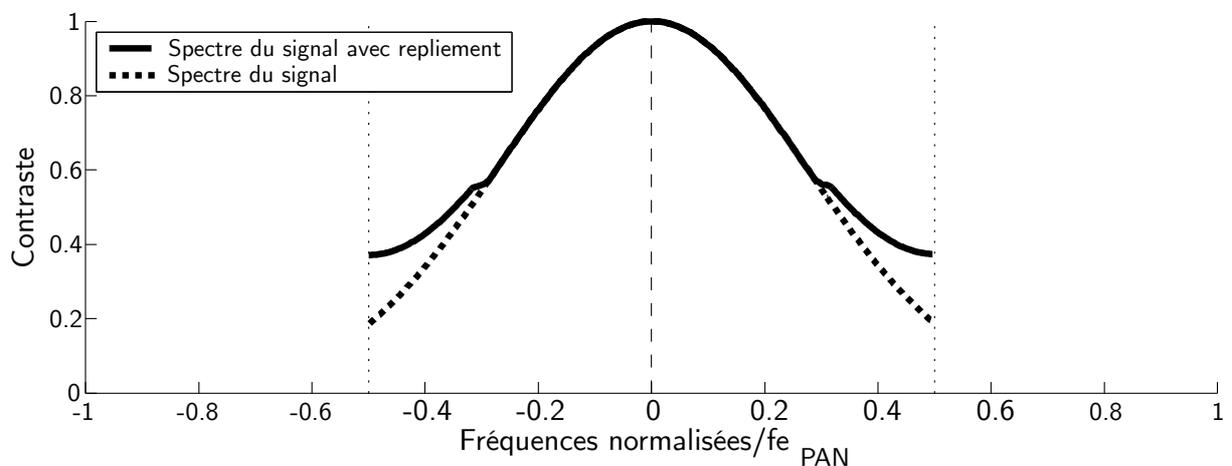


Figure 2.11 : Modification du spectre par repliement

Dans un grand nombre de systèmes d'acquisition numérique du signal audio, il est souvent difficile d'obtenir un échantillonnage suffisant. Pour éviter tout problème d'aliasing, on ajoute dans la chaîne de traitement du signal/image un filtre appelé filtre anti-repliement, dont le rôle est de supprimer les composantes du signal ayant une fréquence supérieure à la fréquence de Nyquist du système. La plupart des systèmes effectuant un échantillonnage sont pourvus de ce type de filtre. Ce filtrage est souvent présent dans la chaîne de traitement du signal même si ce dernier ne semble pas contenir de composantes supérieures à la fréquence de Nyquist. Il est ajouté pour s'assurer que l'échantillonnage respecte le théorème de l'échantillonnage indépendamment du signal d'entrée. Pour un système optique de type imageur, l'ensemble de la chaîne d'acquisition peut être modélisé par la convolution avec un filtre passe-bas et un échantillonnage. Le filtre anti-repliement est obtenu à travers le dimensionnement des éléments optiques et numériques qui constituent l'instrument.

Dans certains cas, il est difficile ou inutile de mettre en place un échantillonnage qui permet de respecter les conditions de Shannon. Au regard des propriétés du signal, mais surtout de son utilisation, il est possible de fixer/déterminer la fréquence maximale utile pour le système. Cette fréquence utile maximale deviendra la fréquence de Nyquist du signal. Dans ce cas, le filtre anti-repliement est alors dimensionné pour que la fréquence de coupure soit égale à la plus haute fréquence « utile ». Dans le cas où le signal comporte des composantes fréquentielles se situant au-delà de la fréquence de Nyquist alors, le filtrage joue son rôle et supprime ou atténue les fréquences se situant au-delà cette fréquence. L'information transportée par le signal est ainsi dégradée du fait que le filtre supprime des harmoniques du signal se trouvant au-delà de la fréquence de Nyquist (Blanchet 2006). Pour une image, un tel filtrage lisserait les structures de petite taille ou texture. Les fréquences ne pouvant pas être représentées par le signal sont filtrées. Le signal conserve une information de bonne qualité jusqu'à la fréquence de Nyquist.

Dans le cas où le filtrage anti-repliement n'est pas effectué, alors, l'échantillonnage va recevoir un signal ayant une plus grande précision (constitué d'une plus large bande de fréquences) car il contiendra toutes les fréquences. Mais comme on vient de le présenter, les hautes fréquences vont créer, et subir, du repliement de spectre. Dans ces conditions, la plus haute fréquence de bonne qualité (sans repliement de spectre) aura une fréquence plus basse que la fréquence de Nyquist du fait de perturbations engendrées par le repliement. Si l'on n'utilise pas de filtre anti-repliement, cela crée une dégradation moins visible, mais au moins aussi importante.

2.2.2 Phénomènes et artefacts liés à l'échantillonnage

Quelques rappels sur les effets de l'échantillonnage seront faits ici. Nous nous sommes inspirés de la thèse de (Blanchet 2006).

Nous considérons ici le cas où le support analogique est plus large que la gamme de fréquences que l'on peut représenter à l'aide de l'échantillonnage de l'instrument. Plusieurs types d'artefacts peuvent apparaître, suivant la manière utilisée pour représenter

l'information. Les trois artefacts principaux liés à l'échantillonnage sont : le flou, repliement de spectre et le ringing.

- **Le flou** : engendré par la limitation de la bande passante avant l'échantillonnage. L'aspect flou est dû à la faible présence de hautes fréquences et au faible contraste des plus hautes fréquences présentes. Le fait de supprimer une partie des hautes fréquences au signal est un processus qui dégrade la qualité visuelle de ce dernier. Ce cas est illustré par la figure 2.12-a).
- **L'aliasing** (repliement de spectre) : Si la bande passante n'est pas limitée avant l'échantillonnage alors des phénomènes d'aliasing apparaîtront. Ce phénomène crée de la fausse information. Visuellement, l'aliasing se manifeste essentiellement par des changements de textures ou le crénelage de segments de droite (figure 2.12-b)).
- **Le ringing** apparaîtra si la bande passante est coupée de façon brutale après l'échantillonnage. Le fait de couper de façon brutale le spectre va créer du ringing (figure 2.12-c)).

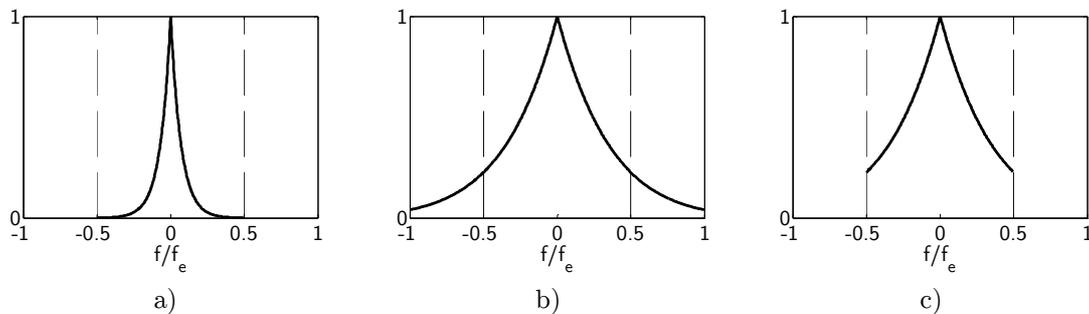


Figure 2.12 : Spectre des différents artefacts d'échantillonnage (Blanchet 2006).

a) flou : le spectre est trop large. b) aliasing : modification des hautes fréquences.

c) ringing : coupure brutale des fréquences.

2.2.3 Le repliement de spectre dans les images MS et la fusion

La partie précédente explique comment se produit le repliement de spectre pour un signal ou une image. Dans cette partie, nous montrons que les images MS sont sujettes au repliement de spectre. Nous présenterons aussi comment le repliement intervient dans la fusion entre les images PAN et MS.

2.2.3.1 Repliement de spectre dans les images MS

Dans un système d'acquisition d'image, les éléments intervenant sur la chaîne image n'ont pas tous un comportement idéal. Ainsi, la somme de leur comportement est équivalente à un filtre passe-bas, qui est exprimé par la FTM, laquelle a été présentée dans le §2.1. La figure 2.13 présente, quant à elle, les FTM du satellite PLÉIADES.

Le pas d'échantillonnage de l'image MS est quatre fois plus grand que celui de l'image PAN. La plus haute fréquence que l'on peut échantillonner devient quatre fois plus petite que pour le cas PAN. Le plus souvent, l'échantillonnage de la modalité PAN est correct du point de vue des conditions de Shannon. Le signal à échantillonner lui, reste le même, car les

capteurs des voies PAN et MS sont placés dans le même télescope et subissent des distorsions optiques pratiquement identiques, l'influence de la variation de la longueur d'onde sur la distorsion optique étant relativement faible.

L'échantillonnage des images MS n'est pas en mesure de respecter les conditions du théorème de l'échantillonnage de Shannon. Par conséquent, les images MS sont susceptibles de contenir des défauts dus au repliement de spectre pour les fréquences plus élevées.

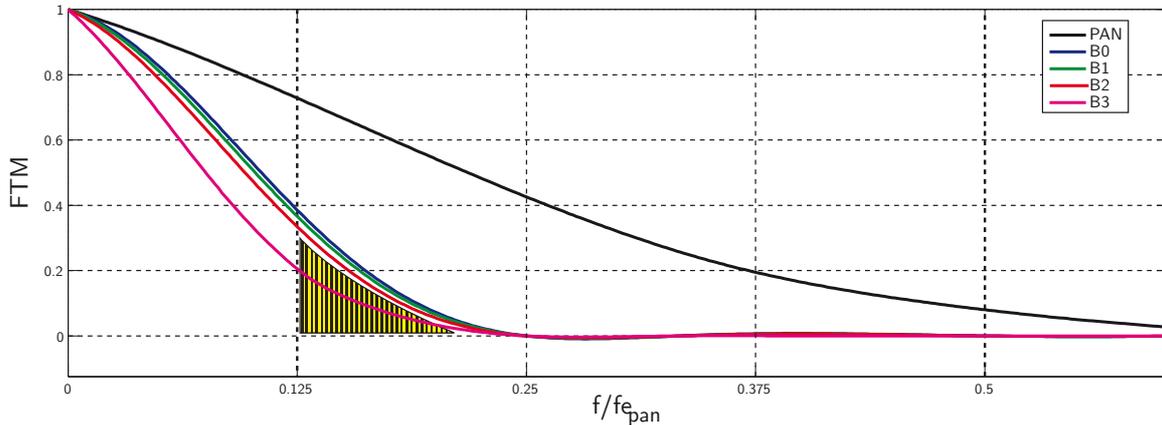


Figure 2.13 : FTM des modalités du satellite PLÉIADES

On peut voir dans la figure 2.13 que le niveau FTM de la modalité PAN est faible au-delà de sa valeur de Nyquist $0,5 f_e$. La fréquence étant normalisée par rapport à la fréquence d'échantillonnage de la modalité PAN, la fréquence de Nyquist pour la modalité MS est dans ce cas $0,125 f_e$. Pour les différents canaux de la modalité MS, la valeur de la FTM pour la fréquence de Nyquist est comprise entre 0,2 et 0,4 selon les canaux.

Le capteur MS reçoit du signal qui peut contenir des fréquences qui se trouvent au-delà de sa fréquence de Nyquist, dans la zone hachurée de la figure 2.13 (Rosak *et al.* 2004). L'énergie va se retrouver sur la gamme de fréquences se situant entre 0 et la fréquence de Nyquist des images MS. Si les niveaux de FTM étaient très faibles comme dans le cas de la modalité PAN, alors il y aurait du repliement, mais dont l'effet serait négligeable. Ce n'est pas tout à fait le cas ici, car la FTM de la gamme de fréquences qui se repliera a un niveau compris entre 0 et 0,4 qui pourrait être suffisamment élevé pour perturber l'information.

Les fréquences qui vont engendrer du repliement de spectre sont des hautes fréquences. Bien que la FTM laisse passer des fréquences qui peuvent se replier, ces fréquences n'existent pas forcément dans le paysage observé. La présence de repliement de spectre dans l'image dépend du paysage observé. Un paysage urbain est constitué de structures créées par l'homme. Pour des résolutions submétriques, un paysage urbain sera constitué d'un plus grand nombre de hautes fréquences que ne le sera un paysage rural (Aiazzi *et al.* 2006) (Wald 2002). Les zones urbaines présentent un risque plus élevé de contenir des fréquences susceptibles de créer du repliement.

Le support de notre discours était le cas de PLÉIADES pour lequel le rapport de résolution est de 4 entre les images PAN et MS. Ce problème de repliement est présent pour les autres satellites et même ceux dont la résolution est plus basse comme SPOT-5. La

preuve en est qu'elle a même été exploitée dans le cas de SPOT-5 avec le SUPERMODE (Latry et Rouge 2003). Les images de la figure 2.14 montrent bien que pour les images à 5 m (figure 2.14-a) il y a du repliement qui est absent dans l'image supermode à 2.5 m (figure 2.14-b).

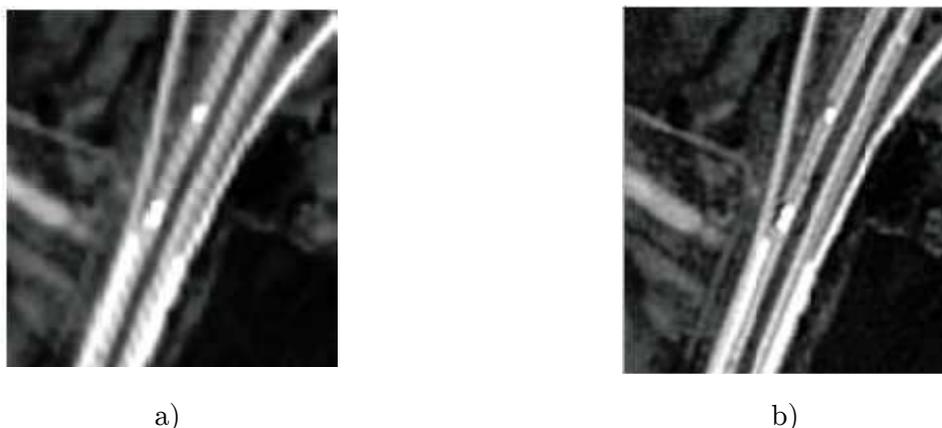


Figure 2.14 : exemple de repliement dans le cas d'image SPOT-5. a) Image SPOT-5 5 m. b) Image SPOT-5 supermode 2.5 m. image (Latry et Rouge 2003)

2.2.3.2 Filtre anti-repliement et satellite d'observation

Dans la partie précédente, nous avons montré que la modalité MS peut contenir du repliement de spectre. La présence de défauts dans des images provenant d'un satellite est surprenante étant donné qu'un satellite est un instrument extrêmement abouti, complexe et coûteux. On s'attend à ce qu'un tel système fournisse des images dépourvues de repliement de spectre. En effet, ce repliement de spectre peut avoir notamment des incidences sur l'exploitation visuelle en introduisant des artefacts au niveau des zones structurées (Coulange et Moisan 2010) ou sur la précision des algorithmes de mesures subpixelles de parallaxe pour les traitements stéréographiques (Delon 2004). Dans cette partie, nous présentons des explications quant à la présence de ce repliement dans les images.

On imagine qu'un tel système pourrait, comme la plupart des systèmes, avoir un filtre anti-repliement de façon à éviter la présence d'aliasing dans ces images. Dans le cas de l'imagerie optique spatiale qui nous intéresse, c'est la FTM qui joue ce rôle de filtre anti-repliement.

Pour une résolution donnée, la susceptibilité d'un système imageur au repliement de spectre est d'autant plus grande que le télescope est de grande taille. Dans le cas de PLEIADES, dans un souci de réduction de coût du système, la taille du télescope est telle que la sensibilité de la voie PAN au repliement est faible, comme on peut l'observer sur la figure 2.13. Par contre, l'acquisition des images MS se faisant à une résolution quatre fois inférieure avec le même télescope, il est impossible d'éviter la présence de repliement de spectre. Il est cependant à noter que la FTM des détecteurs des voies MS jouent, dans une certaine mesure, un rôle de filtre anti-repliement.

Pour éviter la présence de repliement, une solution pourrait être d'avoir deux télescopes sur le satellite, un pour chacune des résolutions. Dans ce cas, chacun des deux instruments optiques pourrait être dimensionné de façon à éviter le repliement de spectre. Une telle configuration évidemment entraîne un surcoût de fabrication de poids et de volume du satellite.

Dans la configuration avec un seul instrument, il pourrait être envisageable d'ajouter un filtre optique avant le détecteur MS pour supprimer les fréquences responsables du repliement de spectre. L'utilisation de ce type de filtre n'est pas faite, car ces derniers dégradent de façon trop importante certaines caractéristiques du cahier des charges comme la FTM ou le SNR. Or le niveau de FTM ou de SNR sont des caractéristiques importantes de la qualité image (QI) alors que le taux de repliement dans l'image n'est pas pour le moment une caractéristique du cahier des charges (Gaudin-Delerieu *et al.* 2008) (Rosak *et al.* 2004).

2.2.3.3 Repliement de spectre et Pansharpening

Nous venons de montrer que les images MS sont susceptibles de contenir du repliement de spectre. Ce repliement de spectre effectivement dans l'image est, par nature, local et fonction des structures du paysage observé. Nous montrons maintenant de quelle façon le repliement de spectre intervient dans la fusion entre les images des modalités PAN et MS.

Les défauts dus au repliement de spectre ont été présentés dans le paragraphe 2.2.2. Ils apparaîtront sur les textures dont l'aspect peut changer ainsi que sur les lignes obliques dont l'aspect peut devenir crénelé. Les zones qui présenteront des défauts sont par exemple les contours avec un fort contraste notamment quand celui-ci est représenté par une structure linéaire ou un toit ondulé (Coulange et Moisan 2010).

De nombreuses méthodes de pansharpening requièrent des opérations de ré-échantillonnage des images MS originales. Le fait de ré-échantillonner peut amplifier l'effet du repliement. Pour obtenir une image synthétisée de bonne qualité, il faut limiter l'apparition d'artefacts durant le ré-échantillonnage. Pour illustrer l'effet possible du repliement de spectre dans un processus de fusion, la figure 2.15 présente une image à haute résolution (la plus à droite) et la même image sous-échantillonnée (la plus à gauche). Cette image BR est ré-échantillonnée tout d'abord sans correction du repliement de spectre (figure 2.15-b) puis avec correction (figure 2.15-c)

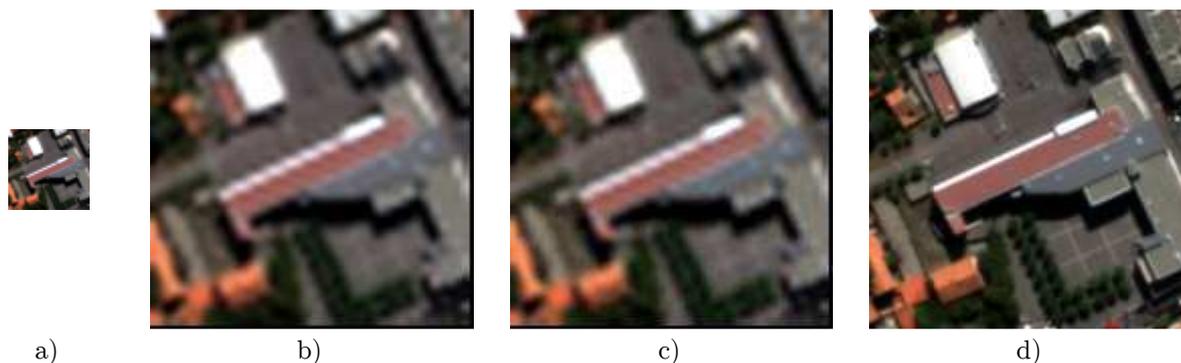


Figure 2.15 : Effet du repliement de spectre dans une image issue de la fusion.
 a) image BR, b) image BR ré-échantillonnée (rééchantillonnage de type spline), c) Image BR ré-échantillonnée (sans repliement de spectre) (rééchantillonnage de type spline), d) image de référence

On voit dans la figure 2.15-b) que le bord du toit a un aspect hachuré (phénomène typique du repliement de spectre). Dans la figure 2.15-c) l'aspect hachuré du toit est moins visible que sur la de figure 2.15-b). L'image de référence figure 2.15-d) est dépourvue de ces hachures. La présence de cet effet « hachuré » est liée au repliement et au ré-échantillonnage. On voit donc dans cet exemple que le repliement de spectre crée une « fausse » information semblant indiquer une structure de bord de toit différente de celle réelle (figure 2.16). Cet exemple est assez visuel du fait qu'il a lieu sur une structure assez grande, mais dans certains cas l'effet peut être beaucoup plus localisé.

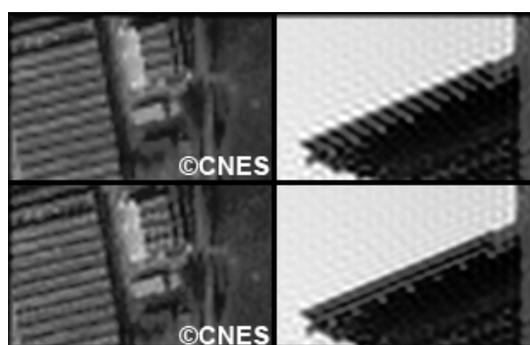


Figure 2.16 : exemple de repliement de spectre. (Coulange et Moisan 2010)

Le fait de rééchantillonner amplifie l'effet du repliement de spectre. Dans un certain nombre de méthodes de *pansharpening* le rééchantillonnage reste nécessaire. Pour obtenir une image synthétisée de bonne qualité, il est nécessaire de limiter l'apparition ou l'amplification de ces artefacts au cours de l'étape de rééchantillonnage.

2.3 Rapport signal à bruit

Le niveau de bruit présent dans les images est quantifié par le rapport signal à bruit nommé « Signal to noise ratio » (SNR) en anglais. Le SNR quantifie le bruit par rapport à l'information contenue dans l'image.

Le SNR est le rapport entre la moyenne du signal et l'écart type du bruit de mesure. Le SNR est une caractéristique importante des images. Plus le bruit est fort dans l'image, plus le

SNR est faible. Le bruit dégrade la lisibilité et l'exploitabilité de l'information contenue dans l'image.

La présence d'un niveau de bruit important sur une zone homogène comportant des structures fines va rendre ces structures invisibles du fait de la seule présence du bruit, sauf si ces dernières ont un contraste important. Avoir un fort contraste pour des structures fines reste peu probable puisque ces structures seront représentées en grande partie par des hautes fréquences dont le contraste est atténué par la FTM. D'ailleurs, du point de vue de la qualité des images d'un système optique, la FTM et le SNR sont souvent liés comme, par exemple, l'introduction de la notion de « *merit factor* » dans le dimensionnement de PLÉIADES (Rosak *et al.* 2004).

Les algorithmes qui utilisent des images satellites sont souvent des processus automatisés. Ces algorithmes peuvent faire par exemple de la détection, du comptage, de la segmentation, de la reconnaissance de cibles. Tous ces traitements sont évidemment sensibles au bruit et sont susceptibles d'aboutir à des résultats non satisfaisants comme par exemple augmenter le taux de fausses alarmes si le niveau de bruit de l'image est trop fort. L'aspect automatique de ces processus les rend le plus souvent sensibles au bruit présent dans les images.

Le niveau de SNR est un paramètre important et ne doit pas être trop dégradé lors de la fusion. À ce titre et compte tenu des traitements que nous allons faire sur les images lors de la fusion, il est apparu nécessaire de considérer ce paramètre pour les images que nous synthétisons. Nous utiliserons les valeurs de SNR typique de PLÉIADES, présentées dans le tableau suivant :

Tableau 2.1 : Niveau de SNR en dB pour les images du satellite PLÉIADES
(mesures au sol) (Gaudin-Delerieu *et al.* 2008)

bandes	B0	B1	B2	B3	PAN
SNR	>130	>130	>130	>130	>147

3 Jeux d'images, concept de fusion, évaluation de qualité, outils

Dans ce chapitre nous présentons tous les éléments nécessaires pour évaluer les propositions d'amélioration que nous allons faire. Nous commençons par les images utilisées. Elles sont aux nombres de trois et sont des simulations d'une image du futur satellite d'observation Pléiade de la ville de Toulouse, et des simulations SPOT-5 pour les villes de Fredericton et Madrid. Les simulations Pléiades sont obtenues à partir d'images aéroportées et celles de SPOT-5 à partir d'images Quickbird.

Puis nous présentons le concept de fusion ARSIS. Celui-ci est basé sur une analyse multi-résolution et s'articule autour de trois modèles ; le modèle multi-échelle permet de décomposer l'information en fonction de la résolution, le modèle inter modalité qui décrit le passage d'une modalité à l'autre et enfin le modèle d'inter modalité à haute résolution qui sert à injecter l'information à haute résolution spatiale dans une autre modalité. La méthode ATWT-M3 correspond à ce concept. Elle est simple, bien connue, et servira de référence pour jauger les différentes améliorations que nous voulons introduire.

Une partie importante de la fusion est l'évaluation des performances de l'algorithme qui se traduit par l'évaluation de la qualité de l'image synthétisée. Dans un premier temps, nous présentons les travaux qui font référence dans l'évaluation de la qualité, et décrivons alors le protocole standard ainsi que les distances et critères calculés. Dans le domaine de la fusion d'image et de l'évaluation de sa qualité il n'y a pas de mesure de qualité liée à la représentation du contraste en fonction de la fréquence, ni de mesure permettant d'évaluer la présence de repliement de spectre dans une image. Ainsi nous avons défini une nouvelle mesure de la qualité des images dans le domaine de la fusion, MTF_{dev} qui permet de calculer une FTM relative entre deux images, et fournit une évaluation de la bonne restitution de la représentation spatiale de l'image synthétisée. C'est la première innovation de la thèse. Pour l'évaluation du repliement de spectre, nous avons évalué la pertinence d'un détecteur de repliement de spectre proposé par Blanchet (Blanchet 2006) qui mesure la variation totale entre une image et cette même image ayant subi des translations subpixeliques.

3.1 Jeux d'images utilisées

Les images que nous utilisons sont des images des villes de Toulouse (France), Madrid (Espagne), et Fredericton (Canada). Les images PAN de ces scènes sont présentées par la figure 3.1. Ces scènes de ville sont des zones pour lesquelles l'urbanisation est dense. Ce type de scène présente beaucoup de hautes fréquences et des signatures spectrales très variées qui en font des zones particulièrement délicates à synthétiser.

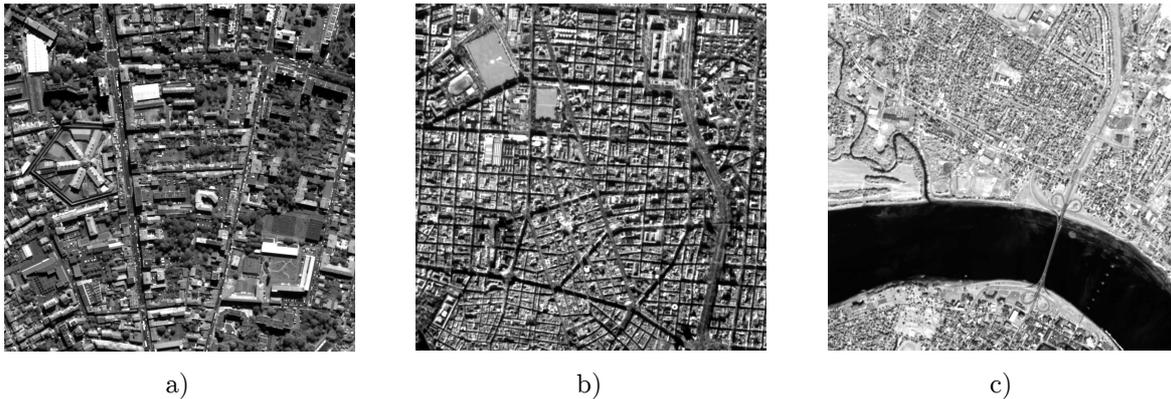


Figure 3.1 : Scène de la ville de : a) Toulouse (PLEIADES). b) Madrid (Quickbird). c) Fredericton(Quickbird).

L'évaluation des performances à « pleine » résolution est impossible, en opérationnelle, du fait de l'absence de l'image MS de référence. Si cette image MS n'existe pas, (Wald *et al.* 1997) ont proposé de dégrader la résolution spatiale des images originales, par exemple de 1 m à 4 m pour PAN, et 4 m à 16 m pour MS. La fusion est faite avec ce nouveau jeu. Les images synthétisées à 4 m sont alors comparées aux images MS originales et on obtient alors un bilan de qualité. On suppose alors que ce bilan reste invariant dans les échelles et qu'il représente ce qu'on aurait obtenu à 1 m. (Thomas 2006) a trouvé par expérimentation que c'était quasiment toujours le cas.

(Alparone *et al.* 2008) ont récemment proposé d'utiliser un indice de qualité qui évalue la qualité sans référence, le QNR. Avec de tels indices, il devient possible d'évaluer les performances de la fusion à pleine résolution, en dehors du cadre d'image de simulation.

Cependant, les travaux de (Khan 2009) montrent qu'une amélioration de la qualité se traduit par une amélioration de la valeur du QNR. Dans ses travaux Khan cherche à optimiser une méthode de fusion par rapport au QNR. Les images obtenues avec cette méthode ne donnent pas des résultats satisfaisants sur toutes les zones. Ainsi, cet indice ne semble pas suffisant pour évaluer l'ensemble de la qualité.

Pour nos travaux, nous avons décidé de travailler essentiellement avec des images issues de simulations. Travailler avec ce type d'images facilite l'évaluation des améliorations proposées de la qualité car dans ce cas nous disposons d'une image de référence à haute résolution spatiale avec une bonne connaissance et une bonne maîtrise de la FTM sous-jacente, du SNR et du repliement de spectre.

3.1.1 Simulation PLEIADES

Le choix des images s'est donc porté sur les simulations d'images du satellite PLEIADES, fournies par le CNES. Ces simulations sont générées à partir d'images aéroportées provenant du capteur PÉLICAN qui fournit des images à une résolution de 35 cm. Les bandes spectrales du capteur PÉLICAN sont proches de celles de PLEIADES (tableau 3.1). Cependant, le capteur PELICAN ne dispose pas de capteur panchromatique identique à PLEIADES. Dans les simulations PLEIADES le canal panchromatique sera simulé par une somme pondérée des différentes bandes spectrales. La simulation PLEIADES de l'image PAN

a ainsi une bande passante spectrale qui diffère légèrement de celle du capteur panchromatique de PLEIADES (tableau 3.1). Les images PAN de PLÉIADES auront une bande spectrale plus large qui devrait permettre de meilleurs résultats spectraux pour la bande B3, car l'image PAN contiendra une plus grande partie du PIR.

Tableau 3.1 : Bandes passantes spectrales (en nm) pour PÉLICAN et PLÉIADES

	PAN	B0	B1	B2	B3
PELICAN		430-550	490-610	600-720	750-950
PLEIADES	480-820	430-550	490-610	600-720	750-950

Pour ces images, sont disponibles pour une même zone :

- le canal PAN à 70 cm de résolution ;
- les 4 canaux MS à 280 cm et 70 cm de résolution ;

Pour ce type d'image nous disposons uniquement de la ville de Toulouse, et plus particulièrement du quartier Saint-Michel. Les images étant issues de simulation, à chaque résolution, les images ont une FTM qui correspond à sa résolution. L'allure des FTM pour les différents canaux a été présentée dans la figure 2.5. La figure 3.2 présente la sensibilité spectrale de ces canaux.

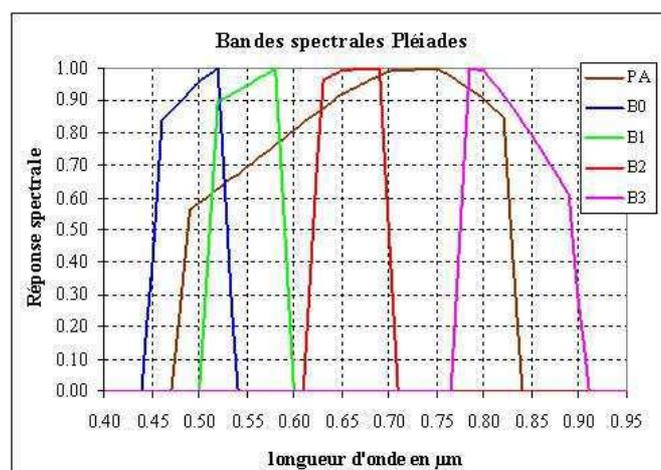


Figure 3.2 : sensibilité spectrale du satellite PLÉIADES
(<http://smc.cnes.fr/PLEIADES/Fr/>)

Nous avons choisi de travailler dans ce cadre de simulation d'images PLÉIADES, car nous avons à notre disposition les caractéristiques détaillées et étayées de ce système, notamment en ce qui concerne la FTM. Avec ces images, nous pourrions synthétiser des images multispectrales à 70 cm (HR), grâce à la fusion des images MS à 280 cm (BR) et de l'image PAN à 70 cm (HR). Les images MS à 70 cm fournies seront les références auxquelles nous pourrions comparer les images issues de la synthèse. Il est donc possible de réaliser la fusion à pleine résolution tout en ayant une image de référence pour évaluer les performances de l'algorithme de fusion. Et ceci, en ayant à chaque résolution une FTM typique.

3.1.2 Simulation SPOT-5

Nous avons réalisé des simulations d'images du satellite SPOT-5 à partir d'images Quickbird. Ceci est rendu possible par un écart de résolution spatiale suffisant entre les images Quickbird et SPOT-5. Les résolutions spatiales selon les bandes sont répertoriées dans le tableau 3.2 pour SPOT-5 et Quickbird. Ces images ont les caractéristiques de résolution spatiale et de FTM des images SPOT-5.

	Quickbird	SPOT-5
PAN	70 cm	5 m
PAN supermode		2.5 m
MS	280 cm	10 m
MS supermode		5 m

Tableau 3.2 : résolutions spatiales des différentes images utilisées

Étant simulées à partir d'images Quickbird et ayant des bandes passantes spectrales différentes de SPOT-5, les bandes spectrales des simulations SPOT-5 conservent les mêmes caractéristiques spectrales que Quickbird (tableau 3.3). De plus, notre approche ne prend pas en compte les caractéristiques spectrales des images. C'est pour ces raisons que nous conservons pour les images simulant les images SPOT-5, les caractéristiques spectrales des images à partir desquelles nous avons simulé ces images. Les images simulées sont donc des images Quickbird à la résolution de SPOT-5 avec des FTM et le SNR de SPOT-5.

Les simulations d'images SPOT-5 sont effectuées pour les villes de Madrid (Espagne) et de Fredericton (Canada). Les détails concernant ces images sont données dans le tableau 3.4 et tableau 3.2.

nm	PAN	Bleu	Vert	Rouge/jaune	PIR	MIR
Quickbird	450-900	450-520	520-600	630-690	760-900	
SPOT-5	480-710		500-590	610-680	780-890	1580-1750

Tableau 3.3 : Bandes passantes spectrales en nm de Quickbird et SPOT-5

	Madrid	Fredericton
Date	29/01/2002	31/08/2002
Size	8604x8336/ 2151x 2084	11784x11812 / 2946x2953
Taille de l'extrait	4096x4096 / 1024x1024	4096x4096 / 1024x1024
Taille de l'image SPOT-5 simulée	1024 (PAN _{supermode}) 512(MS _{supermode})/PAN 256 (MS)	1024 (PAN _{supermode}) 512(MS _{supermode})/PAN 256 (MS)

Tableau 3.4 : informations sur les images Quickbird et les images qui seront simulées

Nous donnons ici quelques détails sur la façon dont nous avons simulé les images. Dans le processus de simulation « normal », il aurait fallu déconvoluer les images Quickbird puis ensuite ajouter la FTM typique des images SPOT-5. Les images Quickbird dont nous disposons semblent visuellement déjà avoir subi une restauration de FTM. Dans le cadre de notre travail, ce qui importe le plus est la différence de FTM entre les images PAN et MS. Le niveau de FTM absolu n'est pas le point le plus critique. Le fait de ne pas corriger la FTM sous-jacente des images Quickbird pour simuler les images SPOT-5 ne fait que modifier le

niveau de FTM absolu, mais n'empêche pas d'appliquer la différence de FTM qui nous intéresse entre les images PAN et MS.

La simulation de ces images commence par un filtrage à l'aide de la FTM de SPOT-5 qui correspond à la résolution que l'on veut lui donner (figure 3.3). Puis suit un ré-échantillonnage qui se fait en plusieurs étapes parce que le rapport de résolution qui permet de passer de la résolution de Quickbird à celle de SPOT-5 n'est pas un entier. Le ré-échantillonnage commence par un sur-échantillonnage de rapport 7 puis un sur-échantillonnage de rapport 5 effectué deux fois.

$$I_{SPOT5} = (I_{QB} * PSF_{SPOT5}) \uparrow 7 \downarrow 5 \downarrow 5 \downarrow 2 \quad 3.1$$

Les images PAN et MS sont générées avec des échantillonnages qui, dans la version finale, seront décalés d'un demi-pixel en ligne et en colonnes. Ces deux échantillonnages sont réellement fournis par SPOT-5. Grâce à ces échantillonnages décalés, il est possible de créer les images SPOT-5 supermode (Latry et Rouge 2003). Le supermode exploite le décalage d'un demi-pixel et crée des images dont l'échantillonnage est deux fois plus fin, mais dont la résolution n'est pas strictement deux fois supérieure mais de l'ordre de $\sqrt{2}$.

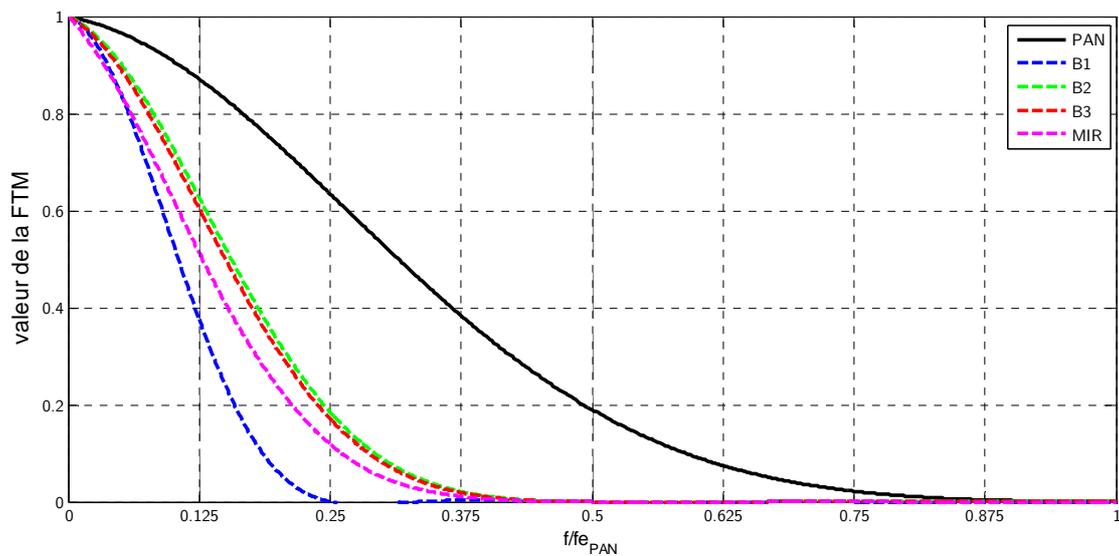


Figure 3.3 : FTM des images SPOT-5.

3.2 Notations

Comme (Thomas 2006) et (Wald 2002) nous définissons l'ensemble des modalités de haute résolution spatiale par A_θ et l'ensemble des modalités de basse résolution par B_l . Les termes de 'haute/basse' résolution spatiale sont des termes relatifs pour désigner l'écart entre les résolutions spatiales des ensembles originaux. Ici, la haute résolution spatiale est notée *res0*. En indice, le θ indique la haute résolution spatiale, par exemple, A_θ . La basse résolution spatiale est notée *res1*, avec l comme indice correspondant, par exemple B_l . La fusion va permettre de synthétiser les modalités B_θ (tableau 3.5).

Chaque ensemble d'images A et B est composé d'une ou de plusieurs images acquises selon différentes modalités. On se place dans le cas où les images ont subi un traitement géométrique préalable pour être parfaitement superposables pixel à pixel. Dans l'ensemble B , l'image B_k désigne l'image acquise dans la bande spectrale (ou modalité) k . Par souci de simplicité, la plupart du temps, l'ensemble A sera réduit à une seule image que nous assimilerons à l'image Pan, à moins qu'il en soit spécifié autrement. Le problème pourra facilement être étendu au cas où A comprendrait plus d'une image ou dans le cas de plus de deux jeux d'images disponibles.

Tableau 3.5 : récapitulatif des notations.

<i>Introduction des notations</i>	<i>PLEIADES (exemple)</i>
$0, 1, 2, 3 \dots$: indices des résolutions spatiales.	
<i>res0, res1, res2, ...</i> : résolutions spatiales successives.	<i>res0</i> = 0.7 m, <i>res1</i> = 2.8 m, <i>res2</i> = 11.2 m, ...
<i>ratio</i> : rapport entre deux résolutions spatiales successives, avec <i>ratio</i> = <i>res1/res0</i> = <i>res2/res1</i> ...	<i>ratio</i> = 4
N : le nombre total de modalités MS.	$N = 4$
A_θ : modalité(s) de haute résolution spatiale.	A_θ = Pan et <i>res0</i> = 0.7 m
$(A_l)^{interp}_\theta$: modalité A à la résolution <i>res1</i> rééchantillonnée pour s'adapter à la taille de A_θ .	$(A_l)^{interp}_\theta$ = Pan à 2.8 m ramenée à la taille (H_1, H_2) de A_θ .
B_{kl} : $k^{\text{ième}}$ modalité de basse résolution spatiale.	B_{kl} = (ex) modalité bleue originale à 2.8 m.
$(B_{kl})^{interp}_\theta$: modalité B_{kl} rééchantillonnée pour s'adapter à la taille de la modalité A_θ .	$(B_{kl})^{interp}_\theta$: modalité bleue rééchantillonnée pour atteindre la taille (H_1, H_2) de la modalité A_θ .

B_{k0}^* : résultat de la fusion : $k^{\text{ième}}$ modalité synthétisée à la haute résolution.	$B_{k0}^* = B_{1,0}^*$ et $res0 = 0.7$ m
B_1 : ensemble des N modalités MS originales.	$B_1 = \{B_{1,1}, B_{2,1}, B_{3,1}, B_{4,1}\}$
$(B_1)^{interp}_0$: ensemble des N modalités MS originales rééchantillonnées.	$(B_1)^{interp}_0 = \{(B_{1,1})^{interp}_0, (B_{2,1})^{interp}_0, (B_{3,1})^{interp}_0, (B_{4,1})^{interp}_0\}$
B_0^* : ensemble des N modalités MS fusionnées.	$B_0^* = \{B_{1,0}^*, B_{2,0}^*, B_{3,0}^*, B_{4,0}^*\}$
$B_1(i, j)$: vecteur multispectral formé par les valeurs des N modalités MS originales au pixel (i, j) .	$B_1(i, j) = \{B_{1,1}(i, j), B_{2,1}(i, j), B_{3,1}(i, j), B_{4,1}(i, j)\}$
$(B_1)^{interp}_0(i, j)$: vecteur multispectral formé par les valeurs des N modalités MS originales rééchantillonnées au pixel (i, j) .	$(B_1)^{interp}_0(i, j) = \{(B_{1,1})^{interp}_0(i, j), (B_{2,1})^{interp}_0(i, j), (B_{3,1})^{interp}_0(i, j), (B_{4,1})^{interp}_0(i, j)\}$
$B_0^*(i, j)$: vecteur multispectral formé par les valeurs des N modalités MS fusionnées au pixel (i, j) .	$B_0^*(i, j) = \{B_{1,0}^*(i, j), B_{2,0}^*(i, j), B_{3,0}^*(i, j), B_{4,0}^*(i, j)\}$
$(B_1)^{ftm-1}$: ensemble des N modalités MS originales déconvoluées à basse résolution.	$(B_1)^{ftm-1} = \{(B_{1,1})^{ftm-1}, (B_{2,1})^{ftm-1}, (B_{3,1})^{ftm-1}, (B_{4,1})^{ftm-1}\}$
$((B_1)^{ftm-1})_0^{ftm}$: ensemble des N modalités MS originales déconvoluées à basse résolution puis convoluée à avec la FTM typique de $res0$.	$((B_1)^{ftm-1})_0^{ftm} = \{((B_{1,1})^{ftm-1})_0^{ftm}, ((B_{2,1})^{ftm-1})_0^{ftm}, ((B_{3,1})^{ftm-1})_0^{ftm}, ((B_{4,1})^{ftm-1})_0^{ftm}\}$

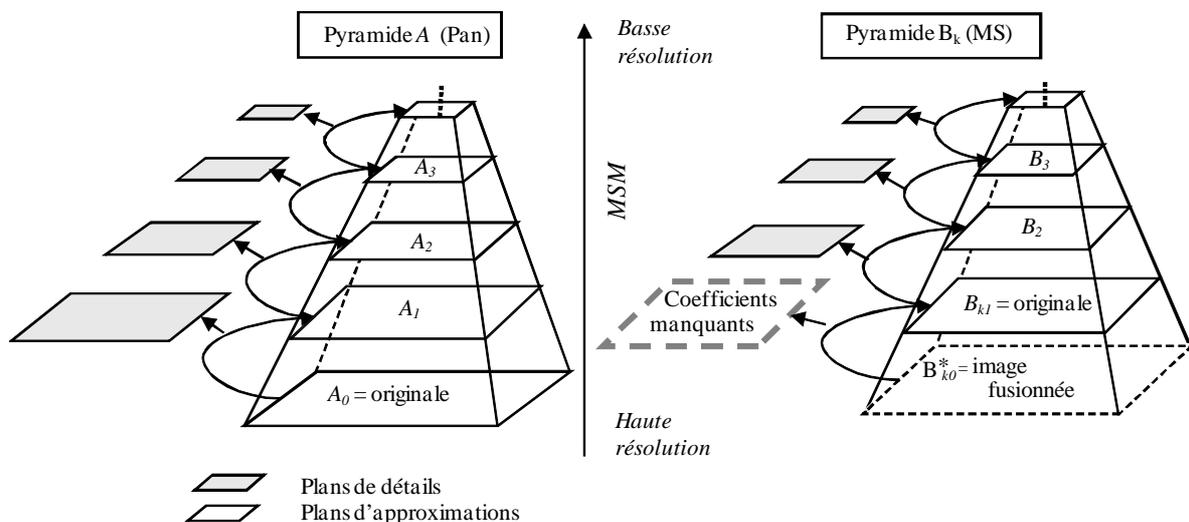
Dans l'ensemble du document pourront être utilisées les abréviations HR et BR pour haute et basse résolution spatiale correspondant respectivement à la résolution spatiale de l'image PAN et de l'image MS.

3.3 Le concept ARSIS

Le concept ARSIS qui signifie "Amélioration de la Résolution Spatiale par Injection de Structure" repose sur l'analyse multi-échelle. Ce concept est très général du fait qu'il ne spécifie pas d'outil mathématique ou informatique (Ranchin et Wald 2000). Ce concept est basé sur le principe que l'information manquante de l'image MS à basse résolution est uniquement composée d'informations hautes fréquences. L'information haute fréquence peut a priori être synthétisée à partir de toutes les données disponibles pour toutes les échelles et toutes les modalités.

Les techniques d'analyse multi-échelle permettent de décomposer l'image en fonction de la résolution spatiale. La décomposition hiérarchique de l'information peut être effectuée à partir d'outils mathématiques basés sur les espaces vectoriels ou de filtrages numériques (Ranchin et Wald 1992). En traitement d'image, la décomposition hiérarchique de l'information est souvent illustrée à l'aide de pyramide, (Mallat 1989) comme dans la figure 3.5, où les différents étages représentent l'image à différentes résolutions spatiales. La plus haute résolution spatiale se trouve en bas de la pyramide alors que la basse résolution se situe au sommet de la pyramide. Le bas de la pyramide MS, représenté en pointillés dans la figure 3.5, constitue les images MS-HR nécessitant l'information en haute fréquence dans les MS d'origine.

Cette information manquante est illustrée en pointillés en bas de la pyramide de l'image MS. Dans le cas général d'ARSIS, la synthèse de ces coefficients manquants est effectuée à l'aide de toutes les informations provenant de toutes les échelles et de toutes les modalités. Le but de la synthèse étant bien sûr d'injecter les informations hautes fréquences permettant d'obtenir l'image MS à haute résolution (synthétisée) qui soit la plus proche possible de l'image qu'aurait acquis le satellite s'il possédait la haute résolution spatiale (Ranchin et Wald 2000).



Le fonctionnement du concept ARSIS est constitué de trois classes de modèles qui sont mis en jeux à travers quatre opérations. Les différents modèles sont :

- **MSM** : Multi Scale Model (modèle multi-échelle). Ce modèle permet d'effectuer la décomposition hiérarchique de l'information de l'image. Les principales transformées utilisées pour établir ce modèle sont : la transformée en Pyramide Laplacienne Généralisée (Aiazzi *et al.* 2002), la transformée en ondelettes (Mallat 1989), les bancs de filtres (Blanc *et al.* 1998) et récemment les curvelets (Garzelli *et al.* 2005)(Wang *et al.* 2007) et *contourlets* (Shah *et al.* 2008) (Do et Vetterli 2005). Le MSM fournit plusieurs plans de détails et un plan d'approximation qui contiendra les informations basse fréquence de l'image. Une propriété nécessaire du MSM est que celui-ci doit être inversible. C'est par cette propriété que la synthèse de l'information à haute résolution spatiale est rendue possible.
- **IMM** : Inter Modality Model. Ce modèle établit une relation qui permet de passer de l'image d'une modalité à celle d'une autre modalité et *vice versa*. Ces modèles sont le plus souvent basés sur des propriétés statistiques des images comme : la moyenne, la variance.... Dans certains cas, des structures peuvent être présentes uniquement dans certaines modalités et invisibles dans d'autres. (Thomas 2006) recense des IMM globaux, c'est-à-dire que le même modèle s'applique à toute l'image, et des IMM locaux, c'est-à-dire que le modèle est défini localement.
- **HRIMM** : High Resolution Inter-Modality Model. Le rôle de ce modèle est d'estimer les coefficients à haute résolution pour une modalité différente de celle dont ils proviennent. Il arrive fréquemment que l'IMM et le HRIMM soient assez proches, voire identiques. Le HRIMM est l'occasion de faire un traitement local pour améliorer le contraste (Thomas 2006) ou aussi de faire une amélioration du SNR au niveau des contours (Lee et Lee 2010).

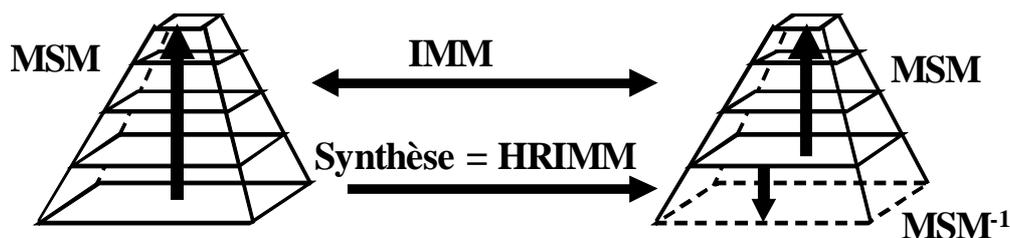


Figure 3.5 : ARSIS, les trois classes de modèles mis en œuvre. (Thomas 2006)

Les modèles intervenant dans le concept ARSIS sont maintenant définis. Les interactions entre ces modèles impliquent une chronologie dans leur mise en œuvre (figure 3.5). Le processus de fusion/synthèse débute par la décomposition hiérarchique de l'information pour les images A et B_k par le MSM. Cette décomposition crée les plans de détails et les plans d'approximations pour ces images et va permettre d'établir le modèle IMM.

Une fois l'IMM établi, il devient possible de calculer le HRIMM. Le HRIMM permet d'estimer les coefficients hautes fréquences manquant dans la pyramide MS. L'étape finale applique l'inverse du MSM qui permet de synthétiser l'image à partir des plans de détails et d'approximations notamment le plan de détails haute fréquence ajouté à la pyramide B_k .

3.4 La méthode de fusion de référence : ATWT-M3

Nos propositions d'amélioration sont mises en pratique à partir de la méthode ATWT-M3. Il est nécessaire de décrire son principe qui se fera à travers la description des trois modèles : MSM, IMM et HRIMM qui la constituent. La méthode ATWT-M3 présentée dans (Ranchin et Wald 2000). Celle-ci est basée sur une analyse multirésolution utilisant des ondelettes à trous, et son nom est l'association des noms du MSM, du couple IMM-HRIMM. Ces deux modèles étant indépendants l'un de l'autre, nous décrivons tout d'abord le modèle ATWT, puis le modèle M3.

3.4.1 ATWT

À Trous Wavelet Transform (ATWT) ou en français transformée en ondelettes à trous. Cette transformée permet de faire une analyse multi-échelle à l'aide d'ondelette. Une propriété importante de ce modèle ou transformée est que celle-ci doit être inversible ce qui est une condition nécessaire pour la fusion. L'utilisation d'ondelette à trous, c'est-à-dire sans sous échantillonnage est plus appropriée à la fusion comme cela a été montré par (Aiazzi *et al.* 2002).

Les ondelettes sont un des outils qui permettent d'effectuer une analyse multi échelle. Les ondelettes à trous sont une version d'ondelette pour lesquelles il n'y a pas de sous-échantillonnage durant la décomposition. (Aiazzi *et al.* 2002) et (Otazu *et al.* 2005) ont montré que les ondelettes non décimées étaient plus appropriées à la fusion d'image que les ondelettes décimées. La transformée en ondelette de la fonction $f(x)$ s'exprime de la façon suivante :

$$ATWT_f(a, b) = \langle f, \Psi_{a,b} \rangle = \frac{1}{\sqrt{|a|}} \int_{-\infty}^{+\infty} f(x) \overline{\Psi\left(\frac{x-b}{a}\right)} \partial x \quad 3.2$$

Où $\Psi_{a,b}$ est la fonction d'ondelette et $\overline{\Psi\left(\frac{x-b}{a}\right)}$ est le complexe conjugué de Ψ
La transformée inverse de l'ondelette à trous s'exprime de la façon suivante:

$$f(x) = \frac{1}{C_\Psi} \iint_{-\infty}^{+\infty} ATWT_f(a, b) \Psi_{a,b}(x) \frac{\partial a \partial b}{a^2} \quad 3.3$$

$ATWT_f(a, b)$ représente les coefficients d'ondelettes, et C_Ψ est la condition d'admissibilité de l'ondelette mère.

Les ondelettes sont fréquemment utilisées pour faire de l'extraction d'information et ce pour du signal (1D) ou dans l'image (2D). Dans le cas d'images satellitales, les ondelettes permettent d'extraire les informations en fonction de la résolution. Même si ces ondelettes n'ont pas la finesse d'analyse d'une transformée de Fourier, leurs capacités de conserver la spatialisation de l'information en font un outil d'analyse efficace pour des images issues de paysage naturel et non stationnaire.

La décomposition de l'information lorsqu'elle est réalisée à l'aide de l'analyse par ondelettes, se fait avec des rapports de résolution le plus souvent entiers multiples de 2, par

décomposition dyadique. La transformée en ondelette est un processus itératif. La première décomposition va séparer dans un rapport de deux les hautes et les basses fréquences. Les basses fréquences sont conservées dans une image dont la résolution spatiale est dégradée par rapport à l'image originale. Cette image est complétée par les coefficients d'ondelettes qui représentent les structures à hautes fréquences de l'image. Ce processus va ainsi se reproduire sur les basses fréquences avec une autre ondelette issue de l'ondelette "mère".

3.4.2 IMM et HRIMM, Modèle : M3

L'IMM est le modèle qui permet de passer de la modalité A à la modalité B et *vice versa*. Le HRIMM est le modèle qui permet d'injecter les coefficients d'ondelettes représentant la haute résolution de modalité A dans la modalité B qui ne dispose pas de la haute résolution. Dans le cas du modèle M3, comme dans la plupart des méthodes dites globales, le HRIMM est identique à l'IMM. Les rôles de ces deux modèles sont illustrés dans le paragraphe 3.3 qui présente le concept ARSIS ainsi que dans la figure 3.6 qui illustre le modèle M3.

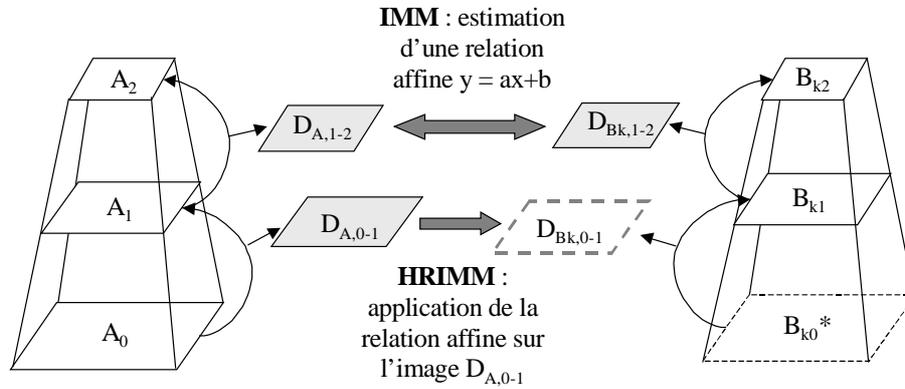


Figure 3.6 : schéma récapitulatif du modèle M3. (source thomas 2006)

Le principe du modèle M3 est d'estimer, par une relation affine, la transformation qui permet de passer des coefficients d'ondelettes de modalité A à ceux de la modalité B et vice-versa. Le coefficient « a » pouvant être calculé de deux façons différentes, soit par la minimisation de l'axe d'inertie (Ranchin et Wald 2000) (Ranchin *et al.* 2003) :

$$a = \frac{\text{var}(D_{Bk,1-2}) - \text{var}(D_{Ak,1-2}) + \sqrt{(\text{var}(D_{Bk,1-2}) - \text{var}(D_{Ak,1-2}))^2 + 4\text{cov}^2(D_{Bk,1-2}, D_{Ak,1-2})}}{2\text{cov}(D_{Bk,1-2}, D_{Ak,1-2})} \quad 3.4$$

où *var*, *covar*, *m* correspondent aux opérateurs variance, covariance et moyenne.

soit par la méthode des moindres carrés (Ranchin et Wald 2000):

$$a = \text{cov}(D_{Bk,1-2}) / \text{var}(D_{Ak,1-2}) \quad 3.5$$

Le coefficient « b » assure, quant à lui, de ne pas introduire de biais. Il est obtenu de la façon suivante :

$$b = m(D_{Bk,1-2}) - a m(D_{A,1-2})$$

La synthèse de la haute résolution spatiale de la modalité B s'effectue en appliquant le HRIMM aux coefficients d'ondelettes qui correspondent à la haute résolution de la modalité A. Ce qui s'exprime de la façon suivante.

$$D_{B_{k,0-1}} = a \times D_{A_{0-1}} + b \quad 3.6$$

Une fois le HRIMM appliqué, on obtient l'image MS synthétisée à haute résolution spatiale par le MSM inverse.

$$(B_{k,1})^* = \frac{1}{C_\psi} \iint_{-\infty}^{+\infty} \left[ATWT_{(B_{k,1})}(a,b) + (a \times ATWT_{(A_{k,0-1})} + b) D_{B_{k,0-1}} \right] \Psi_{a,b}(x) \frac{\partial a \partial b}{a^2} \quad 3.7$$

$ATWT_{(B_{k,1})}$ représente les coefficients d'ondelettes de l'image MS et $ATWT_{(A_{k,0-1})}$ représente les coefficients d'ondelettes des hautes fréquences de l'image PAN, et C_ψ est un terme de normalisation.

Un bilan sur la qualité visuelle typique de cette méthode (ATWT-M3) est établi dans (Thomas 2006) : « les images issues de cette méthode de fusion comportent un léger flou homogène sur toute l'image. On observe une bonne similitude des couleurs comparées aux MS d'origine, mais elles apparaissent souvent un peu ternes. Les hautes radiométries sont souvent entourées d'un halo clair. La perception des petits objets est bonne, même si la couleur de ces petits objets a tendance à diffuser sur les pixels voisins. On note une légère atténuation des objets dans les ombres. »

3.5 Filtres de décomposition et reconstruction des ondelettes à trous

L'implémentation réalisée d'ATWT-M3 utilise des ondelettes à trous pour réaliser la décomposition de l'information. Les ondelettes sont utilisées ici comme analyse multi-échelle.

Néanmoins, le fait d'utiliser des ondelettes à trous semble assez générique par rapport à l'analyse multiéchelle proposée par (Aiazzi *et al.* 2006) qui utilise des bancs de filtres adaptés à la FTM des images dans le cas particulier d'un satellite.

Un avantage à utiliser des ondelettes est qu'il peut devenir plus aisé « d'ajouter » des traitements couramment implémentés avec des ondelettes qui pourraient être effectués en même temps que le processus de fusion, comme du débruitage (Krim *et al.* 1999), (Mallat 1999). Les ondelettes utilisées donnent dans le cas d'une décomposition de profondeur 2 les filtres de décomposition et de reconstruction dont les gabarits sont présentés dans la figure 3.7.

Pour vérifier le caractère générique de ces ondelettes, nous les avons comparées aux ondelettes à trous disponibles dans la wavelet toolbox matlab de Gabriel Peyré (disponible sur matlab central : <http://www.mathworks.com/matlabcentral/fileexchange/5104>). Les gabarits des filtres des ondelettes que nous utilisons sont proches de ceux obtenus avec les ondelettes de Peyré. Les toolbox d'ondelettes de matlab et de wavelab 850 (université de Stanford) ne fournissent pas d'ondelettes à trous.

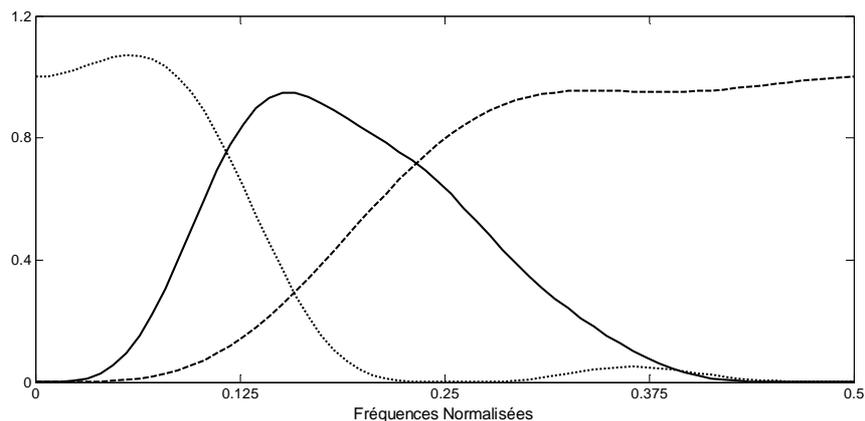


Figure 3.7 : Fonction de transfert des filtres de décomposition de l'ondelette à trous, profondeur 2.

Comme le montre la figure 3.7, les ondelettes utilisées constituent des filtres peu sélectifs en fréquences et peu précis vis-à-vis du rapport de décomposition de $1/4$. Dans la figure 3.7, on peut voir que la bande passante du filtre passe-bas s'étend au-delà de $0,125 f_e$. Le MSM utilisé n'est en aucun cas adapté aux caractéristiques de la FTM de l'instrument d'observation, contrairement aux travaux de (Aiazzi *et al.* 2006)(Lee et Lee 2010) (Khan *et al.* 2009).

La figure 3.8 présente un MSM dont les filtres sont plus sélectifs que ceux des ondelettes que nous utilisons. Ce type de MSM est plus sélectif ou précis en termes de fréquence, car peu d'informations sont conservées au-delà de $0,125 f_e$, ce qui implique une atténuation des fréquences qui se trouvent juste avant $0,125 f_e$ et qui contiennent de l'information spectrale sans forcément contenir de repliement de spectre. Le rapport de résolution de 4 entre les images PAN et MS utilisées dans notre travail fait qu'il ne devrait « théoriquement » pas y avoir d'information au-delà de $0,125 f_e$ dans l'image MS. Dans la pratique on peut voir dans la figure 2.13 que l'image MS contient de l'information au-delà de $0,125 f_e$.

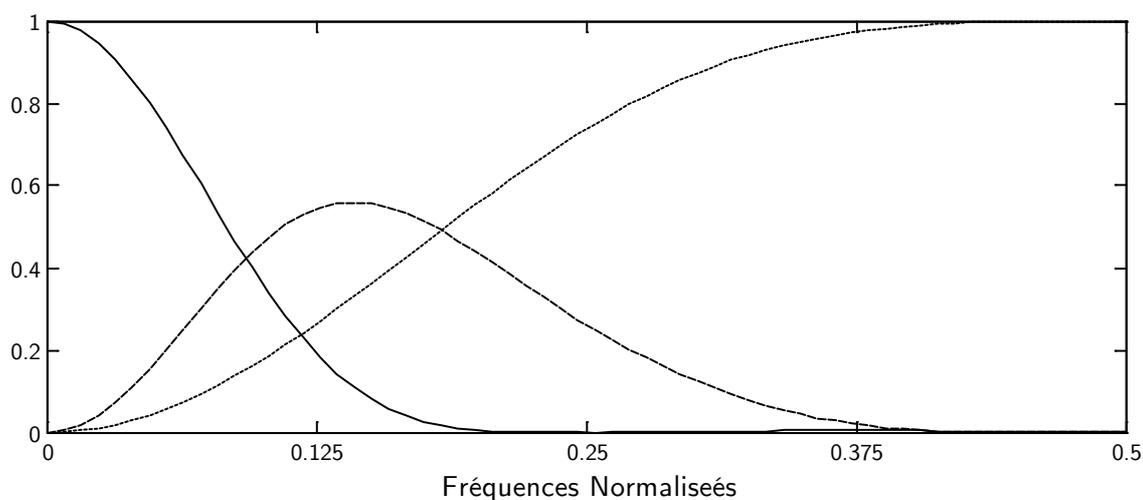


Figure 3.8 : Fonction de transfert MSM plus sélectif en fréquence.

Utiliser une analyse multiéchelle plus sélective et précise au niveau de 0,125 fe peut entraîner l'atténuation/suppression de certaines informations contenues dans les plus hautes fréquences de l'image MS. Ceci peut avoir des conséquences à la fois positives et négatives. En effet, le fait de supprimer ces plus hautes fréquences peut atténuer l'effet de certains artefacts de repliement de spectre. Par contre, ces plus hautes fréquences ne contiennent pas nécessairement de repliement de spectre. Dans ce cas, le fait d'utiliser un filtre sélectif mais non discriminant vis-à-vis du repliement de spectre ne permet pas forcément d'obtenir la meilleure qualité spectrale possible.

Le fait d'utiliser un filtre peu sélectif comme les ondelettes à trous a des conséquences inverses. Il assure la meilleure qualité spectrale possible, mais ne permet pas d'atténuer des artefacts d'aliasing.

3.6 Outils

3.6.1 Protocole d'évaluation de la qualité

Il existe de nombreuses méthodes de fusion aux caractéristiques variées. Certaines méthodes permettant d'obtenir des distorsions spectrales minimales, mais dont les caractéristiques spectrales sont perfectibles. D'autres au contraire, peuvent fournir une meilleure qualité spatiale, mais avec de plus grandes distorsions spectrales. Ainsi, selon le traitement que l'utilisateur souhaite effectuer, il est bon qu'il puisse avoir des informations sur les images qui seront synthétisées.

L'utilisateur de ces méthodes de fusion doit pouvoir se faire une idée des performances de la méthode ainsi que la qualité de l'image synthétisée. Établir un cadre « objectif » à l'évaluation de la qualité n'est pas chose aisée. De nombreux auteurs de méthodes de fusion privilégient les critères de qualité pour lesquels leur méthode est la plus performante. Homogénéiser ce travail et donner un cadre global à l'évaluation de la qualité dans le cas de la fusion PAN/MS représente une bonne partie des travaux de thèse de Claire Thomas (Thomas 2006) dont les travaux se sont basés sur des travaux de référence déjà existants (Wald *et al.* 1997) et (Li 2000).

Pour évaluer la performance de la méthode de fusion, la démarche couramment utilisée en fusion d'image est de comparer l'image synthétisée avec une image de référence. Dans la réalité/pratique, il n'y a pas d'image de référence. Des solutions ont été proposées (Wald *et al.* 1997) et discutées brièvement. Notre cas est cependant très favorable, le processus de simulation permettant de disposer des images de référence.

Le protocole d'évaluation de la qualité mis en place par (Thomas 2006) établit deux propriétés à respecter :

- La « propriété de synthèse » : l'image synthétisée doit être la plus proche possible de ce que serait l'image MS si le satellite possédait la haute résolution spatiale.
- La « propriété de cohérence » implique qu'après dégradation de la résolution de l'image synthétisée à la résolution originale l'image dégradée doit être la plus proche

possible de l'image MS originale. Cette propriété peut paraître triviale, pourtant cette propriété n'est pas respectée par de nombreuses méthodes.

Le déroulement du protocole d'évaluation de la qualité de (Thomas 2006) se trouve légèrement modifié du fait que nous disposons d'une image de référence à haute résolution. Les étapes sont les suivantes :

Étape 1 : la méthode de fusion est appliquée aux deux ensembles d'images initiaux : A_0 et B_1 . Il en résulte ainsi un nouveau jeu d'images B_0^* . Les images B_0^* sont comparées aux images originales B_0 par une évaluation visuelle de la qualité, suivie par le calcul de distances monomodales visant à comparer chacun des termes $B_{k_0}^*$ et B_{k_0} , puis multimodales avec la comparaison multidimensionnelle de B_0^* avec B_0 . Les résultats obtenus pour ces distances sont comparés aux critères définis pour chaque distance.

Étape 2 : vérification de la propriété de cohérence. L'ensemble B_0^* est dégradé jusqu'à la résolution 1 $(B_{k_0}^*)_1$. On analyse visuellement l'écart entre $(B_{k_0}^*)_1$ et B_{k_1} pour chaque modalité k , les images sont en niveaux de gris. De la même façon, on compare les images couleurs $(B_0^*)_1$ et B_1 . Enfin, on mesure la qualité monomodale et multimodale grâce à des distances adaptées. Les résultats obtenus pour ces distances sont comparés aux critères définis pour chaque distance.

3.6.1.1 Propriété de synthèse

La propriété de synthèse stipule que : « **toute image synthétique B_0^* doit être aussi proche que possible de celle qu'observerait le capteur correspondant s'il possédait la haute résolution spatiale. Cette propriété doit être vérifiée pour les modalités prises individuellement ainsi que pour toutes les modalités confondues** » (Thomas 2006).

La propriété de synthèse exprime le but que nous cherchons à atteindre par la fusion, c'est-à-dire une image qui serait issue d'un capteur MS à haute résolution spatiale.

L'évaluation du respect de cette propriété est une façon de quantifier les performances de l'algorithme de fusion. Nous présentons ici comment est évalué le respect de cette propriété. Cette évaluation passe par la comparaison de l'image de référence avec l'image synthétisée. Cette comparaison est quantifiée à l'aide de nombreuses distances de qualité image monomodales et multimodales utilisées couramment en fusion d'image (Nikolakopoulos 2008) (Alparone *et al.* 2007) (Thomas 2006) (Wald 2002) (Li 2000) (Wald *et al.* 1997).

Ces distances sont classées en deux catégories :

- Les distances monomodales : comparent l'image de référence et l'image synthétisée, bande à bande, uniquement. Ce type de distance renseigne sur la qualité spatiale, radiométrique des images.

- Les distances multimodales : comparent de façon conjointe, l'ensemble des bandes de l'image de référence et l'image synthétisée. Ce type de distance renseigne sur la qualité spectrale des images.

Ces distances seront présentées dans la suite lors de la présentation de la plate-forme d'évaluation.

3.6.1.2 Propriété de cohérence

La propriété de cohérence veille à ce que l'augmentation de la résolution spatiale des images MS soit un processus réversible, c'est-à-dire que par la dégradation de la résolution spatiale de l'image synthétisée on retrouve l'image MS originale. Cette propriété est exprimée de la façon suivante : « **toute image synthétique $B_{k_0}^*$, une fois dégradée à sa résolution spatiale originale l , doit être aussi proche que possible de l'image originale B_{k_1} . Cette propriété doit être vérifiée pour les modalités prises individuellement ainsi que pour toutes les modalités confondues.** » (Thomas 2006)

L'évaluation du respect de cette propriété se fait de la même façon que pour la propriété de synthèse, en comparant l'image MS originale et l'image synthétisée pour laquelle sa résolution spatiale a été dégradée à la résolution spatiale initiale.

Si lors de la fusion et de l'augmentation de la résolution, la FTM a été prise en compte, il devient alors nécessaire d'en tenir compte aussi lors de la diminution de la résolution spatiale. Pour ce faire, il faut appliquer à l'image un filtre dont l'allure est proche de la FTM de l'image à basse résolution, puis appliquer le sous-échantillonnage.

3.6.1.3 Plate-forme d'évaluation de la qualité ENVI-IDL

Une plate-forme d'évaluation de la qualité donnant un vaste choix de distance a été mise en place par (Thomas 2006), développé sur la plate-forme ENVI-IDL. C'est à l'aide de cette plate-forme qu'ont été évaluées les méthodes de fusion lors du « Data-fusion Contest 2007 » dont les résultats ont été publiés par (Alparone *et al.* 2007).

Cet outil étant validé et étant une référence c'est donc tout naturellement que nous l'avons utilisé pour nos travaux. L'ensemble des distances disponibles avec cette plate-forme sont présentées dans le tableau 3.6.

Tableau 3.6 : Distances proposées par l'outil d'évaluation de la qualité

Distances monomodales	Distances multimodales
Bias	ERGAS
Relative bias in %	RASE
Difference of variance	Mean SAM
Relative difference of variance in %	Relative bias in norm of spectra
Standard deviation	Standard deviation in norms of spectra
Relative standard deviation in %	Mean value of the resultant vector
RMSE	
Min Max	
Covariance	
Correlation	
Percentile 5 %	
Percentile 95 %	
Universal quality Index (Q)	
Entropy	
Correlation between High frequencies	

- Distances monomodales :

$$\text{Biais : } \text{biais}(B_k, B_k^*) = m(B_k) - m(B_k^*) \quad 3.8$$

$$\text{Relative biais en \% : } \text{biaisrelatifenpourcents} = \frac{100(m(B_k^*) - m(B_k))}{m(B_k)} \quad 3.9$$

$$\text{Différence de variance : } \text{diffVar} = \text{var}(B_k) - \text{var}(B_k^*) \quad 3.10$$

$$\text{Différence de variance relative en \% : } \text{RelDiffVar} = \frac{100(\text{var}(B_k) - \text{var}(B_k^*))}{\text{var}(B_k)} \quad 3.11$$

$$\text{Ecart type : } \sigma_{(B_k - B_k^*)} = \sqrt{\frac{1}{n} \sum_{i=1}^n (x(B_k - B_k^*)_i - \overline{(B_k - B_k^*)})^2} \quad 3.12$$

$$\text{Ecart type relatif en \% : } \sigma_{rel} = \frac{100(\sigma_{(B_k - B_k^*)})}{\sigma_{(B_k)}} \quad 3.13$$

$$\text{RMSE (root mean square error): } \text{RMSE}(B_k, B_k^*) = \frac{1}{nbPix} \sqrt{\sum (B_k(i, j) - B_k^*(i, j))^2} \quad 3.14$$

$$\text{Min : } \min(B_k - B_k^*) \quad 3.15$$

$$\text{Max : } \max(B_k - B_k^*) \quad 3.16$$

$$\text{Covariance : } \sigma_{xy} = \sum_{i=1}^n \sum_{j=1}^m x_i y_j - \bar{x} \bar{y} \quad 3.17$$

Où \bar{x} et \bar{y} sont les moyennes de x et y.

$$\text{Corrélation : } \text{cc}(B_k, B_k^*) = \frac{\sigma_{B_k, B_k^*}}{\sigma_{B_k} \sigma_{B_k^*}} \quad 3.18$$

Où σ_{xy} est la covariance entre x et y

Percentile 5 % : est la valeur qui sépare les 5 % inférieurs des données

Percentile 95 % : est la valeur qui sépare les 95 % inférieurs des données

Universal quality Index (Q) :

$$Q_k(B_k, B_k^*) = cc(B_k, B_k^*) \times \frac{2M(B_k)M(B_k^*)}{M(B_k)^2 + M(B_k^*)^2} \times \frac{2\sigma(B_k)\sigma(B_k^*)}{\sigma(B_k)^2 + \sigma(B_k^*)^2} \quad 3.19$$

Entropie : $H(x) = \sum_{i=1}^n P_i \log_2(P_i)$

Corrélation entre les hautes fréquences :

Cette mesure est identique au coefficient de corrélation présenté plus haut, sauf que cette corrélation est effectuée entre les moitiés supérieures des fréquences, dont l'allure du filtre est présentée dans la figure 3.9

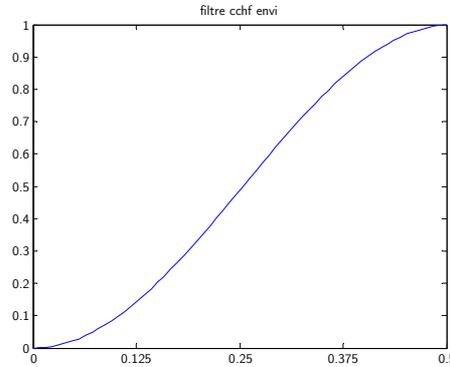


Figure 3.9 : filtre passe haut utilisé pour la corrélation entre les hautes fréquences.

- Distances multimodales:

ERGAS : Erreur relative globale adimensionnelle de synthèse (Wald 2002)

$$ERGAS(B_k, B_k^*) = \frac{100}{rario} \sqrt{\frac{1}{N} \sum_{k=1}^N \left(\frac{RMSE(B_k, B_k^*)}{m(B_k)} \right)^2} \quad 3.20$$

RASE : Relative Average Spectral Error

$$RASE = \frac{100}{m} \sqrt{\frac{1}{N} \sum_{k=1}^N (RMSE(B_k, B_k^*))^2} \quad 3.21$$

M correspond à la moyenne des N modalités

SAM : Spectral angle Mapper

$$SAM(B(i, j), B^*(i, j)) = \cos^{-1} \left(\frac{\langle B^*(i, j), B(i, j) \rangle}{\|B^*(i, j)\|_2 \cdot \|B(i, j)\|_2} \right) \quad 3.22$$

Biais relatif des normes des vecteurs spectraux :

$$diffNorms(i, j) = \frac{\|B(i, j)\| - \|B^*(i, j)\|}{\|B(i, j)\|}$$

Ecart type normes des vecteurs spectraux :

$$\sigma_{(\|B(i,j)\|-\|B^*(i,j)\|)} = \sqrt{\frac{1}{n} \sum_{i=1}^n (x(\|B(i,j)\| - \|B^*(i,j)\|) - \overline{(\|B(i,j)\| - \|B^*(i,j)\|)})^2} \quad 3.23$$

Vecteur résultant- V_{res} :

$$V_{res}(i,j) = \|B(i,j) - B^*(i,j)\| \quad 3.24$$

Certaines de ces distances sont redondantes. Selon les recommandations de (Thomas 2006), nous établirons des bilans de qualité qui comprendront les distances :

- biais relatif (biaisRel (Bk, Bk*)),
- la différence en variance relative (diffVarRel (Bk, Bk*)),
- l'écart-type de l'image de différence (σ Rel (Bk, Bk*)),
- le coefficient de corrélation (cc (Bk, Bk*)),
- coefficient de corrélation des hautes fréquences

ainsi que les distances multimodales :

- ERGAS
- Mean SAM

3.6.2 Outils de mesure du repliement de spectre et de la netteté

Dans nos travaux, nous considérons la prise en compte de la FTM et l'impact du repliement de spectre présent dans les images MS dans l'image MS synthétisée à haute résolution spatiale.

Il est apparu intéressant de voir si des mesures ou détecteurs spécifiques au repliement de spectre et de la netteté permettent de quantifier l'apport visuel de nos travaux. Dans cette partie, nous présentons tout d'abord ces détecteurs basés sur la variation totale qui ont été proposés par (Blanchet 2006). Ces détecteurs présentent un avantage fort du fait qu'il n'y a pas besoin d'image de référence comme le QNR. Ceci est un avantage important dans le domaine de la fusion pour lequel l'image de référence n'est pas disponible, hors cas de simulation. Dans notre cas, où nous cherchons à évaluer l'amélioration due à la prise en compte de l'aliasing ou de la FTM, nous comparerons les images qui auront été synthétisées avec ou sans prise en compte de ces caractéristiques. Puis nous présenterons les résultats de ces mesures qui nous permettront de mettre en évidence de façon quantitative l'amélioration de la qualité des images liée à la netteté ou au taux de repliement.

3.6.2.1 Quantification du repliement

Pour améliorer la qualité de l'image synthétisée, nous supprimons le repliement de spectre dans l'image MS à basse résolution. La fusion n'assure pas de préserver l'amélioration de la qualité obtenue par la suppression du repliement de spectre contenu dans l'image. Le bilan de qualité entre l'image synthétisée et l'image de référence permet de dire quelle image est la plus ressemblante à l'image de référence au vu des mesures effectuées. Le bilan de qualité ne contient pas d'indicateurs spécifiques sur la quantité de repliement de spectre présent dans l'image. Il est apparu nécessaire, en plus du bilan de qualité, de mesurer si la diminution du taux de repliement est conservée par le processus de fusion, et si celle-ci peut être mesurée.

Plusieurs outils ont été mis en place par (Blanchet 2006) pour détecter le repliement de spectre. Nous utiliserons celui qui semble le plus performant au regard des tests présentés dans sa thèse. Ce dernier a d'ailleurs de multiples avantages ; il s'effectue dans le domaine spatial, ne nécessite pas d'image de référence, ni de seuil arbitraire à partir duquel on dit qu'il y a du repliement.

- Description du détecteur :

Cet outil est basé sur la variation totale qui est l'intégrale de la valeur absolue de la dérivée de l'image, et ce, dans les deux directions. L'expression de la variation totale pour une fonction dérivable s'exprime de la façon suivante :

$$TV(f) = \int \sqrt{\left| \frac{df(x,y)}{dx} \right|^2 + \left| \frac{df(x,y)}{dy} \right|^2} dx dy \quad 3.25$$

Selon (Blanchet 2006) : « une image qui contient des transitions brutales a plus de chance de ne pas avoir été échantillonnée dans les conditions de Shannon ». Une image dont

l'échantillonnage ne respecte pas les conditions du théorème de Shannon peut ainsi avoir un niveau de FTM élevé pour des fréquences proches de la fréquence de Nyquist. Ce sont ces conditions qui permettent de représenter des transitions brutales avec un fort contraste. Appliquer une translation subpixelique à une image contenant des transitions brutales va engendrer l'apparition de ringing au niveau de ces dernières (figure 3.10).

Au contraire, une image respectant l'échantillonnage de Shannon aura une FTM nulle pour la fréquence de Nyquist. Il est difficile d'avoir de forts contrastes pour des fréquences élevées dans une image dont l'acquisition respecte les conditions du théorème de Shannon. De ce fait, les hautes fréquences présentes dans l'image auront un faible niveau de contraste, ce qui entraîne des transitions que l'on peut qualifier de « molles ». Au contraire si l'on applique une transition subpixelique à cette transition molle, l'apparition de ringing sera faible.

En appliquant une translation subpixelique, une étape d'interpolation est mise en jeu. Celle-ci va donc engendrer du ringing dans les zones présentant des transitions brusques. Une image respectant Shannon ne générera que très peu de ringing. Plus l'apparition de ringing est forte, plus le niveau de FTM est élevé à la fréquence de Nyquist. Le niveau « élevé » de la FTM est une condition nécessaire pour avoir du repliement. L'apparition de « ringing » dans l'image est donc un indicateur sur les paramètres d'échantillonnage. En pratique, la détection du repliement est obtenue par tirage de Monte-Carlo du déplacement subpixelique en évaluant :

$$G_\varepsilon(u) = -\log_{10}(\mathbb{P}(TV(U_t) \leq TV(u))) \quad 3.26$$

où :

- \mathbb{P} est l'opérateur de probabilité
- u est l'image
- U_t est la série d'images translattées, où les translations correspondent à des tirages aléatoires uniformes sur $[0, 1]$
- TV est l'opérateur de variation totale.

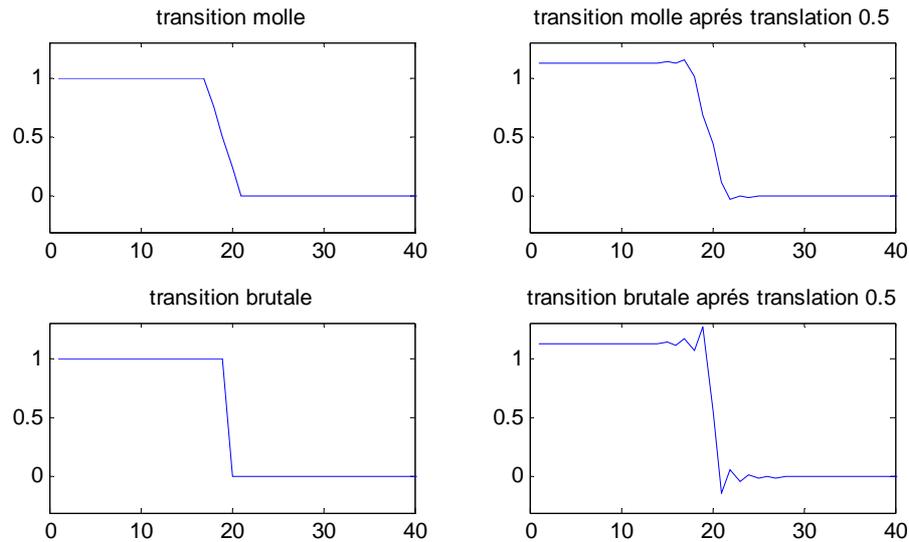


Figure 3.10 : Apparition du ringing selon le type de transition molle/brutale. *In* (Blanchet 2006)

Le ringing fait augmenter la variation totale de l'image puisque cette variation est l'intégrale de la valeur absolue de la dérivée (selon les lignes et les colonnes). En calculant la différence de variation totale entre l'image originale et l'image translatée, nous aurons alors une estimation de l'apparition du ringing dans l'image. Comme nous l'avons dit précédemment, le ringing est lié au fait que l'image n'a pas été bien échantillonnée. Cette différence de variation totale est donc un indicateur des conditions d'échantillonnage et sur le repliement de spectre présent dans l'image.

3.6.2.2 Quantification de la netteté

La caractéristique de FTM de l'image est étroitement liée à la netteté de celle-ci. La prise en compte de la FTM aura un effet sur la netteté de l'image. Nous avons étudié si la prise en compte de la FTM pouvait être quantifiée à travers cette quantification de la netteté.

De la même façon que pour l'indice précédent, celui-ci est basé sur la variation totale. Pour mesurer la netteté de l'image, l'idée est de quantifier grâce à la variation totale, la sensibilité de l'image à un déphasage aléatoire. On note U_ε la série d'images déphasées où le déphasage est issu d'un tirage aléatoire uniforme dans Fourier et u l'image originale. La cohérence de phase est définie par (Blanchet 2006) :

$$F_\varepsilon(u) = -\log_{10}(\mathbb{P}(TV(U_\varepsilon) \leq TV(u))) \quad 3.27$$

Où \mathbb{P} et TV sont les opérateurs de probabilité et de variation totale.

En pratique, le test de cohérence de phase est obtenu en effectuant une perturbation de phase sur la transformée de Fourier de l'image testée. Le terme de perturbation de phase est obtenu par tirage de Monte Carlo. La cohérence de phase est très importante autour d'une transition brutale (figure 3.11-a). Une telle transition peut être considérée comme une somme de sinusoides, lesquelles sont en phase. Dans ce cas, une petite variation de phase va

entraîner l'apparition d'oscillations, comme illustré dans la figure 3.11-c. L'apparition de ces petites oscillations entraîne une augmentation de la variation totale de l'image.

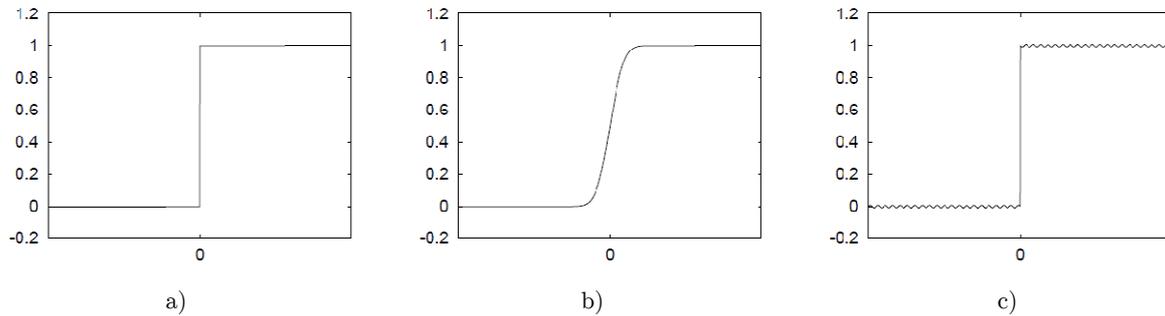


Figure 3.11 : Illustration de la sensibilité de phase d'une transition brusque.
a) transition brutale, b) transition molle, c) exemple pour le déphasage d'une sinusoïde pour la transition brutale. *In* (Blanchet 2006)

Une image très nette va voir sa variation totale augmenter, ce qui entraîne :

$$\mathbb{P}(TV(u_\varepsilon) \leq TV(u)) \simeq 0 \quad 3.28$$

D'où $F_\varepsilon(\mathbf{U}) = \infty$

Plus une image est nette et plus sa cohérence de phase sera importante et plus la valeur de F_ε sera grande.

L'adaptation de FTM a pour but d'améliorer la qualité des images fusionnées notamment au niveau de la qualité géométrique. La qualité de la géométrie dans l'image passe par la netteté des contours. La netteté est liée à la représentation des contrastes et à la cohérence de phase au niveau de ces contours.

L'adaptation de FTM améliore les contrastes par la prise en compte de paramètre de l'instrument d'acquisition. Avec cet indice, nous allons pouvoir quantifier si l'adaptation de FTM permet une meilleure cohérence de phase dans l'image ou si le rehaussement de contraste est relativement bien en phase.

3.6.3 Distance MTF_{dev}

Pour l'ensemble des méthodes de fusion, l'image synthétisée contiendra des informations provenant d'images ayant des FTM différentes. Dans le cas d'une fusion de type ARSIS, le produit fusionné aura pour les basses fréquences l'allure de la FTM de l'image MS. Pour les hautes fréquences, l'image aura l'allure de la FTM de l'image PAN aux modifications de l'HRIMM près.

Il existe dans la littérature quelques indices pour caractériser la FTM (Thomas et Wald 2006) (Lei et Tiziani 1989) (Li 2000). La caractérisation de la FTM est obtenue par l'analyse d'une zone qui présente un fort contraste (edge spread function, ESF). Le fait d'analyser la FTM de façon locale à travers l'ESF n'assure pas d'analyser de manière globale la contribution de la FTM de l'image MS et de l'image PAN.

Nous avons mis en place un outil qui permet de comparer la FTM entre deux images observant la même scène et ceci de façon globale. Cette comparaison est faite par la comparaison des densités spectrales de puissance (DSP) de deux images. Il est alors possible d'observer la différence de contraste de certaines fréquences spatiales (Papoulis 1987).

Pour ce faire, nous cherchons le filtre optimal, au sens des moindres carrés, qui permet de « passer » d'une image à l'autre. Ce filtre modélise donc les différences de FTM entre les deux images comparées et ceci en fonction de la fréquence. Nous cherchons ainsi le filtre h qui permet de minimiser l'erreur/différence entre les images. Ce qui se traduit par :

$$\{h / E[(A * h - B)^2] \min\} \quad 3.29$$

La démarche est présentée en une dimension, elle est identique en deux dimensions. D'après le théorème de l'orthogonalité, on obtient :

$$\forall l \quad E[(A * h(k) - B(k)) (A(k-l))] = 0 \quad 3.30$$

$$E[A * h(k) A(k-l)] = E[B(k)A(k-l)] \quad 3.31$$

$$h * \Gamma_A(l) = \Gamma_{B,A}(l) \quad 3.32$$

Où Γ_R est l'auto-corrélation de R : $\Gamma_R(\delta) = E[r(x) r(x-\delta)]$

* est l'opérateur convolution.

En appliquant la transformée de Fourier (TF) l'équation ci-dessus devient :

$$H(f) \times DSP[A](f) = DSP[B,A](f) \quad 3.33$$

où :

$DSP[A]$ est la Densité Spectrale de Puissance de A définie comme la TF de l'autocorrélation de A.

$DSP[B,A]$ est la Densité Spectrale de Puissance croisée de A et B, définie comme la TF de la corrélation croisée entre A et B.

On obtient l'expression de TF_{dev} (*Transfert Function deviation*):

$$TF_{dev}(A, B) = \frac{DSP[B,A](f)}{DSP[A](f)} = \frac{E[TF(\Gamma(B)) \times TF[\Gamma(A)]^*]}{E[|TF[\Gamma(A)]|^2]} \quad 3.34$$

La transformée de Fourier inverse donne PSF_{dev} :

$$PSF_{dev}(A, B) = TF^{-1}[TF_{dev}(A, B)] \quad 3.35$$

Le module de PSF_{dev} nous donne MTF_{dev} :

$$MTF_{dev}(A, B) = |TF_{dev}(A, B)| \quad 3.36$$

Ces résultats sont obtenus par une estimation de la corrélation de type Welch. Le calcul du filtre H est réalisé par blocs. Le décalage entre deux blocs est optimisé d'après les résultats des travaux de (Welch 1967). Le résultat pour l'image entière est la moyenne des résultats obtenus par blocs, la moyenne des filtres étant une bonne estimation du filtre H global.

Le résultat du calcul donne un filtre $H(f)$ en 2 dimensions. Ce filtre n'est pas isotrope. Le résultat que nous utiliserons est extrait au niveau des axes de symétrie horizontal ou vertical (respectivement longitudinalement et orthogonalement à la trace du satellite).

H/MTF_{dev} peut être vu comme la partie linéaire et stationnaire de la transformation qui lie la FTM des deux images que l'on compare. L'ensemble des différences entre les deux images ne peut pas toujours être modélisé par une convolution. L'estimation linéaire et stationnaire qui est faite à travers MTF_{dev} ne prend pas en compte l'erreur liée à la partie non linéaire. Cette erreur est observée grâce au calcul du résidu de la façon suivante :

$$B_{0k}^* = PSF_{dev} * B_{0k} + \text{residu} \quad 3.37$$

Une amélioration de MTF_{dev} pourrait être contrebalancée par une augmentation de la partie non linéaire. Surveiller l'importance du résidu permet d'assurer la pertinence des résultats fournis par MTF_{dev} .

MTF_{dev} est relié à la corrélation des images, mais dans le domaine de Fourier, cet indice fournit une information bien plus pertinente que le simple rapport des DSP. Par exemple, le rapport des DSP des deux images observant des zones différentes donnera un résultat non nul, au contraire MTF_{dev} donne un résultat très proche de zéros dans ce cas.

Utilisation de MTF_{dev}

Le calcul de MTF_{dev} est effectué entre deux images. L'utilisation pour laquelle ce calcul a été mis en place est la comparaison de la FTM de l'image synthétisée avec la FTM de l'image de référence. Cette mesure nécessite d'avoir une image synthétisée et une image de référence à la même résolution, ceci n'étant rendu possible que dans le cadre de simulation. Le résultat de ce calcul permet de caractériser la FTM de l'image synthétisée par rapport à la FTM d'une image de référence, ce qui en fait un indice de qualité de la FTM de l'image synthétisée.

Cependant, nous avons constaté que le calcul de MTF_{dev} entre une bande de l'image synthétisée et l'image PAN ($MTF_{dev}(PAN, B_k^*)$) permet contre toute attente de caractériser le défaut de FTM autour de 0,125 fe malgré le fait que les bandes passantes spectrales observées soient différentes.

Résultats attendus

La comparaison des FTM fournit donc une courbe fonction de la fréquence normalisée par la fréquence d'échantillonnage de PAN. La valeur de MTF_{dev} est normalisée par rapport à l'image de référence. Le résultat idéal est une droite de valeur unitaire.

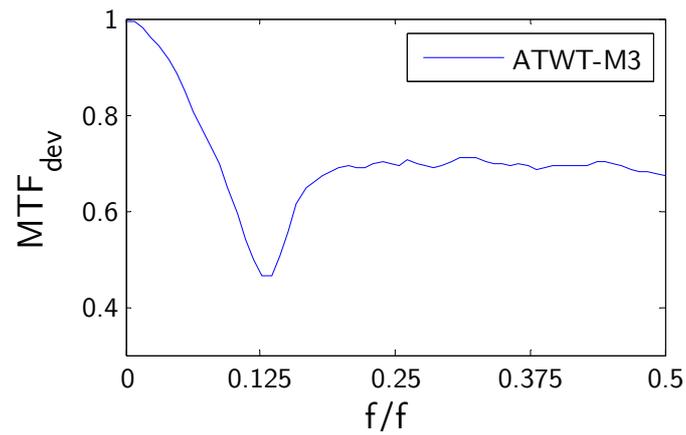


Figure 3.12 : Allure de MTF_{dev} pour la bande B2 synthétisée avec ATWT-M3.

La figure 3.12 présente le type de résultat que l'on obtient pour la synthèse d'un canal avec la méthode ATWT-M3. Ce calcul est effectué pour les quatre canaux. L'allure de ces courbes sera discutée pour les différentes méthodes dans le chapitre 5.

4 Adaptation de FTM dans la fusion

L'objectif de ces travaux est d'améliorer le respect de la propriété de synthèse. On cherche ici plus précisément à améliorer la représentation de la géométrie en modifiant la FTM des images MS afin qu'elle soit le plus proche possible de l'image qu'aurait acquise le capteur multispectral s'il possédait la haute résolution spatiale.

Dans cette partie nous présentons les diverses stratégies possibles pour modifier la FTM des images afin d'obtenir, pour l'image synthétisée, la FTM typique d'un capteur à haute résolution spatiale.

La solution retenue est d'effectuer un prétraitement de l'image MS avant de la fusionner avec l'image PAN. Ce prétraitement consiste tout d'abord à déconvoluer l'image MS, puis à la ré-échantillonner et enfin de la convoluer avec la FTM typique de la haute résolution. Nous avons ajouté l'adaptation de FTM à la méthode de référence ATWT-M3. Cette adaptation de la FTM constitue la deuxième contribution de ma thèse.

Enfin, la deuxième partie de ce chapitre est consacré à l'évaluation des performances de cette nouvelle méthode, ATWT-M3-MTF_{adapted}, qui est effectuée en comparaison des résultats obtenus avec la méthode ATWT-M3. La comparaison des résultats est faite à l'aide du protocole d'évaluation de la qualité présenté dans le chapitre précédent, mais aussi du calcul de MTF_{dev} afin de bien vérifier que notre adaptation de FTM permet bien d'attendre le résultat escompté.

4.1 Influence de la FTM

4.1.1 Objectif

Nous cherchons à améliorer le respect de la propriété de synthèse en termes de géométrie. Pour ce faire, nous prenons en compte la différence de FTM entre les images PAN et MS. Il est ainsi nécessaire de connaître la FTM des images PAN et MS. Ce type d'information est connu avec précision par le maître d'œuvre du satellite d'observation, mais aussi par le fournisseur d'image. Si cette information n'est pas accessible, alors il est possible d'utiliser un algorithme permettant d'en faire l'estimation, comme présentée dans le §2.1.6. En associant les outils de modification de FTM et la connaissance de la FTM des images MS et PAN, il est possible de traiter par déconvolution l'image MS synthétisée de façon à ce qu'elle semble provenir d'un système MS à haute résolution spatiale.

Ceci nous permet d'améliorer les caractéristiques de FTM de l'image synthétisée lors de la fusion et de mieux respecter la propriété de synthèse vis-à-vis de la FTM et donc de la

restitution de la géométrie. Nous voulons que l'image MS synthétisée ait la FTM d'une image à haute résolution et pas la « simple » association non maîtrisée des FTM PAN et MS. Pour cela, un rehaussement de la FTM sera effectué sur les images MS. Ce rehaussement est illustré de manière schématique par les flèches dans la figure 4.1. Ainsi, l'allure de la FTM sous-jacente à l'image synthétisée sera plus proche de l'allure de la FTM d'une image à haute résolution représentée en trait plein dans la figure 4.1.

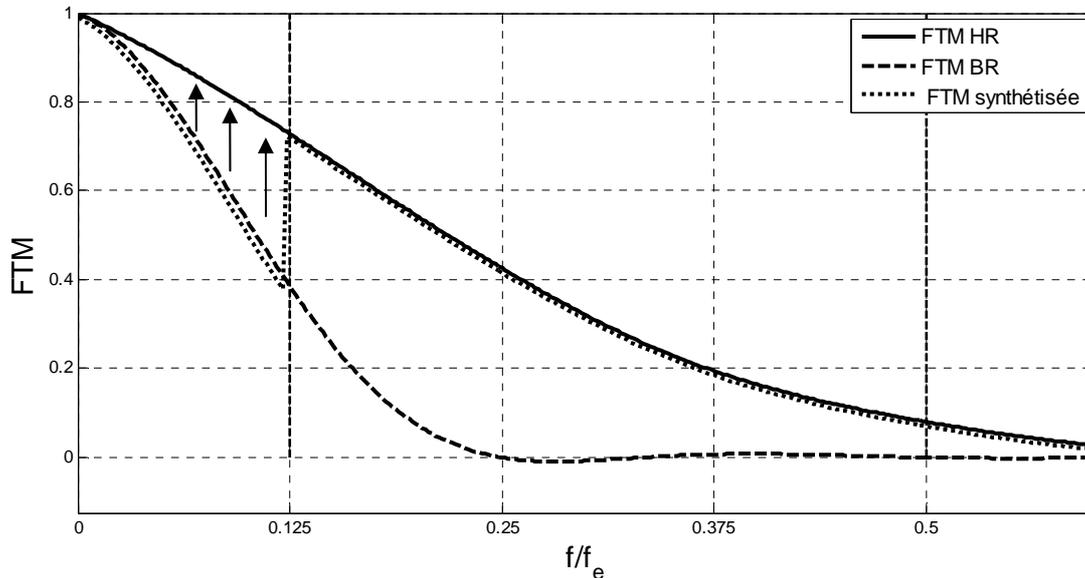


Figure 4.1 : rehaussement de la FTM des images MS pour la synthèse.

Nos travaux seront le plus souvent présentés dans le cas où le rapport de résolution est de 4 et ce qui est le cas actuellement le plus répandu avec les systèmes optiques à haute résolution (Quickbird, PLEIADES, Ikonos). Cependant, il existe des satellites dont le rapport de résolution entre les images MS et PAN est seulement de 2 comme les satellites de la famille SPOT. Nous présentons ce cas dans le paragraphe 4.1.4 en donnant les conclusions liées aux résultats des tests effectués.

4.1.2 La transformation de FTM dans le processus de fusion

Nous avons envisagé différentes solutions pour rehausser la FTM des images MS.

La difficulté à déconvoluer une image est relative à la distance entre la FTM pour la fréquence de Nyquist de l'image (point à $f_e/2$) et sa projection sur l'axe de FTM unitaire. Plus cette distance est grande, plus le niveau de la FTM est bas à la fréquence de Nyquist de l'image. Ceci implique que le contraste est faible pour les fréquences proches de la fréquence de Nyquist. La FTM étant monotone décroissante, lors de la déconvolution les contrastes des hautes fréquences seront d'autant plus fortement rehaussés que la valeur de la FTM à Nyquist est faible. Le bruit présent dans les images est qualifié de blanc, donc en moyenne réparti de manière égale quelle que soit la fréquence spatiale. Le niveau du bruit sera d'autant plus amplifié après déconvolution que la FTM est faible pour les hautes fréquences. Autrement dit, plus la FTM est faible à la fréquence de Nyquist, plus la déconvolution de l'image sera sensible au bruit.

L'écart entre la FTM pour la fréquence de Nyquist et la valeur de FTM unitaire (illustré dans la figure 4.2) est plus grand pour le cas de l'image PAN (flèche de droite) que dans le cas MS (flèche de gauche). La déconvolution de la FTM de l'image PAN est plus sensible au bruit que pour l'image MS (Banham et Katsaggelos 1997) et donc plus délicate à réaliser.

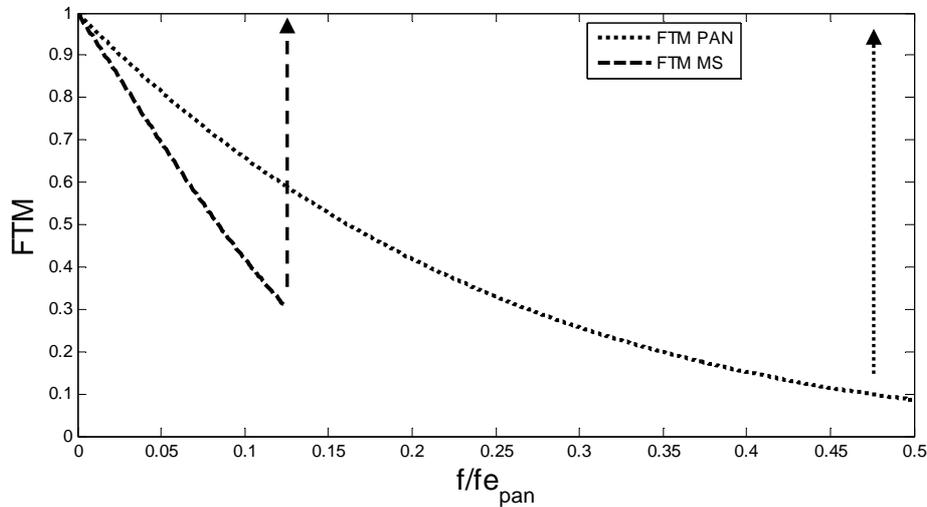


Figure 4.2 : Illustration d'adaptation de FTM.

La modification de la FTM peut être réalisée à différentes étapes du processus de fusion. Nous présentons plusieurs scénarios possibles, et discutons leurs avantages et inconvénients.

- Le premier « scénario » consiste à déconvoluer en premier les images MS et PAN. Puis nous effectuons la fusion de type ARSIS entre les versions déconvoluées de l'image PAN et des images MS. L'image synthétisée est dépourvue de FTM. Nous lui appliquons alors la FTM qui correspond à un capteur à haute résolution spatiale. Le déroulement de ce scénario est illustré dans la figure 4.3.

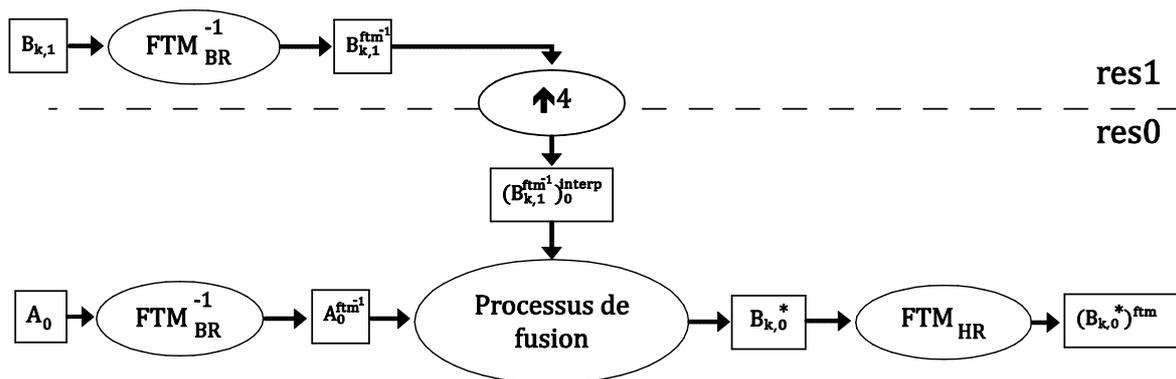


Figure 4.3 : schéma de principe du scénario 1.

Ce scénario implique la déconvolution à la fois de l'image PAN et de tous les canaux de l'image MS. La déconvolution de l'image PAN est assez sensible au bruit, elle entraînerait une amplification importante du bruit et viendrait perturber l'estimation des paramètres des modèles IMM et HRIMM. Pour ce scénario, les hautes fréquences de l'image PAN sont déconvoluées avant la fusion, puis sont de nouveau convoluées après la fusion, avec une FTM

quasiment identique. Les hautes fréquences de l'image synthétisée ont subi deux traitements pour avoir une FTM presque identique.

- Le deuxième « scénario » consiste à modifier la FTM de l'image MS uniquement, pour que la FTM de l'image une fois fusionnée ait la FTM attendue. Le déroulement de ce scénario est illustré sur la figure 4.4. Pour cela, il faut déconvoluer l'image MS puis lui donner la FTM qui correspond à un capteur multimodal à haute résolution une fois rééchantillonnée.

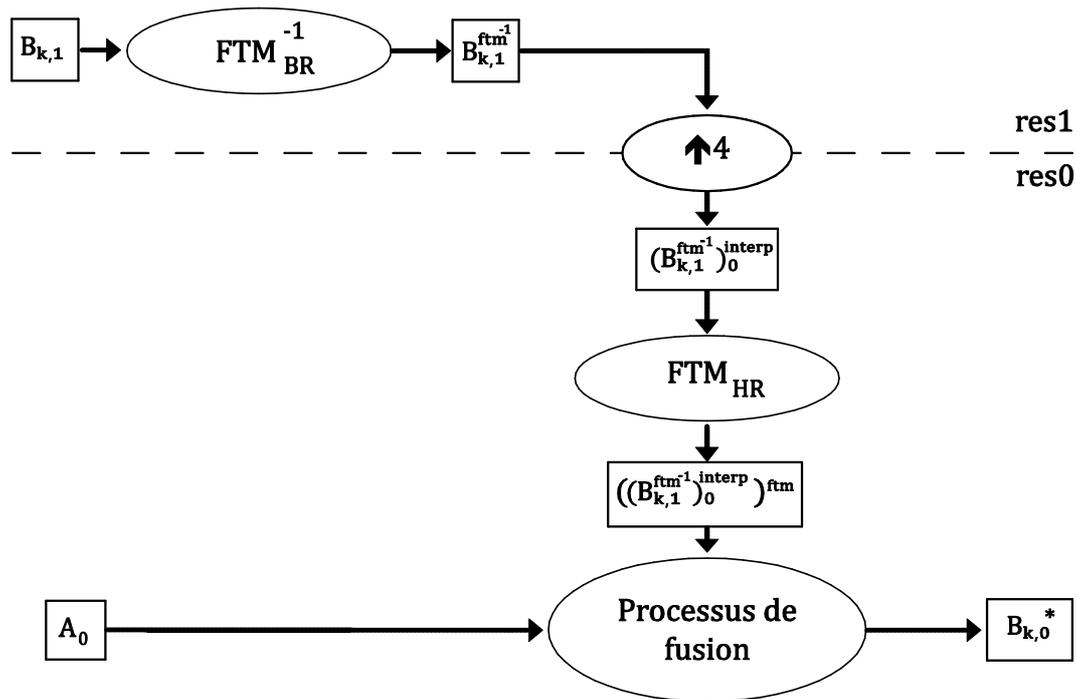


Figure 4.4 : schéma de principe du scénario 2.

Pour ce scénario, seules les images MS seront déconvoluées. La déconvolution des images MS est moins sensible au bruit que pour l'image PAN. Cependant, le fait de déconvoluer les images MS peut amplifier l'importance des artéfacts de repliement de spectre, présent dans les images MS.

- Le troisième « scénario » consiste à effectuer une déconvolution dans les coefficients d'ondelettes, suivant l'idée d'une méthode de fusion de (Thomas 2006) qui est de rehausser les coefficients en ondelettes. La déconvolution a lieu dans le processus de fusion comme présenté dans la figure 4.5.

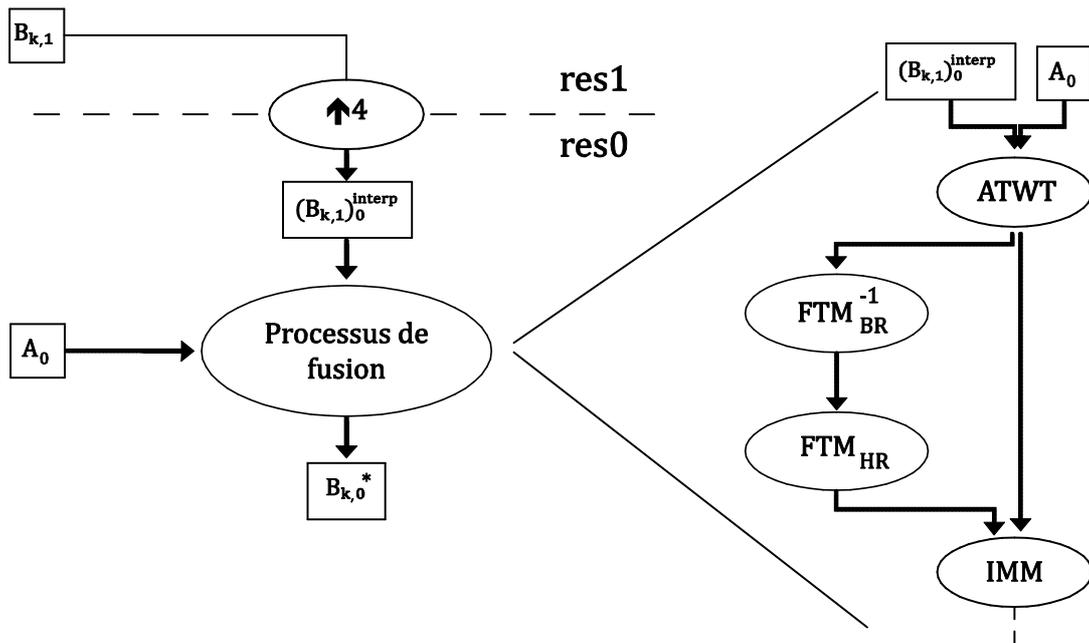


Figure 4.5 : schéma de principe du scénario 3

Les ondelettes que nous utilisons sont des ondelettes à trous (sans décimation). Déconvoluer dans les ondelettes oblige à déconvoluer chaque plan d'ondelette dont la taille est identique à l'image complète. Ceci entraîne de nombreuses déconvolutions coûteuses en temps de calcul. De plus, déconvoluer dans les ondelettes n'assure pas spécialement un meilleur résultat. Le point intéressant de ce scénario pourrait être d'associer à la déconvolution un traitement de débruitage après la déconvolution.

- Le quatrième « scénario » consiste à déconvoluer l'image MS avec la FTM de l'image PAN de façon à rehausser la FTM.

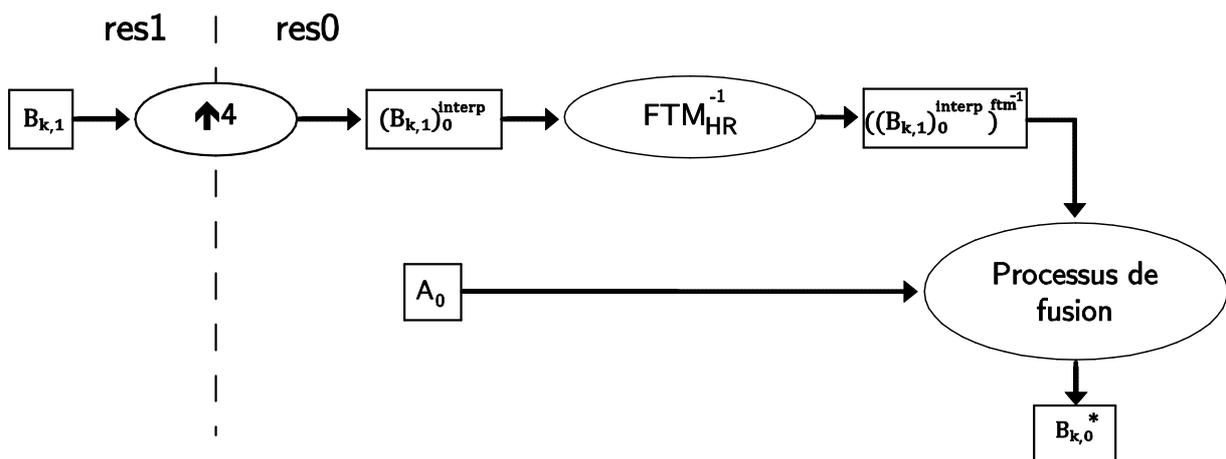


Figure 4.6 : schéma de principe du scénario 4

Cette solution présente a priori des avantages importants. Le fait de déconvoluer une image avec la FTM de l'image PAN permet de modifier la FTM en une seule étape contrairement au deuxième scénario. Déconvoluer l'image MS avec la FTM de l'image PAN permet une déconvolution plus « douce » et pourrait permettre une amplification du bruit

moins importante que dans le cas du deuxième scénario. Ce scénario a été implanté : les expérimentations préliminaires ont montré que cette solution rehausse la FTM, mais n'assure pas en terme de FTM des résultats aussi satisfaisants que le deuxième scénario. L'annexe A étaye les raisons qui ont fait que ce scénario n'a pas été conservé au profit du scénario 2.

- **Solution retenue :**

La solution retenue est le deuxième scénario. Celui-ci consiste à déconvoluer l'image MS pour obtenir une image dépourvue de FTM. Une fois ré-échantillonnée au pas d'échantillonnage de l'image PAN, cette image sera convoluée avec la PSF d'un instrument à haute résolution. Les traitements appliqués à l'image MS pour obtenir l'image MS ré-échantillonnée et adaptée en FTM (noté $MS_{\text{rééch,MTFadapted}}$) sont exprimés à travers l'équation suivante :

$$MS_{\text{rééch,MTFadapted}} = ((MS *^{1-} PSF_{MS}) \uparrow 4) * PSF_{PAN} \quad 4.1$$

Le second scénario se différencie du processus de fusion classique par la modification des données d'entrée : $(B_{k,1})_0^{\text{interp}}$ au lieu $((B_{k,1}^{FTM^{-1}})^{\text{interp}})^{FTM}$. La figure 4.7 reprend le scénario 2 et montre le schéma classique de fusion.

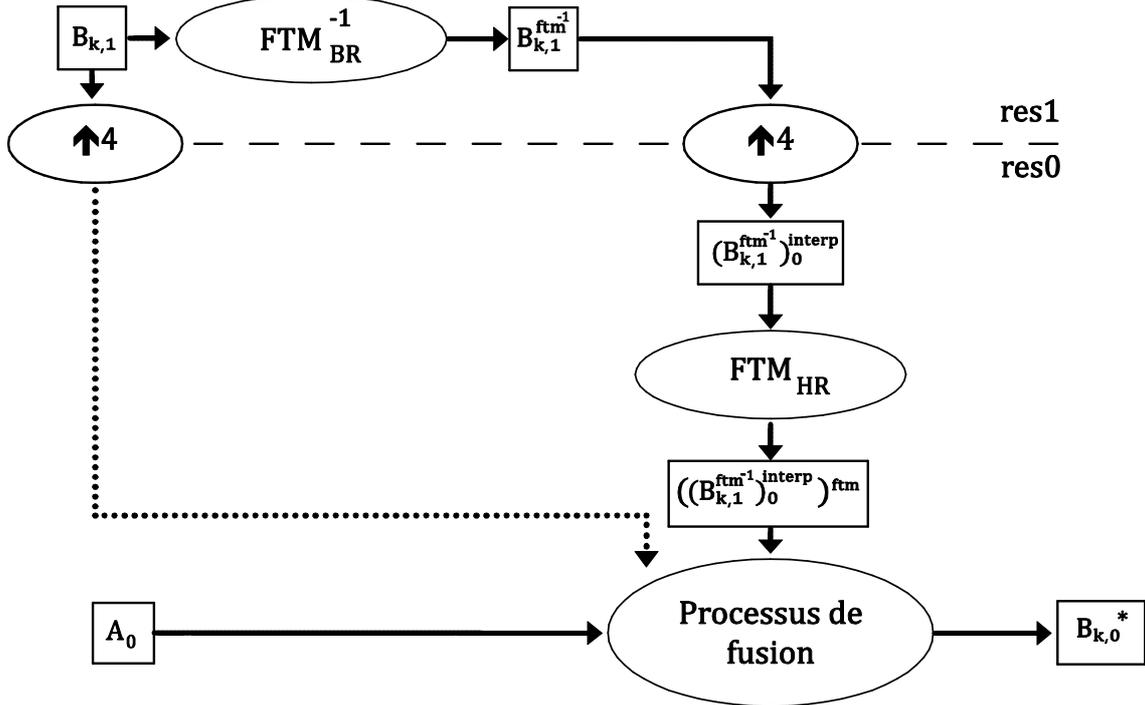


Figure 4.7 : Schéma classique (traits tirets) et scénario 2 (traits pleins)

La figure 4.8-a présente l'image $(B_{k,1})_0^{\text{interp}}$ et la figure 4.8-b l'image $((B_{k,1}^{FTM^{-1}})^{\text{interp}})^{FTM}$. L'image de droite (b) offre plus de détails et de piqué que celle de gauche (a). On pressent qu'en augmentant la qualité des images d'entrée d'un point de vue géométrique, on obtiendra une synthèse de meilleure qualité géométrique.

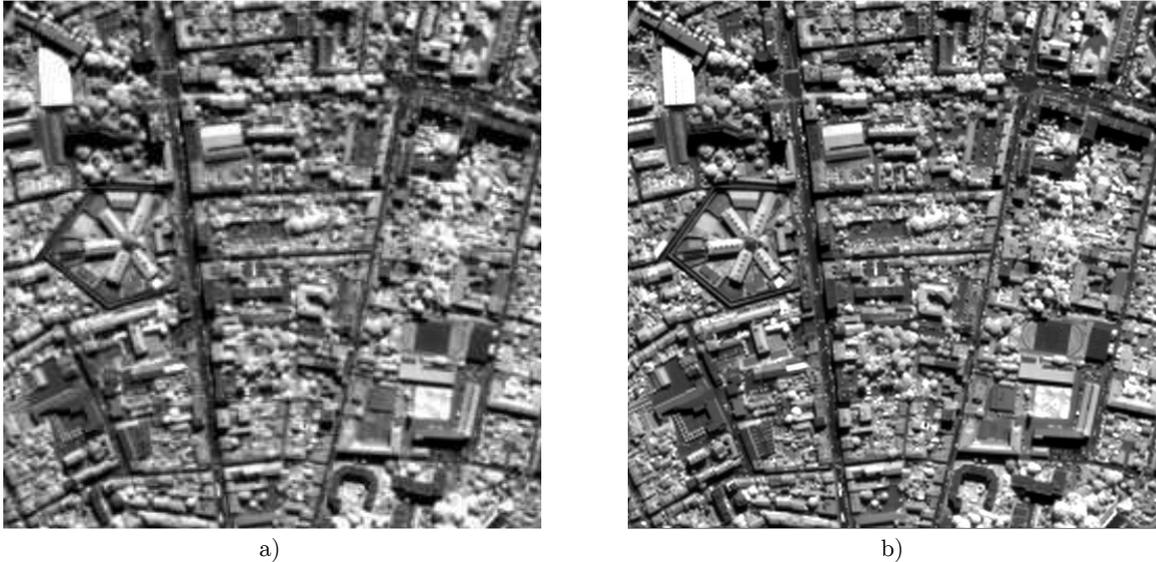


Figure 4.8 : Effet de l'adaptation de FTM. a) MS_{BR} (originale) ré-échantillonnée.
b) $MS_{rééch, MTFadapted}$.

Le fait d'utiliser des ondelettes non décimées (à trous) nécessite une étape de ré-échantillonnage. L'importance de cette étape de ré-échantillonnage dans la qualité du produit de fusion est connue (Khan *et al.* 2008) (Blanc *et al.* 1998) (Wald *et al.* 1997). L'adaptation de FTM pourrait gagner à utiliser une étape de ré-échantillonnage optimisée pour la fusion c'est-à-dire en minimisant la présence d'artefact de type ringing ou repliement de spectre. Une partie des travaux de (Khan *et al.* 2008) va dans ce sens. Ceci pourrait être intégré dans le scénario.

4.1.3 Adaptation de FTM sans déconvolution optique

Le problème de fusion qui consiste en la synthèse d'une image qui aurait pu être vue par le capteur s'il avait la haute résolution spatiale, amène à s'intéresser à des caractéristiques uniquement liées au capteur et à les modifier. La partie optique des images PAN et MS est identique du fait que les capteurs de ces images sont placés derrière la même ou quasiment la même configuration optique si les images proviennent du même satellite. Comme nous l'avons présenté dans le §2.1, la FTM d'une image provient à la fois de la partie optique et de la partie détecteur.

$$FTM = FTM_{optique} * FTM_{détecteur} \quad 4.2$$

Nous voulons synthétiser une image qui aurait pu être observée par un détecteur multimodal possédant la haute résolution. Dans la partie 4.1.2, la déconvolution des images MS considérait la FTM totale. L'image MS était déconvoluée de la FTM optique et de la FTM détecteur. Dans cette partie, nous montrons qu'il est possible de prendre en compte la FTM sans déconvoluer la partie de la FTM liée à la partie optique.

Notre but étant de prendre en compte la différence de FTM entre les voies PAN et MS, le fait de déconvoluer la partie liée à la partie optique commune n'apporte rien à notre objectif.

Dans ce cas, pour modifier la FTM, nous considérons uniquement la FTM liée au détecteur et non la FTM totale de l'instrument. En ne déconvoluant que la FTM liée à la partie détecteur de l'image, la déconvolution devient moins importante et pourrait entraîner une amplification du bruit légèrement moins grande. Cette proposition apporte un avantage important du fait que les images MS seront déconvoluées par la même FTM quelque soit la bande spectrale considérée ; la FTM détecteur est facilement modélisable par la relation proposée par (Delvit *et al.* 2003):

$$FTM_{détecteur}(f_x, f_y) = \text{sinc}(f_x) \text{sinc}(\alpha f_x) \quad 4.3$$

4.1.4 Effet de l'adaptation de FTM selon le rapport de résolution : 2 ou 4

Nous avons présenté précédemment le cas où le rapport de résolution est de quatre entre PAN et MS. Certains satellites fournissent des images ayant un rapport de résolution seulement de deux, comme le satellite Landsat ou les satellites de la famille SPOT. L'effet de l'adaptation de la FTM lors de la fusion n'a pas forcément la même importance si le rapport de résolution est de 2 ou 4.

Nous cherchons à évaluer la différence d'effet selon le rapport de résolution. La figure 4.9 présente de façon schématique l'allure de la FTM dans le cas d'une fusion de type ARSIS selon que le rapport de résolution est 2 ou 4.

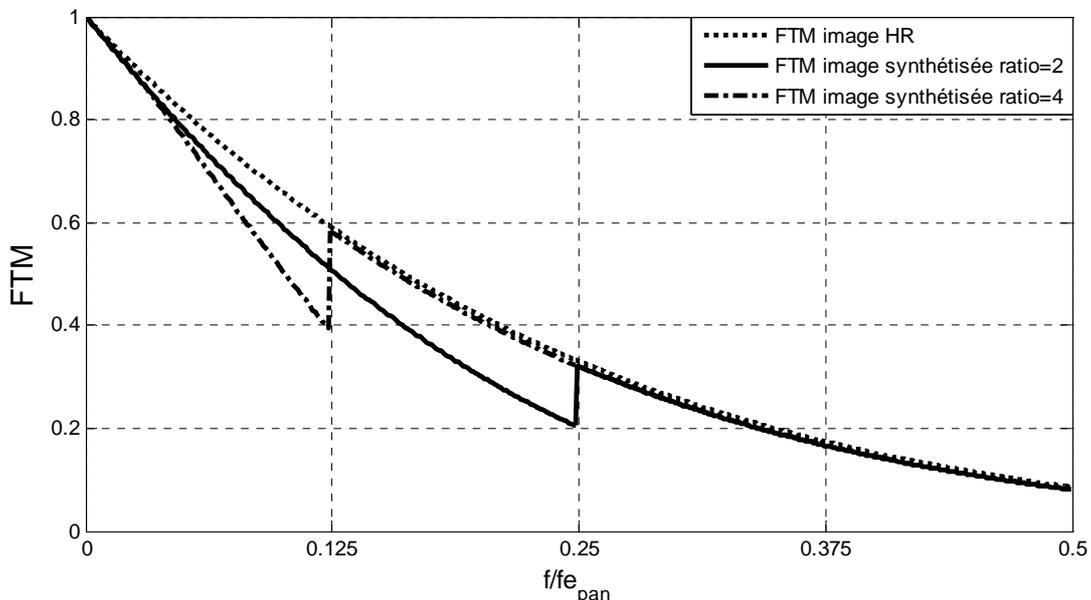


Figure 4.9 : Différence de FTM de l'image synthétisée selon le rapport de résolution: 2 ou 4

Les différences principales sont l'amplitude de la différence de FTM et la proportion de l'image dont la FTM sera rehaussée. Dans le cas où le rapport de résolution est de 4, l'amplitude du rehaussement de la FTM est plus importante que dans le cas où le rapport de résolution est de 2. Par ailleurs, dans le cas où le rapport de résolution est de 2, le

rehaussement de la FTM se fait sur la moitié inférieure des fréquences alors que celui-ci ne touche que le quart inférieur des fréquences dans le cas où le rapport de résolution est de 4.

Le jeu d'images utilisé dans nos travaux regroupe des images simulant les images des satellites SPOT-5 et PLEIADES. En appliquant la prise en compte de la FTM à ces deux types d'images, nous pourrions constater si l'effet de la prise en compte de la FTM apporte une amélioration plus importante pour un rapport de résolution de 2 ou de 4.

4.2 ATWT-M3-MTFadapted

Pour mesurer l'effet de l'adaptation de FTM, il est nécessaire de l'appliquer à une méthode de fusion. Le choix de la méthode à laquelle on associe l'adaptation de FTM est donc important. Ce choix s'est fait parmi les méthodes performantes et remplissant les critères qui sont les suivants :

- Basée sur le concept ARSIS
- IMM-HRIMM global
- Décomposition non décimée

En effet, l'adaptation de FTM a été plutôt pensée pour les méthodes basées sur le concept ARSIS. Il semble donc plus intéressant d'évaluer son effet dans ce cadre. Certaines méthodes de fusion, qui utilisent des traitements locaux et globaux combinés, obtiennent des performances légèrement supérieures aux méthodes n'utilisant que des traitements globaux (Thomas *et al.* 2008). Le caractère global de l'adaptation de FTM nous a poussé à l'associer à une méthode ayant, elle aussi, ce caractère global. De plus, les traitements locaux entraîneraient certainement une plus grande difficulté d'interprétation, et pourraient se révéler trompeuses pour l'évaluation de l'effet apporté par l'adaptation de FTM.

La méthode ATWT-M3 est donc choisie comme méthode de base pour l'implantation du scénario 2. Elle a été présentée en §3.4.

4.2.1 Description de la méthode

La méthode résultant de cette implantation est appelée ATWT-M3-MTFadapted. L'image MS est déconvoluée pour obtenir une image dépourvue de FTM (ou de FTM unitaire sur l'ensemble des fréquences). Puis, une fois ré-échantillonnée au pas d'échantillonnage de l'image PAN par une fonction spline de degré 7 (Unser *et al.* 1992);(Unser *et al.* 1992) cette image est convoluée avec la FTM d'une image à haute résolution (FTM détecteur HR).

$$MS_{réech,FTM\ adapted} = ((MS_{BR} *^{1-} PSF_{MS}) \uparrow 4) * PSF_{PAN} \quad 4.4$$

La déconvolution est effectuée par un filtrage inverse dont l'expression est la suivante :

$$G = \frac{\bar{H}}{\max(\|H\|^2 + \epsilon^2)} \quad 4.5$$

où H est la FTM que l'on veut de l'image à déconvoluer, et ϵ est le terme de régularisation afin de s'assurer de ne pas diviser par zéro.

La déconvolution utilisée ici est assez simple. Des méthodes de déconvolution plus complexe existent allant d'une approche basée sur le filtrage de Wiener (Papoulis 1987) à des approches de type variationnelles (Samson *et al.* 2000). Comme nous cherchons à montrer l'apport de la modification de la FTM, nous nous sommes contentés d'une méthode de déconvolution de type pseudo-inverse. En fonction des résultats obtenus avec cette méthode, une déconvolution plus adaptée pourrait être mise en place. Il sera bon de choisir une méthode de déconvolution permettant une amplification du bruit la plus faible possible.

4.2.2 *Evaluation des performances de la méthode*

Nous ne cherchons pas à évaluer la méthode ATWT-M3-MTFadapted en elle-même mais plutôt le gain que permet d'obtenir l'adaptation de FTM par rapport à ATWT-M3. La différence de performance mesurée entre ces deux méthodes pourra donc être attribuée à la prise en compte de la FTM lors de la fusion. Les images synthétisées par ces deux méthodes seront comparées aux images de références. Les comparaisons entre ces images seront de plusieurs types : analyse visuelle (qualitatif), analyse quantitative (bilan qualité), et en terme de FTM (MTF_{dev}).

La comparaison quantitative est réalisée à l'aide de l'outil présenté dans le chapitre 3. Les distances ou critères sélectionnés sont des distances couramment utilisées dans le domaine de la fusion.

L'analyse visuelle est réalisée de façon globale pour évaluer l'aspect général du piqué ou du flou, mais aussi pour des zones qui présentent encore des artefacts. Les structures les plus difficiles à représenter sont d'une part les structures qui présentent de très forts contrastes et d'autre part les petites structures aux signatures spectrales variées. C'est grâce à ces structures que nous pourrions évaluer la qualité visuelle des images synthétisées. Les exemples visuels les plus pertinents seront présentés dans la partie sur les résultats visuels §4.2.3.2.

Les FTM sous-jacentes des images synthétisées sont évaluées à l'aide de l'outil MTF_{dev} présenté dans le §3.6.3. Cet outil fournit une comparaison de la FTM de l'image synthétisée avec celle de l'image de référence en fonction de la fréquence. Le résultat de cet outil est une courbe dont la valeur idéale est 1 quelle que soit la fréquence. En comparant ces courbes pour les deux images, il est possible d'identifier visuellement l'image synthétisée dont la FTM sous-jacente est la plus proche de l'image de référence.

4.2.3 *Les résultats*

Les images que nous synthétisons font 1024x1024 pixels. Pour éviter tout effet de bord, nous calculons les résultats de l'analyse quantitative sur une vignette centrée de 512x512 pixels. Nous avons bien sûr vérifié que nous obtenions des conclusions identiques en modifiant la taille de cette vignette.

4.2.3.1 *Analyses quantitatives*

Dans cette partie relative aux analyses des résultats des différentes méthodes fusions, nous présentons les bilans de qualité obtenus pour les images dont nous disposons. Nous

commençons par les images qui présentent un rapport de résolution de 4 entre les images PAN et MS (tableau 4.1, tableau 4.2, tableau 4.3), puis pour les images dont le rapport de résolution 2 (tableau 4.4, tableau 4.5). Enfin, nous présenterons les résultats obtenus avec le détecteur de netteté. Dans les tableaux présentant les bilans de qualité, nous avons mis en gras les valeurs significativement améliorées.

Tableau 4.1 : Bilan qualité image PLEIADES pour Toulouse

Modalité	Indice	Valeur idéale	ATWT-M3	ATWT-M3-MTFadapted
Bleu	biasRel (%)	0	-0,001	0,004
	diffVarRel (%)	0	14	6
	σ Rel (%)	0	8,7	7,4
	cc	1	0,983	0,987
	ccHF	1	0,920	0,922
Vert	biasRel (%)	0	0,0	0,0
	diffVarRel (%)	0	16	7
	σ Rel (%)	0	8,6	6,6
	cc	1	0,988	0,991
	ccHF	1	0,961	0,963
Rouge	biasRel (%)	0	0,0	0,0
	diffVarRel (%)	0	16	8
	σ Rel (%)	0	9,3	6,5
	cc	1	0,991	0,995
	ccHF	1	0,969	0,973
PIR	biasRel (%)	0	0,0	0,0
	diffVarRel (%)	0	34	28
	σ Rel (%)	0	13,7	11,4
	cc	1	0,967	0,977
	ccHF	1	0,879	0,887
Indices multimodaux	ERGAS	0	2,6	2,0
	mean SAM	0	3,1	2,9

Tableau 4.2 : Bilan qualité image SPOT-5 supermode pour Madrid (Espagne)

Modalité	Indice	Valeur idéale	ATWT-M3	ATWT-M3-MTFadapted
Bleu	biasRel (%)	0	-0,003	0,000
	diffVarRel (%)	0	22	10
	σ Rel (%)	0	2,9	2,2
	cc	1	0,980	0,986
	ccHF	1	0,908	0,915
Vert	biasRel (%)	0	0,0	0,0
	diffVarRel (%)	0	23	11
	σ Rel (%)	0	3,9	2,8
	cc	1	0,983	0,989
	ccHF	1	0,934	0,940
Rouge	biasRel (%)	0	0,0	0,0
	diffVarRel (%)	0	22	10
	σ Rel (%)	0	5,0	3,4
	cc	1	0,985	0,992
	ccHF	1	0,948	0,955
PIR	biasRel (%)	0	0,0	0,0
	diffVarRel (%)	0	20	9
	σ Rel (%)	0	5,1	3,5
	cc	1	0,986	0,992
	ccHF	1	0,939	0,947
Indices multimodaux	ERGAS	0	1,1	0,8
	mean SAM	0	0,8	0,7

Tableau 4.3 : Bilan qualité image SPOT-5 supermode pour Fredericton (Canada)

Modalité	Indice	Valeur idéale	ATWT-M3	ATWT-M3-MTFadapted
Bleu	biasRel (%)	0	-0,007	0,001
	diffVarRel (%)	0	12	7
	σ Rel (%)	0	5,5	4,8
	cc	1	0,978	0,983
	ccHF	1	0,741	0,758
Vert	biasRel (%)	0	0,0	0,0
	diffVarRel (%)	0	10	6
	σ Rel (%)	0	7,2	6,1
	cc	1	0,983	0,987
	ccHF	1	0,797	0,811
Rouge	biasRel (%)	0	0,0	0,0
	diffVarRel (%)	0	10	5
	σ Rel (%)	0	10,4	8,9
	cc	1	0,984	0,988
	ccHF	1	0,792	0,807
PIR	biasRel (%)	0	0,0	0,0
	diffVarRel (%)	0	5	3
	σ Rel (%)	0	8,6	7,1
	cc	1	0,992	0,995
	ccHF	1	0,893	0,904
Indices multimodaux	ERGAS	0	2,0	1,7
	mean SAM	0	1,9	1,6

Les valeurs des indices obtenues pour la méthode ATWT-M3-MTFadapted sont plus proches des valeurs idéales que ne le sont celles obtenues pour ATWT-M3. Ceci qui signifie que l'adaptation de FTM permet d'améliorer l'ensemble des critères. C'est un point important, car il n'est pas forcément évident d'obtenir une amélioration de l'ensemble des distances de façon aussi « unanime ».

L'amélioration apportée par l'adaptation de FTM se fait dans des proportions différentes selon les distances ou critères. Pour certains critères, l'amélioration peut être importante avec des gains d'environ 25 % alors que pour d'autres, du fait des résultats déjà très proches de la valeur idéale pour la méthode de référence, l'amélioration ne sera que de l'ordre du pourcent.

- **Biais relatif** : Cette valeur est déjà très faible, voir quasiment idéale sans adaptation de la FTM. L'adaptation de la FTM n'apporte presque pas de modification, le biais reste très faible.
- **Coefficient de corrélation** : les valeurs obtenues par la méthode ATWT-M3 sont supérieures à 0,97, ce qui constitue une corrélation déjà très importante. La marge d'amélioration pour cette distance est donc faible. Cependant, l'adaptation de FTM apporte une amélioration.
- **Coefficient de corrélation des hautes fréquences** : les remarques faites sur cet indice suivent celle du coefficient de corrélation bien que la marge d'amélioration soit plus importante.

- **Différence en variance relative** : cette distance permet de quantifier l'innovation et d'évaluer si celle-ci est trop ou pas assez importante. Si l'innovation est trop grande, alors la valeur de cet indice devient négative. Ici, cet indice est positif dans les deux cas. Il est donc possible d'ajouter plus de hautes fréquences afin de s'approcher de l'image de référence. L'adaptation de FTM permet de réduire cette distance d'un facteur souvent proche de 2.
- **Écart-type de l'image de différence** : cette distance est améliorée de façon non négligeable puisqu'on obtient un gain compris entre 10 % et 30 %.
- **ERGAS** : la valeur de l'ERGAS est réduite de 10 à 30 %, l'amélioration de cette distance est assez importante. Cette distance mesure la qualité spectrale de l'image synthétisée, l'adaptation de FTM permet l'amélioration de la qualité spectrale de l'image synthétisée.
- **Mean SAM** : de la même façon que l'ERGAS, la valeur de cette distance de type multimodale est réduite grâce à l'adaptation de FTM. L'amélioration de cette distance est comprise entre 5 et 10 %.

Les résultats sur les distances monomodales montrent que l'adaptation de FTM permet une amélioration notable des distances mesurées. L'adaptation de FTM permet d'améliorer la netteté de l'image. Ce sont les distances monomodales qui sont les plus sensibles à cette caractéristique. L'amélioration des distances monomodales est un résultat attendu. Au contraire, le comportement des distances multimodales (ERGAS, Mean SAM) n'est pas un résultat attendu : par l'adaptation de FTM nous cherchions à diminuer le flou dans l'image, mais pas spécifiquement la qualité spectrale.

Plusieurs explications à ce résultat : le fait d'améliorer la qualité monomodale de chaque bande peut avoir comme conséquence d'améliorer la qualité spectrale de l'image synthétisée. Deuxième explication, l'adaptation de FTM rehausse le contraste de certaines fréquences. Le rehaussement pertinent de ces fréquences permet d'améliorer la qualité spectrale de l'image synthétisée.

- **Cas SPOT-5 avec un rapport de résolution de 2**

Voyons maintenant les cas où le rapport de résolution est de 2 entre les images PAN et MS. Ce cas concerne les images des satellites de la famille SPOT (tableau 4.4, tableau 4.5).

Tableau 4.4 : Bilan qualité image SPOT-5 pour Fredericton (Canada)

Modalité	Indice	Valeur idéale	ATWT-M3	ATWT-M3-MTFadapted
Bleu	biasRel (%)	0	-0,005	-0,003
	diffVarRel (%)	0	3,3	1,6
	σ Rel (%)	0	3,1	3,0
	cc	1	0,993	0,993
	ccHF	1	0,748	0,757
Vert	biasRel (%)	0	-0,008	-0,005
	diffVarRel (%)	0	2,3	0,7
	σ Rel (%)	0	4,0	3,8
	cc	1	0,994	0,995
	ccHF	1	0,770	0,777
Rouge	biasRel (%)	0	-0,012	-0,007
	diffVarRel (%)	0	2,0	0,5
	σ Rel (%)	0	6,0	5,7
	cc	1	0,995	0,995
	ccHF	1	0,766	0,773
PIR	biasRel (%)	0	-0,014	-0,006
	diffVarRel (%)	0	0,4	-0,6
	σ Rel (%)	0	4,4	4,4
	cc	1	0,998	0,998
	ccHF	1	0,841	0,842
Indices multimodaux	ERGAS	0	1,13	1,09
	mean SAM	0	1,07	1,02

Tableau 4.5 : Bilan qualité image SPOT-5 pour Madrid (Espagne)

Modalité	Indice	Valeur idéale	ATWT-M3	ATWT-M3-MTFadapted
Bleu	biasRel (%)	0	-0,002	-0,002
	diffVarRel (%)	0	0,6	-3,4
	σ Rel (%)	0	1,5	1,5
	cc	1	0,993	0,993
	ccHF	1	0,807	0,808
Vert	biasRel (%)	0	-0,002	-0,002
	diffVarRel (%)	0	0,0	-4,3
	σ Rel (%)	0	1,8	1,9
	cc	1	0,995	0,995
	ccHF	1	0,813	0,813
Rouge	biasRel (%)	0	-0,001	-0,001
	diffVarRel (%)	0	-0,7	-5,1
	σ Rel (%)	0	2,1	2,3
	cc	1	0,996	0,996
	ccHF	1	0,809	0,808
PIR	biasRel (%)	0	-0,001	0,808
	diffVarRel (%)	0	-0,7	-4,7
	σ Rel (%)	0	2,3	2,5
	cc	1	0,996	0,996
	ccHF	1	0,799	0,996
Indices multimodaux	ERGAS	0	0,49	0,53
	mean SAM	0	0,43	0,43

Les deux tableaux ci-dessus montrent que l'adaptation de FTM n'apporte pas d'amélioration dans le cas d'un rapport de résolution spatiale de 2, effectué ici pour les images MS et PAN de SPOT-5.

En conclusion, le gain de l'adaptation de la FTM dans le cas d'un rapport de résolution de 2 est nettement moins flagrant que dans le cas d'un rapport 4. Ceci peut s'expliquer par la constatation que nous avons faite précédemment sur la différence de niveau de FTM à Nyquist des MS à compenser : dans le cas d'un facteur 2 de résolution, la différence étant plus faible, l'adaptation de FTM est moins pertinente.

- **Détection de netteté**

Nous présentons ici les résultats de ce détecteur appliqué à nos images. Comme indiqué dans le chapitre précédent, ce détecteur évalue la cohérence de phase de l'image. Plus l'image est nette et plus la phase dans l'image est cohérente, plus une perturbation de cette phase entraînera une augmentation de la valeur de variation totale calculée par le détecteur. Ce détecteur a deux paramètres d'entrées : la taille de la vignette sur laquelle le calcul est effectué et l'amplitude de la perturbation de phase que l'on applique.

Le calcul est effectué sur des vignettes de rayon 64 pixels, et ce, sur l'ensemble des images. Nous appliquons ce détecteur aux différentes bandes spectrales de chacune des images obtenues par ATWT-M3 et ATWT-M3-MTFadapted, le résultat pour chaque vignette est la somme des valeurs obtenues pour chaque bande.

L'amélioration globale sur l'image obtenue par l'adaptation de FTM est de 0,5 %. L'amélioration de la netteté mesurée avec ce détecteur est relativement faible. L'adaptation de FTM modifie de façon significative une bande de fréquence de proportion assez faible dans l'image synthétisée. C'est pour cette raison que de l'amélioration mesurée de la netteté est faible.

Ainsi, nous avons décidé d'effectuer la détection sur l'image synthétisée sous-échantillonnée d'un rapport de deux. De cette façon, la modification apportée par la modification de FTM occupe une proportion de fréquence plus importante dans les hautes fréquences de l'image sous-échantillonnée. Dans ces conditions, il est plus probable de pouvoir mesurer une amélioration de la netteté avec ce détecteur. Le calcul est fait dans ces conditions pour des vignettes de taille deux fois inférieure soit de rayon 32 pixels. L'amélioration obtenue dans ce cas passe à 4,5 %. L'utilisation du détecteur de netteté n'est pas tout à fait pertinente lorsqu'elle est appliquée à l'image synthétisée à pleine résolution.

4.2.3.2 Analyse visuelle des images synthétisées

Les résultats des analyses quantitatives montrent que l'adaptation de FTM tend à améliorer la qualité des images synthétisées par rapport à de l'image de référence. Dans cette partie, nous allons présenter les résultats issus de l'analyse visuelle des images MS synthétisées.

La majeure partie des résultats visuels est présentée dans le cas des images PLÉIADES, c'est avec ces images que le gain de qualité est le plus visible. L'amélioration apportée sur les images est nette quand elle est observée sur un écran. Sur la version imprimée du manuscrit, l'amélioration apportée par l'adaptation de FTM naturellement moins visible.

- Halo plus faible



Figure 4.10 : Extrait de l'image de Toulouse, cas du halo de la voiture blanche.

a) ATWT-M3, b) ATWT-M3-MTFadapted, c) référence.

La figure 4.10 représente un véhicule sur une route délimitée par le terre-plein central de couleur rouge. Le blanc du véhicule bave ou crée un halo autour de celui-ci. On peut observer dans les différentes images de la figure 4.10, que ce halo n'est pas présent dans l'image de référence (figure 4.10-c) et qu'il est moins important dans la figure 4.10-b que dans la figure 4.10-a.

L'adaptation de FTM permet de réduire l'intensité/importance de ce halo et d'obtenir une représentation plus proche de l'image de référence.

Pour mieux comprendre ce qui a permis d'atténuer l'effet de halo, nous faisons une coupe perpendiculaire à l'axe de la route. Le tracé du profil dans une seule direction permet de mieux comprendre ce que l'on perçoit sur les extraits de la figure 4.10. Le halo a une teinte légèrement bleutée, le phénomène sera plus prononcé pour cette couleur. Le profil de l'image sera donc effectué pour la bande spectrale du bleu.

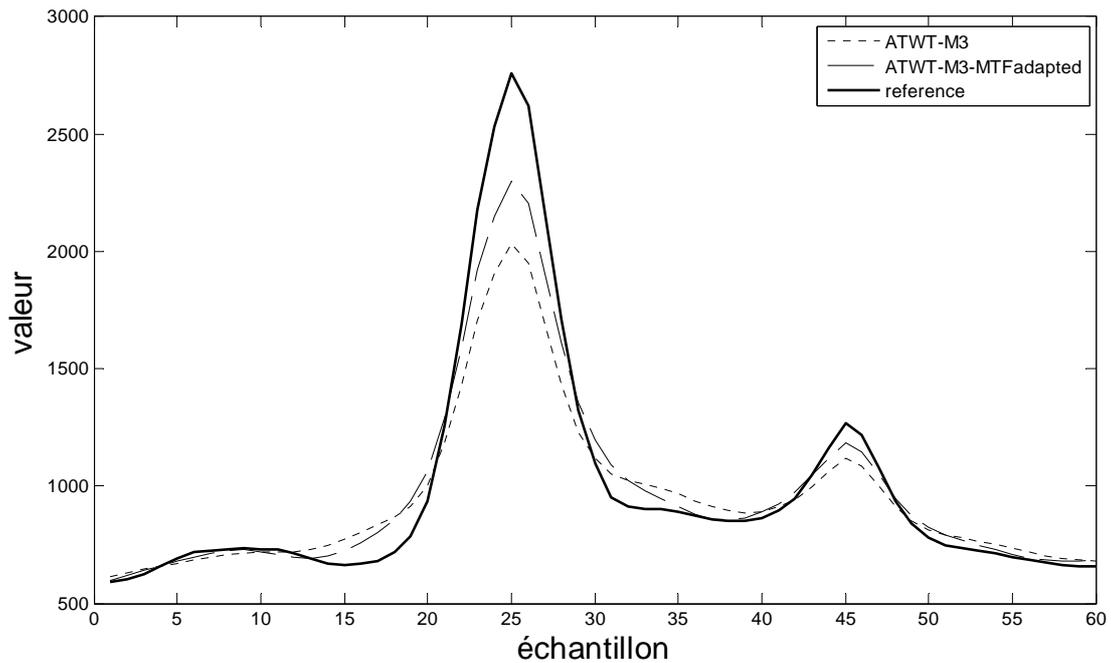


Figure 4.11 : profil de l'image perpendiculairement à la route.

La figure 4.11 représente le profil horizontal de la bande spectrale du bleu pour les trois cas présentés dans la figure 4.10. Ce qui est constaté pour cette bande spectrale bleue l'est aussi pour les autres modalités, mais dans une moindre mesure.

La voiture blanche dans le profil est représentée par le pic qui se situe au niveau de la graduation 25 de l'axe des abscisses. Le maximum de ce pic atteint des valeurs assez différentes pour les trois cas représentés ici. L'image de référence atteint la valeur la plus importante suivie par l'image ATWT-M3-MTFadapted dont la valeur est supérieure à celle de l'image ATWT-M3.

Une propriété de la FTM est de conserver la moyenne. Par conséquent, si la valeur du maximum augmente, alors le pic va avoir tendance à devenir plus étroit. On observe bien la différence de halo entre les échantillons 30 et 40 dans la figure 4.11. Dans cette zone on retrouve bien les observations visuelles avec halo qui s'étend sur 5 pixels dans le cas de ATWT-M3-MTFadapted alors qu'il s'étend sur 10 pixels dans le cas de ATWT-M3. La prise en compte de la FTM permet d'obtenir un maximum plus élevé et un pic légèrement plus étroit. C'est la combinaison de ces deux phénomènes qui permet de réduire l'effet de halo.

Dans ce paragraphe, nous montrons que l'adaptation de FTM permet une meilleure représentation de la géométrie dans l'image. Pour illustrer nos propos, nous allons, comme dans le paragraphe précédent, extraire en une dimension le profil (horizontal) d'une partie de l'image de Toulouse. La figure 4.12 présente le contexte de l'extraction du profil, la ligne blanche représente la zone où nous effectuons le profil.

Le profil de l'image va sur son passage trouver essentiellement des toits aux radiométries différentes. Les toits sont en majorité dans cette région constituée de tuiles rouges, c'est pour cette raison que l'extrait sera réalisé pour la modalité rouge.



Figure 4.12 : image de contexte pour un profil.

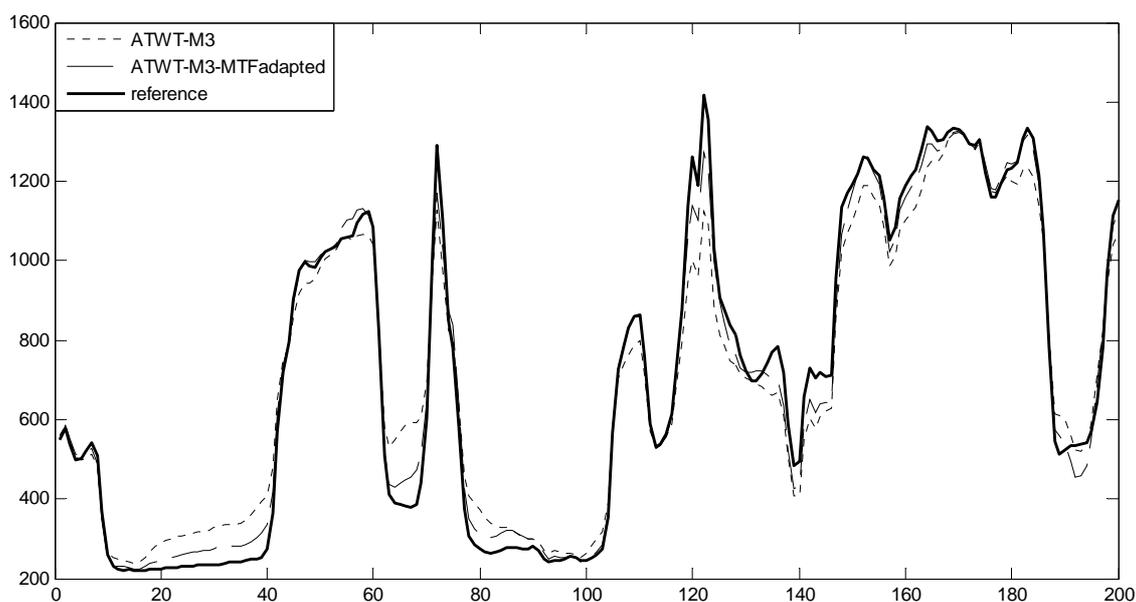


Figure 4.13 : profil pour la bande B2 dont le contexte du profil est présenté dans la figure précédente.

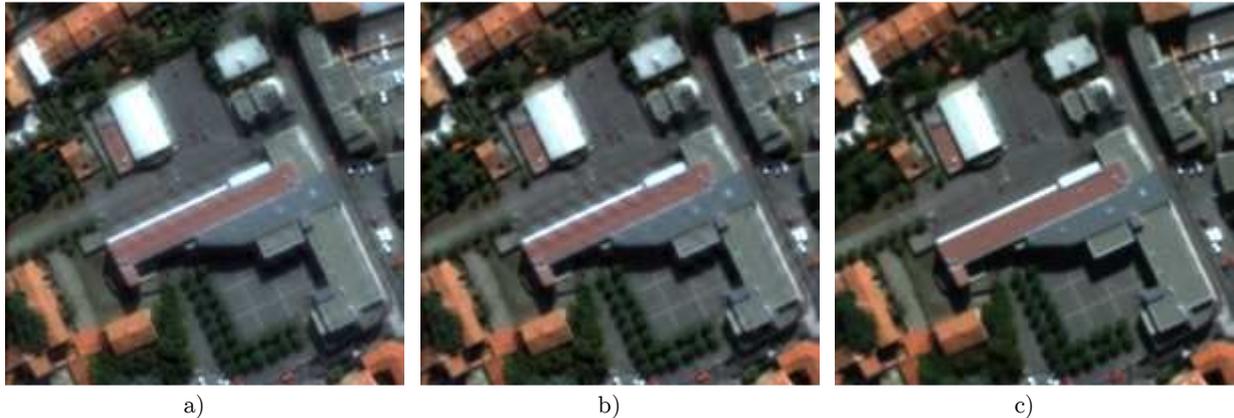
Dans la figure 4.13 on voit entre les échantillons 10 et 100 que le contraste est différent selon les images. Autour de l'échantillon 70 on peut voir que l'adaptation de FTM permet d'obtenir un contraste plus important qui permet une représentation de la géométrie plus proche de l'image de référence.



Figure 4.14 : Extrait de l'image de Toulouse, cas bâtiment blanc. a) ATWT-M3, b) ATWT-M3-MTFadapted, c) référence.

La figure 4.14 présente le toit d'un bâtiment blanc proche de la prison St Michel. Sur ce toit se trouvent des ouvertures que l'on distingue nettement dans la figure 4.14-c. Ces ouvertures sont invisibles dans la figure 4.14-a, alors que la prise en compte de la FTM permet d'identifier la rangée d'ouvertures se trouvant le plus à gauche. De plus, le contour du bâtiment est aussi plus net dans la figure 4.14-b que dans la figure 4.14-a. On voit dans cet exemple que l'adaptation de la FTM permet d'obtenir plus d'information dans l'image synthétisée, ici avec « l'apparition » des ouvertures situées sur le toit et un meilleur rendu du bord du bâtiment très contrasté avec son environnement.

Cette méthode permet visuellement une amélioration importante vis-à-vis du contraste de la netteté. Mais celle-ci présente cependant un défaut important. On peut observer dans la figure 4.15 que l'adaptation de FTM dans le cas présent amplifie l'importance de l'artefact de repliement de spectre. Ceci n'est pas dû spécifiquement à l'adaptation de FTM, mais aussi aux ondelettes utilisées ici qui laissent passer l'information légèrement au-delà de $0,125 f_e$ et donc ne suppriment pas les fréquences susceptibles de contenir du repliement de spectre comme c'est le cas ici.



a) b) c)
Figure 4.15 : Extrait de l'image de Toulouse, cas bâtiment rouge sujet au repliement de spectre. a) ATWT-M3, b) ATWT-M3-MTFadapted, c) référence.

Les exemples présentés précédemment montrent des améliorations notables sur certaines zones. Ces améliorations sont présentes dans l'ensemble de l'image de façon plus ou moins importante en fonction du contexte. Plus les structures présentent des contrastes importants et plus l'amélioration est notable. Globalement l'image est plus nette et présente de meilleurs contrastes qui permettent une meilleure lisibilité de l'image.

4.2.3.3 Analyse des FTM images synthétisées

L'outil MTF_{dev} présenté dans le §3.6.3 permet de caractériser la FTM de l'image synthétisée. Dans cette partie nous allons présenter les résultats obtenus dans deux situations:

- Situation dite "oracle", nous disposons de l'image MS de référence
- Situation non oracle nous n'avons pas d'image de référence.

Nous commençons avec la situation "oracle". Cette situation est rendue possible du fait que nous utilisons des images aéroportées qui simulent précisément ce que seront les images du satellite, PLEIADES dans ce cas. La simulation permet bien entendu de générer les images MS à haute résolution. Ces images ne seront pas fournies par le satellite, mais constituent l'image que nous cherchons à atteindre par le processus de fusion.

Les résultats de MTF_{dev} sont présentés par la figure 4.16 pour les différentes bandes spectrales et comparent la FTM de l'image ATWT-M3 et l'image ATWT-M3-MTFadapted à celle de l'image de référence. On rappelle qu'une image avec la FTM idéale donnerait une droite de valeur unitaire.

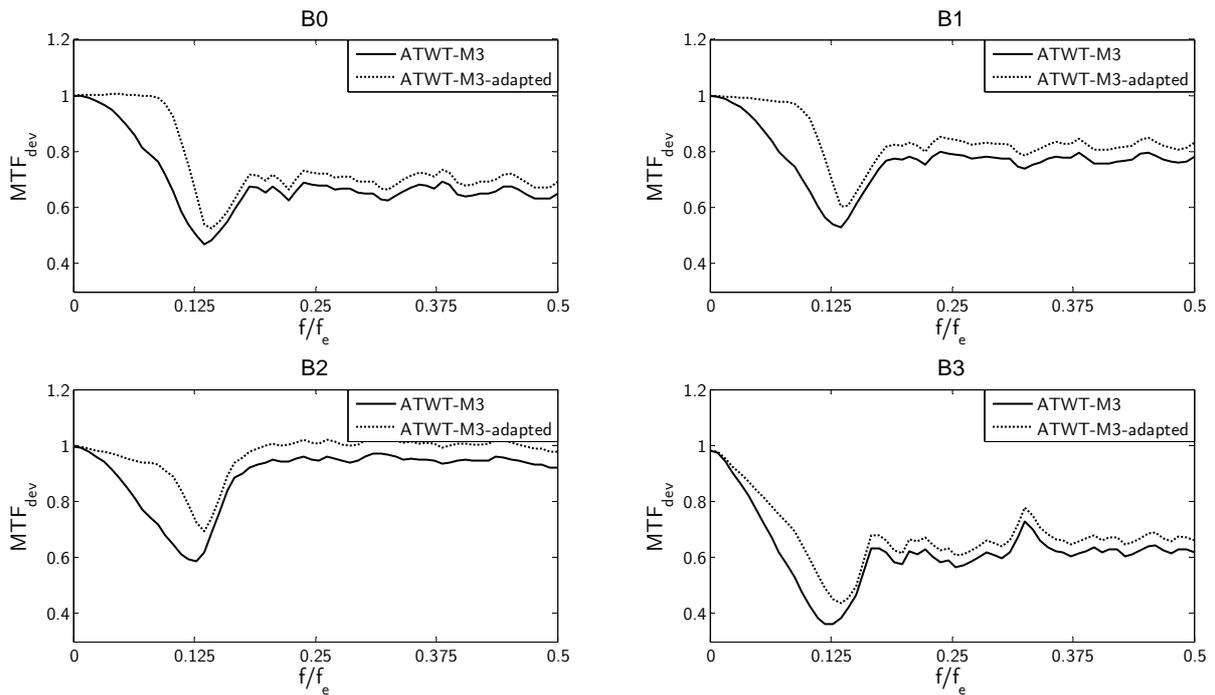


Figure 4.16 : MTF_{dev} calculé par rapport aux images de références, images PLEIADES de Toulouse.

Comme nous le spécifions lors de la présentation de MTF_{dev} il est nécessaire de contrôler la valeur du résidu pour assurer la pertinence des résultats obtenus avec MTF_{dev} . Dans le tableau 4.6 nous présentons la valeur de la norme du résidu lié au calcul de MTF_{dev} présenté dans la figure 4.16. On constate dans ce tableau que la valeur du résidu est à peu près équivalente pour ATWT-M3 et ATWT-M3-MTFadapted. Ainsi, l'amélioration constatée avec MTF_{dev} n'est pas transférée dans le résidu. L'interprétation des courbes de MTF_{dev} est donc pertinente.

Tableau 4.6 : norme du résidu pour le calcul de MTF_{dev} normalisée en pourcent par la norme de l'image de référence, image de Toulouse

	bleu	vert	rouge	PIR
ATWT-M3	3,6	3,25	2,97	4,80
ATWT-M3-MTFadapted	3,53	3,18	2,92	4,66

L'adaptation de FTM se fait sur les images MS. Par conséquent, son effet aura des conséquences uniquement pour les fréquences comprises entre 0 et $0,125 f_e$. On constate bien dans la figure 4.16 que le fait d'adapter la FTM permet de relever la FTM et de lui donner une allure plus proche de la FTM idéale du fait que l'allure de MTF_{dev} est plus proche de 1. Le tableau 4.6 présente un résultat très important, car il montre que, par une transformation linéaire mise en place par l'adaptation de FTM, on obtient à la fois une amélioration de la partie linéaire observée avec MTF_{dev} , mais aussi de la partie non linéaire du fait que le résidu est moins important lors de la prise en compte de la FTM. Ceci qui prouve bien l'intérêt de ce processus.

On constate que l'adaptation de FTM n'a pas un comportement tout à fait identique selon les bandes spectrales. Cette différence de comportement est due notamment aux informations contenues dans l'image. On observe une amélioration plus faible pour la bande B3 qui fait l'acquisition des longueurs d'ondes du PIR. Pour le canal PIR, la végétation a des valeurs importantes alors que la plupart des structures humaines ont pour ce canal des valeurs faibles. La déconvolution va relever de façon plus importante les hautes fréquences. Or la végétation est plutôt constituée de basses fréquences. Ainsi, la déconvolution va moins rehausser la FTM pour ce canal que pour les autres canaux.

La figure 4.17 et la figure 4.18 présentent les résultats des calculs de MTF_{dev} pour les images SPOT-5. La prise en compte de la FTM se fait sur les images MS. L'adaptation de la FTM aura donc un effet uniquement sur les fréquences allant de 0 à $0,125 f_e$. On observe bien une amélioration de la FTM sur cette bande de fréquence car la courbe ATWT-M3- $MTF_{adapted}$ est plus proche de 1 que la courbe de ATWT-M3. Cependant, l'amélioration est plus faible que dans le cas de PLEIADES, du fait que la différence de FTM entre les images MS et PANsupermode dans le cas de SPOT-5 est plus faible que dans le cas PLEIADES.

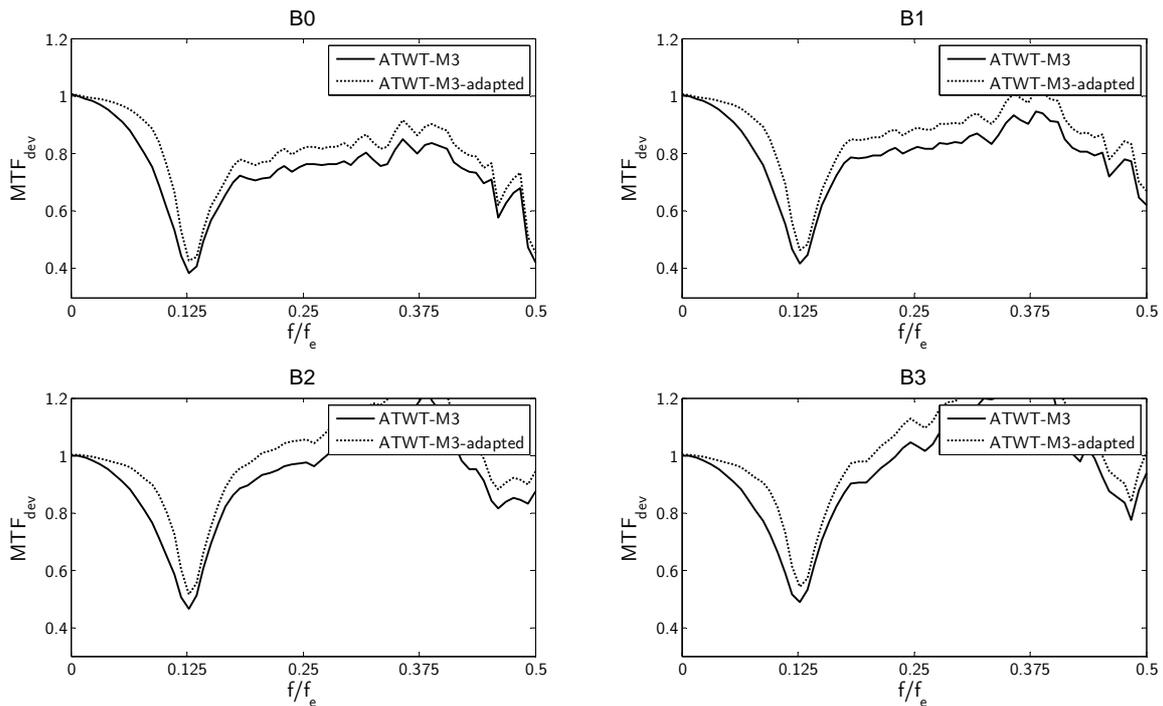


Figure 4.17 : MTF_{dev} calculé par rapport aux images de références, images SPOT-5 de Madrid.

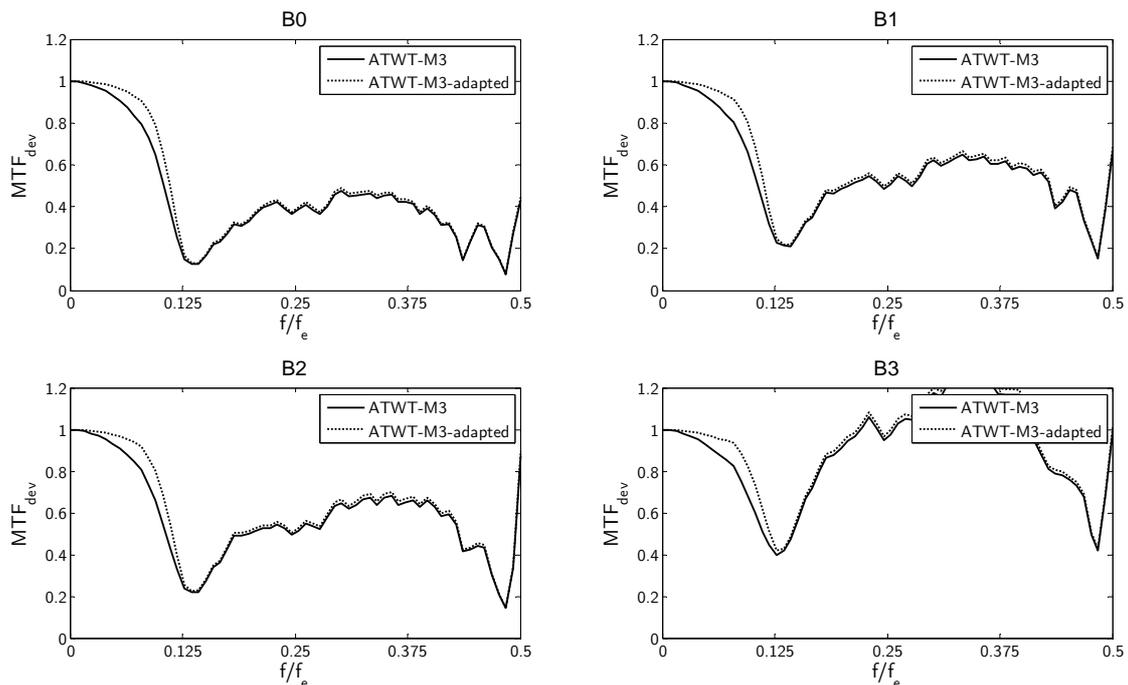


Figure 4.18 : MTF_{dev} calculé par rapport aux images de références, images SPOT-5 de Fredericton.

Passons au cas où MTF_{dev} est calculé dans une situation « non oracle ». Dans ce cas, MTF_{dev} est calculé entre chaque canal de l'image synthétisée et l'image PAN. Le fait de comparer des bandes spectrales différentes fournit une information moins pertinente que lorsque le calcul est effectué entre des bandes spectrales identiques.

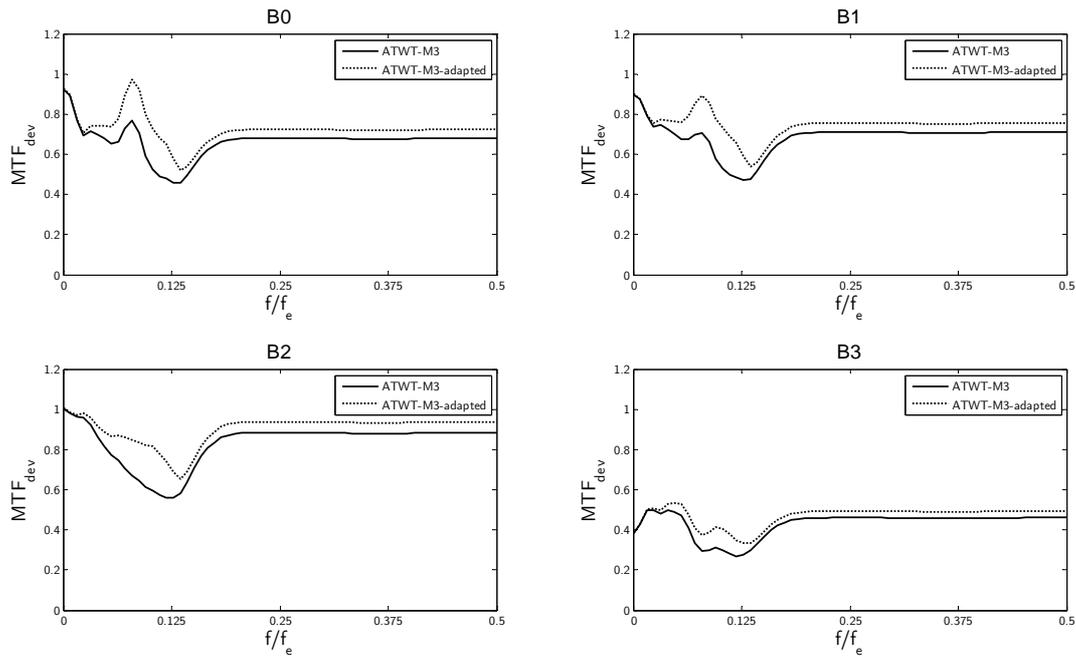


Figure 4.19 : MTF_{dev} calculé par rapport à l'image PAN MTF_{dev} , images PLEIADES de Toulouse.

Cependant, on observe dans la figure 4.19 un comportement assez semblable. Le défaut de FTM que l'on constate à $0,125 f_e$ est représenté de façon similaire dans le cas « oracle » (figure 4.16) que dans le cas « non-oracle » (figure 4.19). Les informations fournies par MTF_{dev} dans cette situation « non-oracle » restent pertinentes.

4.3 Conclusion

Cette nouvelle méthode de fusion ATWT-M3- $MTF_{adapted}$ ajoute la prise en compte de la FTM à la méthode ATWT-M3, prise comme référence. Avec MTF_{dev} nous avons vérifié que la FTM de l'instrument virtuel qui aurait acquis les images que nous synthétisons avec ATWT-M3- $MTF_{adapted}$ est plus proche de celle d'un capteur à haute résolution spatiale que ne l'est celle issue de la méthode de référence. La prise en compte de la FTM permet ainsi de mieux restituer la géométrie des structures présentes dans l'image. Nous avons montré par exemple, que certaines structures étaient rendues visibles grâce à l'adaptation de FTM. Cette dernière permet aussi une amélioration de la netteté dans l'image ; des halos sont encore présents mais leur intensité est sensiblement atténuée.

Ces observations visuelles sont confirmées par les bilans qualité établis par rapport à l'image de référence. Toutes les distances calculées présentent des valeurs plus proches de la valeur idéale lorsque la FTM est prise en compte. L'adaptation de FTM permet bien une amélioration de la qualité spatiale et spectrale des images synthétisées.

Cependant, cette méthode est sensible au bruit radiométrique. Le fait d'appliquer une déconvolution lors de l'adaptation de la FTM entraîne une amplification du bruit radiométrique dans l'image synthétisée.

Par ailleurs, le repliement de spectre se trouve essentiellement dans les hautes fréquences des images MS et induit des artefacts visuels. Le fait de déconvoluer améliore le contraste des hautes fréquences, mais aussi celui des fréquences responsables des artefacts de repliement de spectre.

Cette méthode présente des résultats globalement intéressants. Pour que cette méthode soit vraiment performante en tous points, il faudrait éviter d'amplifier le bruit radiométrique ainsi que les artefacts liés au repliement de spectre.

5 Compensation du rapport de FTM (MTF_{dev})

Ce chapitre propose une méthode, ATWT-M3-MTFequalized qui est la troisième contribution de mes travaux. Cette méthode est conçue de façon à prendre en compte la différence de FTM, en évitant de toute amplification du repliement de spectre et du bruit.

Cette méthode compense la différence de FTM entre les voies PAN et MS en injectant dans les basses fréquences, de l'information provenant de l'image PAN. Le point important est le fait qu'en calculant MTF_{dev} entre l'image synthétisée (ici par la méthode ATWT-M3) et l'image PAN, il est possible d'observer le défaut de FTM que l'on constate autour de la fréquence de Nyquist de l'image MS dans l'image synthétisée. Ainsi la quantité d'information provenant de l'image PAN à réinjecter est déduite du résultat de MTF_{dev} entre l'image ayant déjà été synthétisée (ici par la méthode ATWT-M3) et l'image PAN pour la fréquence $0,125 f_e$.

5.1 ATWT-M3-MTFequalized

5.1.1 Principe

Cette méthode se déroule en deux temps. Dans un premier temps, la fusion ATWT-M3 est effectuée entre les images MS et PAN. Le deuxième temps consiste alors à mesurer le défaut de FTM au niveau de $0,125 f_e$. Pour ce faire, nous calculons MTF_{dev} entre chaque bande de l'image synthétisée avec ATWT-M3 et l'image PAN. La façon de mesurer le défaut de FTM est présentée par la figure 5.1 pour un canal arbitraire.

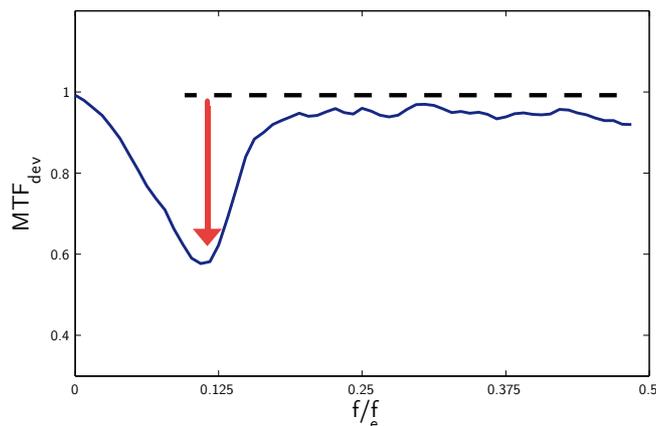


Figure 5.1 : Mesure du défaut de FTM.

La valeur de ce défaut de FTM va varier pour chaque bande. Une fois mesuré, il faut compenser ce défaut de MTF_{dev} . Nous effectuons alors une deuxième « injection » de coefficients d'ondelettes PAN dans l'image synthétisée. Les coefficients d'ondelettes injectés

seront les coefficients d'ondelettes qui se trouvent autour de 0,125 fe. L'ajout de ces coefficients d'ondelette sera pondéré par la mesure de MTF_{dev} à la fréquence 0,125 fe afin de réinjecter des coefficients d'ondelette pour chaque bande avec les bons niveaux d'amplitude. La figure 5.2 présente le schéma de principe de la méthode ATWT-M3-MTFequalized.

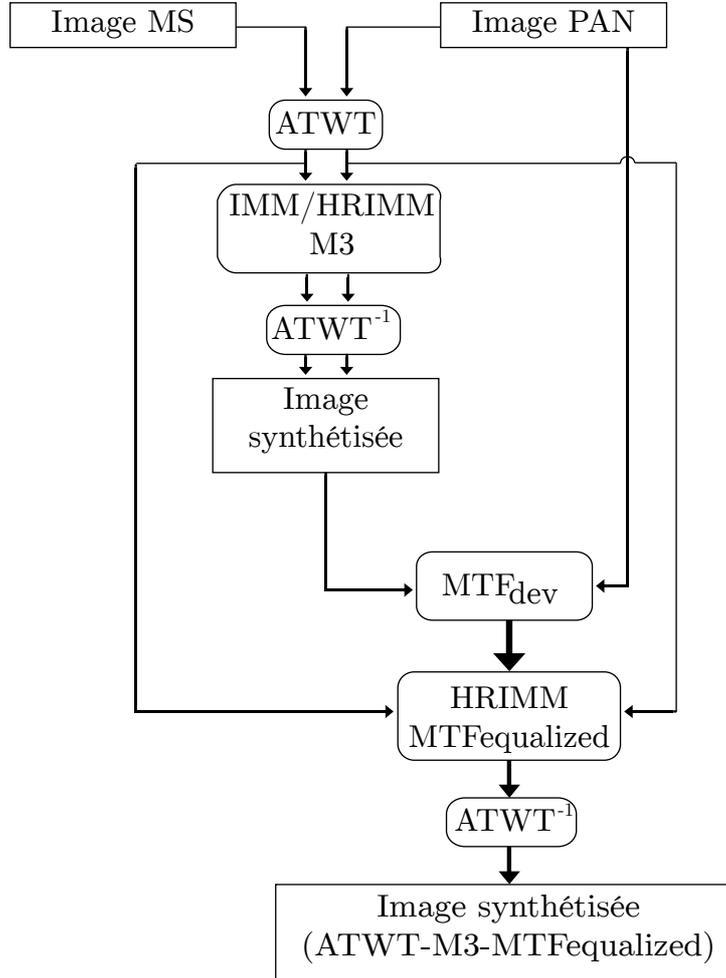


Figure 5.2 : Schéma de principe de la méthode ATWT-M3-MTFequalized

5.1.2 Description de la méthode

La méthode ATWT-M3 ayant déjà été présentée, nous nous contentons de décrire le HRIMM MTFequalized. Ce modèle a besoin comme information d'entrée, du résultat du calcul de MTF_{dev} entre l'image synthétisée par la méthode ATWT-M3 et l'image PAN à 0,125 fe. Cette valeur sera notée :

$$MTF_{dev} \left((B_{k,1})^{ATWT-M3}, PAN \right) (0,125 fe)$$

où $(B_{k,1})^{ATWT-M3}$ est l'image synthétisée avec la méthode ATWT-M3 et k indique le numéro de la bande spectrale considérée. On exprime alors le HRIMM MTFequalized :

$$D_{B_{k,0-1}} = (B_{k,1})^{ATWT-M3} + MTF_{dev} \left((B_{k,1})^{ATWT-M3}, PAN \right) (0,125 fe) (D_{A_{0,5-1}} + D_{A_{1-1,5}}) \quad 5.1$$

où $D_{B_{k,0-1}}$ représente les coefficients d'ondelettes (plans de détails) de l'image B pour la bande spectrale k et pour les fréquences spatiales comprises entre res_0 et res_1 .

La compensation du niveau faible de MTF_{dev} se fait ici en exploitant la transformée en ondelette déjà effectuée à travers la méthode ATWT-M3. La transformée en ondelette n'est pas une décomposition précise en fréquence. La mise en œuvre de ce principe pourrait être plus performante en effectuant la compensation du défaut de MTF_{dev} avec une transformation différente, plus précise en fréquence autour de $f_e/8$, comme des bancs de filtres ou des paquets d'ondelettes.

5.1.3 Les résultats

Nous ne cherchons pas spécifiquement à évaluer les performances intrinsèques de la méthode ATWT-M3- $MTF_{equalized}$, mais plus l'apport de performance par rapport à ATWT-M3.

L'évaluation des performances se fera de la même façon que pour la méthode ATWT-M3- $MTF_{adapted}$. Les images synthétisées par les méthodes ATWT-M3 et ATWT-M3- $MTF_{equalized}$ seront comparées aux images de références. Les comparaisons entre ces images seront de plusieurs types : analyse visuelle (qualitative), analyse quantitative (bilan qualité), et en terme de FTM (MTF_{dev}).

Analyses quantitatives

De la même façon que nous l'avons fait pour ATWT-M3- $MTF_{adapted}$, nous calculons les bilans de qualité entre les images synthétisées et les images de références. Ici nous allons évaluer l'apport de la prise en compte de la FTM dans la méthode ATWT-M3- $MTF_{equalized}$ par rapport à ATWT-M3. Les bilans sont présentés dans les tableaux 5.1 à 5.3 pour respectivement les images de Toulouse, Madrid et Fredericton.

Tableau 5.1 : Bilan qualité image PLEIADES pour Toulouse

Bande spectrale	Indice	Valeur idéale	ATWT-M3	ATWT-M3-MTFequalized
Bleu	biasRel (%)	0	-0,001	0,001
	diffVarRel (%)	0	14	1
	σ Rel (%)	0	8,7	7,7
	cc	1	0,983	0,986
	ccHF	1	0,920	0,915
Vert	biasRel (%)	0	0,0	0,0
	diffVarRel (%)	0	16	0
	σ Rel (%)	0	8,6	6,5
	cc	1	0,988	0,991
	ccHF	1	0,961	0,958
Rouge	biasRel (%)	0	0,0	0,0
	diffVarRel (%)	0	16	-1
	σ Rel (%)	0	9,3	7,7
	cc	1	0,991	0,992
	ccHF	1	0,969	0,966
PIR	biasRel (%)	0	0,0	0,0
	diffVarRel (%)	0	34	26
	σ Rel (%)	0	13,7	10,7
	cc	1	0,967	0,979
	ccHF	1	0,878	0,886
Indices multimodaux	ERGAS	0	2,6	2,1
	mean SAM	0	3,1	3,5

Tableau 5.2 : Bilan qualité image SPOT-5 supermode pour Madrid (Espagne)

Bande spectrale	Indice	Valeur idéale	ATWT-M3	ATWT-M3-MTFequalized
Bleu	biasRel (%)	0	-0,003	-0,003
	diffVarRel (%)	0	21	-5
	σ Rel (%)	0	2,8	2,7
	cc	1	0,981	0,980
	ccHF	1	0,902	0,900
Vert	biasRel (%)	0	0,0	0,0
	diffVarRel (%)	0	22	-5
	σ Rel (%)	0	3,8	3,1
	cc	1	0,984	0,986
	ccHF	1	0,929	0,935
Rouge	biasRel (%)	0	0,0	0,0
	diffVarRel (%)	0	21	-6
	σ Rel (%)	0	4,8	3,6
	cc	1	0,986	0,990
	ccHF	1	0,942	0,955
PIR	biasRel (%)	0	0,0	0,0
	diffVarRel (%)	0	19	-7
	σ Rel (%)	0	5,0	4,2
	cc	1	0,987	0,989
	ccHF	1	0,933	0,949
Indices multimodaux	ERGAS	0	1,0	0,9
	mean SAM	0	0,8	0,7

- **Biais relatif** : cette valeur est déjà très faible dans le cas d'ATWT-M3, $MTF_{equalized}$ ne modifie pas la valeur de cet indice et permet de conserver un biais extrêmement faible.
- **Coefficient de corrélation** : les valeurs obtenues par la méthode ATWT-M3 sont proches de la valeur idéale. La compensation de la différence de FTM permet d'améliorer encore un peu le coefficient de corrélation et ceci pour la majorité des cas. Dans certains cas, il arrive que cette valeur soit seulement conservée.
- **Coefficient de corrélation des hautes fréquences** : cet indice n'est pas amélioré de manière générale.
- **Différence en variance relative** : cette distance permet de quantifier l'innovation. $MTF_{equalized}$ permet de réduire de façon importante la valeur de cette distance. Dans certains cas, cette valeur devient légèrement négative ce qui signifie une injection trop importante de hautes fréquences.
- **Écart-type de l'image de différence** : cette distance est améliorée de façon non négligeable puisqu'on obtient un gain compris entre 10 % et 30 %.

Distances multimodales

- **ERGAS** : la valeur de l'ERGAS est réduite de 10 à 30 %, l'amélioration de cette distance est assez importante. L'adaptation de FTM permet l'amélioration de la qualité spectrale de l'image synthétisée.
- **Mean SAM** : de la même façon que l'ERGAS, la valeur de cette distance est réduite grâce à l'adaptation de FTM. L'amélioration de cette distance est comprise entre 5 et 10 %.

Les bilans qualité présentés à travers les tableaux ci-dessus montrent que la majorité des distances est améliorée par $MTF_{equalized}$. Le gain apporté par $MTF_{equalized}$ est légèrement plus faible pour les images SPOT-5 (tableau 5.2) que pour PLEIADES puisque la différence de FTM entre les images MS et PAN est plus faible dans le cas de SPOT-5.

Tableau 5.3 : Bilan qualité image SPOT-5 supermode pour Fredericton (Canada)

Bande spectrale	Indice	Valeur idéale	ATWT-M3	ATWT-M3-MTFequalized
Bleu	biasRel (%)	0	-0,006	0,005
	diffVarRel (%)	0	12	12
	σ Rel (%)	0	5,5	6,4
	cc	1	0,978	0,963
	ccHF	1	0,740	0,676
Vert	biasRel (%)	0	0,0	0,0
	diffVarRel (%)	0	10	8
	σ Rel (%)	0	7,1	7,6
	cc	1	0,983	0,970
	ccHF	1	0,796	0,751
Rouge	biasRel (%)	0	0,0	0,0
	diffVarRel (%)	0	10	6
	σ Rel (%)	0	10,4	10,3
	cc	1	0,984	0,974
	ccHF	1	0,790	0,764
PIR	biasRel (%)	0	0,0	0,0
	diffVarRel (%)	0	5	-5
	σ Rel (%)	0	8,6	8,1
	cc	1	0,992	0,961
	ccHF	1	0,890	0,833
Indices multimodaux	ERGAS	0	2,0	2,1
	mean SAM	0	1,9	2,8

Les résultats pour Fredericton sont moins bons que pour Madrid. Certaines distances sont non améliorées, d'autres sont légèrement dégradées. Ceci est dû à une zone moins urbaine et plus périurbaine avec plus de végétation. Une autre raison peut être un recalage perfectible.

En ce qui concerne les distances multimodales, l'ERGAS est amélioré pour toutes les images, mais le mean SAM est dégradé dans le cas de l'image PLEIADES.

- **Détection de la netteté**

La détection de la netteté est effectuée de la même façon que dans le chapitre précédent en calculant sur des vignettes de 64 pixels et en faisant la somme des valeurs obtenues pour chaque bande spectrale.

L'amélioration globale sur l'image obtenue par l'adaptation de FTM est de 6 %. Ce résultat est un peu meilleur pour ATWT-M3-MTFequalized que pour ATWT-M3-MTFadapted, car la proportion de fréquence rehaussée est plus importante.

5.1.3.1 Analyse visuelle

Pour les résultats visuels, nous utiliserons les mêmes exemples que de la méthode ATWT-M3-MTFadapted.

Le premier exemple est présenté dans la figure 5.3. On voit dans la figure 5.3-a un halo autour de la voiture, lequel est inexistant dans l'image de référence (figure 5.3-c). Ce halo est nettement atténué dans la figure 5.3-b par rapport à la figure 5.3-a. En plus de réduire le halo autour de la voiture, la figure 5.3-b présente des contrastes un peu plus importants et des couleurs plus vives.

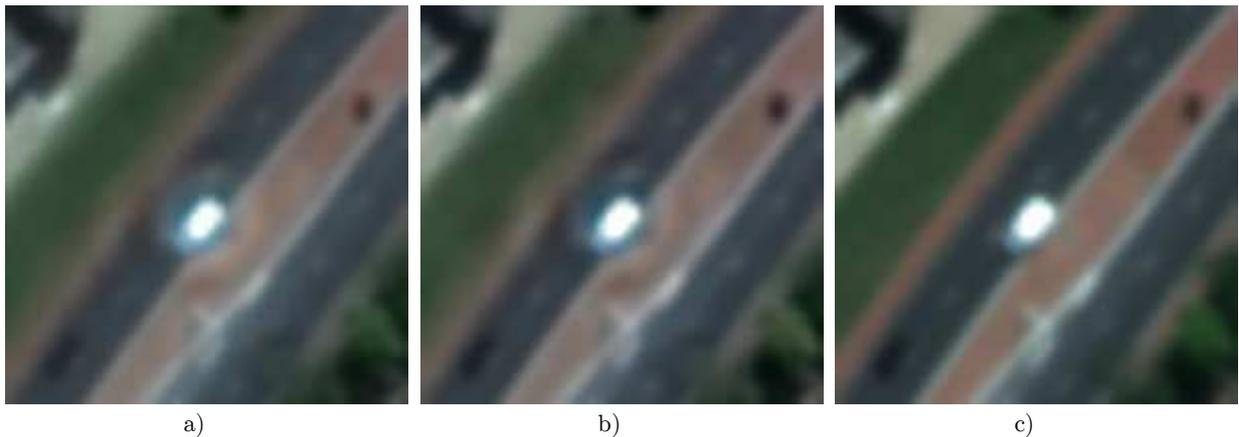


Figure 5.3 : Extrait de l'image de Toulouse, cas de la voiture blanche. a) ATWT-M3, b) ATWT-M3-MTFequalized, c) référence.

Le deuxième exemple dans l'image de Toulouse (figure 5.4) illustre le gain en qualité géométrique. La figure 5.4-b offre un aspect plus net et de meilleurs contrastes que la figure 5.4-a. Pour ce cas, nous avons constaté que la méthode ATWT-M3-MTFadapted avait tendance à amplifier l'importance de l'artefact de repliement de spectre que l'on distingue dans la figure 5.4-a, mais qui n'est pas présent dans l'image de référence (figure 5.4-c). On voit dans la figure 5.4-b que l'artefact de repliement de spectre est atténué de façon assez importante. Avec cette méthode, nous arrivons à obtenir des améliorations assez similaires à la méthode ATWT-M3-MTFadapted tout en diminuant l'artefact de repliement de spectre.



Figure 5.4 : Extrait de l'image de Toulouse, cas bâtiment rouge sujet au repliement de spectre. a) ATWT-M3, b) ATWT-M3-MTFequalized, c) référence.

Dans le troisième exemple (figure 5.5) on constate que la figure 5.5-a ne donne pas accès à la totalité de l'information, car il est impossible de distinguer les rangées d'ouvertures présentes sur le toit bien visibles dans la figure 5.5-c. Par contre, on distingue bien ces ouvertures dans la figure 5.5-b. Cependant, la figure 5.5-b présente une légère distorsion spectrale au niveau de ces ouvertures qui ont un aspect bleuté plus important que dans l'image de référence (figure 5.5-c). On observe aussi que le contour du bâtiment est plus net dans la figure 5.5-b que dans la figure 5.5-a. Ici encore on constate que visuellement la méthode ATWT-M3-MTFequalized améliore de façon importante la qualité de l'image synthétisée. Nous avons noté une légère distorsion spectrale. Celle-ci ne se produit qu'au niveau des ouvertures et ne fait qu'exagérer un peu l'aspect bleuté déjà présent dans l'image de référence.

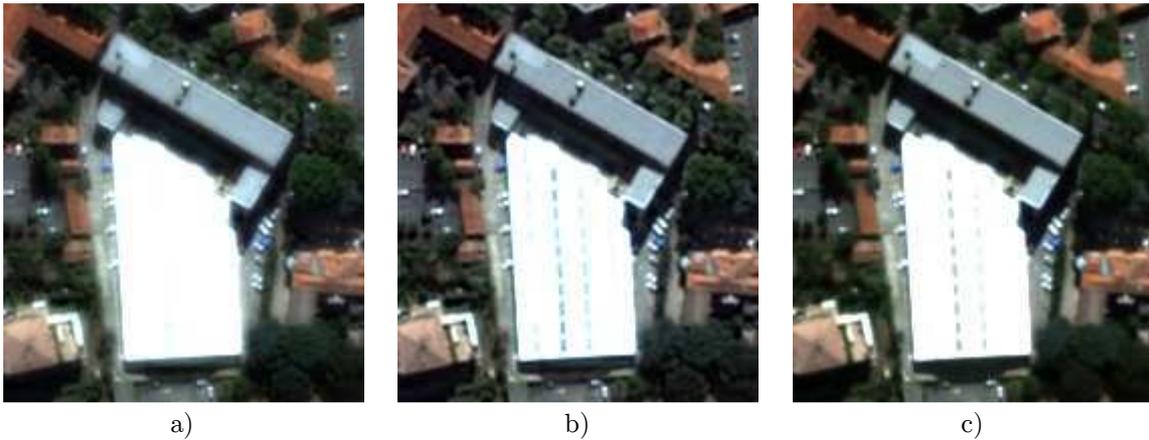


Figure 5.5 : Extrait de l'image de Toulouse, cas bâtiment blanc. a) ATWT-M3, b) ATWT-M3-MTFequalized, c) référence.

5.1.3.2 FTM des images synthétisées

Les trois figures ci-dessous offrent des comparaisons de MTF_{dev} pour ATWT-M3, ATWT-M3-MTFequalized et ATWT-M3-MTFadapted.

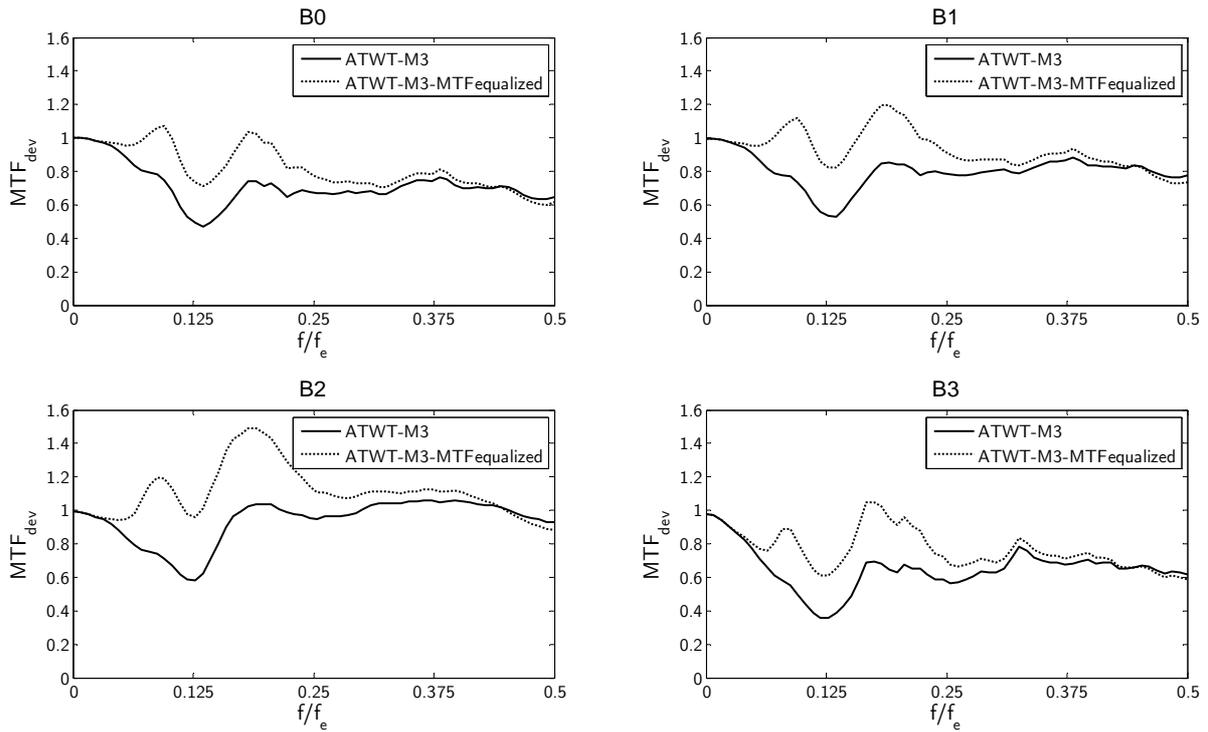


Figure 5.6 : MTF_{dev} pour les méthodes : ATWT-M3, ATWT-M3-MTFadapted, ATWT-M3-MTFequalized. Image PLEIADES de Toulouse.

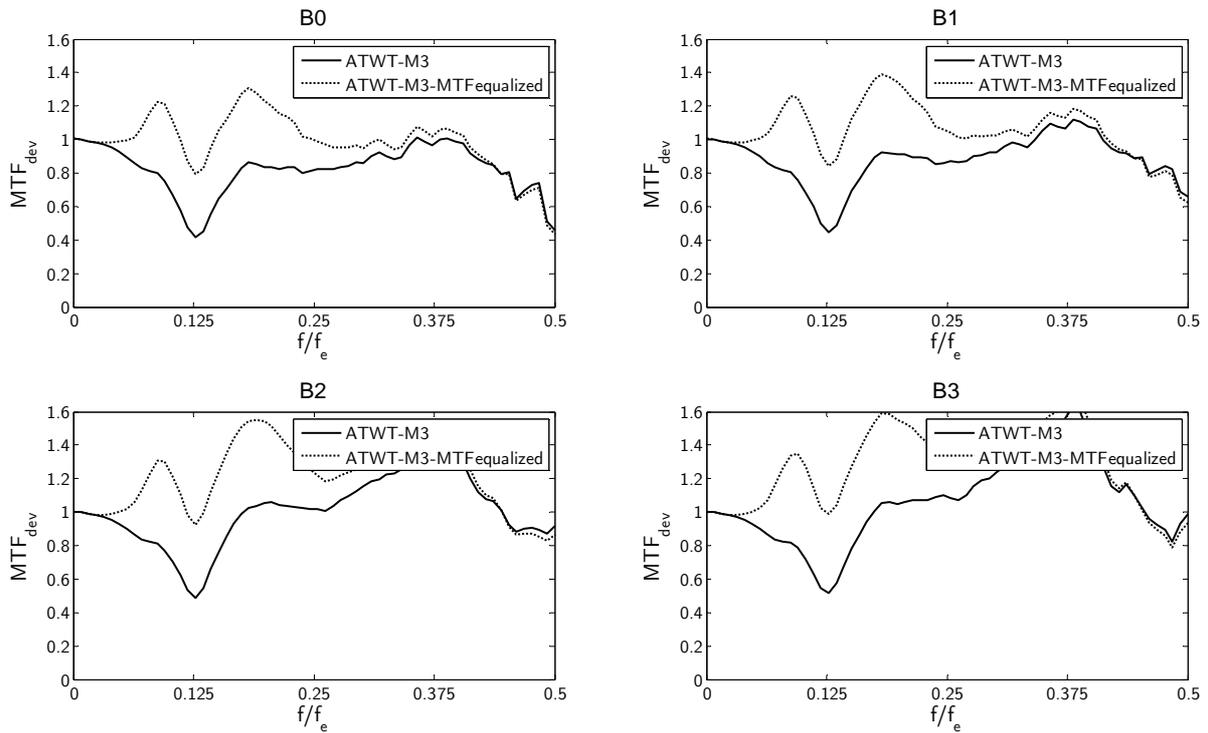


Figure 5.7 : MTF_{dev} pour les méthodes : ATWT-M3, ATWT-M3-MTFadapted, ATWT-M3-MTFequalized. Image SPOT-5 Madrid.

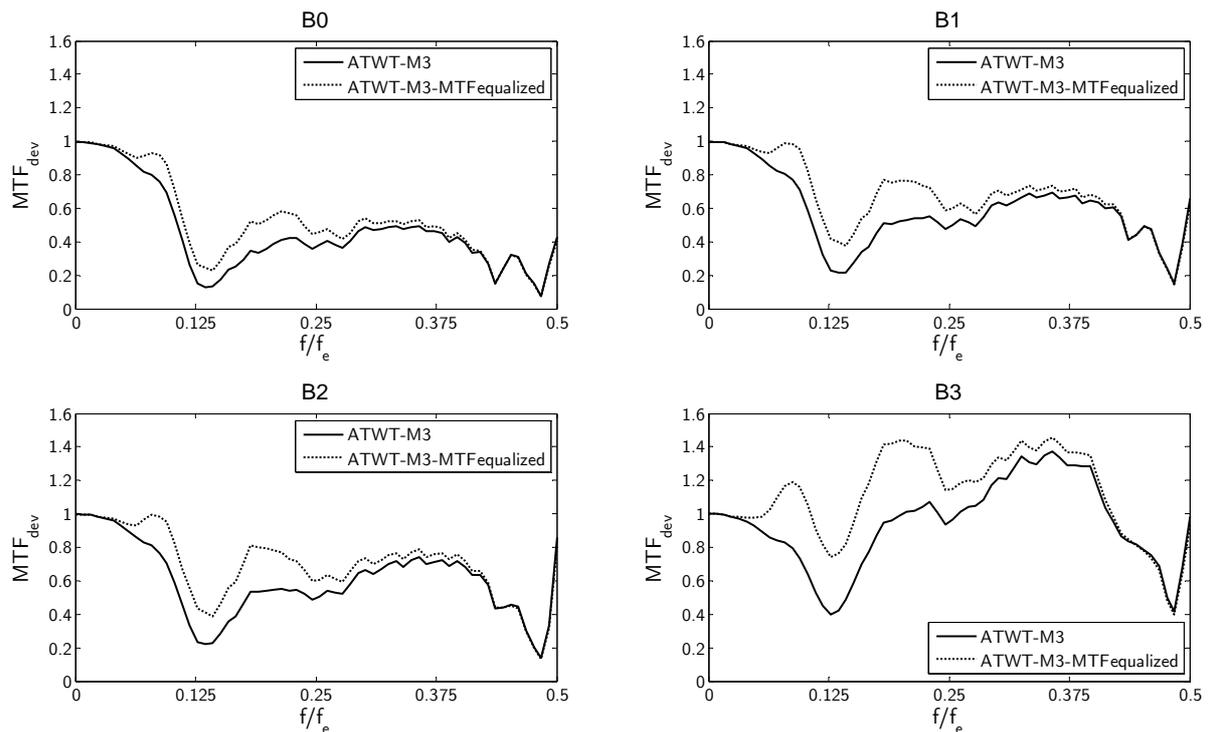


Figure 5.8 : MTF_{dev} pour les méthodes : ATWT-M3, ATWT-M3-MTFadapted, ATWT-M3-MTFequalized. Image SPOT-5 Fredericton.

On peut observer dans ces figures que dans la zone comprise entre 0 et $0,125 f_e$, la FTM est bien rehaussée et le défaut de FTM est atténué par la méthode ATWT-M3-MTFequalized. Pour ce premier quart des fréquences, la méthode ATWT-M3-MTFequalized permet une réelle amélioration de la FTM de l'image synthétisée par rapport à l'image synthétisée avec ATWT-M3.

Cependant, les résultats ne sont pas aussi bons au-delà de $0,125 f_e$. On observe après $0,125 f_e$ un rebond important pour toutes les modalités. Ce rebond n'est pas désirable notamment du fait que pour certains canaux, MTF_{dev} dépasse parfois 1. Il est dû à l'utilisation des ondelettes à trous lors de la deuxième injection de coefficients d'ondelettes PAN. MTF_{dev} apporte ici une information importante en nous permettant d'observer ce rebond. Pour éviter au maximum ce rebond, il faut être plus précis en fréquence autour de $0,125 f_e$. Utiliser une transformée en ondelettes pour effectuer la deuxième injection de coefficient de l'image PAN n'est peut-être pas la meilleure solution. Les ondelettes ne sont peut-être pas l'outil le plus approprié pour effectuer le HRIMM-MTFequalized. Un banc de filtre optimisé pour ce cas pourrait permettre d'éviter ce rebond.

5.2 Conclusion

Cette méthode permet une amélioration de la représentation de la géométrie. L'outil MTF_{dev} met bien en évidence que la FTM de l'instrument virtuel qui aurait acquis l'image synthétisée avec ATWT-M3-MTFequalized est plus proche de celle d'un instrument à haute résolution spatiale que ne l'est l'image issue de la méthode de référence (ATWT-M3). Cette méthode parvient à améliorer la netteté dans l'image et l'importance des halos est, quant à elle, diminuée. Les bilans qualité confirment les observations avec une amélioration des distances liées à la qualité spatiale. Les distances qui évaluent la qualité spectrale sont par contre légèrement dégradées. Ceci est dû au fait que la compensation du défaut de MTF_{dev} est réalisée à l'aide d'information provenant de l'image PAN, c'est-à-dire de moins bonne qualité spectrale. Cependant, cette méthode permet de conserver le niveau de bruit de la méthode de référence, car cette prise en compte de la FTM n'entraîne pas d'amplification du bruit radiométrique. Enfin, comme l'information introduite pour compenser le défaut de MTF_{dev} provient de l'image PAN, elle ne contient quasiment pas de repliement de spectre pour les fréquences considérées et ceci permet une atténuation d'artefact de repliement de spectre.

6 Influence du repliement de spectre et du SNR

Le fait d'utiliser des images issues de simulation permet de créer des images MS ne contenant pas de repliement de spectre. Dans ce chapitre nous évaluons l'effet du repliement de spectre sur la qualité des images synthétisée.

L'amélioration permise par la suppression du repliement de spectre est faible au regard des bilans de qualité, mais on perçoit visuellement l'atténuation des artéfacts liés à ce phénomène.

Un détecteur spécifique au repliement de spectre proposé par Blanchet a été mis en place afin de constater si ce détecteur était plus sensible que les bilans de qualités à la présence de repliement de spectre dans les images.

Enfin, nous montrons l'importance de la méthode de rééchantillonnage sur la qualité de l'image synthétisée, et la tendance de la méthode ATWT-M3-MTFadapted à amplifier le bruit radiométrique dans l'image synthétisée.

6.1 Repliement de spectre

Dans cette partie, nous allons montrer l'influence du repliement de spectre dans les images. Il est possible dans le cadre d'images de simulation de connaître le repliement de spectre que contiennent les images MS, et ainsi d'évaluer son impact sur la qualité de l'image synthétisée, que ce soit visuellement ou au travers de différentes distances ou critères de qualité image.

6.1.1 Images de simulation et repliement de spectre

Au cours du traitement simulant les images du satellite PLEIADES, à partir d'images du capteur aéroporté PELICAN, il est possible d'ajouter un filtre anti-repliement quand les images ont encore la haute résolution des images PELICAN.

$$I_{PLEIADES,sans\ repliement} = \left((I_{PELICAN} * f_{anti-repliement}) * FTM_{MS} \right) \downarrow 8 \quad 6.1$$

$$I_{PLEIADES,sans\ repliement} = (I_{PELICAN} * FTM_{MS}) \downarrow 8 \quad 6.2$$

Le filtre anti-repliement supprime les informations dont la fréquence dépasse celle de Nyquist de l'image MS. Le processus de simulation permet donc d'obtenir des images MS avec ou sans repliement de spectre.

6.1.2 Suppression du repliement de spectre et fusion

L'étape de ré-échantillonnage a été identifiée depuis longtemps comme une étape ayant un fort impact sur la qualité de l'image synthétisée (Khan *et al.* 2008) (Blanc *et al.* 1998) (Wald *et al.* 1997). Le ré-échantillonnage communément rencontré en fusion est de rapport 2 ou 4. Si l'échantillonnage est effectué en respectant le théorème de Shannon, alors il est considéré comme réversible. Dans ce cas, l'image reconstruite à partir des échantillons est identique en théorie à l'image originale.

Généralement, les conditions du théorème de Shannon ne sont pas respectées scrupuleusement par les imageurs optiques et l'échantillonnage ne permet pas de revenir à l'image initiale. Bien sûr, chaque méthode de ré-échantillonnage a des caractéristiques propres et certaines sont mieux adaptées à la fusion, comme celle utilisée par (Khan *et al.* 2008).

Les reproches que l'on fait à l'échantillonnage dans le domaine de la fusion sont notamment l'apparition d'artefacts. La figure 6.1 présente dans une zone urbaine un exemple d'artefacts dus au repliement de spectre. Les artefacts de repliement de spectre sont présents dans les images MS originales, mais leur faible taille fait qu'ils sont peu visibles comme on peut le constater dans la figure 6.1-a. Ils deviennent plus visibles dans l'image synthétisée (figure 6.1-b). Ainsi, le fait de suréchantillonner une image contenant du repliement de spectre a tendance à accentuer l'impact visuel, notamment dans le cas de la fusion où l'on ajoute de l'information haute fréquence qui est dépourvue de repliement de spectre.



Figure 6.1 : artefact de repliement de spectre. a) image MS originale à basse résolution spatiale. b) image synthétisée (ATWT-M3).

La présence de ce type d'artefact dépend aussi de la zone observée. Pour certaines zones plutôt homogènes, comme des grandes étendues agricoles, il y aura peu ou pas de zones affectées par du repliement de spectre. Les zones urbaines sont les zones les plus difficiles du fait de la grande variété des signatures spectrales et des hautes fréquences, ces dernières entraînant la présence d'artefact de repliement de spectre.

Dans la figure 6.2, nous présentons deux images obtenues par la même méthode de fusion, l'une contenant du repliement de spectre (figure 6.2-a) et l'autre dans laquelle le repliement de spectre a été supprimé (figure 6.2-b).



Figure 6.2 : Effet du repliement dans l'image synthétisée avec ATWT-M3. a) image MS originale. b) images MS sans repliement de spectre

Dans la figure 6.2-a, on distingue des stries blanches. Ces dernières sont moins présentes dans la figure 6.2-b. Le fait de supprimer le repliement de spectre dans les images MS permet bien de réduire les artefacts dans l'image synthétisée.

6.1.3 Amélioration de la qualité par suppression de repliement de spectre

Nous observons visuellement sur certaines zones que l'absence de repliement permet de supprimer des artefacts (figure 6.2). Le calcul de bilan qualité entre les images synthétisées et l'image de référence confirme les constatations visuelles (tableau 6.1).

Tableau 6.1 : Bilan qualité avec ou sans repliement de spectre dans les images MS

Modalité	Indice	Valeur idéale	ATWT-M3	ATWT-M3 sans repliement
Bleu	biasRel (%)	0	0,406	1,7
	diffVarRel (%)	0	11	13
	σ Rel (%)	0	10,2	10,1
	cc	1	0,981	0,982
	ccHF	1	0,900	0,901
Vert	biasRel (%)	0	0,4	1,7
	diffVarRel (%)	0	13	14
	σ Rel (%)	0	9,8	9,8
	cc	1	0,986	0,986
	ccHF	1	0,947	0,948
Rouge	biasRel (%)	0	0,3	1,6
	diffVarRel (%)	0	14	15
	σ Rel (%)	0	10,5	10,6
	cc	1	0,989	0,989
	ccHF	1	0,951	0,952
PIR	biasRel (%)	0	0,5	1,2
	diffVarRel (%)	0	31	31
	σ Rel (%)	0	15,4	15,4
	cc	1	0,954	0,955
	ccHF	1	0,821	0,821
Indices multimodaux	ERGAS	0	2,9	3,0
	mean SAM	0	3,3	3,3

Les distances les plus pertinentes vis-à-vis du repliement de spectre et des artefacts sont celles qui caractérisent les aspects de qualité spatiale, le coefficient de corrélation et le coefficient de corrélation des hautes fréquences. On voit dans le tableau 6.1 que l'image synthétisée à partir d'une image MS sans repliement de spectre permet une légère amélioration de ces coefficients par rapport à l'image synthétisée à partir de l'image MS contenant du repliement de spectre. L'amélioration reste très faible, car, les artefacts dus au repliement de spectre représentent seulement de petites zones dans l'image.

- **Détection de repliement de spectre par variation totale.**

Nous avons appliqué le détecteur de repliement de spectre présenté dans le chapitre 3. Nous l'avons tout d'abord appliqué aux images MS (tableau 6.2) et nous constatons que cette distance permet bien à basse résolution spatiale de déterminer l'image contenant le plus de repliement de spectre.

Tableau 6.2 : Détection du repliement de spectre dans les images MS

Quantité de repliement	MS sans repliement de spectre	MS originale avec repliement de spectre
Bleu	1,42	1,90
Vert	1,37	1,65
Rouge	1,22	1,49
PIR	1,26	1,52

Les résultats de la détection du repliement de spectre dans les images synthétisées sont présentés dans les tableaux ci-dessus. La détection est calculée pour différentes zones de l'image et des tailles de vignettes différentes. Ce détecteur met en jeu le logarithme d'une différence. Si cette différence est trop faible, alors le résultat est infini.

Tableau 6.3 : détection du repliement de spectre image synthétisée bande B0

Centre	taille	sans repliement	avec repliement
512;512	256x256	1,5218	1,902
300;300	256x256	0,5146	0,5291
300;700	256x256	0,7377	0,7377
700;300	256x256	-∞	-∞
700;300	100x100	0,0212	0,0223
700;700	256x256	2,6	-∞
800;260	200x200	0,3579	0,5111
260;260	200x200	-∞	2,2999
260;800	200x200	0,181	0,1996
512;512	480x480	1,8228	1,902
210;210	200x200	0,4051	0,4106

Tableau 6.4 : détection du repliement de spectre image synthétisée bande B1

Centre	taille vignette	sans repliement	avec repliement
512;512	256x256	1,601	1,9989
300;300	256x256	0,7317	0,8375
300;700	256x256	0,6767	0,7317
700;300	256x256	-∞	-∞
700;300	100x100	0,0399	0,0411
700;700	256x256	-∞	-∞
800;260	200x200	1,3705	-∞
260;260	200x200	-∞	2,6010
260;800	200x200	0,3199	0,4023
512;512	480x480	1,4249	1,3969
210;210	200x200	0,4674	0,5254

Tableau 6.5 : détection du repliement de spectre image synthétisée bande B2

Centre	taille vignette	sans repliement	avec repliement
512;512	256x256	1,902	1,9989
300;300	256x256	0,7089	0,6925
300;700	256x256	0,8375	0,9108
700;300	256x256	2,6010	-∞
700;300	100x100	0,0507	0,0532
700;700	256x256	2,6010	-∞
800;260	200x200	1,6467	2,1239
260;260	200x200	-∞	-∞
260;800	200x200	0,5254	0,5798
512;512	480x480	1,186	1,1386
210;210	200x200	0,6716	0,7881

Tableau 6.6 : détection du repliement de spectre image synthétisée bande B3

centre	taille	sans repliement	avec repliement
512;512	256x256	1,5218	1,5218
300;300	256x256	0	0
300;700	256x256	0,0495	0,1167
700;300	256x256	2,601	2,601
700;300	100x100	0,0246	0,0246
700;700	256x256	1,6467	1,5218
800;260	200x200	0,0223	0,0258
260;260	200x200	1,0447	1,1096
260;800	200x200	0,0972	0,0972
512;512	480x480	2,2999	2,2999
210;210	200x200	0,2546	0,2871

Au vu des résultats, le détecteur ne donne pas toujours une valeur plus élevée pour l'image qui contient du repliement de spectre. A posteriori, le fait que ce détecteur ne fonctionne pas parfaitement sur nos images n'est pas très étonnant, car le principe de détection est lié aux hautes fréquences de l'image. Or celles-ci proviennent de l'image PAN qui ne contient pas de repliement du fait de la valeur faible de la FTM à sa fréquence de Nyquist. L'utilisation de ce détecteur de repliement de spectre n'est donc pas tout à fait adaptée, du fait que l'information haute fréquence provient de l'image PAN, qui est bien moins peu sensible au repliement de spectre, comparée à l'image MS.

6.1.4 Variation de la FTM en fonction de la suppression de repliement

Nous venons de montrer que l'absence de repliement de spectre dans les images MS permet la suppression d'artefacts visuels et une amélioration très faible des bilans de qualité. Nous allons étudier maintenant si la suppression du repliement de spectre permet une amélioration de la FTM de l'image synthétisée.

Le repliement de spectre contenu dans les hautes fréquences de l'image MS peut contribuer au défaut de FTM constaté autour de 0,125 fe. Pour vérifier cette idée, nous comparons les images synthétisées selon que les images MS fusionnées contiennent ou non du repliement de spectre. Un exemple de résultat est présenté dans la figure 6.3.

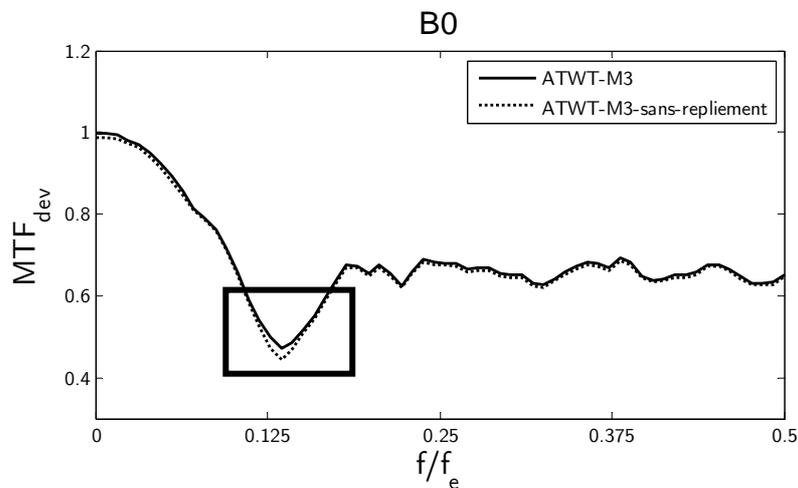


Figure 6.3 : Modification de la FTM due à la suppression du repliement de spectre.

La figure 6.3 représente MTF_{dev} par rapport à la FTM de l'image de référence, laquelle ne contient pas de repliement de spectre. Nous pensions que le fait de supprimer le repliement permettrait de réduire l'amplitude du défaut de FTM à 0,125 fe. C'est l'inverse du résultat que nous obtenons. La seule différence significative se trouve à 0,125 fe et la valeur de MTF_{dev} est plus faible quand l'image MS ne contient pas de repliement. Cet abaissement de la FTM a lieu pour la gamme de fréquences [0,1:0,15] identifiée précédemment comme présentant un risque de contenir du repliement de spectre.

Pour obtenir l'image sans repliement de spectre il a fallu appliquer un filtre anti-repliement dont la fréquence de coupure correspond à 0,125 fe. Le filtre appliqué n'étant pas parfait, l'application de ce dernier a inévitablement atténué des fréquences dans sa bande de transition, qui ne constituaient pas de repliement de spectre. Cette légère dégradation de MTF_{dev} serait donc la conséquence du filtre anti-repliement appliqué lors de la simulation.

Autrement dit, MTF_{dev} n'est pas sensible à la présence de repliement de spectre dans les images synthétisées. Ceci peut s'expliquer par le fait que le repliement de spectre effectif dans une image est un phénomène local et donc difficilement décelable par un outil d'analyse basé sur des moyennes glissantes de transformées de Fourier plutôt adaptées à des signaux stationnaires.

6.1.5 Ré-échantillonnage et artefact de repliement de spectre

Le fait de prendre une méthode de ré-échantillonnage de type spline donne la possibilité de choisir le degré utilisé. Plus le degré est important, plus le ré-échantillonnage aura une allure proche du ré-échantillonnage parfait sans considérer les rebonds.

Selon les bilans de qualité images, la qualité de l'image synthétisée est d'autant meilleure que le degré de la fonction spline est grand. Ceci est confirmé par : l'ERGAS, le coefficient de corrélation, ainsi que le coefficient de corrélation des hautes fréquences. Avec un ré-échantillonnage de type spline avec un degré élevé, il y aura une meilleure représentation du contraste des hautes fréquences de l'image à ré-échantillonner.

- Du point de vue visuel :

Degré élevé : bon contraste et couleurs plutôt vives, aspect net, mais présence d'artefacts de repliement de spectre, meilleure restitution de la FTM dans l'image synthétisée.

Degré faible : aspect flou, couleurs moins vives, mais pas de présence d'artefact de repliement de spectre, car les fréquences supportant ce phénomène ont été coupées par le ré-échantillonnage.

En effet, plus la méthode de ré-échantillonnage est « molle » (fonction spline de degré faible), moins les fréquences contenant du repliement seront représentées et moins les artefacts de repliement de spectre seront visibles. Plus le degré de la fonction spline utilisé pour le ré-échantillonnage est grand, meilleurs seront les contrastes des hautes fréquences, mais plus le contraste des fréquences contenant du repliement sera important. Ceci entraîne l'apparition d'artefacts de repliement de spectre. On peut observer dans la figure 6.4 que plus le degré augmente et plus les stries dues au repliement de spectre deviennent visibles.

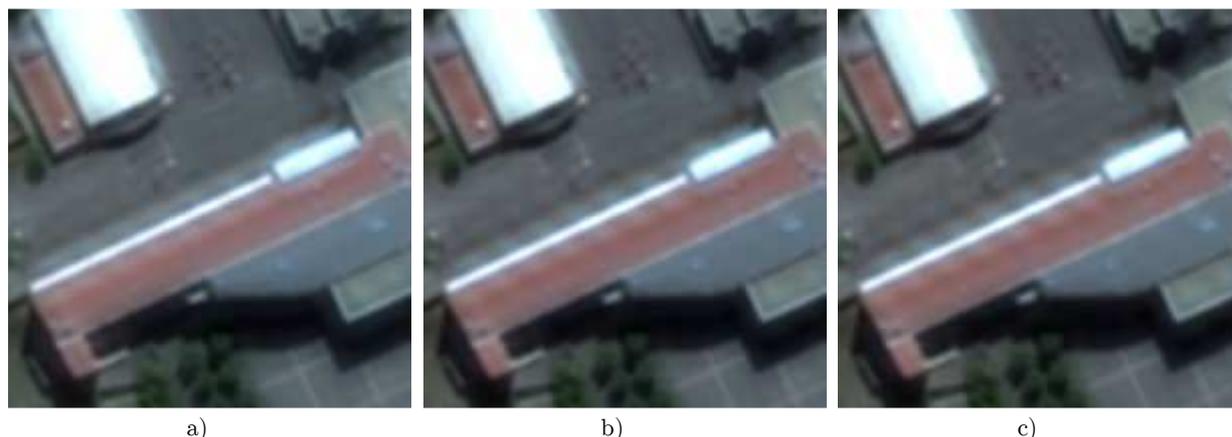


Figure 6.4 : Artefact de repliement de spectre en fonction du degré de la fonction spline de ré-échantillonnage dans les images synthétisées. a) spline degré 1. b) spline degré 3. c) spline degré 7.

Faut-il faire un compromis ? Du point de vue des performances, il est préférable de prendre un degré élevé autant pour l'ensemble des distances du bilan qualité que du point de vue FTM. Au contraire, par rapport aux artefacts visuels, il est préférable de prendre un

degré plus bas bien qu'un meilleur rendu des couleurs soit obtenu par des fonctions splines de degré élevé.

En conclusion, en supprimant le repliement de spectre, il est possible de prendre une spline de degré élevé et ainsi obtenir de meilleures performances statistiques. Le fait d'avoir supprimé le repliement permet malgré la spline de degré élevé de ne pas avoir d'artefacts de repliement de spectre dans l'image synthétisée.

6.1.6 Influence de la méthode d'interpolation sur la FTM de l'image synthétisée

Pour montrer l'impact de la méthode de ré-échantillonnage nous faisons varier le degré de la fonction spline utilisée dans la méthode de fusion ATWT-M3. Les degrés les plus bas étant équivalents à des méthodes d'interpolation connues ont été répertoriés dans le tableau 6.7.

Tableau 6.7 : équivalence de l'interpolation spline pour certains degrés

Degrés spline	0	1	2	3
Interpolation équivalente	Plus proche voisin	linéaire	Bi-linéaire	Bi-cubique

Pour constater l'effet de la méthode d'échantillonnage sur la FTM de l'image synthétisée nous synthétisons des images avec la méthode ATWT-M3 ; le degré de la fonction spline utilisée variera entre 1 et 7.

Visuellement plus le degré de la fonction spline est élevé et plus l'image paraît nette (figure 6.5). L'impression de plus grande netteté de l'image s'apprécie mieux sur une zone plus vaste. Le contraste contribue aussi à l'aspect de netteté, celui-ci est meilleur quand le degré de la fonction spline est élevé.

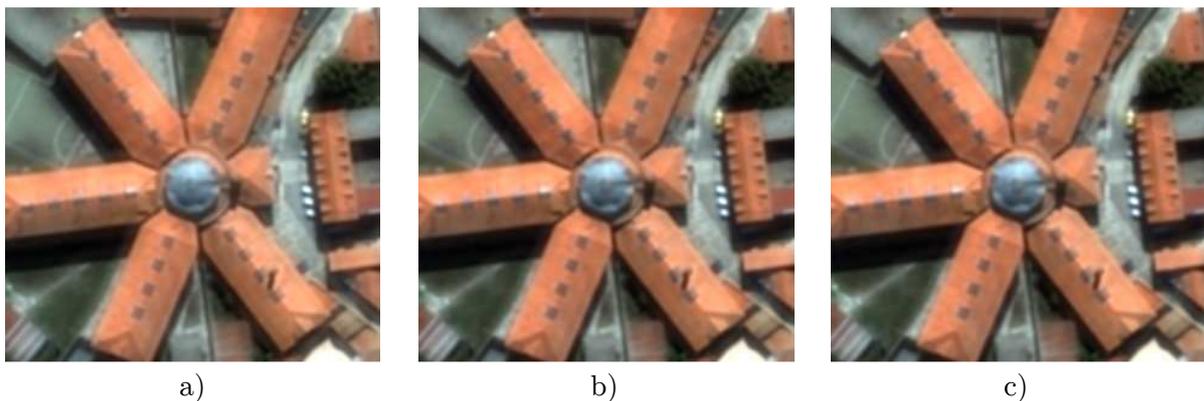


Figure 6.5 : extrait Toulouse ATWT-M3 pour différents degrés de la fonction spline d'interpolation, a) spline degré 1, b) spline degré 4, c) spline degré 7.

Les observations visuelles sont confirmées par MTF_{dev} dont les courbes sont présentées dans la figure 6.6. On voit dans cette figure que plus le degré de la fonction spline augmente et plus MTF_{dev} est proche de 1. En termes de FTM, on voit dans la figure 6.6 qu'il y a une grande différence entre une fonction spline de degré 1 et une fonction spline de degré 3. Par

la simple observation de MTF_{dev} , il paraît nécessaire de prendre un degré supérieur à 2 ou 3, le bénéfice obtenu en prenant de degré important devient plus faible au-delà du degré 4.

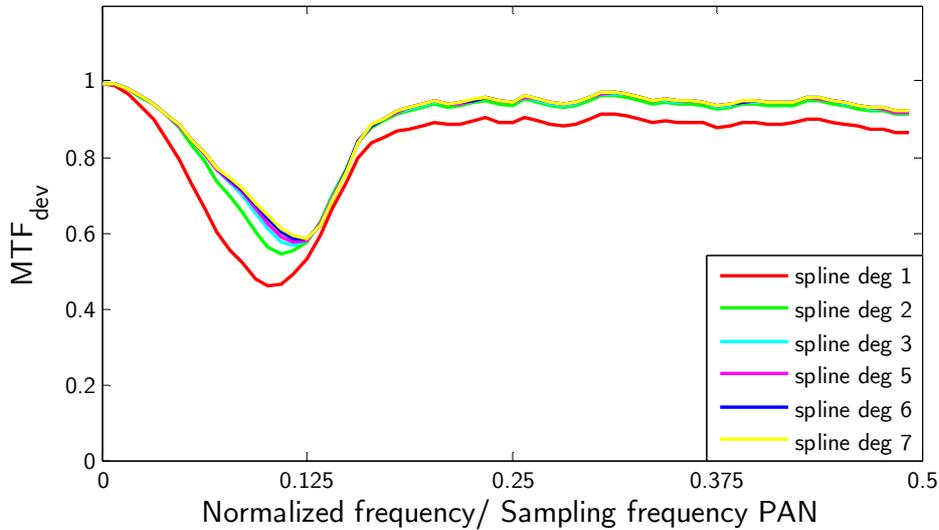


Figure 6.6 : Effet du degré de la fonction spline d'interpolation (canal rouge).

Nous venons de constater que visuellement et vis-à-vis de la FTM il y a un avantage à prendre un degré élevé. Voyons si ces constatations sont confirmées par des distances couramment utilisées en fusion comme l'ERGAS, ou le coefficient de corrélation. Ces résultats sont présentés dans le tableau 6.8 pour la méthode ATWT-M3 et dans le tableau 6.9 pour la méthode ATWT-M3-MTFadapted.

Tableau 6.8 : évolution de l'ERGAS, du coefficient de corrélation et de mean SAM en fonction du degré de la fonction spline utilisée pour la méthode ATWT-M3.

	Degré de la fonction spline d'interpolation				
	1	2	4	6	7
ERGAS	3,38	2,74	2,62	2,58	2,57
Mean SAM	3,41	3,12	3,09	3,07	3,07
CC bleu	0,9721	0,9809	0,9825	0,9831	0,9832

Tableau 6.9 : évolution de l'ERGAS, du coefficient de corrélation et de mean SAM en fonction du degré de la fonction spline utilisée pour la méthode ATWT-M3-MTFadapted.

	Degré de la fonction spline d'interpolation				
	1	2	4	6	7
ERGAS	2,707	2,066	1,982	1,957	1,957
Mean SAM	3,125	2,896	2,886	2,89	2,89
CC bleu	0,98212	0,98722	0,98755	0,98747	0,98747

L'évolution du coefficient de corrélation avec le degré de la fonction spline montre toujours le gain de qualité important entre un degré 1 et 4. L'amélioration de la qualité devient plus faible pour les degrés supérieurs à 3.

Par contre, l'ERGAS qui est un indice de qualité multimodale montre un intérêt à utiliser un degré au moins égal à 5. On obtient pour ERGAS une amélioration de plus de

25% en utilisant une fonction spline de degrés supérieur à 5 par rapport à une fonction spline de degré 1. Le gain obtenu au-delà de 5 devient plus faible.

Nous venons ainsi de montrer que la méthode d'interpolation a un impact sur la FTM de l'image synthétisée. Voyons maintenant si l'interpolation et l'adaptation de FTM peuvent avoir un impact cumulé. Dans la figure 6.7, nous avons représenté MTF_{dev} pour l'image de Toulouse, synthétisées avec et sans adaptation de FTM, pour différents degrés de la fonction spline. On observe que l'adaptation de FTM améliore la FTM, quel que soit le degré de la fonction de ré-échantillonnage.

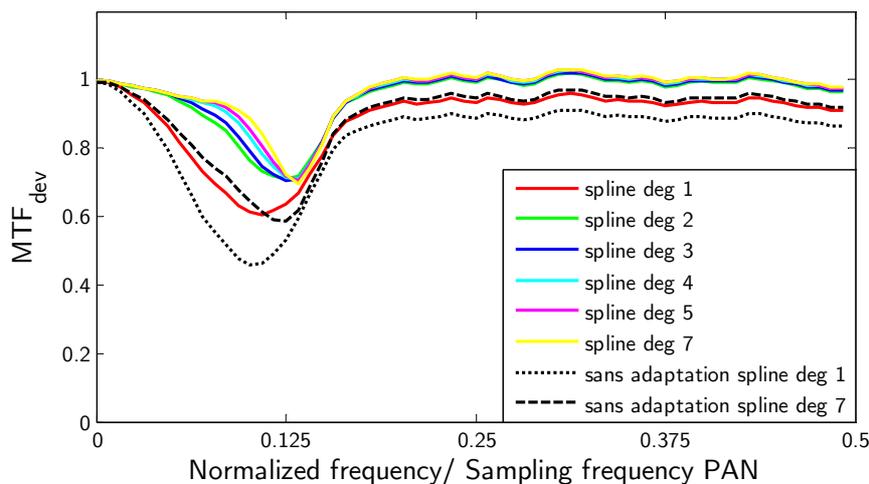


Figure 6.7 : MTF_{dev} fonction du degré de la fonction spline de ré-échantillonnage.

Du point de vue de l'analyse par MTF_{dev} , on obtient des résultats équivalents entre la méthode ATWT-M3- $MTF_{adapted}$ utilisant une fonction spline de degré 1 et la méthode d'ATWT-M3 utilisant une fonction spline de degré 7. Le gain apporté par l'adaptation de FTM est équivalent pour la méthode ATWT-M3 au passage d'un degré 1 à un degré 7. Cependant, les améliorations de la restitution de la FTM obtenues par ATWT-M3- $MTF_{adapted}$ et un degré élevé de la méthode de rééchantillonnage se combinent. Ainsi, la solution permettant la meilleure restitution de la FTM est la combinaison de la méthode ATWT-M3- $MTF_{adapted}$ avec un degré élevé pour la fonction spline de rééchantillonnage.

Dans l'article de (Lee et Lee 2010) l'interpolation est une simple interpolation bilinéaire (pour des raisons de rapidité). Le choix de ce type d'interpolation est cohérent avec le fait que ces auteurs cherchent à faire une méthode rapide. Cependant en termes de FTM, mes expériences ont montré que ce type d'interpolation n'est pas bon pour obtenir une bonne FTM pour l'image synthétisée.

Plus le degré de la fonction spline est élevé et plus l'interpolation est coûteuse en temps de calcul. Mes travaux n'ont pas pour but de faire une méthode rapide, mais de donner à l'image synthétisée une FTM proche de la FTM typique d'une image à haute résolution. Plus le degré de la fonction spline est élevé et plus la FTM de l'image synthétisée est proche de la FTM de l'image de référence. Les méthodes qui sont implémentées pour ces travaux privilégieront des fonctions spline de degré élevé.

6.2 SNR et nouvelles méthodes de fusion

Dans le cas de simulations PLÉIADES, les images sont dépourvues de bruit. Le SNR des images du satellite PLÉIADES a été dimensionné de façon à pouvoir faire de la restauration et donc de la déconvolution (Gaudin-Delerieu *et al.* 2008).

Nous avons effectué des tests avec les niveaux de bruits nominaux mesurés au sol sur l'instrument (Gaudin-Delerieu *et al.* 2008), mais aussi dans des cas plus défavorables où le SNR est multiplié par 0,5 de façon à évaluer le comportement de nos méthodes vis-à-vis du bruit. Les résultats de ces mesures sont présentés dans le tableau 6.10.

Tableau 6.10 : Niveau de SNR des images synthétisées pour les méthodes ATWT-M3, ATWT-M3-MTFadapted et ATWT-M3-MTFequalized (Toulouse)

	nominal SNR				0,5 x nominal SNR			
	SNR images avant fusion	SNR images synthétisées			SNR images avant fusion	SNR images synthétisées		
		ATWT-M3	ATWT-M3-MTFadapted	ATWT-M3-MTFequalized		ATWT-M3	ATWT-M3-MTFadapted	ATWT-M3-MTFequalized
PAN	147	--	--	--	74	--	--	--
B0	130	120	93	115	65	60	46	58
B1		112	90	107		56	44	53
B2		102	84	97		51	41	48
B3		136	102	132		68	51	67

On remarque dans le tableau ci-dessus que la méthode ATWT-M3-MTFadapted entraîne une baisse du SNR. Cette méthode effectue l'adaptation de FTM en utilisant les outils de convolution et déconvolution. Lors de la mise en place de cette méthode, nous nous doutions que le bruit serait amplifié. Ce tableau confirme cette amplification, celle-ci est d'autant plus importante que la méthode de déconvolution est simple. Une méthode de déconvolution plus complexe pourrait créer une amplification moindre du bruit et un débruitage pourrait y être ajouté.

7 Conclusion et perspectives

La propriété de synthèse que nous cherchons à respecter au mieux, stipule que l'image synthétisée doit être la plus proche possible de l'image MS si le capteur possédait la haute résolution spatiale. Ainsi, l'idée forte et directrice de l'ensemble de nos travaux est de considérer certains paramètres liés à l'instrument d'acquisition et leurs variations avec la résolution spatiale. Dans ces travaux, nous avons cherché à prendre en compte la FTM, ainsi que le repliement de spectre présent dans les images MS, tout en veillant au SNR des images synthétisées.

Il n'est pas toujours facile de connaître les traitements effectués sur les images par le fournisseur d'images : algorithme de restauration de FTM, débruitage, etc. Ces différents traitements modifient la FTM des images fournies. C'est pourquoi nous avons adopté une méthodologie fondée sur la simulation d'images de satellites. Ce cadre permet une bonne connaissance de différents paramètres de l'image notamment la FTM et le repliement des images que nous manipulons et une évaluation des performances plus aisée et fine.

Mes recherches ont eu comme point de départ une constatation forte : la différence de FTM qui existe entre les images PAN et MS, différence généralement liée essentiellement à la FTM de détection, les voies PAN et MS partageant le même télescope en entrée du système. Cette constatation a été appuyée par des publications montrant l'influence de la FTM sur la qualité des images fusionnées.

Les travaux actuellement menés dans le domaine de la fusion et qui traitent de la FTM cherchent à effectuer une décomposition de l'information suivant les caractéristiques de FTM de l'instrument. Notre objectif est légèrement différent puisque nous cherchons à prendre en compte la différence de FTM entre les images PAN et MS, de façon à ce que la qualité de l'image synthétisée ne soit pas desservie par cette différence.

Dans cette thèse, nous faisons l'hypothèse qu'il est possible d'améliorer la qualité des images synthétisées par un respect plus poussé de la propriété de synthèse. Nous avons établi une propriété de synthèse exprimée en termes de géométrie dans le sens défini dans ce document et relatif à la restitution des contrastes par fréquences spatiales. Cette définition de la géométrie est ainsi fortement liée à la FTM sous-jacente. Nous cherchons à donner à l'image synthétisée la FTM typique d'une image MS ayant la haute résolution spatiale, ce qui permet un meilleur respect de la propriété de synthèse en termes de géométrie.

Cette hypothèse a été l'objet d'expérimentations qui ont donné lieu à deux nouvelles méthodes de fusion. Elles ont été implémentées à partir d'un concept – le concept ARSIS – et d'une méthode, tous deux bien connus et éprouvés. Nous avons ajouté à cette méthode connue ATWT-M3 la prise en compte de la différence de FTM de deux façons. La première

cherche à améliorer la qualité géométrique des images MS en entrée du processus de fusion. Nous faisons l'hypothèse que l'amélioration des données en entrée mènera à une meilleure qualité de l'image en sortie. La deuxième méthode cherche à compenser la différence de FTM par une utilisation d'une plus grande proportion issue de l'information de l'image PAN dans les images MS synthétisées.

Puis en utilisant les travaux d'évaluation de la qualité validés et éprouvés, nous avons évalué l'amélioration apportée par la prise en compte de la différence de FTM sur la qualité des images synthétisées. Les résultats obtenus révèlent globalement une amélioration de la qualité des images synthétisée. Cette amélioration est observée visuellement et à travers de nombreux indices de qualité image : amélioration des distances et critères à la fois monomodales et multimodales et diminution de certains artefacts visuels.

La littérature sur l'évaluation de la qualité image dans le domaine de la fusion ne possédait pas d'outil permettant la caractérisation de la FTM de l'image synthétisée. Nous proposons un tel outil MTF_{dev} afin de pouvoir évaluer l'amélioration de FTM obtenue pour les images synthétisées mettant en jeu la prise en compte de la différence de FTM. Cet outil a permis de constater les conséquences de la différence de FTM des images PAN et MS sur l'image synthétisée. Il sera ajouté aux analyses visuelles et aux bilans qualité qui permettent d'évaluer la qualité.

Les méthodes mises en place donnent des résultats plutôt satisfaisants. Cependant, il est possible de raffiner encore leur implémentation, ce qui pourrait permettre d'améliorer les résultats.

Pour la méthode ATWT-M3-MTFequalized, le principe du modèle HRIMM MTFequalized semble très pertinent. Cependant, l'implémentation faite à travers les ondelettes à trous n'est pas assez précise et sélective en fréquence. Ainsi, utiliser des paquets d'ondelettes ou des bancs de filtre pour ce HRIMM permettrait certainement de meilleurs résultats. La compensation du défaut de MTF_{dev} serait plus précise et permettrait d'éviter le rebond observé après $0,125 fe$. En plus de la précision fréquentielle, il semble envisageable d'ajouter une contrainte de réinjection basée sur le contexte suivant l'idée de (Thomas 2006) avec la méthode ATWT-SharpnedM3 utilisant la corrélation des structures pour ce HRIMM, d'autant plus que cet HRIMM se fait pour des résolutions communes aux images PAN et MS ou relativement proches, ce qui permet une bonne efficacité de la corrélation.

La méthode ATWT-M3-MTFadapted entraîne une dégradation du SNR. Pour atténuer ce défaut, il est possible d'utiliser des algorithmes de déconvolution permettant une amplification réduite du bruit ou d'ajouter une étape de débruitage basée sur un seuillage ou l'atténuation des faibles coefficients d'ondelettes.

Cette méthode amplifie les artefacts de repliement de spectre. En utilisant un MSM pour lequel la fréquence de coupure se trouvera avant la fréquence de Nyquist de l'image MS, ce qui n'était pas forcément le cas avec le MSM utilisé ici, il est possible d'atténuer les fréquences responsables du repliement de spectre. En ayant atténué ces fréquences, nous

pourrions conserver les bénéfices d'ATWT-M3-MTFadapted tout en évitant d'amplifier les artefacts de repliement de spectre.

La deuxième constatation que nous avons faite est également liée aux caractéristiques de l'instrument et à la géométrie. Nous avons étudié l'impact du repliement de spectre dont la présence est relative à la FTM et à l'échantillonnage mis en œuvre.

La plupart des satellites d'observation qui observent la surface terrestre à des résolutions spatiales différentes fournissent des images à basse résolution spatiales qui présentent un risque de contenir du repliement de spectre. Dans nos travaux, nous avons cherché à mesurer l'impact de ce repliement sur la qualité des images synthétisée. Ces travaux ne pouvaient être mis en place que dans le cadre d'images de simulation, qui permet de connaître les zones touchées par ce phénomène.

Ces travaux ont montré que la suppression du repliement permet de supprimer des artefacts visuels importants, mais les outils usuels d'évaluation de la qualité des images de fusion ne permettent pas la quantification des performances. Nous avons utilisé des outils basés sur la variation totale proposée par Blanchet (2006), mais leur application aux images synthétisée ne s'est pas révélée pertinente et ces indices sont sensibles aux bruits.

Les travaux menés sur les techniques de SUPER-XS (Douillet et Blanc 2009), SUPERMODE (Letry et Rouge 2003) ou encore les travaux plus généraux sur la super-résolution (Almansa *et al.* 2004) (Malgouyres 1999), qui exploitent le repliement de spectre pour améliorer la résolution d'une série d'images nous font penser qu'il y a nécessairement dans la fusion un moyen d'exploiter le repliement de spectre pour améliorer la qualité des images synthétisées.

Nous connaissons et maîtrisons le processus visant à transformer l'image MS à haute résolution (qui constitue notre objectif) en l'image MS réelle. Ce processus passe par une étape de filtrage et un sous-échantillonnage régulier ou pas. Une façon de tirer parti du repliement de spectre peut être d'ajouter à l'image MS de l'information haute fréquence sans repliement de spectre qui proviendrait de l'image PAN, puis d'effectuer un processus itératif visant à assurer la propriété de cohérence avec un terme d'attache aux données et un terme de régularisation basée, par exemple, sur la variation totale. En plus de la nécessité d'une parfaite connaissance des décalages géométrique entre les images PAN et MS, la mise en place de l'algorithme numérique de minimisation de la fonction de coût résultant de la somme pondérée du critère d'attache aux données et du critère de régularisation est très délicate. (Almansa *et al.* 2004). Nous pensons de plus que le choix d'une image fusionnée en initialisation d'un tel processus itératif pourrait permettre une meilleure et plus rapide convergence d'un tel algorithme de super-résolution.

De manière générale, nous retenons de notre travail deux choses principales. Nous avons défini une quantité : MTF_{dev} , qui permet de quantifier l'écart de FTM entre deux images ayant la même résolution spatiale. Cette quantité semble assez précise pour être utilisée dans les bilans qualité des images synthétisées et même hors du cadre de simulation en évaluant MTF_{dev} entre les images MS-HR et les images PAN.

Nous avons développé et testé deux façons de prendre en compte la différence de FTM. Nous les avons exploitées sur une méthode particulière ATWT-M3. Le principe de la prise en compte de la FTM pour la méthode ATWT-M3-MTFadapted est rappelé dans la figure 7.1.

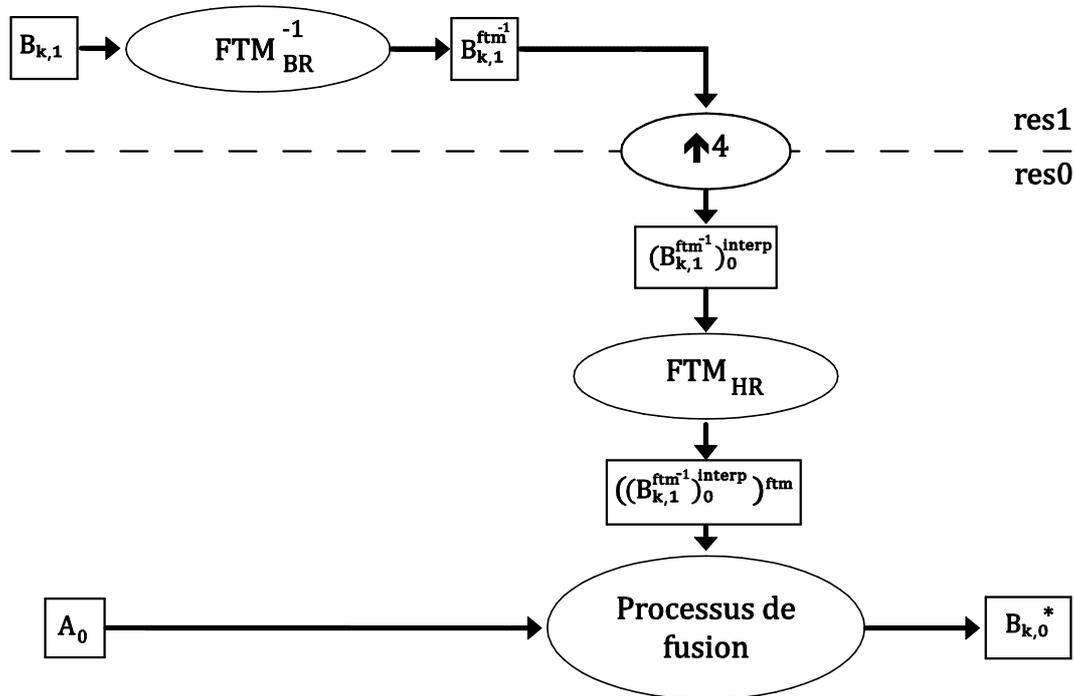


Figure 7.1 : implantation de MTFadapted. A_0 représente l'image PAN, et l'image $B_{k,0}$ représente une bande de l'image MS. Cette image est d'abord déconvoluée par une FTM de type BR puis sur-échantillonnée d'un facteur 4 puis convoluée par une FTM de type HR pour enfin être fusionnée avec l'image A.

Nous estimons que ce schéma peut s'appliquer à quantité de modèles IMM, MSM et HRIMM. En d'autres termes, il est généralisable aux méthodes appartenant au concept ARSIS.

Dans ce schéma, la prise en compte de la FTM faite par MTFadapted est réalisée en modifiant la FTM des images MS en amont du processus de fusion. Par conséquent, cette façon de prendre en compte la FTM peut aussi se généraliser aux méthodes des familles « projection-substitution » et « contribution spectrale relative ». Les principes de ces deux familles sont discutés ci-après.

Contrairement à MTFadapted, la compensation de la différence de FTM réalisée à travers MTFequalized corrige la différence de FTM dans le cas d'ARSIS, mais en aval du processus de fusion (figure 7.2). Nous avons étudié si cette prise en compte de la FTM peut s'appliquer à des méthodes de la famille « projection-substitution » ainsi qu'à la famille « contribution spectrale relative ».

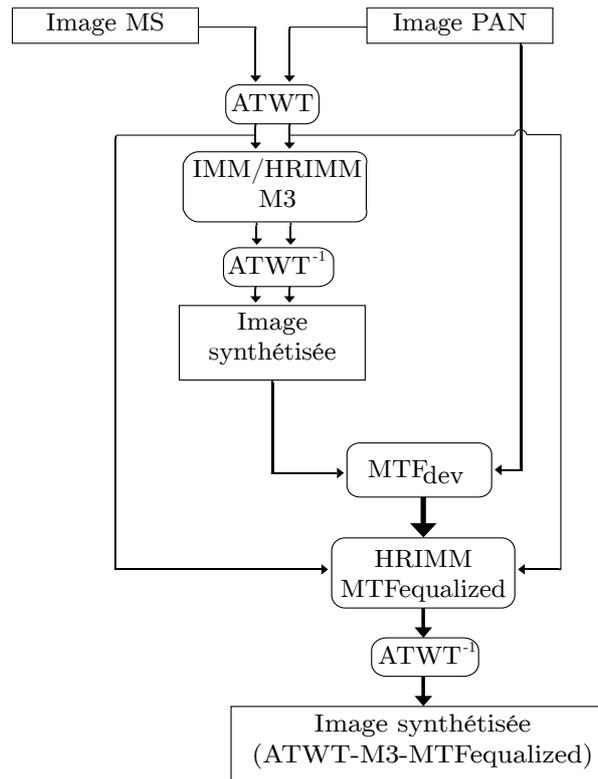


Figure 7.2 : Schéma de principe de la méthode ATWT-M3-MTFequalized.

Le principe des méthodes de la famille « projection-substitution » consiste à projeter les bandes MS dans un espace colorimétrique différent ou sur l'espace des vecteurs propres, puis à substituer le plan de plus grande variance par l'image PAN et à revenir dans l'espace de départ. Pour ce concept de fusion, la transformation se fait dans l'espace des couleurs et non dans celui des fréquences spatiales (résolution). Il n'y a donc pas de raison de trouver un défaut de FTM pour une fréquence donnée du fait que la fusion est réalisée en changeant d'espace colorimétrique et non par une analyse multirésolution. Les images synthétisées possèdent une FTM proche de celle de PAN. La prise en compte de la FTM est alors inutile.

Le principe de la famille « contribution spectrale relative » est que la somme des images MS est égale à l'image PAN dans une version où la résolution spatiale est dégradée. L'alliance des données est effectuée dans le domaine spatial en multipliant l'image MS rééchantillonnée par l'image PAN pondérée en fonction de la bande spectrale considérée. Effectuer l'alliance des données dans le domaine spatial peut entraîner des comportements non désirés dans les fréquences et donc pour la FTM de l'image synthétisée. Nous avons réussi dans le cas d'ARSIS à corriger un défaut de FTM lié au concept ARSIS. Le principe des méthodes de la famille « contribution spectrale relative » n'assure pas une bonne restitution, ni une maîtrise de la FTM pour l'image synthétisée. La FTM de l'image synthétisée sera dépendante de la corrélation locale des structures présente dans les images PAN et MS. Ainsi, nous ne pourrions rencontrer comme dans le cas d'ARSIS un défaut de FTM dont l'allure est indépendante de l'image. Dans ces conditions, il est difficile d'appliquer le principe de MTFequalized aux méthodes de la famille « contribution spectrale relative ».

Ainsi, si notre travail s'achève en pratique sur la proposition de deux nouvelles méthodes, il ouvre une nouvelle opportunité d'amélioration de toutes les méthodes existantes.

Annexe A

Commentaires sur le scénario 4 pour la méthode ATWT-M3-MTFadapted

ATWT-M3 est une méthode de référence à laquelle nous avons ajouté une prise en compte de la différence de FTM entre l'instrument MS et PAN. L'adaptation de FTM peut être réalisée de plusieurs façons. Dans le manuscrit principal, nous avons présenté quatre scénarios possibles. Le scénario que nous avons finalement retenu est le scénario 2. La justification visant à préférer le scénario 2 plutôt que le scénario 4 est abordée de manière résumée dans ce chapitre. Cette annexe a pour objectif de la développer.

Pour rappel, le scénario 4 présente un avantage certain par le fait que l'adaptation de FTM ne se fait que par une seule déconvolution, alors que le scénario 2 nécessite d'abord de déconvoluer l'image MS puis de la convoluer avec la FTM d'un capteur à haute résolution, une fois que celle-ci a été rééchantillonnée. Pour ce scénario 4, le traitement consiste à déconvoluer l'image MS par la FTM d'un capteur HR une fois que l'image a été rééchantillonnée (fig. A.1).

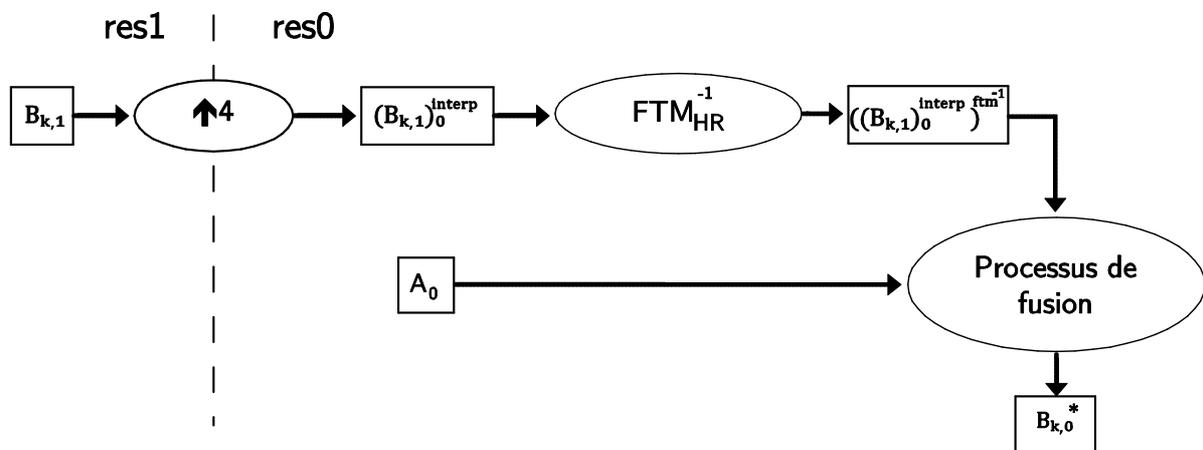


Figure A.1 : schéma de principe du scénario 4

Nous avons réalisé le scénario 4 et avons trouvé des résultats moins convaincants que ceux du scénario 2. Les figures suivantes montrent quelques exemples et illustrent les différences de résultats par rapport au scénario 2.

La figure A.2 montre l'exemple d'un toit blanc avec les rangées d'ouvertures, présenté en fig. 4.14. On note bien dans cet exemple que les images a) et b) sont assez semblables. Dans ces deux images, il est difficile de distinguer les ouvertures présentes sur le toit. Ces ouvertures sont visibles dans l'image c) issue du scénario 2 ainsi que dans l'image de référence. De la même façon le contour du bâtiment blanc est plus net dans l'image c) que dans les images a) et b) qui sont proches au niveau du rendu visuel.

Au vu de cet exemple visuel, on voit déjà que la prise en compte de la FTM réalisée avec le scénario 4 est nettement moins significative que ne l'est celle du scénario 2.

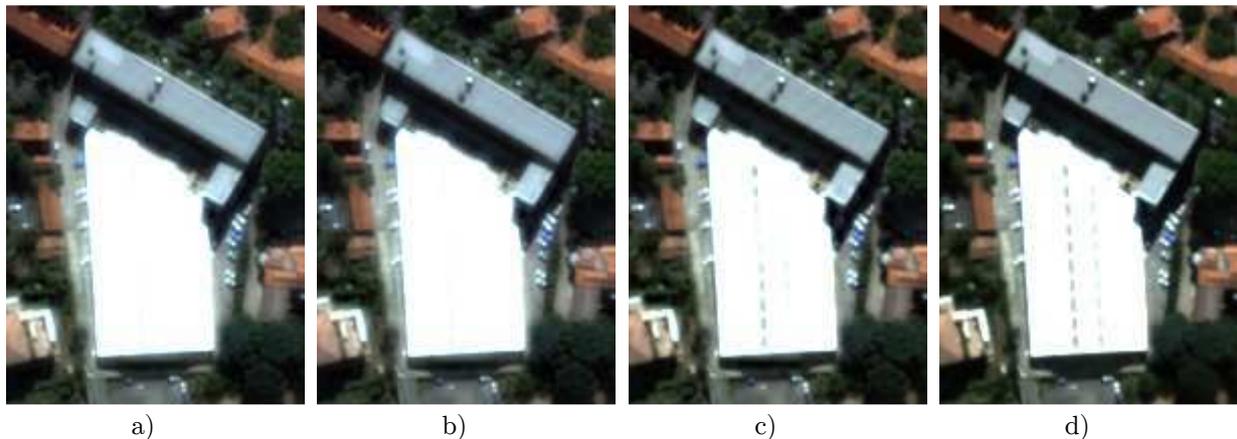


Figure B.2 : Extrait de l'image de Toulouse, cas bâtiment blanc. a) ATWT-M3, b) ATWT-M3-MTFadapted-scénario 4, c) ATWT-M3-MTFadapted-scénario 2, d) référence.

Pour appuyer ces premières observations, nous présentons dans la figure a.3 les résultats du calcul de MTF_{dev} pour l'image de Toulouse pour les fusions de type : ATWT-M3, ATWT-M3-MTFadapted-scénario 4, et ATWT-M3-MTFadapted-scénario 2.

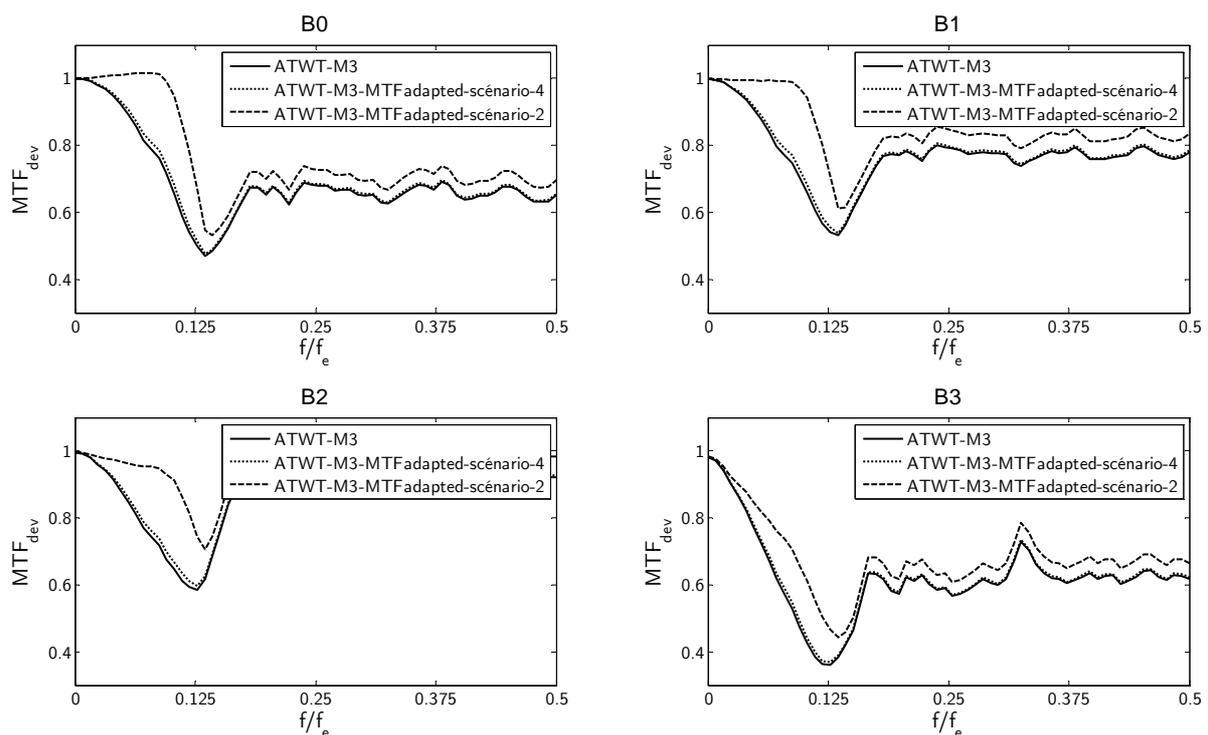


Figure A.3 : MTF_{dev} calculé pour l'image de Toulouse et pour les méthodes de fusion : ATWT-M3, ATWT-M3-MTFadapted-scénario4, ATWT-M3-MTFadapted-scénario2.

On observe bien pour le scénario 4 que l'allure de MTF_{dev} est presque identique entre ATWT-M3 et ATWT-M3-MTFadapted-scénario4. Cette adaptation de FTM ne modifie presque pas la représentation du contraste dans l'image. Ce n'est pas le cas pour le scénario 2. Les observations visuelles et en termes de FTM coïncident et montrent que ce scénario 4

n'est pas très performant dans notre cas. C'est pourquoi nous avons retenu le scénario 2 plutôt que le scénario 4.

Annexe B :

Communication pour le 29th EARSeL Symposium, Chania, Greece, 18-20 June 2009

Data fusion: taking into account the modulation transfer function in ARSIS-based pansharpening methods

Pierre MASSIP^a, Philippe BLANC^a and Lucien WALD^{a,1}

^a*Center of Energy and Processes, Mines ParisTech /Armines, 06904 Sophia Antipolis, France*

Abstract. Multispectral images provided by satellite have a poor spatial resolution while panchromatic images (PAN) exhibit a spatial resolution two or four times better. Data fusion is a mean to synthesize MS images at higher spatial resolution than original by exploiting the high spatial resolution of the PAN. This process is often called pan-sharpening. The synthesized multispectral images should be as close as possible to those that would have been acquired by the corresponding sensors if they had this high resolution. The methods based on the concept “Amélioration de la Résolution Spatiale par Injection de Structures” (ARSIS) concept are able to deliver synthesized images with good spectral quality but whose geometrical quality can still be enhanced. We propose to consider the characteristics of the sensor to improve the geometrical quality. We take explicitly into account the modulation transfer function (MTF) of the sensor in the fusion process. Though this study is limited in methods and data, we observe a better restitution of the geometry and an improvement in the majority of quality indices classically used in pan-sharpening. The communication also presents a means to assess the respect of the synthesis property from a MTF point of view.

Keywords. Image fusion, pan-sharpening, multispectral, panchromatic, PSF, MTF, ARSIS, remote sensing

Introduction

Image fusion is a vivid topic in remote sensing, among other domains. Fusion was defined by the EARSeL Working Group [1] as “a formal framework in which are expressed means and tools for the alliance of data originating from different sources. It aims at obtaining information of a greater quality, although the exact definition of ‘greater quality’ will depend on the application”. Here we specifically focus on the synthesis of multispectral (MS) images at a higher spatial resolution than original by exploiting the high spatial resolution of another image.

Current satellites provide images of two kinds with opposite characteristics. Multispectral (MS) images have a good spectral quality but a poor spatial resolution whereas panchromatic (PAN) images have on the contrary a high spatial resolution but with a poorer spectral quality. These two types of images allow to well identify observed structures through different information by geometrical information on the one hand and by the spectral signature on the other hand. The interest of combining both types of images has been demonstrated by many authors.

Data fusion is a means to synthesize MS images at higher spatial resolution than original by exploiting the high spatial resolution of the PAN. This process is often called pan-sharpening. Here, the aim of fusion is to perform a high-quality transformation of the MS content when increasing the spatial resolution from the original one to that of the PAN image. The problem may be seen as the inference of the information that is missing to the original MS images for the construction of the MS images synthesized at a better resolution [5-7].

¹ Corresponding Author: L. Wald, Mines ParisTech/Armines, Sophia Antipolis, 06904, France.
E-mail: Lucien.wald@mines-paristech.fr

[2-4] establish that the fused images should obey several properties, that were synthesized by the EARSeL Working Group [6-7]. The consistency property expresses that the fused images should offer a strong consistency with the original data set: if a fused product is downsampled to its original low resolution, the original MS image should be retrieved. The synthesis property deals with the quality of the high-quality transformation: any synthetic image should be as identical as possible to the image that the corresponding sensor would observe with the highest spatial resolution, if existent. This holds for each individual modality MS and the multi-modality set.

Recent published works tend to demonstrate that the methods based on the ARSIS concept better synthesize images with respect to the expected spectral properties [2, 7-12]. However, the need to refine the spatial/geometrical quality of synthesized image is often underlined [2] [7] [12-13]. Geometry can be defined as the spatial organization of the radiometry in the image [14]. Characteristics of the imaging instrument are important with respect to the restitution of the geometry of the landscape. Besides its spectral and spatial resolutions, an important characteristic of optical instruments is the impulse response which is named point spread function (PSF). The PSF has an associate value, the modulation transfer function (MTF). The MTF is the modulus of the Fourier transform of the PSF. According to [15], the MTF determines the upper limits to the image quality, *i.e.*, the image resolution or sharpness and describes the image quality in terms of contrast as a function of spatial frequency. The difference in sampling rates between MS and PAN images implies a difference in the MTF of these images. This difference must be taken into account when synthesizing MS images at the PAN spatial resolution. A few fusion methods already include this difference in MTF through the multiscale model (MSM) in the ARSIS concept [2] [13] [16-18].

Our manner to treat the MTF is different from these previous works. We take explicitly into account the MTF of the sensor in the fusion process for each modality, independently from the MSM.

- Objective

We now express the synthesis property in terms of geometry: the restitution of the radiometric organization of the observed landscape by the synthesized image should be as identical as possible to the organization that the corresponding sensor would observe with the highest spatial resolution, if existent. The work presented here aims at including the MTF into an existing fusion method in order to demonstrate that taking into account the difference in MTF leads to a better respect of the synthesis property from the geometrical point of view without degrading the quality of the other aspects of the synthesized images.

As said before, we selected an existing method belonging to the ARSIS concept. This method is called ATWT-M3, where ATWT denotes the MSM and means “à trous” wavelet transform, and M3 is the inter-band structure model described by [5], [10]. This method has been shown to provide good results in most cases; it already includes the difference in MTF in its MSM. Therefore, it is considered as an efficient method against which we will compare our own results. Whereas the results will not be the same than those obtained here, our way of taking into account the difference in MTF may easily apply to other methods belonging to the ARSIS concept and likely less easier, to other methods.

By explicitly integrating the difference in MTF, the aim is to give a 'real' MTF to the synthesized images. In figure 1, are drawn the MTFs for the four MS channels of the Pleiades sensor, discussed later, and that for the PAN, as a function of the frequency f normalized with the sampling frequency of the PAN. In this case, the spatial resolution of the PAN image (70 cm) is 4 times better than the MS (280 cm). Note that in this figure, the MTF of the MS are zero for relative frequencies equal to, or greater than, 0.25 ($=1/4$). It means that one cannot distinguish details less than 280 cm in size in the MS images. On the contrary, the MTF of PAN is equal to 0.4. The difference in MTF between low and high spatial resolution images is evidenced in this figure: the MTF of MS has a sharper decrease with frequency than that of PAN. Of particular interest here is the relative Nyquist frequency for MS, equal to 0.125. It denotes the smallest frequency that can be detected in MS images, *i.e.*, twice the pixel size.

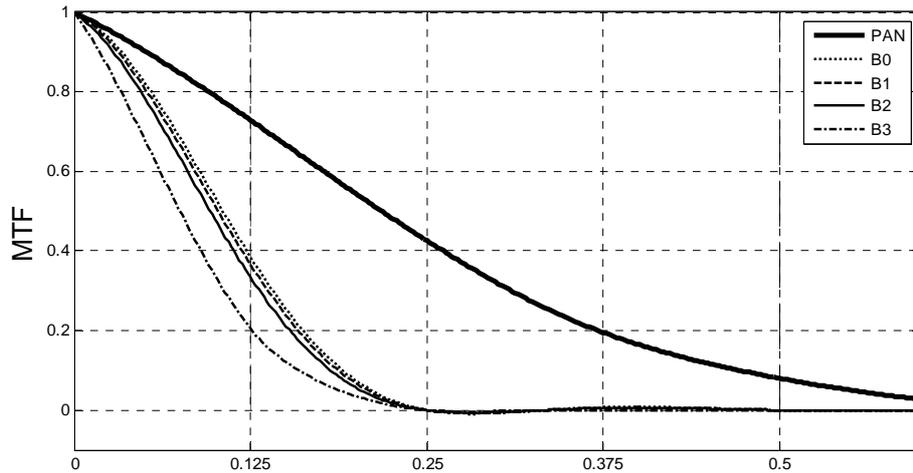


Figure 1. Modulation transfer function for MS and PAN images.

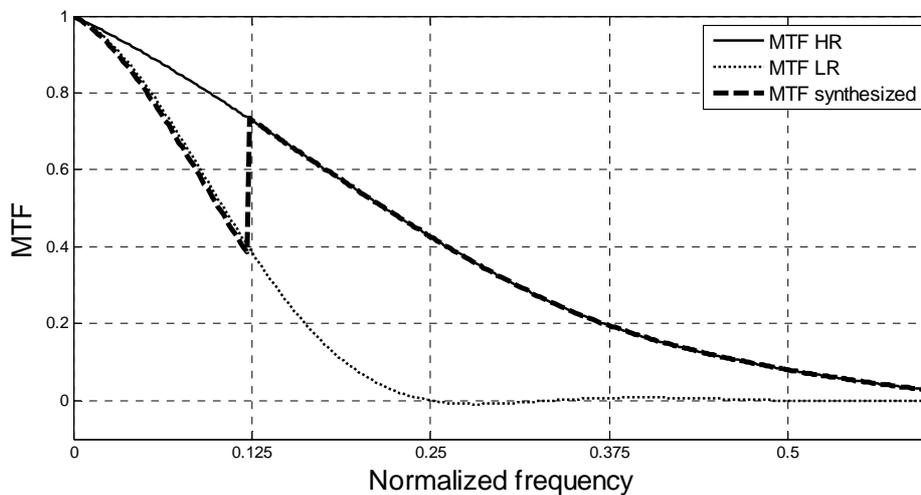


Figure 2. Schematic representation of the MTF of the synthesized MS images. LR and HR, respectively, mean low and high resolution.

When increasing the spatial resolution by injecting high-frequencies into the original MS images, the MTF of the synthesized image should look like that of PAN. However, without modification of the MTF during the fusion process, the resulting MTF exhibits a discontinuity for the relative frequency 0.125 as schematically drawn in Fig. 2. To attenuate the gap between the value of MS and PAN at this frequency, a solution consist in “raising” the MTF of the MS frequencies in the range $[0, 0.125]$ so that it is close to that of a high resolution image. Doing so provides a MTF closer to a “real” MTF and similar contrast in the image for a same frequency without considering its origin: PAN or MS.

- Data set

The data set used for this work is a set of images resulting from the simulation of the Pleiades mission of the Centre National d'Etudes Spatiales (CNES). These images are simulated from images at a higher resolution from the airborne sensor Pelican. CNES kindly provides us with MS images at both high (70 cm) and low (280 cm) spatial resolution and a PAN image at high resolution (70 cm). The tests have been done on an excerpt of the city of Toulouse in France.

These images result from simulation; it means that we have the so-called original MS images at 280 cm, the PAN image at 70 cm and the actual MS images at 70 cm which constitute the reference that should be attained. This is not the common situation in which the actual MS images at high resolution are unknown. However, it is a very convenient situation to assess the influence of a parameter such as the MTF because all parameters are known. During the creation of the simulated Pleiades images, the MTF is controlled by CNES for each image and resolution. Consequently, the MTF of MS, PAN and actual MS images are known for this study.

- Modification of MTF for fusion in ARSIS based methods

As said before, we exploit an existing method: ATWT-M3, and we create a new one: ATWT-M3-FTMadapted by taking the MTF explicitly into account in the method. We transform the MTF before the decomposition of the MS image by the undecimated wavelet transform (“à trous” wavelet transform). The application of the undecimated transform implies a resampling of the original MS images to the high resolution. The modification of the MTF is done on this resampled MS image.

The transformation starts with a deconvolution of the original MS images at low resolution. The deconvolution kernel corresponds to the detector part of the MTF—the optical MTF is not taken into account.— Then, we perform the resampling of this image that has now only the optical MTF. The resampling is of spline type. Then we convolute the resampled images with the MTF of a high resolution detector.

After that, the method ATWT-M3 is applied which is based on undecimated wavelet transform [5]. For each modality, the high frequencies wavelet coefficients from the PAN image are transformed into MS high frequencies wavelet coefficients by applying an affine function. The parameters of this function are determined by a least-square fitting between the PAN and MS low-frequencies wavelet coefficients [5]. Finally, the inverse wavelet transform is applied to obtain the synthesized MS images at high resolution.

- Results

The results of the two methods: ATWT-M3 and ATWT_M3_FTMadapted are compared to the reference image. As recommended by several authors [3-4] [6] [12], we perform statistical comparisons and a visual analysis of the possible discrepancies.

- *Statistical results*

Quality budgets are obtained by the very same tool than that used for the "data fusion contest 2006" [12]. Table 1 exhibits the result of mono-modal indices: relative bias (biasRel), relative difference variance in percent (diffVarRel), relative standard deviation (σ Rel), correlation coefficient (cc), and correlation coefficient of high frequency (ccHF), as well multimodal indices: ERGAS and mean SAM.

The ideal values are presented for each index. This table shows that the bias for each method is very close to the ideal value; this is characteristic of the methods based on the ARSIS concept [5] [7] [12]. The bias is smaller for ATWT-M3 than for ATWT-M3-FTMadapted but in both cases, it remains very close to 0. For the other parameters, one notes a large increase in quality budget when adapting the MTF.

Table 1. Comparison of quality budget for ATWT-M3 and ATWT-M3-FTMadapted

modalities	Indexes	Ideal	ATWT-M3	ATWT-M3-FTMadapted	Delta in percent
bleu	biasRel	0	-0,001	0,006	-670
	diffVarRel	0	14,4	3,3	77
	σ Rel	0	8,7	7,2	18
	Cc	1	0,9833	0,9875	0,4
	ccHF	1	0,9203	0,9224	0,2
green	biasRel	0	-0,001	0,007	-1100
	diffVarRel	0	15,5	4,3	72
	σ Rel	0	8,6	6,3	26
	Cc	1	0,9875	0,9921	0,5
	ccHF	1	0,9605	0,9628	0,2
red	biasRel	0	0,001	0,012	-1836
	diffVarRel	0	15,8	6,0	62
	σ Rel	0	9,3	6,0	35
	Cc	1	0,9905	0,9953	0,5
	ccHF	1	0,9690	0,9734	0,5
NIR	biasRel	0	0,016	0,018	-11
	diffVarRel	0	34,2	26,7	22
	σ Rel	0	13,7	10,9	21
	Cc	1	0,9669	0,9793	1,3
	ccHF	1	0,8783	0,8886	1,2
multimodal indexes	ERGAS	0	2,57	1,96	24
	mean SAM	0	3,07	2,89	6

All these indices demonstrate improvements. Nevertheless, they do not inform on the aspect of the MTF of the synthesized images. Consequently, we perform a comparison of MTF between the reference and the synthesized images in order to evaluate if the modification of the MTF meets the expectations. This comparison is based on the cross correlation in the Fourier domain (of the PSD of the images) and is made relative to the MTF of the reference image:

$$mtf_{rel} = \frac{cross\ psd2d(xs_{ref}, xs^*)}{psd(xs_{ref})} \quad (0.1)$$

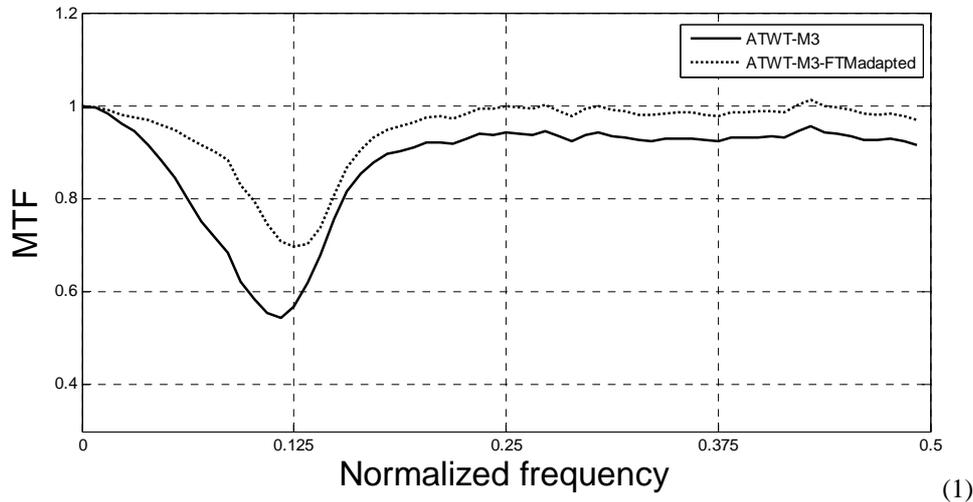


Figure 3. MTF of ATW-M3 and ATWT-3-FTM adapted relative to the MTF reference.

The MTF_{rel} denotes the behaviour of the MTF of the synthesized image compared to that of the reference image, as a function of the relative frequency (Fig. 3). The ideal result is a straight line for a value of one. Fig. 3 displays the results for band B2, corresponding to the wavelength of the red.

In this figure, the MTF_{rel} with adaptation of MTF (ATWT-M3-FTMadapted) is closer from the ideal value 1 than that of the method ATWT-M3, and this for all frequencies. Consequently, the improvement brought by the adaptation of the MTF is demonstrated in this case. One may note in this figure that the MTF_{rel} exhibits a trough between 0.1 and 0.15. It is believed that this is due to the fact that only the low frequency of the resampled image are injected in the synthesized image and also because information between the PAN and each MS band are not the same [10] which make that the PAN information is not fully exploited to create the high frequencies wavelet coefficients. This is under investigation and should be confirmed.

- *Visual analysis*

Figure 4 exhibits three images: one without MTF adaptation (a), one with MTF adaptation (b) and the image of reference (c). In this excerpt we observe that the contour line of the white building is sharper in (b) than in (a) and that the image (b) is the closest to (c). This holds also for color composites and here again, we find an improvement when adapting the MTF.

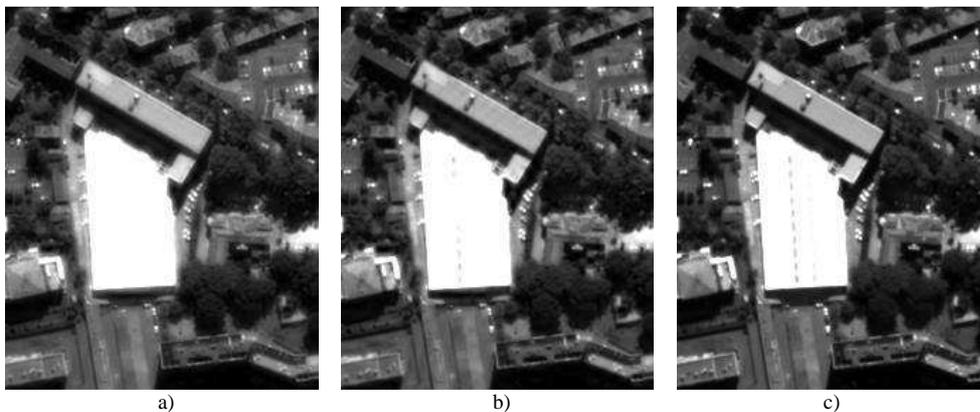


Figure 4. Excerpt of Pleiades images over Toulouse, France. Red modality synthesized with a) ATWT-M3, b) ATWT-M3-FTMadapted, c) reference image. Copyright CNES 2000.

- Conclusion

We have shown how the difference in MTF between low and high resolution can be taken into account into a “classical” fusion method based on the ARSIS concept. In a given case, we have observed a better restitution of the geometry and an improvement in the majority of quality indices classically used in pan-sharpening. A new measure has been proposed to evaluate the respect of the synthesis property from a geometrical point of view. Take into account the difference of MTF between MS and PAN images allow to have synthesized images of better quality. This work can be apply to others concept. It could be interesting to observe the difference of behavior of this technique depending on the concept it's applied on. The MTF relative that is given by our tool let us observe improvement and weakness of the MTF of the fuse image.

Though this study is limited in methods and data, the present results are encouraging and may constitute a new way to improve the restitution of geometrical features by already efficient fusion methods.

References

- L. Wald, Some terms of reference in data fusion. *IEEE Transactions on Geosciences and Remote Sensing*, 37 (1999), 1190-1193.
- T. Ranchin, B. Aiazzi, L. Alparone, S. Baronti, and L. Wald, Image fusion—The ARSIS concept and some successful implementation schemes. *ISPRS J. Photogramm. Remote Sens.*, vol. 58, no. 1/2, pp. 4–18, Jun. 2003.
- L. Wald, T. Ranchin, and M. Mangolini, Fusion of satellite images of different spatial resolutions: Assessing the quality of resulting images. *Photogrammetric Engineering & Remote Sensing*, 63 (1997), 691-699.
- J. Li, Spatial quality evaluation of fusion of different resolution images. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.*, 33 (2000), B2-2, 339–346.
- T. Ranchin and L. Wald, Fusion of high spatial and spectral resolution images: The ARSIS concept and its implementation. *Photogrammetric Engineering and Remote Sensing*, 66 (2000), 49-61.
- C. Thomas and L. Wald, Assessment of the quality of fused products. In Proc. 24th EARSeL Symp. New Strategies for Eur. Remote Sens., Dubrovnik, Croatia, Oluic, Ed. Rotterdam, The Netherlands: Millpress, May 25–27, 2004, pp. 317–325.
- C. Thomas, T. Ranchin, L. Wald, and J. Chanussot, Synthesis of multispectral images to high spatial resolution: A critical review of fusion methods based on remote sensing physics. *IEEE Transactions on Geosciences and Remote Sensing*, 46 (2008), 1301-1312.
- D. A. Yocky, Multiresolution wavelet decomposition image merger of Landsat Thematic Mapper and SPOT panchromatic data. *Photogramm. Eng. Remote Sens.*, 62 (1996), 1067–1074.
- Aiazzi B., L. Alparone, S. Baronti and A. Garzelli, Context-driven fusion of high spatial and spectral resolution images based on oversampled multiresolution analysis. *IEEE Transactions on Geosciences and Remote Sensing*, 40 (2002), 2300-2312.
- Z. Wang, D. Ziou, C. Armenakis, D. Li, and Q. Li, A comparative analysis of image fusion methods. *IEEE Trans. Geosci. Remote Sens.*, 43 (2005), 81–84.
- P. Terretaz, Comparison of different methods to merge SPOT P and XS data : Evaluation in an urban area. In Proceedings, EARSeL Symposium 1997 “ Future Trends in Remote Sensing,” Lyngby, Denmark, P. Gudmansen Ed., A. A. Balkema Publ., 1998, 435-443.
- L. Alparone, L. Wald, J. Chanussot, C. Thomas, P. Gamba, and L. M. Bruce, Comparison of pansharpening algorithms: Outcome of the 2006 GRS-S data fusion contest. *IEEE Trans. Geosci. Remote Sens.*, 45 (2007), 3012–3021.
- C. Thomas, Fusion d’images de résolutions spatiales différentes (Fusion of images of different spatial resolutions). Ph.D Thesis, Ecole des Mines de Paris, Paris, France, 11 December 2006, 165 p. and annexes.
- G. Peyré, Géométrie multi-échelles pour les images et les textures. Report Ecole Polytechnique, Centre de Mathématiques Appliquées, Palaiseau, France, 2005.
- D. Grois, I. Shcherback, T. Danov, and O. Yadid-Pecht, Theoretical approach to CMOS APS PSF and MTF modeling – evaluation. *Sensors Journal*, IEEE 6 (2006), 118-124.
- B. Aiazzi, L. Alparone, S. Baronti, A. Garzelli, M. Selva, A MTF-based spectral distortion minimizing model for pan-sharpening of very high resolution multispectral images of urban areas. In Proceedings 2nd GRSS/SPRS Joint Workshop on "Data Fusion and Remote Sensing over Urban Areas", 2003, 90-94.
- A. Garzelli, F. Nencini, Interband structure modeling for Pan-sharpening of very high-resolution multispectral images. *Information Fusion*, 6 (2005), 213 - 224
- B. Aiazzi, L. Alparone, S. Baronti, A. Garzelli, M. Selva, MTF-tailored multiscale fusion of high-resolution MS and Pan imagery. *Photogrammetric Engineering & Remote Sensing*, 72 (2006), 591-596.

Bibliographie

- [Aiazzi *et al.*, 2002] Aiazzi, B., Alparone, L., Baronti, S. et Garzelli, A. (2002). Context-driven fusion of high spatial and spectral resolution images based on oversampled multiresolution analysis. *IEEE Transactions on Geoscience and Remote Sensing*, 40(10):2300–2312.
- [Aiazzi *et al.*, 2003] Aiazzi, B., Alparone, L., Baronti, S., Garzelli, A. et Selva, M. (2003). An mtf-based spectral distortion minimizing model for pansharpening of very high resolution multispectral images of urban areas. In *2nd GRSS/ISPRS Joint Workshop on "Data Fusion and Remote Sensing over Urban Areas"*, pages 90–94.
- [Aiazzi *et al.*, 2006] Aiazzi, B., Alparone, L., Baronti, S., Garzelli, A. et Selva, M. (2006). Mtf-tailored multiscale fusion of high-resolution ms and pan imagery. *Photogrammetric Engineering & Remote Sensing*, 72(5):591–596.
- [Aiazzi *et al.*, 2009] Aiazzi, B., Baronti, S., Lotti, F. et Selva, M. (2009). A comparison between global and context-adaptive pansharpening of multispectral images. *IEEE Geoscience and Remote Sensing Letters*, 6(2):302–306.
- [Almansa *et al.*, 2004] Almansa, A., Durand, S. et Rougé, B. (2004). Measuring and improving image resolution by adaptation of the reciprocal cell. *Journal of Mathematical Imaging and Vision*, 3(3):235–279.
- [Alparone *et al.*, 2008] Alparone, L., Aiazzi, B., Baronti, S., Garzelli, A., Nencini, F. et Selva, M. (2008). Multispectral and panchromatic data fusion assessment without reference. *Photogrammetric Engineering & Remote Sensing*, 74(2):193–200.
- [Alparone *et al.*, 2004] Alparone, L., Baronti, S., Garzelli, A. et Nencini, F. (2004). A global quality measurement of pan-sharpened multispectral imagery. *IEEE Geoscience and Remote Sensing Letters*, 1(4):313–317.
- [Alparone *et al.*, 2007] Alparone, L., Wald, L., Chanussot, J., Thomas, C., Gamba, P. et Bruce, L. (2007). Comparison of pansharpening algorithms: Outcome of the 2006 grs-s data-fusion contest. *IEEE Transactions on Geoscience and Remote Sensing*, 45(10):3012–3021.

- [Banham et Katsaggelos, 1997] Banham, M. et Katsaggelos, A. (1997). Digital image restoration. *Signal Processing Magazine, IEEE*, 14(2):24–41.
- [Blanc *et al.*, 1998] Blanc, P., Blu, T., Ranchin, T., Wald, L. et Aloisi, R. (1998). Using iterated rational filter banks within arsis concept for producing landsat multispectral images. *International Journal in Remote Sensing*, 19(12):2331–2343.
- [Blanchet, 2006] Blanchet, G. (2006). *Etude des artefacts de flou, ringing et aliasing en imagerie numérique. Applications à la restauration*. Thèse de doctorat, Ecole Normale Supérieure de Cachan.
- [Chen *et al.*, 2009] Chen, X.-b., Yang, S.-z., Wang, X.-h. et Qiao, Y.-l. (2009). Satellite image blind restoration based on surface fitting and iterative multishrinkage method in redundant wavelet domain. *Optik - International Journal for Light and Electron Optics*, In Press, Corrected Proof:–.
- [Chesnel, 2008] Chesnel, A.-L. (2008). *Quantification de dégâts sur le bâti liés aux catastrophes majeures par images satellite multimodales très haute résolution*. Thèse de doctorat, Mines Paristech.
- [CNES/ONERA/IGN, 2008] CNES/ONERA/IGN (2008). *Imagerie spatiale des principes d'acquisition au traitement des images optiques pour l'observation de la Terre*. Numéro ISBN 978.2.85428.844.5. Cépaduès.
- [Coulange et Moisan, 2010] Coulange, B. et Moisan, L. (2010). An aliasing detection algorithm based on suspicious colocalizations of fourier coefficients. *In proceedings of the International Conference on Image Processing*.
- [Delon, 2004] Delon, J. (2004). *Comparaison fine d'images et autres problèmes*. Thèse de doctorat, École normale supérieure de Cachan.
- [Delvit *et al.*, 2003] Delvit, J.-M., Leger, D., Roques, S. et Valorge, C. (2003). Modulation transfer function and noise assessment. *In Proc. IEEE International Geoscience and Remote Sensing Symposium IGARSS '03*, volume 7, pages 4500–4502.
- [Do et Vetterli, 2005] Do, M. et Vetterli, M. (2005). The contourlet transform: an efficient directional multiresolution image representation. *IEEE Transactions on Image Processing*, 14(12):2091 –2106.
- [Douillet et Blanc, 2009] Douillet, F. et Blanc, P. (2009). Method for increasing the resolution of multi-spectral images. patent : WO 2009/030698 A1.

-
- [Dubois *et al.*, 1997] Dubois, J.-M., Donnay, J. P. et Ozer, A. e. a. (1997). *Téledétection des milieux urbains et périurbains*. Numéro ISBN 2-9200-2174-5. Agence universitaire de la Francophonie.
- [Estribeau et Magnan, 2004] Estribeau, M. et Magnan, P. (2004). Fast mtf measurement of cmos imagers at the chip level using iso 12233 slanted-edge methodology. In Meynart, R., Neeck, S. P. et Shimoda, H., éditeurs : *Sensors, Systems, and Next-Generation Satellites VIII*, volume 5570, pages 557–567. SPIE.
- [Gamba et Chanussot, 2008] Gamba, P. et Chanussot, J. (2008). Foreword to the special issue on data fusion. *IEEE Transactions on Geoscience and Remote Sensing*, 46(5):1283 –1288.
- [Garzelli et Nencini, 2005] Garzelli, A. et Nencini, F. (2005). Interband structure modeling for pan-sharpening of very high-resolution multispectral images. *Information Fusion*, 6(3):213 – 224.
- [Gaudin-Delerieu *et al.*, 2008] Gaudin-Delerieu, C., Lamard, J.-L., Cheroutre, P., Bailly, B., Dhucq, P. et Puig, O. (2008). The high resolution optical instruments for the pleiades hr earth observation satellites. In *Proceedings of the International Conference on Space Optics, 14-17 October 2008, Toulouse, France*.
- [Girard et Girard, 1999] Girard, M.-C. et Girard, C.-M. (1999). *Traitement des Données de Téledétection*. Dunod.
- [Grois *et al.*, 2006] Grois, D., Shcherback, I., Danov, T. et Yadid-Pecht, O. (2006). Theoretical approach to cmos aps psf and mtf modeling - evaluation. *Sensors Journal, IEEE*, 6(1):118–124.
- [Khan *et al.*, 2009] Khan, M., Alparone, L. et Chanussot, J. (2009). Pansharpening quality assessment using the modulation transfer functions of instruments. *IEEE Transactions on Geoscience and Remote Sensing*, 47(11):3880 –3891.
- [Khan *et al.*, 2008] Khan, M., Chanussot, J., Condat, L. et Montanvert, A. (2008). Indusion: Fusion of multispectral and panchromatic images using the induction scaling technique. *IEEE Geoscience and Remote Sensing Letters*, 5(1):98 –102.
- [Khan, 2009] Khan, M. M. (2009). *Fusion and Quality Assesment of Multispectral, Hyperspectral and Panchromatic Satellite Images*. Thèse de doctorat, Institut Polytechnique de Grenoble.

- [Kohm, 2004] Kohm, k. (2004). Modulation transfer function measurement method and results from the orbview-3 high resolution imaging satellite. *In Proceedings of ISPRS 2004*.
- [Krim *et al.*, 1999] Krim, H., Tucker, D., Mallat, S. et Donoho, D. (1999). On denoising and best signal representation. *IEEE Transactions on Information Theory*, 45(7):2225 –2238.
- [Kundur et Hatzinakos, 1996] Kundur, D. et Hatzinakos, D. (1996). Blind image deconvolution. *Signal Processing Magazine, IEEE*, 13(3):43 –64.
- [Lamard *et al.*, 2004] Lamard, J.-L., Gaudin-Delrieu, C., Valentini, D., Renard, C., Tournier, T. et Laherrere, J.-M. (2004). Design of the high resolution optical instrument for the pleiades hr earth observation satellites. *In B. Warmbein. ESA SP-554, Noordwijk, N., éditeur : Proceedings of the 5th International Conference on Space Optics (ICSO 2004)*.
- [Laporterie-Déjean *et al.*, 2005] Laporterie-Déjean, F., de Boissezon, H., Flouzat, G. et Lefèvre-Fonollosa, M.-J. (2005). Thematic and statistical evaluations of five panchromatic/multispectral fusion methods on simulated pleiades-hr images. *Information Fusion*, 6(3):193 – 212.
- [Latry *et al.*, 2001] Latry, C., Rougé, B. et Baillarin, S. (2001). La chaîne image spot5 thr : un exemple d’optimisation globale. *In 18° Colloque sur le traitement du signal et des images, 2001, Toulouse, pages 184–187. GRETSI*.
- [Latry et Rouge, 2003] Latry, C. et Rouge, B. (2003). Super resolution: quincunx sampling and fusion processing. *In Geoscience and Remote Sensing Symposium, 2003. IGARSS '03. Proceedings. 2003 IEEE International*, volume 1, pages 315 – 317 vol.1.
- [Le Pennec, 2002] Le Pennec, E. (2002). *Bandelettes et représentations géométriques des images*. Thèse de doctorat, École Polytechnique (CMAP).
- [Lee et Lee, 2010] Lee, J. et Lee, C. (2010). Fast and efficient panchromatic sharpening. *IEEE Transactions on Geoscience and Remote Sensing*, 48(1):155 –163.
- [Leger *et al.*, 1994] Leger, D., Duffaut, J. et Robinet, F. (1994). Mtf measurement using spotlight. *In Geoscience and Remote Sensing Symposium, 1994. IGARSS '94. Surface and Atmospheric Remote Sensing: Technologies, Data Analysis and Interpretation.*, volume 4, pages 2010 –2012 vol.4.

-
- [Lei et Tiziani, 1989] Lei, F. et Tiziani, H. (1989). Modulation transfer function obtained from image structures. In Linkwithz, K. et Hangleiter, U., éditeurs : *Proceedings of the International Workshops on High precision navigation, Stuttgart, 1988*, pages 366–377. Springer, Heidelberg.
- [Li, 2000] Li, J. (2000). Spatial quality evaluation of fusion of different resolution images. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 33:339–346.
- [Luxen et Förstner, 2002] Luxen, M. et Förstner, W. (2002). Characterizing image quality: Blind estimation of the point spread function from a single image. In *PCV02*, page A: 205.
- [Malgouyres, 1999] Malgouyres, F. (1999). *Augmentation de la résolution d'images satellitales: Théorie variationnelle et applications*. Thèse de doctorat, Université de Paris IX – Dauphine.
- [Mallat, 1989] Mallat, S. (1989). A theory for multiresolution signal decomposition: the wavelet representation. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 11(7):674 –693.
- [Mallat, 1999] Mallat, S. (1999). *A Wavelet Tour of Signal Processing*. Academic Press.
- [Marion, 97] Marion, A. (97). *Acquisition & visualisation des images*. Eyrolles.
- [Nikolakopoulos, 2008] Nikolakopoulos, K. G. (2008). Comparison of nine fusion techniques for very high resolution data. *Photogrammetric Engineering & Remote Sensing*, 74(5):647–659.
- [Nunez et al., 1999] Nunez, J., Otazu, X., Fors, O., Prades, A., Pala, V. et Arbiol, R. (1999). Multiresolution-based image fusion with additive wavelet decomposition. *IEEE Transactions on Geoscience and Remote Sensing*, 37(3):1204–1211.
- [Otazu et al., 2005] Otazu, X., Gonzalez-Audicana, M., Fors, O. et Nunez, J. (2005). Introduction of sensor spectral response into image fusion methods. application to wavelet-based methods. *IEEE Transactions on Geoscience and Remote Sensing*, 43(10):2376 – 2385.
- [Papoulis, 1987] Papoulis, A. (1987). *Signal Analysis*. Numéro ISBN 0-07-048460-0. McGraw Hill International Editions.
- [Patra et al., 2002] Patra, S., K., Mishra, N. et Chandrakanth, R. (2002). Image quality improvement through mtf compensation - a treatment to high resolution data. *Indian Cartographer*, DAPI-15:86–93.

- [Piella, 2003] Piella, G. (2003). A general framework for multiresolution image fusion: from pixels to regions. *Information Fusion*, 4(4):259 – 280.
- [Pohl et van Genderen, 1998] Pohl, C. et van Genderen, J.-L. (1998). Multisensor image fusion in remote sensing: concepts, methods and applications. *International Journal in Remote Sensing*, 19(5):823–854.
- [Péteri et Ranchin, 2004] Péteri, R. et Ranchin, T. (2004). Potentialités des nouveaux capteurs à très haute résolution spatiale pour l'extraction des réseaux de rues urbains. *European Journal of GIS and Spatial Analysis*, 14(3-4):485–504.
- [Ranchin et al., 2003] Ranchin, T., Aiazzi, B., Alparone, L., Baronti, S. et Wald, L. (2003). Image fusion - the arsis concept and some successful implementation schemes. *International Journal of Photogrammetry & Remote Sensing*, 58:4–18.
- [Ranchin et Wald, 1993] Ranchin, T. et Wald, L. (1993). The wavelet transform for the analysis of remotely sensed images. *International Journal of Remote Sensing*, 14(3):615–619.
- [Ranchin et Wald, 2000] Ranchin, T. et Wald, L. (2000). Fusion of high spatial and spectral resolution images: The arsis concept and its implementation. *Photogrammetric Engineering and Remote Sensing*, 66(1):49–61.
- [Rosak et al., 2004] Rosak, A., Latry, C., Pascal, V. et Laubier, D. (2004). From spot5 to pleiades-hr : Evolution of the instrumental specifications. In *Proceedings of the 5th International Conference on Space Optics 2004 (ISCO), 30 March - 2 April 2004, Toulouse, France*, numéro ISBN 92-9092-865-4, pages 141 – 148.
- [Ryan et al., 2003] Ryan, R., Baldrige, B., Schowengerdt, R. A., Choi, T., Helder, D. L. et Blonski, S. (2003). Ikonos spatial resolution and image interpretability characterization. *Remote sensing of environment*, 88(1-2):37–52.
- [Samson et al., 2000] Samson, C., Blanc-Feraud, L., Aubert, G. et Zerubia, J. (2000). A variational model for image classification and restoration. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(5):460 –472.
- [Shah et al., 2008] Shah, V., Younan, N. et King, R. (2008). An efficient pan-sharpening method via a combined adaptive pca approach and contourlets. *IEEE Transactions on Geoscience and Remote Sensing*, 46(5):1323 –1335.

-
- [Shcherback et Yadid-Pecht, 2001] Shcherback, I. et Yadid-Pecht, O. (2001). Cmos aps mtf modeling. *Electron Devices, IEEE Transactions on*, 48(12):2710 – 2715.
- [Sroubek et Flusser, 2003] Sroubek, F. et Flusser, J. (2003). Multichannel blind iterative image restoration. *IEEE Transactions on Image Processing*, 12(9):1094–1106.
- [Thomas, 2006] Thomas, C. (2006). *Fusion d'images de résolutions spatiales différentes*. Thèse de doctorat, Ecole des Mines de Paris.
- [Thomas et al., 2008] Thomas, C., Ranchin, T., Wald, L. et Chanussot, J. (2008). Synthesis of multispectral images to high spatial resolution: A critical review of fusion methods based on remote sensing physics. *IEEE Transactions on Geoscience and Remote Sensing*, 46(5):1301–1312.
- [Thomas et Wald, 2006] Thomas, C. et Wald, L. (2006). A mtf-based distance for the assessment of geometrical quality of fused products. In *Proceedings of FUSION06, 10-13 July, Florence (Italy)*, pages 1–7.
- [Unser et al., 1992] Unser, M., Aldroubi, A. et Eden, M. (1992). Polynomial spline signal approximations: filter design and asymptotic equivalence with shannon's sampling theorem. *IEEE Transactions on Information Theory*, 38(1):95 –103.
- [Viallefont-Robinet et Léger, 2010] Viallefont-Robinet, F. et Léger, D. (2010). Improvement of the edge method for on-orbit mtf measurement. *Opt. Express*, 18(4):3531–3545.
- [Wald, 1999] Wald, L. (1999). Some terms of reference in data fusion. *IEEE Transactions on Geoscience and Remote Sensing*, 37(3):1190–1193.
- [Wald, 2000] Wald, L. (2000). Quality of high resolution synthesised images: is there a simple criterion? In SEE/URISCA, éditeur : *the third conference "Fusion of Earth data: merging point measurements, raster maps and remotely sensed images"*, Sophia Antipolis, France, January 26-28, 200, pages 99–103.
- [Wald, 2002] Wald, L. (2002). *Data Fusion: Definitions and Architectures. Fusion of Images of Different Spatial Resolutions*. Numéro 2-911762-38-X. Les Presses de l'Ecole des Mines.
- [Wald et al., 1997] Wald, L., Ranchin, T. et Mangolini, M. (1997). Fusion of satellite images of different spatial resolutions: Assessing the quality of resulting images. *Photogrammetric Engineering & Remote Sensing*, 63:691–699.

- [Wang *et al.*, 2005] Wang, Z., Ziou, D., Armenakis, C., Li, D. et Li, Q. (2005). A comparative analysis of image fusion methods. *IEEE Transactions on Geoscience and Remote Sensing*, 43(6):1391–1402.
- [Welch, 1967] Welch, P. (1967). The use of fast fourier transform for the estimation of power spectra: A method based on time averaging over short, modified periodograms. *IEEE Transactions Audio and Electroacoustics*, 15:70–73.
- [Yadid-Pecht, 2000] Yadid-Pecht, O. (2000). Geometrical modulation transfer function for different pixel active area shapes. *Optical Engineering*, 39(4):859–865.

Fusion de données : prise en compte des caractéristiques liées à l'imageur lors de la synthèse d'images multispectrales à haute résolution spatiale

RESUME : Les satellites d'observation de la Terre (Quickbird, Ikonos, SPOT-5, PLÉIADES) fournissent des images de deux types : soit une image panchromatique à haute résolution spatiale mais basse résolution spectrale, soit une image multispectrale dont la résolution spatiale est inférieure d'un rapport 2 ou 4, mais avec une bonne résolution spectrale. Des méthodes permettant la fusion de ces deux types d'images existent, cependant il est encore possible d'améliorer la qualité des images synthétisées.

Nos travaux visent à améliorer la qualité des images synthétisées par une meilleure prise en compte des caractéristiques de l'imageur comme ; la Fonction de Transfert de Modulation (FTM) et l'échantillonnage.

Les travaux que nous avons menés sur l'impact du repliement de spectre présent dans les images MS ont mis en évidence que l'absence de repliement permet la suppression des artefacts visuels importants, mais les outils usuels d'évaluation de la qualité des images de fusion ne permettent pas la quantification des performances.

La FTM caractérise la capacité de l'instrument à représenter un contraste en fonction de la fréquence spatiale. Celle-ci diffère entre les images PAN et MS. Nous avons proposé à partir de méthodes existantes, deux méthodes de fusion permettant de mieux prendre en compte la FTM et ainsi d'améliorer la qualité et les caractéristiques de l'image synthétisée.

Mots clés : fusion d'images, modalités panchromatique et multispectrales, concept ARSIS, infrarouge, télédétection, analyse multiéchelle et multirésolution, transformée en ondelettes, fonction de transfert de modulation, repliement de spectre.

Data fusion: integrating imaging sensor characteristics in the synthesis of multispectral images at high spatial resolution

ABSTRACT: Earth observation satellites (Quickbird, Ikonos, SPOT-5, PLÉIADES) provide two types of images: the panchromatic image at high spatial resolution with a poor spectral resolution, and the multispectral image with a good spectral resolution but at a spatial resolution 2 or 4 times lower than each of the panchromatic. Image fusion methods for this to types of images already exist, whereas we still can add/provide amelioration that aim to improve the quality of synthesized images.

My work look for improving the quality of synthesized images thanks to a better consideration of the sensor characteristics like the Modulation Function Transfer (MTF), and the sampling.

The work on the evaluation of the effect of the aliasing of the MS on the synthesized MS show that the suppression of aliasing suppress visual artifact but it reveal that it not significant with actual quality analysis.

MTF express the sensor capacities to represent a contrast in function of the spatial frequency. Based on an existing method we proposed two new methods that allowed to better take into account the MTF difference that exist between PAN and MS images.

Two new methods had been proposed that are based on existing method and allowed to better take into which take into account the difference of MTF between PAN and MS images.

Keywords : image fusion, panchromatic and multispectral modalities, ARSIS concept, infrared, remote sensing, Multiscale multiresolution analysis, wavelet transform, modulation function transfer, aliasing.

