HAL
open science

Estimation de l'attitude d'un satellite à l'aide de caméras pushbroom et de capteurs stellaires

Régis Perrier

To cite this version:

Régis Perrier. Estimation de l'attitude d'un satellite à l'aide de caméras pushbroom et de capteurs stellaires. Mathématiques générales [math.GM]. Université de Grenoble, 2011. Français. NNT : 2011GRENM040 . tel-00637592

HAL Id: tel-00637592

https://theses.hal.science/tel-00637592

Submitted on 2 Nov 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.

UNIVERSITÉ DE GRENOBLE

THÈSE

Pour obtenir le grade de

DOCTEUR DE L'UNIVERSITÉ DE GRENOBLE

Spécialité : **Mathématiques et Informatique**

Arrêté ministériel : 7 août 2006

Présentée par

Régis Perrier

Thèse dirigée par **Peter STURM**
et codirigée par **Elise ARNAUD**

préparée au sein du **Laboratoire Jean Kuntzmann**
et de l'école doctorale **MSTII**

Estimation de l'attitude d'un satellite à l'aide de caméras *pushbroom* et de capteurs stellaires

---

Thèse soutenue publiquement le **27 septembre 2011**,
devant le jury composé de :

Mr Augustin LUX
Professeur Institut National Polytechnique de Grenoble, Président
Mme Josianne ZERUBIA
Directrice de Recherche INRIA, Rapporteur
Mr Adrien BARTOLI
Professeur Université d'Auvergne, Rapporteur
Mr Patrick BOUTHEMY
Directeur de Recherche INRIA, Examinateur
Mr André JALOBEANU
Chargé de Recherche CNRS, Examinateur
Mr Peter STURM
Directeur de Recherche INRIA, Directeur de thèse
Mme Elise ARNAUD
Maître de Conférences Université Joseph Fourier, Co-Directeur de thèse
Mr Mathias ORTNER
Ingénieur EADS Astrium, Co-Directeur de thèse

Résumé

Les caméras *pushbroom* sont omniprésentes en imagerie satellitaire. Ce capteur linéaire enregistre des images 1-D et utilise le défilement du satellite autour de la terre pour construire des bandeaux d'image ; son principe de fonctionnement est identique aux scanners et photocopieurs que l'on peut utiliser tous les jours. Les avantages liés à cette technologie sont principalement une résolution d'image étendue qui va bien au delà des caméras perspectives, un coût d'exploitation faible et une robustesse au contexte spatial. Pour reconstruire des images couleur, le plan focal d'un satellite embarque plusieurs caméras *pushbroom* sensibles à différentes bandes spectrales de la lumière.
Ce mode d'acquisition dépendant du temps suppose que l'orientation du satellite, également appelée attitude dans cette étude, ne varie pas au cours du survol d'une scène. Les satellites ont jusqu'à maintenant été considérés comme stables du fait de leur inertie. Cependant les technologies récentes développées dans la recherche spatiale tendent à réduire leur taille et alléger leur poids pour les rendre plus agiles et moins coûteux en énergie lors de leur mise en orbite. La résolution des capteurs a également été améliorée, ce qui rend nettement plus critique la moindre oscillation de l'imageur. Ces facteurs cumulés font qu'un changement d'attitude de quelques microradians peut provoquer des déformations géométriques notables dans les images.
Les solutions actuelles utilisent les capteurs de positionnement du satellite pour asservir son attitude et rectifier les images, mais elles sont coûteuses et limitées en précision. Les images contiennent pourtant une information cohérente sur les mouvements du satellite de par leurs éventuelles déformations. Nous proposons dans cette étude de retrouver les variations d'attitude par recalage des images enregistrées par le satellite. Nous exploitons la disposition des caméras *pushbroom* dans le plan focal ainsi que la nature stationnaire des oscillations pour conduire l'estimation. Le tout est présenté dans un cadre bayesien, où les données images peuvent se mêler avec une information *a priori* sur le mouvement ainsi que des mesures exogènes fournies par un capteur stellaire couramment appelé *star tracker*. Différentes solutions sont décrites et comparées sur des jeux de données satellitaires fournis par le constructeur de satellite EADS Astrium [1].

Mots clés : caméra *pushbroom*, *star tracker*, satellite, variation d'attitude, recalage d'image, modèle temporel, estimation bayésienne.

1. Ce travail a été entièrement financé par EADS Astrium (constructeur européen de satellite).

Abstract

Linear pushbroom cameras are widely used for earth observation applications. This sensor acquires 1-D images over time and uses the straight motion of the satellite to sweep out a region of space and build 2-D image ; it operates in the same way as a usual flatbed scanner. Main advantages of such technology are : robustness in the space context, higher resolution than classical 2-D CCD sensors and low production cost. To build color images, several pushbroom cameras of different modalities are set in parallel onto the satellite's focal plane.

This acquisition process is dependent of the time and assumes that the satellite's attitude remains constant during the image recording. However, the recent manufacture of smaller satellites with higher sampling resolution has weakened this assumption. The satellite may oscillates around its rotations axis, and an angular variation of a few microradians can result in noticeable warps in images.

Current solutions use inertial sensors on board the satellite to control the attitude and correct the images, but they are costly and of limited precision. As warped images do contain the information of attitude variations, we suggest to use image registration to estimate them. We exploit the geometry of the focal plane and the stationary nature of the disturbances to recover undistorted images. To do so, we embed the estimation process in a Bayesian framework where image registration, prior on attitude variations and measurements of a star tracker are fused to retrieve the motion of the satellite. We illustrate the performance of our algorithm on four satellite datasets provided by EADS Astrium [2].

Keywords : pushbroom camera, star tracker, satellite, attitude variation, image registration, temporal model, bayesian estimation.

2. This work was funded by EADS Astrium (European aerospace company and satellite manufacturer).

Remerciements

Plusieurs personnes ont contribué au bon déroulement de ce travail et je tiens à les remercier. L'exercice n'est pas simple car la thèse permet de côtoyer tellement de gens passionnants qu'on peut difficilement les oublier. Je commencerai par les personnes dont le nom ne figure pas ici mais qui se reconnaîtront pour m'avoir aidé à progresser d'une façon ou d'une autre au cours de ces trois dernières années.

Merci aux membres de mon jury pour avoir accepté de lire mon manuscrit, jugé de façon constructive mon travail et validé l'ensemble en m'accordant le titre de docteur.

Ce travail n'aurait jamais existé sans le soutien financier de la société EADS Astrium ; je les remercie pour m'avoir confié ce sujet d'étude vraiment passionnant.

Elise et Peter ont été plus que de simples encadrants dans cette thèse. Ils m'ont aidé à clarifier mes idées et donné l'opportunité de m'épanouir dans l'enseignement et l'organisation de la conférence ORASIS ; ces deux expériences m'ont enrichi à tous les niveaux. Ils m'ont fait confiance dans la gestion de mon temps de travail et je les en remercie. Du coté de EADS Astrium, Mathias a fait preuve de patience et a su m'appuyer auprès de sa direction pour me permettre de publier. Il a également inspiré plusieurs idées dans ce travail de par sa connaissance du problème, un grand merci.

L'équipe PERCEPTION a été un environnement de travail très stimulant au cours de ces années. C'était un plaisir de pouvoir discuter avec les permanents sur de nombreux sujets divers et variés pendant les repas et les pauses cafés. Une mention spéciale aux deux assistantes de choc de l'équipe, Anne et Marie-Eve, qui nous font oublier toute formalité administrative par la qualité de leur travail et qui m'ont permis de voyager dans des lieux magiques tels qu'Hawaii ou la Nouvelle-Zélande...

Je remercie tous les membres non permanents de l'équipe, ils sont malheureusement trop nombreux pour tenir en quelques lignes entre les nouveaux et les anciens. Je pense aux petits cakes au beurre que sont Amaël, Antoine, Jamil, Gaëtan, Michel et Benjamin : toujours prêts pour l'aventure, plus efficaces qu'une recherche google... bref les collègues et amis que l'on souhaite avoir tout le temps. Bonne continuation également à mes camarades de bureau qu'étaient Mauricio, Kiran, Andrei ; on est tous docteurs maintenant ! Un grand merci à notre chef italien Simone qui a rendu possible tellement de barbecue près du garage à vélo de l'INRIA. Enfin, une mention pour Lamiae et Ramya qui ont apporté une touche féminine dans ce monde singulier de la recherche en informatique.

Un clin d'oeil aux petits gars du LGIT puisque je n'ai pas fait de thèse dans leur laboratoire, notamment à Anthonio Matyniani, Tom le rider tecktonik et les Causses' brothers. Au-delà des moments hilarants que j'ai pu vivre en leur présence, je pense avoir appris beaucoup de chacun d'eux et je les en remercie.

Je termine ces remerciements en les adressant à mes amis toulonnais Bastien et Christophe, à ma famille, et à Carole pour m'avoir soutenu à tout moment dans cette thèse. C'est avec une petite pointe de tristesse que je clos ces belles années, mais le meilleur est à venir.

Table des matières

I Introduction **13**
- I.1 Introduction au capteur *pushbroom* . . . 13
- I.2 Contexte d'utilisation . . . 13
- I.3 Avantages et contraintes . . . 15
- I.4 Plan du manuscrit . . . 17

II Problématique et état de l'art **19**
- II.1 Introduction . . . 19
- II.2 Configuration du plan focal . . . 19
- II.3 Problématique . . . 21
- II.4 Etat de l'art . . . 23
  - II.4.1 Géométrie . . . 23
  - II.4.2 Estimation de l'attitude . . . 24
  - II.4.3 Recalage d'image . . . 25
- II.5 Conclusion . . . 27

III Formalisme bayesien du problème **29**
- III.1 Introduction . . . 29
- III.2 Cadre probabiliste . . . 29
- III.3 Formulation du maximum *a posteriori* . . . 32
- III.4 Vraisemblance des images . . . 33
  - III.4.1 Statistique de bruit des images . . . 34
  - III.4.2 Fonction de déformation . . . 35
  - III.4.3 Expression d'une convolution . . . 36
  - III.4.4 Ecriture de la vraisemblance . . . 37
  - III.4.5 Linéarisation du terme image . . . 38
  - III.4.6 Maximum de Vraisemblance (MV) . . . 39
  - III.4.7 Exemple d'estimation par MV . . . 41
- III.5 *A priori* sur l'attitude . . . 43
  - III.5.1 Notations . . . 45
  - III.5.2 Amplitude . . . 46
  - III.5.3 Dérivée seconde . . . 49
  - III.5.4 Processus Gaussiens . . . 52
- III.6 Processus autorégressif . . . 57
- III.7 Discussion sur les estimateurs . . . 60
- III.8 Contraintes sur l'évolution temporelle . . . 63
  - III.8.1 Modèle polynômial par morceaux . . . 63

III.8.2 Expression de la vraisemblance 66
III.8.3 Estimateur du MV 66
III.8.4 Résultats sur les données 67
III.8.5 Discussion 67
III.9 Fusion de données 69
III.9.1 Extension du modèle probabiliste 70
III.9.2 Bruit du capteur et vraisemblance 70
III.9.3 Expression du MAP 72
III.9.4 Résultats sur les données 73
III.9.5 Modèle polynômial basse fréquence 74
III.10 Conclusion 76

IV Algorithmes d'estimation et résultats 79
IV.1 Introduction 79
IV.2 Estimation des hyperparamètres 79
IV.2.1 Erreur de prédiction 80
IV.2.2 Validation Croisée (VC) 81
IV.2.3 Erreur photométrique 87
IV.2.4 Vraisemblance marginale 91
IV.2.5 Coefficients du modèle polynômial basse fréquence 92
IV.2.6 Coefficients du modèle autorégressif 92
IV.3 Recalage multimodal 94
IV.3.1 Images brutes 98
IV.3.2 Images de contours 98
IV.3.3 Correction radiométrique 102
IV.3.4 Discussion 104
IV.4 Estimation itérative 105
IV.5 Résultats sur des données 106
IV.5.1 Jeu de données ASTRO 109
IV.5.2 Jeu de données MLF 116
IV.5.3 Jeu de données Reunion 122
IV.5.4 Jeu de données ALMF 128
IV.6 Expériences complémentaires 133
IV.6.1 Pourcentage de pixels dans l'estimation 133
IV.6.2 Influence du choix de l'hyperparamètre 136
IV.6.3 Estimation du bruit image 137
IV.6.4 Signal non stationnaire 138
IV.7 Conclusion 139

V Conclusion 141

A Matrices et développements MV/MAP 145
A.1 Matrice de décalage 145
A.2 Matrice de convolution 146
A.3 Matrice d'*a priori* sur la dérivée seconde 146
A.4 Matrice d'*a priori* autoregressif 147
A.5 Matrice de temps du modèle polynômial par morceaux 147
A.6 Matrice de contraintes du modèle polynômial par morceaux 147

A.7 Matrice de filtrage basse fréquence 148
A.8 Solution du maximum de vraisemblance 148
A.9 Développement du maximum *a posteriori* 149

CHAPITRE I

Introduction

Cette thèse s'intéresse à l'estimation des variations d'attitude d'un satellite à partir des images de la terre qu'il photographie. Elle a été financée par EADS Astrium qui est un acteur majeur de l'aérospatiale en Europe. Ce chapitre introductif présente les motivations de ce travail en commençant par présenter le fonctionnement d'une caméra *pushbroom* qui est à l'origine de notre travail. Nous étudions son principe d'acquisition, ainsi que ses avantages sur des capteurs plus traditionnels de type *pinhole*. La fin de cette partie effleure la problématique, qui sera détaillée au chapitre suivant, pour nous amener au plan de travail de ce manuscrit.

I.1 Introduction au capteur *pushbroom*

Dans son principe, la caméra *pushbroom* est un capteur linéaire qui enregistre des images selon une dimension dans l'espace. A l'opposé de la caméra *pinhole* présentée en figure I.1(a) qui projette en deux dimensions une scène observée, ce système optique se contente de la photographier sur une seule ligne de pixel.
Lorsqu'il est fixé sur une plateforme mobile dont le déplacement est rectiligne et orthogonal à son axe, le capteur parcourt la scène et acquiert plusieurs images 1-D au cours du temps. Les lignes de pixel accolées entre elles forment une bande qui peut être découpée en plusieurs images 2-D ; le résultat obtenu donne l'impression que la prise de vue a été faite par une caméra standard, alors même que le capteur n'est constitué que d'une ligne de photosites. Le procédé d'acquisition est résumé par la figure I.2, et son opposition au modèle *pinhole* est décrite en figure I.1.
Nous pouvons définir un terme important pour la suite de notre étude d'après la figure I.2 : nous désignons par attitude les trois angles de rotation de la caméra que sont le tangage, le roulis et le lacet, respectivement les rotations autour de l'axe x, y et z ; ce nom et cette convention d'orientation sont très usuels dans l'ingénierie aérospatiale.

I.2 Contexte d'utilisation

L'idée de photographier une scène par défilement à la manière de la caméra *pushbroom* n'est pas récente, elle a été employée dans l'imagerie aérienne autour des années 1930. Un film argentique embarqué dans un avion était étiré lorsque l'engin survolait une zone

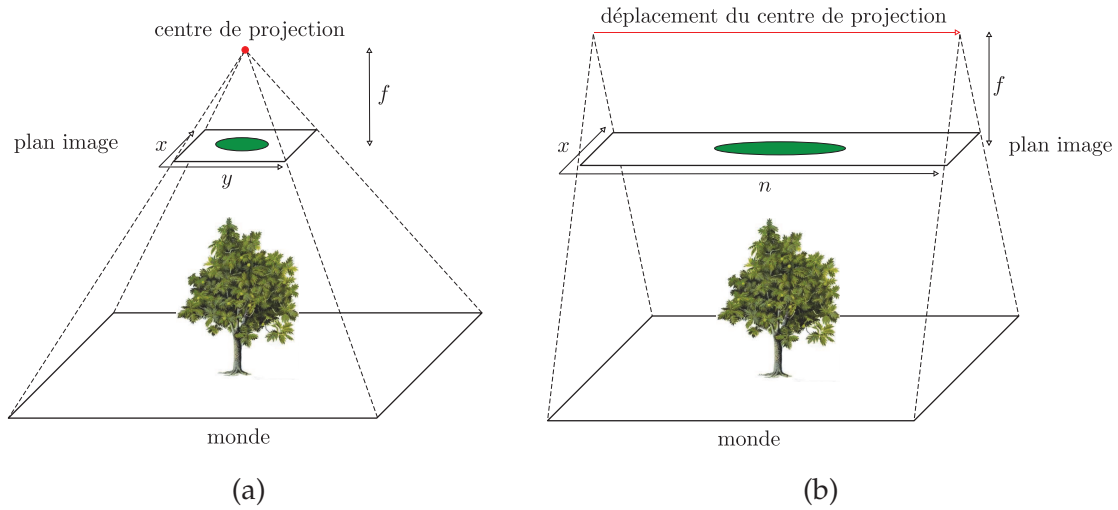


FIGURE I.1 – (a) Représentation du modèle d'acquisition *pinhole*, où la projection est perspective selon les deux axes du plan image x et y. Le résultat est une image 2-D, ou matrice de pixel, de dimension dépendante de la taille du capteur. La variable f désigne la distance focale. (b) Principe d'acquisition *pushbroom*, où la projection est perspective selon l'axe x et orthographique selon l'axe n. L'image 2-D se construit au fil du déplacement du capteur orthogonalement à son axe x ; sa taille est liée à la dimension du capteur selon x, à l'échantillonnage temporel de la caméra et à la durée d'acquisition selon l'axe n.

d'intérêt, ce concept portait alors le nom de Sonne cameras. Par la suite, ce système a été largement déployé sur des vols de reconnaissance pendant les guerres, jusqu'à l'apparition des premiers capteurs CCD avec un support optique adapté dans les années 1970, toujours pour des applications militaires. A cette époque, des plateformes stéréovisions étaient déjà disponibles pour obtenir une information de relief sur la scène. Plusieurs caméras disposées en parallèle sur un même plan focal et pointant au sol à des angles légèrement différents du nadir - verticale descendante entre le sol et l'engin d'observation - permettaient par comparaison des images d'estimer une élévation du terrain.
L'utilisation de cette technologie dans le contexte spatial a démarré en 1983, avec l'utilisation d'une caméra MOMS (*Modular Optoelectronic Multispectral Scanner*) à bord d'une navette spatiale américaine *Challenger* [MOMS 83] ; cet imageur a été développé par une société allemande du nom de Messerschmitt-Boelkow-Blohm. Ensuite, plusieurs satellites ont adopté des solutions variées sur le principe du premier modèle allemand. On peut notamment citer la série française des satellites SPOT depuis 1986 [SPOT 86], le satellite américain QuickBird en 2001 [QuickBird 01], jusqu'à GeoEye-1 [GeoEye 10] qui propose une des meilleures résolutions pour un satellite d'observation à application civile [1]. Nous reportons le lecteur à l'article [Gordon 05] pour une revue intéressante et très détaillée de toutes les technologies de type *pushbroom* qui ont été déployées pour l'observation de la terre.
A une autre échelle, les capteurs *pushbroom* sont beaucoup employés dans les scanners et photocopieurs personnels, où la caméra est fixée sur un rail motorisé pour parcourir le document ; ce système a l'avantage d'être peu coûteux. Certains appareils photographiques nommés *digital scan back* proposent également d'utiliser des CCD linéaires à ba-

1. 1.65 mètre en multispectral et 0.41 mètre en panchromatique

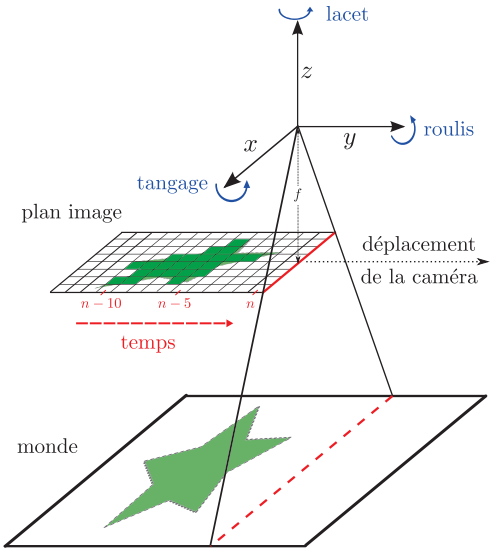


FIGURE I.2 – Principe simplifié de la construction d'une image 2-D par déplacement de la caméra *pushbroom*. Au cours du temps, les lignes de pixels accumulées forment une photographie de perception presque identique à ce que pourrait obtenir une caméra *pinhole*. L'axe de la caméra est représenté par x, son déplacement est rectiligne selon l'axe y et l'axe z est orthogonal au plan du monde. Au cours de notre étude, nous ferons plus souvent référence à l'axe n du plan image qui est identique à l'axe du repère spatial y ; notons également que dans le modèle *pushbroom*, le temps défile selon l'axe n.

layage, plutôt qu'une matrice CCD classique 2-D. Les constructeurs de ces technologies - parmi les plus connus Better Light et Kigamo - mettent en avant la qualité des couleurs et la résolution des images jusqu'à (12000×15990) pixels. Ils sont avant tout utilisés pour des photographies de paysage où la scène est statique.
Dans d'autres domaines, le principe du balayage de la surface observée selon le modèle *pushbroom* est largement utilisé dans les aéroports et les ports pour observer aux rayons X les contenus de bagages et de marchandises. Dans l'imagerie médicale, la technique d'échographie, les scanners à rayons X ou raisonnance magnétique fonctionnent aussi selon ce mode d'acquisition. Les lecteurs de code barre sont également des types spécifiques de capteur linéaire. Enfin ajoutons que certains imageurs 3-D utilisent un balayage laser pour estimer la profondeur des objets analysés.

I.3 Avantages et contraintes

Certains avantages associés à l'utilisation de caméras *pushbroom* ont déjà été dits, le prix de fabrication et la résolution des images sont les deux éléments les plus importants. Ceci est particulièrement vrai dans le contexte spatial, où tout composant d'un satellite doit être validé au préalable pour supporter l'espace. Les produits électroniques grand public tels que nous les connaissons comme les appareils photo numérique ont des capteurs de résolution importante, mais aucun d'entre eux ne peut supporter les conditions extrêmes de température, les rayonnements nocifs et les flux de particules dans l'environnement spatial. Les caméras *pushbroom* développées pour les satellites sont robustes, et

proposent des résolutions nettement supérieures à ce qu'acquièrent nos appareils photo ; en comparaison le satellite GeoEye-1 datant de 2008 enregistre plus de 37000 pixels [2] sur une seule ligne alors que l'appareil professionnel Nikon D3X de 2009 a une matrice de (6048×4032) pixels sur son capteur. Nous pouvons préciser qu'il existe des systèmes d'acquisition en imagerie aérienne comme la Ultracam Eagle [Microsoft 10] dont la matrice a une taille de (20000×13000) pixels, mais qui ne peuvent pas répondre aux spécificités de l'espace.
Pour obtenir des images couleurs, les systèmes d'acquisition *pushbroom* doivent utiliser plusieurs caméras linéaires associées chacune à un filtre pour enregistrer des composantes spectrales spécifiques de la lumière. On retrouve assez souvent quatre bandes spectrales sensibles dans les satellites d'observation [Gordon 05] : rouge, vert, bleu et panchromatique. Cette dernière bande intègre l'ensemble des trois précédentes avec la composante infrarouge. Les capteurs linéaires sont fixés en parallèle sur un même plan focal, en conséquence l'acquisition d'une même scène n'est pas simultanée pour les quatre caméras, mais légèrement décalée dans le temps proportionnellement à l'espacement entre elles. Les images sont donc recalées suivant leur position respective dans le plan focal pour composer une image couleur, mais on peut comprendre également qu'une petite variation de l'attitude au cours du temps d'acquisition décale toutes les lignes dans chacune des images multispectrales. Cette observation est à l'origine même de notre étude, dont nous développerons complètement la problématique au chapitre suivant. Remarquons ici que contrairement aux capteurs matriciels qui utilisent des filtres pour enregistrer des images couleurs [Bayer 76], l'étape de dématriçage n'est pas nécessaire dans l'acquisition *pushbroom* car les capteurs de chaque bande de couleur sont de même taille. L'imageur est donc supposé stable lors du balayage de la scène par les caméras *pushbroom*, ce qui est rarement le cas dans le contexte de l'imagerie aérienne qui préfère utiliser des caméras de type *pinhole* [Gordon 05], ou qui utilise intensivement des capteurs inertiels pour mesurer les dérives en attitude et les rectifier dans les images [Gruen 03]. Le satellite a longtemps été considéré comme une plateforme stable de par son poids, de plus ses capteurs [3] permettent d'enregistrer des dérives basses fréquences pour contrôler son attitude et rectifier les images si nécessaire. Seulement, les avancées technologiques poussent au design de satellites plus petits, plus agiles, moins coûteux, et plus performants en terme de résolution d'image [SPOT 86]. Ces éléments cumulés fragilisent l'hypothèse de stabilité et rendent plus critiques les variations d'attitude comme analysé dans [de Lussy 08]. En effet, un satellite petit et léger sera aisément mis en orbite, mais son moment d'inertie sera plus réduit et il oscillera plus facilement. De même, une résolution accrue rendra plus visibles les déformations géométriques dans les images. Enfin une façon intéressante de réduire le coût d'un satellite peut consister à supprimer certains capteurs inertiels, mais peut-on palier leur absence alors qu'ils sont utiles dans l'asservissement en positionnement du satellite, et dans la rectification des images ? Nous touchons ici plusieurs points qui justifient notre travail : les variations d'attitude sont devenues un problème dans l'acquisition d'images satellitaires, les capteurs inertiels de type *star tracker* ou gyroscope ne sont pas suffisamment précis dans le contexte spatial pour enregistrer finement les mouvements selon les trois angles, et nous cherchons à nous affranchir en partie de ceux-ci. Tout le travail développé dans cette thèse cherche à estimer les variations d'attitude directement à partir des images en s'appuyant sur la géométrie de l'acquisition

2. Taille de la caméra panchromatique.
3. *Star tracker*, gyroscope et *sun tracker* sont des exemples de capteurs embarqués dans le satellite.

pushbroom et du plan focal.

I.4 Plan du manuscrit

L'exposé de notre travail se déroule comme suit en cinq chapitres :

- Le premier chapitre introductif a permis d'expliquer le fonctionnement d'une caméra *pushbroom*, de justifier son utilisation dans l'imagerie satellitaire par rapport au plus traditionnel capteur *pinhole*, et d'introduire des points essentiels qui motivent cette étude : estimer les variations d'attitude à partir des images.

- Le second chapitre développe entièrement la problématique. Les conséquences des variations d'attitude sont explicitées directement sur les images, un état de l'art est fait sur les thèmes de la calibration de caméras *pushbroom* et de l'estimation de l'attitude telle qu'elle a été envisagée jusqu'à maintenant, pour se focaliser sur les méthodes de recalage d'image qui sont à la base de nos travaux.

- Le troisième chapitre se consacre aux expressions formelles mathématiques résolvant notre problème d'estimation de l'attitude. Nous justifions dans un premier temps l'utilisation d'un cadre bayesien qui regroupe un terme de recalage d'image et un *a priori* sur l'évolution temporelle de l'attitude. Nous proposons ensuite différents modèles prenant en compte la nature physique du problème et nous ajoutons des données *star tracker* au système pour guider l'inférence.

- Le quatrième chapitre présente les algorithmes d'estimation de toutes les variables mathématiques introduites au chapitre précédent. Quatre jeux de données satellitaires sont traités et les méthodes proposées sont comparées selon différents critères de performance.

- Le cinquième et dernier chapitre conclut sur la faisabilité de l'approche proposée, et ouvre différentes perspectives liées au problème que nous aurons présenté.

L'essentiel des contributions de cette thèse sont présentées dans les chapitres 3 et 4.

CHAPITRE II

Problématique et état de l'art

II.1 Introduction

Le principe de fonctionnement du capteur *pushbroom* a été exposé au chapitre précédent. Dans cette partie nous nous focalisons sur la problématique et les solutions employées dans la littérature pour y faire face. Dans un premier temps, la configuration typique du plan focal d'un satellite est détaillée ainsi que ses conséquences sur l'acquisition des images. Nous cherchons ensuite à expliquer l'impact de phénomènes oscillatoires - que nous nommerons variations d'attitude - sur les images qu'enregistre le satellite, tout en justifiant leur présence. Dans un second temps, une revue des travaux de recherche liés aux caméras *pushbroom* et plus particulièrement à notre thématique d'estimation de l'attitude est faite. Nous verrons notamment que très peu d'études ont envisagé d'attaquer le problème de la façon que nous le souhaitons. Nous finirons par étendre cette revue par un état de l'art sur les méthodes de recalage d'image qui sont à la base de notre travail en justifiant nos choix.

II.2 Configuration du plan focal

Nous avons expliqué au chapitre précédent que le satellite embarque plusieurs caméras *pushbroom* pour restituer des images couleurs ; chaque caméra est équipée d'un filtre qui sélectionne une certaine bande spectrale de la lumière. Ces capteurs sont généralement assemblés sur un même plan focal et en parallèle pour optimiser l'acquisition [Gordon 05] ; la figure II.1 donne un aperçu schématique d'un plan focal usuel.

On comprend rapidement qu'une même scène est visualisée à différents instants par chaque caméra, τ_{ij} étant l'écart en temps entre les caméras i et j. Ainsi en se référant au plan focal de la figure II.1 pour composer une ligne de pixels en couleur, l'image 1-D rouge au moment n est associée à la verte au temps $n + \tau_{34}$ et à la bleue à l'instant $n + \tau_{24}$. Ce système suppose que le satellite est parfaitement stable lors de l'acquisition temporelle des images. Imaginons que l'attitude du satellite soit modifiée uniquement à l'instant n, les quatre lignes de visée sont déplacées en conséquence, modifiant les quatre scènes qui devaient être photographiées. La scène vue à l'instant n par la caméra rouge ne sera pas comme prévu celle que verra la caméra verte au temps $n + \tau_{34}$, et ainsi de suite. Nous touchons de près le problème, observons tout d'abord les conséquences géométriques des variations d'attitude sur les images.

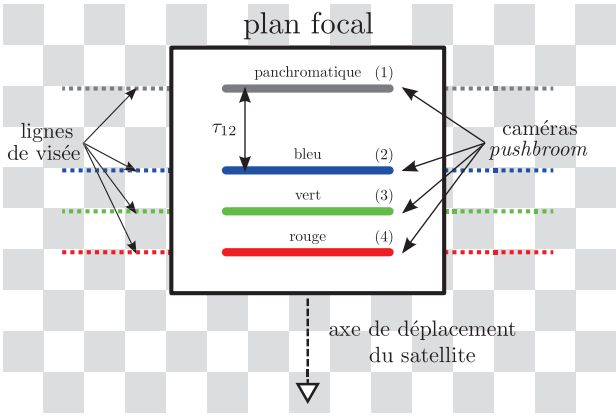


FIGURE II.1 – Schéma d'un plan focal standard de satellite d'observation. Les caméras sont disposées en parallèle et le satellite se déplace orthogonalement à leur axe ; les lignes de visée correspondent aux projections des caméras au sol selon le modèle *pushbroom* de la figure I.1(b) p. 14. Cette configuration implique que la caméra bleue (2) voit à l'instant n ce que la caméra panchromatique (1) verra à l'instant $n + \tau_{12}$.

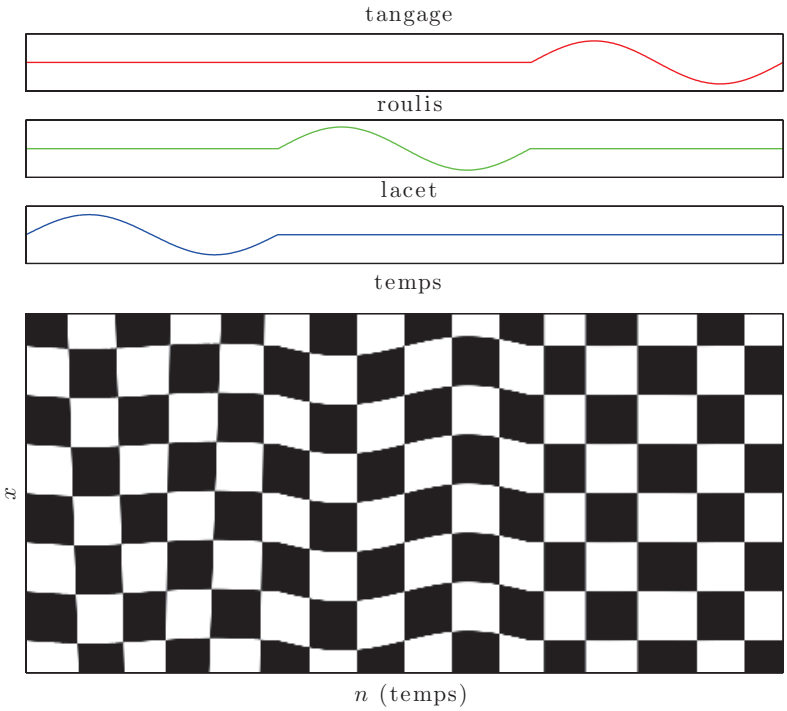


FIGURE II.2 – Simulation de l'acquisition d'un damier par une caméra *pushbroom* dont les angles de rotation (le lacet, le roulis et le tangage) varient au cours du temps. Les déformations du damier sont caractéristiques de chaque angle : le lacet incline plus ou moins les lignes verticales, le roulis fait osciller les lignes horizontales et le tangage augmente ou diminue la largeur des cases.

II.3 Problématique

Replaçons-nous tout d'abord dans le contexte d'une caméra *pushbroom* dont l'attitude varie au cours du temps. L'exemple synthétique de la figure II.2 décrit précisément les déformations qu'on peut attendre au sein des images.
Si la caméra avait été fixe lors de la prise de vue du damier, les cases de celui-ci auraient été régulières. L'oscillation pour le lacet, qui d'après la figure I.1(b) p.14 est une rotation dans le plan image autour de l'axe z, joue sur l'inclinaison des lignes verticales de l'image ; les déformations sont plus importantes sur les bords. La variation de roulis est une rotation autour de l'axe de déplacement de la caméra y et fait apparaitre ici une oscillation sur les lignes horizontales du damier. Enfin le tangage [1] est un mouvement autour de l'axe x de la caméra, il réduit ou augmente la largeur des cases du damier en modifiant l'échantillonnage temporel de l'image. L'amplitude des déformations en unité de pixel semble la même entre les trois angles, cependant pour ce test synthétique le lacet est 200 fois supérieur au roulis et au tangage. En effet cet angle de rotation s'effectue dans le plan image, contrairement au roulis et tangage qui sont des transformations amplifiées par la distance focale du satellite. Ainsi pour une même amplitude, l'impact de ces deux derniers sera largement majoritaire sur les images.
Lorsqu'on transpose ces variations d'attitude aux acquisitions multiples faites par les caméras *pushbroom* d'un satellite, le problème prend une tournure particulière. Toutes les images sont déformées à la façon de l'exemple de la figure II.2 et la correspondance temporelle entre les pixels des images selon le plan focal n'est plus assurée. La figure II.3 fait état de ce problème en montrant des images provenant de trois caméras appartenant à un même plan focal. Elles ont été affectées par des oscillations du satellite lors de la prise de vue, nous ne représentons ici que l'angle de roulis. Ces photographies ont été alignées conformément au décalage temporel que provoque l'espacement entre les caméras - l'image de modalité rouge a servi de référence aux deux autres - et on remarque que le roulis devient aussi décalé en temps.
L'image couleur construite à partir des trois autres présente des artefacts sur ses contours qui s'apparentent à des bavures colorées ; la variation de roulis a désynchronisé les images 1-D entre chaque bande spectrale. Les images 2-D ont un échantillonnage spatial irrégulier et les scènes observées par les caméras *pushbroom* ne se correspondent plus. La figure II.3 montre l'effet de décalage pour le roulis, mais il est aussi valable pour les deux autres angles que sont le lacet et le tangage. Au delà du simple fait que les images ne reflètent plus la vraie géométrie de la scène, la perte de connaissance de la géolocalisation des pixels et le problème d'alignement des trois bandes spectrales sont des conséquences importantes d'un satellite instable. Ce travail de thèse est focalisé sur ce dernier problème et recherche une façon d'estimer les variations d'attitude pour rectifier les images.

Nous avons commencé à parler des raisons pour lesquelles l'attitude d'un satellite peut varier au cours du chapitre introductif. La première génération des imageurs spatiaux était considérée comme très stable en raison de leur poids et de leurs caméras de faible résolution qui ne pouvaient pas faire transparaitre de déformations dans les images. A cela s'additionne les nombreux capteurs inertiels qui équipent chaque satellite

1. Le lecteur confus par ces rotations, ou qui ne souhaite pas se reporter à la figure du chapitre précédent peut se repérer de la façon suivante avec les yeux rivés sur la figure II.2 du manuscrit : l'axe du tangage est le même que celui du hochement de tête pour dire non, celui du roulis est le même que celui pour dire oui, et le lacet correspond à l'axe autour duquel on dodine de la tête lorsqu'on hésite.

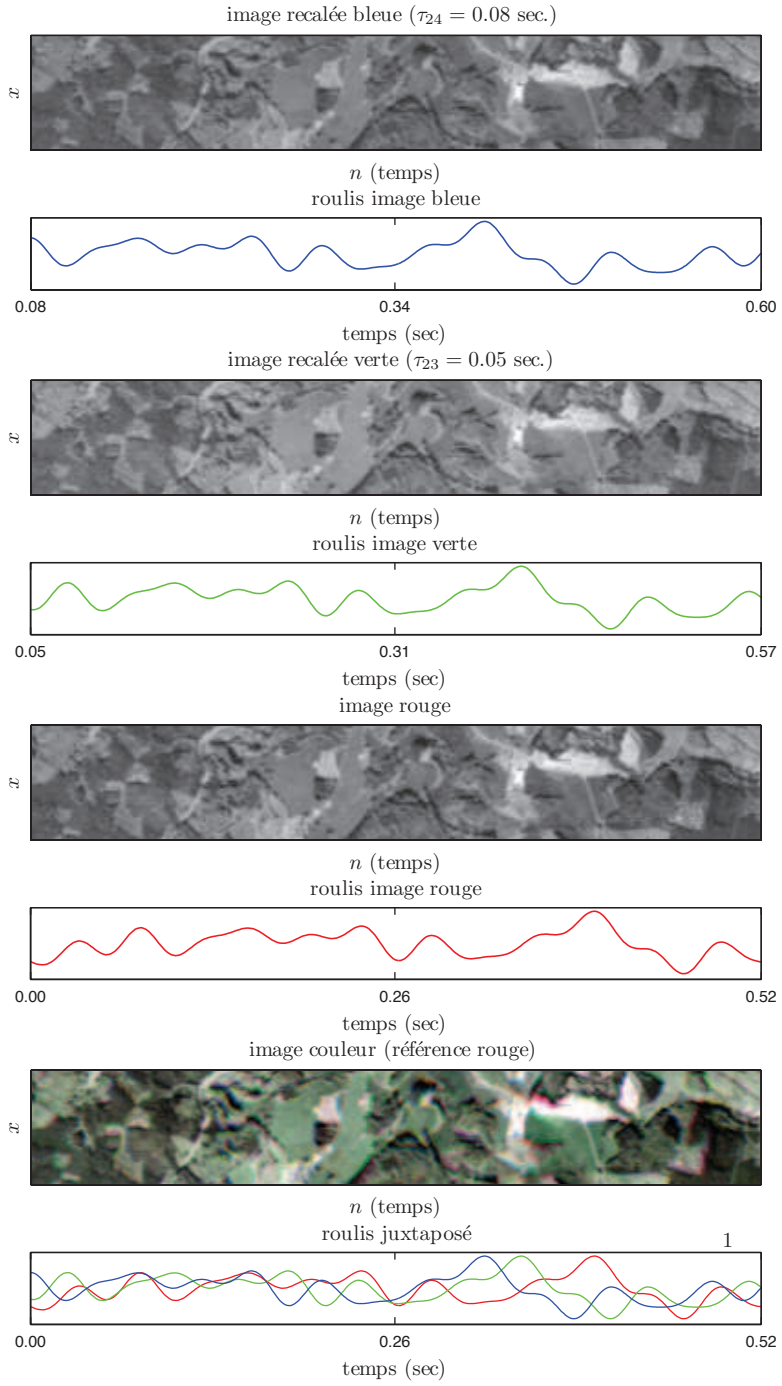


FIGURE II.3 – Extraits d'images de satellite de taille (60×400) pixels (la caméra linéaire contient donc 60 photosites) selon trois modalités : rouge, verte et bleue. Pour construire l'image couleur, chaque image est recalée en translation suivant n selon une référence - qui est ici la rouge - proportionnellement à l'écart entre les deux caméras dans le plan focal. Si l'acquisition est parasitée par une oscillation de roulis, les images 1-D sont désynchronisées dans les images et des artéfacts, que nous nommerons bavures, apparaissent sur les contours dans l'image couleur. Une bavure est particulièrement visible sur la zone blanche dans la partie droite de l'image.

pour enregistrer ses dérives et corriger ses trajectoires. La nouvelle génération des satellites est équipée de capteur *pushbroom* de haute résolution. De plus pour faciliter leur mise en orbite, ces satellites sont construits à partir de matériaux composites pour réduire leur poids. Ils naviguent plus facilement et les coûts énergétiques sont par la même occasion réduits. Dans la fabrication, les coûts peuvent également être diminués si certains capteurs sont retirés, mais ceci se fait au risque de perdre certaines informations de positionnement. En revanche, ces nouveaux satellites sont plus sensibles à de légères perturbations : les turbulences spatiales, les mouvements des roues à inertie, ainsi que les moteurs du satellite qui comme la plupart des systèmes motorisés vibrent selon des fréquences propres. Suivant la configuration du satellite, un simple écart de quelques microradians dans son attitude est suffisant pour observer les problèmes de la figure II.3. Enfin jusqu'à maintenant, des solutions qui s'appuient sur des capteurs inertiels ont été utilisées. Sont-elles réellement bien adaptées au problème et peut-on s'en affranchir ? Nous allons détailler ces points par la suite.

II.4 Etat de l'art

Nous détaillons l'état de l'art en trois parties, en commençant par parler des modèles géométriques de caméra *pushbroom* qui ont parfois été occultés par des modèles plus traditionnels, puis nous discuterons des méthodes d'estimation de l'attitude telles qu'elles ont été envisagées jusqu'à maintenant, pour terminer sur les techniques de recalage d'image qui sont aux fondements de notre approche.

II.4.1 Géométrie

Les modèles géométriques d'acquisition d'images par une caméra *pushbroom* existent depuis plusieurs années, [Gupta 95] est peut-être un des premiers travaux qui s'attaque au domaine satellitaire. Il est particulièrement révélateur de la difficulté du problème. Les auteurs décrivent un modèle orbital complet de l'imageur et tente d'estimer ses paramètres par des correspondances entre points 2-D et points 3-D, elles-mêmes fournies par des points de contrôle au sol - plus couramment nommés *Ground Control Points* (GCP) - connus à l'avance. Les auteurs montrent notamment que la précision des valeurs nominales de paramètres données par les capteurs inertiels du satellite est limitée et que la rectification selon leur approche est effective sur une expérimentation. L'attitude du satellite est supposée fixe lors de l'acquisition des images. Les mêmes auteurs proposent dans [Gupta 97] un travail qui est fondateur dans le domaine, le modèle d'acquisition est simplifié et peut s'étendre en dehors du contexte satellitaire. Ils s'appuient sur l'hypothèse supplémentaire que le capteur se déplace de façon rectiligne lors de l'acquisition des images, ce qui est valide à l'échelle d'une prise de vue de quelques secondes. L'accomplissement majeur de ces travaux est l'expression d'une matrice semblable à la matrice essentielle des caméras perspectives pour relier les points de deux images prises par des capteurs *pushbroom*. Un résultat concluant de reconstruction 3-D est présenté sur des données du satellite SPOT, une nouvelle fois l'attitude du satellite est supposée fixe. Une autre approche applicable aux caméras *pushbroom* est présentée dans [Drareni 08], elle n'utilise pas de correspondance de points 2-D 3-D comme précédemment, mais la transformation homographique produite par les images d'objets plans. Des tests sont conduits à partir d'un scanner à plat, les hypothèses sont les mêmes que pour les travaux décrits plus haut.

En dehors des deux premiers travaux que nous avons cités, peu d'études se sont focalisées sur les satellites car les données sont moins accessibles et l'environnement spatial est plus difficile à caractériser. En revanche l'imagerie aérienne est un domaine plus actif où des solutions impressionnantes de reconstruction 3-D à partir de caméras *pushbroom* existent depuis quelques années [Hirschmüller 05, Gordon 05]. Le design des capteurs est plus flexible et moins critique que dans le contexte spatial, au point d'être concevable par des laboratoires de recherche [Gruen 03]. Malgré des oscillations plus importantes de l'imageur - un hélicoptère ou un petit avion civil -, des capteurs GPS (Global Positioning System) et inertiels de haute qualité permettent d'avoir l'information de mouvement et de rectifier les images. Une fois les images corrigées, la disparité entre les images peut être mesurée le long des lignes épipolaires et un modèle 3-D estimé [Hirschmüller 05].
Dans le domaine de la vision par ordinateur, les caméras *pushbroom* n'ont pas suscité le même engouement que les modèles *pinhole* du fait d'un mode d'acquisition dépendant du temps plus difficile à appréhender. Pourtant dans un contexte parfaitement calibré où la pose et l'orientation de la caméra sont connues à chaque instant, les auteurs de [Rademacher 98] montrent par des simulations qu'elles sont mieux adaptées à l'échantillonnage de la fonction plénoptique d'une scène pour le rendu d'image [2]. Toute l'information est contenue dans une unique image dont chaque ligne possède un centre de projection différent, contrairement aux solutions classiques avec plusieurs images prises par une caméra perspective et où l'information est très redondante. La synthèse d'une nouvelle image 2-D s'en trouve facilitée. Les études géométriques citées au début peuvent difficilement être ajustées à notre contexte, inclure une incertitude sur les trois angles à l'acquisition de chaque ligne de l'image en tentant d'estimer la géométrie 3-D de la scène observée est un problème très mal contraint. Dans le travail de cette thèse, nous nous contentons uniquement d'estimer l'attitude du satellite. Nous ne développerons pas de modèle géométrique complet d'acquisition, nous nous appuyons sur la configuration du plan focal en faisant les hypothèses suivantes : le satellite se déplace de façon rectiligne à une hauteur et une vitesse constantes, la scène observée est suffisamment loin pour limiter les effets d'occultation [3], enfin le satellite sans variation d'attitude est supposé observer dans la direction de son nadir. Ce dernier terme couramment utilisé en astronomie désigne la verticale descendante située sous le capteur et orthogonale à la surface observée.

II.4.2 Estimation de l'attitude

Le problème de variation d'attitude est très critique en imagerie aérienne ; les avions sont sensibles aux vents et la proximité de la scène observée rend les déformations particulièrement visibles au sein des images. L'environnement reste cependant très bien contrôlé, tous les mouvements oscillatoires de l'appareil peuvent être enregistrés par des capteurs de positionnement très précis [Cramer 00, Gruen 03]. A cela peuvent s'ajouter des GCP qui par interaction avec l'utilisateur associent des pixels représentant une même scène entre les différentes images. On peut notamment citer [Poli 02] qui propose des solutions pour combiner données GPS et GCP pour le recalage des images.
Dans le cas des satellites, les capteurs de type *star tracker* et gyroscope ont longtemps été

2. à partir d'une collection d'images sur une scène, cette méthode permet de synthétiser une nouvelle image d'un nouveau point de vue sans que la caméra ait été présente physiquement à cet endroit.

3. Nous disposons en réalité d'une information numérique sur l'élévation du terrain, comme nous l'expliquons au chapitre 4 dans la partie sur les résultats.

utilisés pour asservir le positionnement de l'engin et rectifier les images. Ils ont cependant une précision limitée, rappelons que les contraintes spatiales sont très exigeantes vis-à-vis de l'électronique embarquée. Leur fréquence d'échantillonnage est généralement inférieure à 16Hz, ce qui est très bas en comparaison d'un capteur *pushbroom* qui acquiert des images à plus de 900Hz. Ceci implique qu'en théorie, tout mouvement de la caméra de plus de 8Hz ne peut pas être enregistré. Si quelques modèles récents de *star tracker* présentent des précisions inférieures au microradian avec un échantillonnage à 50Hz tel que le *High Accuracy Star Tracker* [BallAerospace 07], leur utilisation augmente considérablement le prix du satellite. Nous considérons que ces solutions ne sont pas bien adaptées aux variations d'attitude qui peuvent influencer les nouveaux satellites. Si les aléas de l'espace provoquent des lentes dérives basses fréquences, les moteurs auront tendance à vibrer sur des moyennes fréquences que ne peuvent pas restituer les capteurs de positionnement. Enfin, précisons que ces derniers étant très coûteux dans le design du satellite, s'en affranchir peut être une bonne solution économique.

A notre connaissance, une seule étude récente recense le problème dont nous parlons et envisage une solution similaire à la nôtre dans le sens où elle exploite les images en tant que données pour estimer les variations d'attitude [de Lussy 08] ; la société EADS Astrium est également détentrice d'un brevet exploitant cette même idée [Crombez 08]. En effet, malgré l'apparent désordre dans l'image couleur de la figure II.3, toutes les images multispectrales contiennent par leurs déformations une information cohérente sur les variations d'attitude. L'auteur propose d'extraire des fenêtres autour de points caractéristiques dans les images et de les comparer par corrélation. Chaque paire d'images est comparée, une servant de référence fixe - nous avons déjà insisté sur le fait que toutes les images étaient déformées -, la déformation est estimée et de cette façon une information relative sur l'attitude est obtenue. En intégrant toutes les mesures relatives sur l'attitude données par les correspondances entre images, les auteurs remontent à l'information d'attitude absolue par déconvolution.
Nous pensons que cette approche souffre de plusieurs manques. Premièrement elle ne tient pas compte de la géométrie de la caméra linéaire car elle extrait et compare des points alors que l'image 1-D est rigide et devrait être utilisée dans sa totalité. De plus, l'utilisation d'une fenêtre de corrélation se limite à l'estimation d'une translation, ce qui néglige la déformation à l'intérieur même de la fenêtre. Ensuite elle n'inclut pas dès le départ l'information de déformation absolue entre les images, mais doit procéder en plusieurs étapes avec essentiellement une estimation relative du mouvement lié à l'attitude pour chaque paire et une déconvolution par la suite. Enfin, cette méthode ne permet pas facilement d'intégrer des mesures exogènes de type *star tracker* lorsqu'elles sont disponibles, ou même une information *a priori* que l'on aurait sur les variations d'attitude. Nous souhaitons aborder le problème en intégrant tous ces éléments pour optimiser l'inférence, pour cela nous devons choisir une technique de recalage d'image plus adaptée.

II.4.3 Recalage d'image

Les méthodes de mise en correspondance d'images à partir d'un modèle paramétrique de mouvement sont nombreuses et ce thème de recherche était encore très actif il y a quelques années dans le milieu de la vision par ordinateur. Aujourd'hui la recherche se focalise beaucoup sur des problèmes de recalage non rigide, notamment dans le do-

maine médical. Nous ne chercherons pas à faire une revue de toutes celles-ci mais plutôt à identifier nos besoins pour déterminer une solution adaptée. Dans notre contexte, les déformations observées sur les images peuvent être de quelques pixels en amplitude. Le déplacement de chaque ligne de pixels est paramétrée par trois angles à estimer. Nous souhaitons obtenir une précision d'alignement sous-pixellique et une convergence suffisamment rapide en quelques minutes.

Les algorithmes de recalage d'images peuvent être séparés grossièrement en deux classes [Szeliski 06] : les méthodes qui s'appuient sur des éléments caractéristiques et les méthodes fondées sur les intensités des pixels. Les premières utilisent des éléments saillants comme des points ou des contours repérés sur les deux images par un détecteur - on peut citer [Harris 88, Lowe 04] qui sont très connus, mais loin d'être exhaustifs - , ces points sont mis en lien par une métrique et les correspondances trouvées permettent d'estimer les paramètres du modèle de déformation géométrique. Ces approches sont très adaptées à l'estimation de déplacement large entre les deux images, elles sont notamment beaucoup utilisées dans les logiciels de construction de mosaïque à partir de plusieurs images [d'Angelo 10].
Les secondes procèdent par une recherche de correspondance pixel par pixel et sont souvent connues sous le nom de méthodes directes [Irani 99, Szeliski 06]. Après avoir choisi une métrique pour comparer les intensités des pixels entre les deux images - par exemple la somme des carrés des différences d'intensité [Lucas 81], ou le produit de corrélation entre les images [Zitov 03] - , les paramètres du modèle géométrique qui optimisent ce critère sont recherchés. Cette recherche peut être exhaustive sur une petite zone comme pour les méthodes par corrélation, ou guidée par le gradient de l'image à la façon d'une descente de gradient comme le propose l'algorithme de Lucas-Kanade [Lucas 81, Baker 04]. Quelles que soient les combinaisons choisies entre métriques et algorithmes de recherche, ces approches supposent que la déformation est de l'ordre de quelques pixels pour pouvoir converger correctement et rapidement vers un optimum.
Si la première classe fondée sur des éléments caractéristiques est activement utilisée pour plusieurs applications comme la reconnaissance d'objet [Lowe 04], les méthodes directes restent majoritairement employées pour affiner les alignements d'image, par exemple dans la construction d'une mosaïque où les images ont été au préalable assemblées grossièrement par des éléments caractéristiques [Szeliski 06]. Elles font notamment une utilisation optimale de tous les pixels des images et assurent une précision sous-pixellique [Zitov 03, Baker 04]. Pour cette raison, nous optons pour les méthodes directes dans l'idée d'exploiter efficacement l'ensemble des données disponibles dans les images.
Les méthodes fonctionnant par corrélations locales sont très utilisées dans l'imagerie satellitaire [Inglada 04] ; elles sont robustes, précises et relativement rapides [Zitov 03]. En revanche elles ne peuvent estimer localement qu'un mouvement de translation ; elles font donc une hypothèse de mouvement local rigide ce dont nous souhaitons nous affranchir.
L'algorithme de Lucas-Kanade est une solution plus flexible qui permet de définir une déformation paramétrée indépendante pour chaque ligne de pixel. C'est une des premières méthodes de recalage [Lucas 81, Baker 04, Szeliski 06], elle s'appuie sur le gradient des images pour trouver les paramètres optimaux ce qui évite une recherche exhaustive. Elle utilise tous les pixels des images en minimisant le résidu au carré de leur intensité. Ces remarques en font une méthode de choix sur laquelle notre étude va se construire.

II.5 Conclusion

La problématique a été exposée dans cette partie et les objectifs ont été définis : estimer les variations d'attitude du satellite par mise en correspondance des images qu'il a acquis. Le choix de l'algorithme de Lucas-Kanade a été justifié par le contexte, il est aux fondements de la méthode que nous allons exposer au chapitre suivant.
Nous souhaitons insister sur le fait que le problème investi ici est très récent et pour le moment peu connu. Ceci explique le peu de travaux qui ont été faits sur le sujet. Les enjeux sont pourtant réels et bien précis, les solutions existent et doivent être réadaptées au contexte. Le chapitre suivant expose la méthode choisie dans son formalisme mathématique.

CHAPITRE III

Formalisme bayesien du problème

III.1 Introduction

Les chapitres précédents ont permis d'introduire la problématique et de décrire les solutions qui ont été mises en œuvre jusqu'à maintenant. Le problème d'estimation des variations d'attitude à partir des images acquises est mal posé, car toutes les images sont déformées. En effet, l'essentiel des techniques de recalage d'image s'appuient sur une image de référence, à laquelle une autre image est comparée pour estimer un déplacement. Dans notre contexte, aucune référence n'est disponible pour conduire le recalage d'image ; l'utilisation d'informations *a priori* et de mesures exogènes est nécessaire pour restreindre le champ des déplacements qui sont solutions. La formulation du problème dans un cadre bayesien permet de regrouper ces différents éléments dans une expression unique, où l'estimateur du *maximum a posteriori* peut être utilisé pour déduire les variations d'attitude.
Ce chapitre introduit les notations mathématiques et formalise dans un premier temps le problème de façon probabiliste dans une expression globale, dont chaque terme est développé. Plusieurs modèles sont proposés pour décrire l'évolution temporelle de l'attitude, jusqu'à l'intégration en fin de chapitre de données *star tracker* en plus des images pour accompagner l'estimation de l'attitude. Les différentes solutions sont analysées et comparées sur un même jeu de données.

III.2 Cadre probabiliste

L'utilisation d'un cadre probabiliste dans l'estimation d'un processus stochastique peut se justifier de différentes façons. La plupart des signaux mesurés par des capteurs sont bruités, et ce pour des signaux de différente nature : photographie d'un paysage, enregistrement sonore, mesure de secousse sismique, etc. Ils peuvent pourtant partager une même représentation mathématique de ce phénomène :

$$f(n) = \theta(n) + \epsilon_f(n), \text{ pour } n \in E, \tag{III.1}$$

où $f(n)$ désigne le signal bruité en sortie de capteur, $\theta(n)$ le signal original tel qu'on souhaiterait l'avoir enregistré, $\epsilon_f(n)$ le bruit de mesure spécifique au capteur, et n une variable définie dans un espace E pouvant correspondre au temps, à l'espace, ou autre

suivant la nature du signal.
Cette représentation simple, et très utilisée, de la chaîne d'acquisition d'un signal est un modèle mathématique qui ne peut prendre en compte tous les phénomènes qui ont engendré la mesure $f(n)$. Cependant, l'équation (III.1) est simple à manipuler, et peut être considérée comme un bon modèle si en connaissance du signal $\theta(n)$, elle peut reproduire des mesures bruitées proches de celles que l'on observe réellement.
L'équation (III.1) contient deux processus aléatoires entre le signal $\theta(n)$ et le bruit $\epsilon(n)$. Le caractère aléatoire se manifeste par une difficulté à prédire dans le temps le futur état du processus, en conséquence il est souvent décrit à l'aide de distributions de probabilité. Sans information sur la nature du signal, le bruit d'un capteur peut en revanche être plus ou moins bien défini. Il est souvent représenté statistiquement par une gaussienne de moyenne nulle et de variance σ_f^2 :

$$\epsilon_f(n) \sim \mathcal{N}(0, \sigma_f^2). \tag{III.2}$$

L'objectif sera alors de trouver un estimateur tel que la différence entre les mesures $f(n)$ et le signal estimé $\hat{\theta}(n)$, que l'on nomme résidu, suive la loi de probabilité définie dans l'expression (III.2). Ce principe est exposé dans le schéma III.1, où sont représentés successivement un jeu de 100 mesures, un signal estimé, les résidus associés, et la distribution des résidus comparée à la distribution théorique.

Il n'est pas rare cependant que les données ne suffisent pas à elles seules à restaurer le signal d'origine car la représentation (III.1) n'est pas parfaite et le problème inverse peut être mal posé. L'utilisation d'informations supplémentaires, ou d'un modèle paramétrique sur le signal $\theta(n)$, permet de restreindre les solutions du problème d'estimation. Supposons que le signal puisse s'écrire sous la forme d'une fonction linéaire selon des paramètres a_i pour $i \in \{0, 1, 2, 3\}$:

$$\theta(n) = a_0 + a_1 n + a_2 n^2 + a_3 n^3 + \epsilon_\theta(n), \tag{III.3}$$

où $\epsilon_\theta(n)$ est un bruit associé au modèle, qu'il existe une contrainte liant trois échantillons consécutifs du signal selon des coefficients b_j pour $j \in \{0, 1, 2\}$:

$$b_0\theta(n-1) + b_1\theta(n) + b_2\theta(n+1) \sim \mathcal{N}(0, \sigma_b^2), \tag{III.4}$$

et qu'un capteur supplémentaire fournisse une seconde mesure bruitée de θ :

$$g(m) = \theta(m) + \epsilon_g(m), \text{ pour } m \in F, \tag{III.5}$$

où F désigne un espace inclu E, on peut par exemple supposer que les mesures $g(m)$ sont sous-échantilonnées par rapport à $f(n)$. Le problème consiste maintenant à déterminer un $\hat{\theta}(n)$ en connaissance des relations précédentes : (III.1), (III.3), (III.4) et (III.5) ; il peut se résumer schématiquement à la figure III.2.

Nous verrons par la suite que l'utilisation d'un cadre probabiliste associé à la loi de Bayes permet de répondre de façon élégante à ce problème d'inférence. Le rapport [Punska 99] en fait une démonstration très précise et détaillée. En pratique, malgré l'apparente simplicité de ces expressions, l'estimation dans ce cas-ci reste complexe. Il est difficile d'affirmer par avance l'utilité d'un modèle comme celui défini en (III.3) lorsqu'il

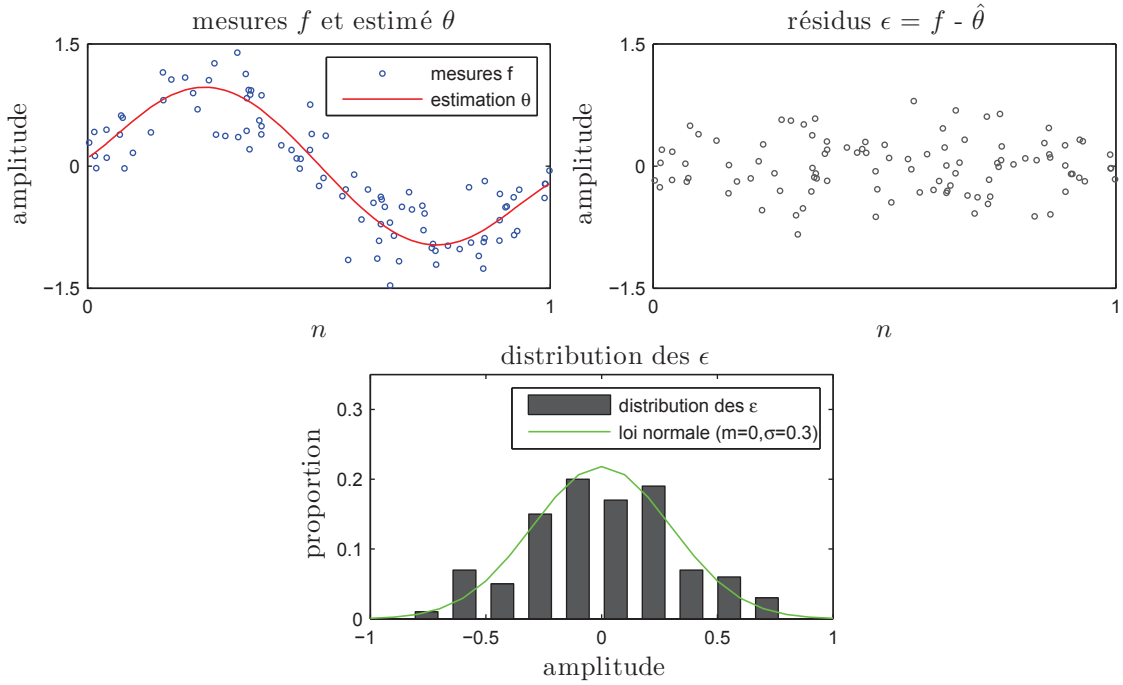


FIGURE III.1 – Dans cet exemple, 100 mesures sont générées à partir d'une période de sinusoide bruitée par une loi normale centrée et d'écart-type 0.3 ; ces mesures sont représentées par les points bleus en haut à gauche. Un modèle polynômial de degré 5 est estimé par moindres carrés sur les mesures, il est tracé en rouge et correspond à $\hat{\theta}$. Sur la figure du haut à droite, les points noirs montrent les valeurs prisent par les résidus $f(n) - \hat{\theta}(n)$. En bas, la distribution des résidus est comparée à leur loi théorique tracée en vert.

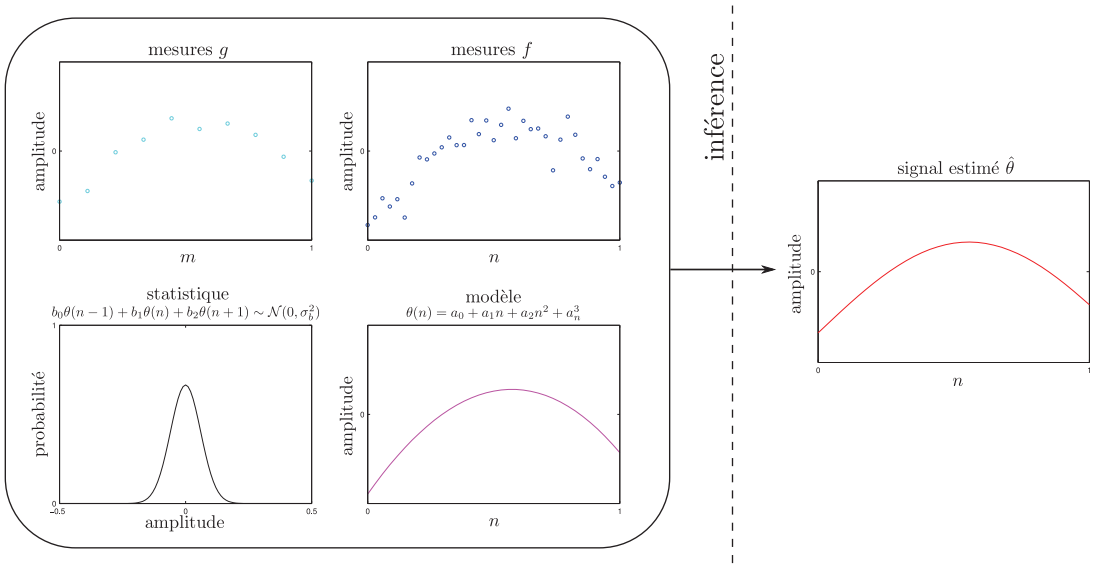


FIGURE III.2 – Représentation de la problématique d'estimation où mesures, contrainte statistique et modèle se mélangent pour déterminer un signal en sortie.

existe déjà une relation du type (III.4). De même dans ce formalisme, de nouveaux paramètres tels que les a_i sont à déterminer alors que seul $\theta(n)$ nous intéresse. Chaque information à apporter au problème d'estimation doit être judicieusement choisie, au risque de dégrader les performances de l'estimateur. Ces remarques justifient l'utilisation d'un cadre probabiliste autour duquel la suite de ce chapitre va s'articuler.

III.3 Formulation du maximum *a posteriori*

Soit $\mathcal{I} = \{I_1, I_2, I_3, I_4\}$ l'ensemble des quatre images acquises par le satellite : panchromatique, rouge, verte et bleue respectivement. Nous appelons θ le vecteur contenant les variations d'attitude qui ont parasité les images, et que nous cherchons à déterminer. La règle de Bayes décompose la loi *a posteriori* de l'attitude en connaissance des images de la façon suivante :

$$p(\theta|\mathcal{I}) = \frac{p(\mathcal{I}|\theta)p(\theta)}{p(\mathcal{I})}. \tag{III.6}$$

Le premier terme du numérateur $p(\mathcal{I}|\theta)$ constitue la vraisemblance de nos images, c'est la densité de probabilité de nos images en supposant connaitre l'attitude. Le second terme $p(\theta)$ est l'*a priori*, il exprime nos croyances sur les variations d'attitude avant même d'avoir observé les images. Le terme du dénominateur $p(\mathcal{I})$ est connu sous le nom de vraisemblance marginale - ou également évidence -, il est indépendant de l'inconnue θ et tient son nom de la propriété de marginalisation suivante :

$$p(\mathcal{I}) = \int p(\mathcal{I}|\theta)p(\theta)d\theta. \tag{III.7}$$

On peut supposer que l'expression (III.7) ne joue aucun rôle dans l'inférence, étant indépendante de la variable que nous recherchons. De plus, elle nécessite le calcul d'une intégrale qui dans de nombreux cas ne dispose pas d'une expression analytique ; il faut alors recourir à des méthodes d'approximation parfois complexes pour pouvoir l'évaluer. Cependant, la vraisemblance marginale possède des propriétés intéressantes de sélection de paramètres que nous observerons au chapitre suivant.
Nous omettrons donc dans un premier temps l'expression de $p(\mathcal{I})$. La loi *a posteriori* intègre toute notre connaissance sur l'attitude θ, et peut se réécrire de la façon suivante :

$$p(\theta|\mathcal{I}) \propto p(\mathcal{I}|\theta)p(\theta). \tag{III.8}$$

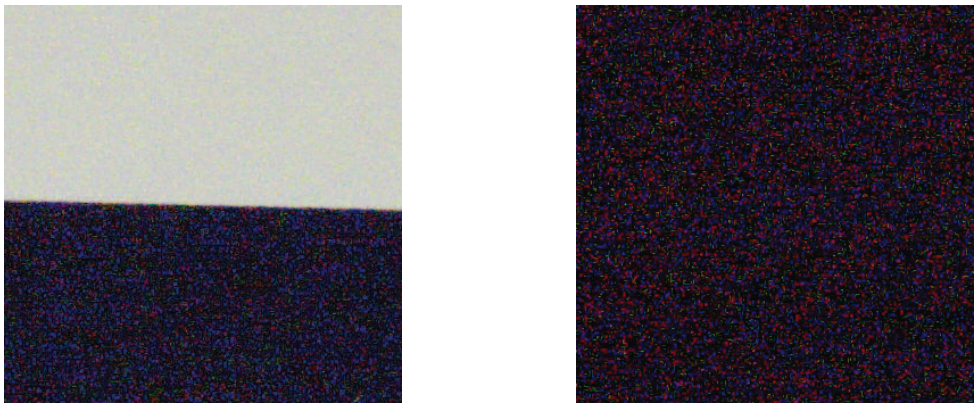
L'estimateur du maximum *a posteriori* (MAP) correspond au mode de la distribution de probabilité *a posteriori* :

$$\hat{\theta}_{MAP} = \underset{\theta}{\operatorname{argmax}}\ p(\theta|\mathcal{I}). \tag{III.9}$$

Nous nous attacherons à développer cette équation au cours des paragraphes suivants. Nous nous intéressons tout d'abord au premier terme de l'expression (III.8). Nous allons proposer au paragraphe suivant deux formulations pour la vraisemblance des images, que nous distinguerons dans la suite de notre étude par les qualificatifs de premier cas et second cas de vraisemblance.

III.4 Vraisemblance des images

Le calcul de la vraisemblance d'une mesure présuppose une connaissance sur les statistiques de bruit qui lui sont associées. Dans notre contexte de photographie, le capteur quantifie l'irradiance d'une scène en comptant le nombre de photons qu'il reçoit sur chaque photosite, en convertissant ceux-ci en électrons puis en bits. Les sources de bruit sont multiples dans les images obtenues : elles peuvent dépendre de la chaleur environnant le capteur, de la quantité de lumière émise par la scène, du temps d'exposition, de la taille des photosites et de la distance entre eux - le lecteur peut se référer à [McHugh 05, Wong 03] pour un exposé simple du problème. A l'échelle électronique seule, un bruit permanent existe indépendamment de la présence de source lumineuse, il est principalement lié à l'agitation thermique des composants et est nommé *dark current* ; un exemple typique est visible en figure III.3(b). Au niveau de la scène observée, le bruit photonique se caractérise par une dépendance à la quantité de lumière perçue ; sur une image, les régions les plus sombres contiennent en proportion de signal perçu plus de bruit que les plus claires. La figure III.3(a) illustre bien ce phénomène. Enfin en considérant à plus haut niveau la chaîne d'acquisition complète d'une image, d'autres éléments comme l'optique peuvent être une source de nuisance ; le chapitre 2 de [Jalobeanu 01] ainsi que [Rougé 97] détaillent clairement le procédé de formation des images dans le contexte spatial avec tous les phénomènes complexes qui lui sont associés.



(a) (b)

FIGURE III.3 – (a) Photo d'une interface entre deux surfaces noire et blanche, on distingue à l'oeil un bruit plus prononcé sur la zone sombre par rapport à la zone claire. C'est la caractéristique du bruit photonique, auquel s'est ajouté d'autres sources de bruit dans l'image (photo RAW, DMC-LX3 f3.2 - 1/500 - 3200 ISO). (b) Photo prise dans le noir complet révélant le bruit électronique du capteur, aussi connu sous le nom de *dark current* (photo RAW, DMC-LX3 f8.0 - 4 - 3200 ISO).

Récemment, plusieurs modèles de bruit liés à l'enregistrement d'une image par un appareil photo numérique ont été proposés. Ils peuvent servir à débruiter des images [Liu 08, Luisier 11], ou à estimer des paramètres de sensibilité pour l'acquisition et la fusion de plusieurs images HDR (High Dynamic Range) en une seule [Hasinoff 10]. En l'absence de données constructeur précises sur le capteur, ces méthodes doivent estimer les paramètres de leur modèle à partir d'images, ce qui ajoute un échelon supplémentaire d'incertitude avant le traitement final de l'image et peut limiter leur utilisation pratique.

Pour cette raison, les algorithmes les plus connus de suppression de bruit sur les images supposent que ce bruit suit globalement une loi normale, centrée et identiquement distribuée selon les pixels [Starck 02, Chang 00, Buades 05]. Ce choix de loi normale se justifie d'abord par le théorème central limite, stipulant que la somme de plusieurs petites sources de bruit s'approche d'une distribution gaussienne, mais également par sa commodité mathématique. Il est intéressant de noter que les méthodes citées se distinguent essentiellement dans la façon dont elles spatialisent le traitement, en débruitant plus fortement les régions homogènes, tout en maintenant les contours et les textures dans l'image. Elles font donc une hypothèse intrinsèque de dépendance du niveau de bruit aux données observées.

III.4.1 Statistique de bruit des images

Les statistiques de bruit utilisées en recalage d'image font également une hypothèse de bruit blanc dans les mesures, essentiellement pour les raisons précitées. Ce choix peut être explicite [Robinson 04], ou implicite par la fonction de coût choisie et la façon de la minimiser [Lucas 81, Irani 99, Baker 04, Szeliski 06]. Rappelons que le recalage désigne pour nous la mise en correspondance de deux images d'une même scène prise en deux endroits différents, et donc avec des matrices de projection différentes pour chaque caméra ; les pixels entre les deux images sont liés par un modèle géométrique de déformation paramétré. Dans ce contexte, d'autres sources de bruit peuvent s'ajouter : occultation entre les deux images, variations de luminosité, apparition d'un nouvel élément dans une des images, etc. Ces types de bruit étant complexes à caractériser, les méthodes de recalage font généralement des hypothèses simples sur le bruit, tout en diminuant, ou supprimant, dans le processus d'estimation les contributions de mesures aberrantes qui ne suivent pas la statistique choisie. Plusieurs algorithmes existent, l'article [Black 96] est une très bonne référence dans ce domaine. L'estimateur devient robuste, compensant la simplicité du modèle de bruit choisi. Le lecteur peut se reporter à [Rousseeuw 87] pour plus de détails sur l'estimation robuste.
Le problème satellitaire auquel nous nous attachons a deux caractéristiques importantes : les photos d'une même scène sont prises à des intervalles de temps très proches, limitant l'apparition ou la disparation d'éléments entre deux images, et la scène observée est très loin du satellite, diminuant les phénomènes d'occultations. L'hypothèse de bruit blanc dans la différence entre deux images est donc une approximation *a priori* correcte, bien que parfois limitée comme nous le verrons au chapitre suivant lors du traitement des données.
Pour développer un modèle mathématique associé à notre problème, nous ferons plusieurs hypothèses au cours de ce chapitre dont les suivantes :

- nous supposerons que les images sont de même modalité ; ceci revient à dire que les caméras du plan focal de la figure II.1 (p.20) ont une même sensibilité au spectre de lumière,
- la scène observée est plane ; la distance entre le satellite et la terre est très grande comme nous l'avons mentionné un peu plus haut,
- le satellite se déplace de façon rectiligne uniforme ; nous en avons parlé au chapitre précédent.

Nous discuterons de chacune de ces conditions par la suite dans ce manuscrit. Selon la géométrie du plan focal, et en l'absence d'oscillation du satellite, nous espérons observer

la statistique de bruit suivante sur les images :

$$I_i(x,n) - I_j(x,n+\tau_{ij}) \sim \mathcal{N}\big(0,\sigma_I^2\big), \tag{III.10}$$

où σ_I^2 est la variance de ce bruit ; I_i et I_j désignent deux images prises par les caméras *pushbroom* i et j. Le couple $(x,n) \in \mathbb{R}^2$ est le domaine dans lequel la fonction bidimensionnelle I prend ses valeurs. Ce domaine se compose d'un nombre fini de positions de pixels dans les images ; ces dernières sont définies sur une grille discrète. Les valeurs prisent par la fonction image en dehors de cette grille peuvent être obtenues par interpolation. La variable x indique la position de chaque photosite le long du capteur *pushbroom*, la variable n correspond aux instants échantillonnés dans le temps où la caméra a enregistré une ligne de pixels. Ainsi chaque image est de taille $(L \times N)$ où L correspond au nombre de photosites du capteur, et N au nombre d'échantillons en temps permettant l'acquisition de l'image complète. Le décalage temporel τ_{ij} entre l'image i et l'image j est la conséquence directe de la géométrie du plan focal, telle qu'elle a été observée au chapitre précédent. Une même scène sur terre est observée à différents instants de temps par chaque caméra, ces écarts temporels sont donc proportionnels aux distances entre capteurs. Afin de positionner chaque image sur un même référentiel, et dans un souci de simplification, nous adopterons l'écriture suivante par la suite :

$$I^{\tau}(x,n) = I(x,n+\tau). \tag{III.11}$$

Selon nos hypothèses, l'opérateur $.^{\tau}$ peut être compris comme une fonction de translation de l'image d'une valeur τ selon l'axe n.

III.4.2 Fonction de déformation

Lors de l'acquisition, des petites variations de l'attitude du satellite produisent des déformations géométriques au sein de l'image. Nous émettons les hypothèses suivantes :

- il existe une fonction de déformation $\mathcal{W}$ qui, à une position de pixel donnée dans une image, associe sa nouvelle position dans l'image suivant une variation d'attitude donnée. Nous la définissons de la façon suivante :

$$\begin{aligned} \mathcal{W} : \mathbb{R}^2 \times \mathbb{R}^3 &\rightarrow \mathbb{R}^2 \\ \mathcal{W}(x,n;\theta(n)) &= (x',n'). \end{aligned} \tag{III.12}$$

  Le vecteur réel $\theta(n) = [\theta^l(n)\ \theta^r(n)\ \theta^t(n)]^T$ désigne l'attitude du satellite à l'instant n ; il est de taille (3×1) et concatène respectivement le lacet, le roulis et le tangage. Cette fonction peut être vue comme la transformation d'une image acquise de façon stable par le satellite, en une image pour laquelle l'attitude du satellite a varié. Nous supposerons de plus que cette fonction est inversible et possède la propriété de composition suivante :

$$\mathcal{W}(\mathcal{W}(x,n;\theta_1);\theta_2) = \mathcal{W}(x,n;\theta_1+\theta_2) \tag{III.13}$$

- Le décalage temporel, tel que nous l'avons défini à l'équation (III.10), est constant au cours du temps.

Concernant le premier point, il est très difficile de construire la fonction $\mathcal{W}$ d'un point de vue analytique. S'il est possible de projeter la ligne de visée de la caméra *pushbroom*

au sol assez simplement, le problème inverse - qui consiste à déterminer le lieu d'intersection d'un point du sol sur l'image acquise au cours du temps en tenant compte des variations d'attitude - est nettement plus complexe [Gupta 95, Gupta 97]. La cinématique du satellite est de même difficile à appréhender étant donné les effets de couplage qui existent entre les trois angles de l'attitude. Le lecteur peut se rendre au paragraphe IV.5 (p.106) pour plus de détails. Nous nous contenterons ici de remarquer qu'étant donné l'amplitude des variations d'attitude - qui sont de l'ordre du microradian -, la fonction $\mathcal{W}$ peut être vu au premier ordre comme une petite rotation géométrique de chaque ligne. Au sein même des images, la transformation s'apparente à des petites translations pour le roulis et le tangage, raison pour laquelle nous admettrons la relation (III.13).
Le second point est une conséquence des remarques que nous venons faire, bien qu'en réalité sa valeur dépend de la géométrie de la scène, ainsi que de l'effet de tangage et de lacet qui introduisent un échantillonnage temporel irrégulier dans les images.
Dans ce contexte oscillatoire, et en prenant en compte ces hypothèse et la notation proposée en (III.11), l'équation (III.10) devient :

$$I_i\big(\mathcal{W}(x,n;\theta(n))\big) - I_j^{\tau_{ij}}\big(\mathcal{W}(x,n;\theta(n+\tau_{ij}))\big) \sim \mathcal{N}\big(0,\sigma_I^2\big). \tag{III.14}$$

Il est important de noter une nouvelle fois que l'attitude est décalée en temps entre les deux images d'un intervalle de τ_{ij}. Ceci est la conséquence de la géométrie du plan focal, et nous permet d'estimer les mouvements de rotations du satellite. L'équation (III.14) est à l'origine de notre premier cas de vraisemblance que nous développerons ensuite.

III.4.3 Expression d'une convolution

La relation (III.14) nous indique que chaque ligne des deux images I_i et I_j est déformée par un mouvement de rotation. D'après la relation (III.13), une conséquence intéressante est la réécriture du terme de gauche de l'équation (III.14), en appliquant au préalable une rotation d'angle $-\theta(n)$ aux deux images :

$$I_i\big(\mathcal{W}(x,n;\theta(n))\big) - I_j^{\tau_{ij}}\big(\mathcal{W}(x,n;\theta(n+\tau_{ij}))\big) = I_i(x,n) - I_j^{\tau_{ij}}\big(\mathcal{W}(x,n;\theta(n+\tau_{ij})-\theta(n))\big). \tag{III.15}$$

I_i devient une image fixe de référence qui est comparée à l'image I_j, elle-même transformée par une composition de rotations $\theta(n+\tau_{ij})-\theta(n)$. Enfin, nous pouvons remarquer le développement suivant :

$$\theta(n+\tau_{ij}) - \theta(n) = k_{ij} * \theta(n), \tag{III.16}$$

qui correspond sous forme matricielle à :

$$\begin{bmatrix} \theta^l(n+\tau_{ij}) - \theta^l(n) \\ \theta^r(n+\tau_{ij}) - \theta^r(n) \\ \theta^t(n+\tau_{ij}) - \theta^t(n) \end{bmatrix} = \begin{bmatrix} \theta^l(n) * k_{ij}(n) \\ \theta^r(n) * k_{ij}(n) \\ \theta^t(n) * k_{ij}(n) \end{bmatrix}. \tag{III.17}$$

L'opérateur $*$ réalise la convolution de chaque angle de l'attitude θ par le noyau k_{ij} ; ce dernier permet de différentier θ suivant l'écart en temps τ_{ij}. Supposons que cinq échantillons en temps séparent l'acquisition de l'image i et de l'image j, nous pouvons déduire la forme discrétisée du noyau ci-dessous :

$$\tau_{ij} = 5, \text{ et } k_{ij} = \begin{bmatrix} 1 & \underbrace{0\ 0\ 0\ 0}_{\tau_{ij}-1 \text{ zeros}} & -1 \end{bmatrix}^T. \tag{III.18}$$

Notons que la variable τ_{ij}, fixée par la géométrie du plan focal, ne prend pas nécessairement que des valeurs entières. Auquel cas le noyau adoptera une forme légèrement différente, selon l'exemple suivant :

$$\tau_{ij} = 5.7, \text{ et } k_{ij} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & -0.3 & -0.7 \end{bmatrix}^T. \tag{III.19}$$

Le noyau effectue alors une combinaison linéaire des deux échantillons les plus proches pour prendre en compte une valeur non entière de l'écart, à la façon d'une interpolation linéaire. Par conséquent l'expression (III.14), qui est notre second cas de vraisemblance image, peut également s'écrire :

$$I_i(x,n) - I_j^{\tau_{ij}}\big(\mathcal{W}(x,n;k_{ij} * \theta(n))\big) \sim \mathcal{N}\big(0,\sigma_I^2\big), \tag{III.20}$$

l'estimation de l'attitude θ à partir d'une paire d'images I_i et I_j peut être vue comme une opération de déconvolution par le noyau k_{ij}. Les deux expressions (III.14) et (III.20) pourtant équivalentes ont des formulations matricielles assez différentes, et n'aboutissent pas au même résultat après optimisation selon θ, comme nous l'observerons dans les paragraphes suivants.

III.4.4 Ecriture de la vraisemblance

Nous pouvons déduire l'expression de la vraisemblance pour une paire d'images dans le premier cas de l'équation (III.14) :

$$\begin{aligned} p(I_i, I_j|\theta) &= \prod_{n=1}^{N}\prod_{x=1}^{L} p\Big(I_i\big(\mathcal{W}(x,n;\theta(n))\big) - I_j^{\tau_{ij}}\big(\mathcal{W}(x,n;\theta(n+\tau_{ij}))\big)\Big|\theta\Big) \\ &= \prod_{n=1}^{N}\prod_{x=1}^{L} \frac{1}{\sqrt{2\pi}\sigma_I} \exp\left(-\frac{\big(I_i\big(\mathcal{W}(x,n;\theta(n))\big) - I_j^{\tau_{ij}}\big(\mathcal{W}(x,n;\theta(n+\tau_{ij}))\big)\big)^2}{2\sigma_I^2}\right). \end{aligned} \tag{III.21}$$

Le développement est sensiblement le même pour la seconde vraisemblance, associée à l'équation (III.20) :

$$\begin{aligned} p(I_i, I_j|\theta) &= \prod_{n=1}^{N}\prod_{x=1}^{L} p\Big(I_i(x,n) - I_j^{\tau_{ij}}\big(\mathcal{W}(x,n;k_{ij} * \theta(n))\big)\Big|\theta\Big) \\ &= \prod_{n=1}^{N}\prod_{x=1}^{L} \frac{1}{\sqrt{2\pi}\sigma_I} \exp\left(-\frac{\big(I_i(x,n) - I_j^{\tau_{ij}}\big(\mathcal{W}(x,n;k_{ij} * \theta(n))\big)\big)^2}{2\sigma_I^2}\right). \end{aligned} \tag{III.22}$$

La vraisemblance globale dans les deux cas se déduit aisément, elle associe toutes les vraisemblances de chaque paire d'images de façon symétrique :

$$p(\mathcal{I}|\theta) = \prod_{i,j;i\neq j} p(I_i, I_j|\theta). \tag{III.23}$$

Dans leur forme actuelle, les logarithmes des deux vraisemblances sont non linéaires selon θ. En effet, les intensités des images prises en chaque pixel ne sont généralement pas linéaires selon la position du pixel. Ceci ne facilite pas la recherche d'un $\hat{\theta}$, mais ces deux

expressions pourraient être optimisées en recherchant de façon exhaustive la variation d'attitude la plus probable. Cette approche facile à mettre en place sera également coûteuse en temps, et elle ne prendra pas avantage du principe de petites déformations que nous avons vu au chapitre précédent. Notons également que ces équations ne peuvent pas être développées facilement sous forme matricielle : le vecteur d'attitude dépend de l'indice n dans les images et il est de plus décalé, ou convolué, en temps selon les images.

III.4.5 Linéarisation du terme image

Parmi les nombreuses solutions de recalage d'image existantes - le lecteur peut se référer à l'article [Szeliski 06] pour une très bonne vue d'ensemble de celles-ci - la méthode de Lucas-Kanade [Lucas 81, Baker 04] s'est particulièrement distinguée au cours des années 80 pour sa simplicité et sa robustesse. Elle a donné lieu à une famille d'algorithmes de recalage fondée sur le gradient de l'image [Robinson 04]. L'idée consiste à linéariser au premier ordre le terme image déformée selon sa série de Taylor autour de 0 :

$$I\big(\mathcal{W}(x,n;\theta(n))\big) \approx I(x',n') + \nabla I(x',n')\ \frac{\partial \mathcal{W}}{\partial \theta}(x',n';\theta(n))\ \theta(n). \tag{III.24}$$

Dans cette expression,

$$\nabla I(x,n) = \left[\frac{\partial I}{\partial x}(x,n)\ \ \frac{\partial I}{\partial n}(x,n)\right] \tag{III.25}$$

est un vecteur (1×2) correspondant au gradient de l'image déformée évalué en (x,n), et

$$\frac{\partial \mathcal{W}}{\partial \theta}(x,n;\theta(n)) = \begin{bmatrix} \frac{\partial \mathcal{W}_x}{\partial \theta^l}(x,n;\theta(n)) & \frac{\partial \mathcal{W}_x}{\partial \theta^r}(x,n;\theta(n)) & \frac{\partial \mathcal{W}_x}{\partial \theta^t}(x,n;\theta(n)) \\ \frac{\partial \mathcal{W}_n}{\partial \theta^l}(x,n;\theta(n)) & \frac{\partial \mathcal{W}_n}{\partial \theta^r}(x,n;\theta(n)) & \frac{\partial \mathcal{W}_n}{\partial \theta^t}(x,n;\theta(n)) \end{bmatrix} \tag{III.26}$$

est une matrice de taille (2×3) contenant les dérivées partielles des déformations selon chaque angle et évaluées en $(x,n;\theta)$; $\mathcal{W}_x$ et $\mathcal{W}_n$ sont respectivement les fonctions de déformations selon les axes x et n. Nous expliquerons la procédure de calcul des matrices de dérivées partielles de la transformation au chapitre suivant (IV.5 p.106). Ce type de développement est très clairement présenté formellement et schématiquement dans [Baker 04].

Il est important de comprendre que cette approximation n'est valable que pour des valeurs petites du vecteur $\theta(n)$; en d'autres termes, on suppose que les déformations géométriques sur l'image I ne sont que de l'ordre du pixel. Ceci peut constituer une faiblesse de la méthode comme [Szeliski 06] le précise, mais notre contexte de petites déformations répond bien à cette hypothèse. En pratique pour des déformations de quelques pixels, cette approximation est réalisée itérativement : un incrément $\Delta\theta(n)$ est estimé à chaque itération, et par sommation forme l'estimation $\hat{\theta}(n)$. Le schéma algorithmique complet sera précisé au chapitre suivant. Pour simplifier la lecture du cadre théorique, nous supposerons par la suite que $\theta(n)$ satisfait l'approximation faite en equation (III.24).

Le terme de droite de l'égalité (III.24) est maintenant linéaire selon $\theta(n)$, nous pouvons le reformuler en terme de matrice et vecteur, de la façon suivante :

$$\mathbf{i} + J\theta. \tag{III.27}$$

La variable θ a déjà été introduite auparavant au paragraphe III.3 comme le vecteur d'attitude au cours du temps, voici sa forme :

$$\theta = [\theta^l(n_1)\ \ldots\ \theta^l(n_N)\ \ \theta^r(n_1)\ \ldots\ \theta^r(n_N)\ \ \theta^t(n_1)\ \ldots\ \theta^t(n_N)]^T. \tag{III.28}$$

Ce vecteur est de taille $(3N \times 1)$ et concatène les trois angles sur l'ensemble du temps. La variable $\mathbf{i}$ désigne le vecteur image de taille $(NL \times 1)$ tel que :

$$\mathbf{i} = [I(x_1, n_1) \ \dots \ I(x_L, n_1) \ \ I(x_1, n_2) \ \dots \ I(x_L, n_2) \ \dots \ I(x_1, n_N) \ \dots \ I(x_L, n_N)]^T, \quad \text{(III.29)}$$

où les indices sous les variables x et n indexent les positions discrètes des pixels de l'image. En pratique, chaque colonne de l'image est mise bout à bout dans le vecteur.
La variable J, que nous nommerons jacobienne, est une matrice creuse de dimension $(NL \times 3N)$. Elle se construit comme suit pour les indices entiers $p \in \{1, 2, \dots, N\}$ et $q \in \{1, 2, \dots, L\}$:

$$\begin{aligned}
J(p + (q-1)L, p) &= \frac{\partial I}{\partial x}(x_q, n_p) \, \frac{\partial \mathcal{W}_x}{\partial \theta^l}(x_q, n_p; \theta(n_p)) + \frac{\partial I}{\partial n}(x_q, n_p) \, \frac{\partial \mathcal{W}_n}{\partial \theta^l}(x_q, n_p; \theta(n_p)) \\
J(p + (q-1)L, N + p) &= \frac{\partial I}{\partial x}(x_q, n_p) \, \frac{\partial \mathcal{W}_x}{\partial \theta^r}(x_q, n_p; \theta(n_p)) + \frac{\partial I}{\partial n}(x_q, n_p) \, \frac{\partial \mathcal{W}_n}{\partial \theta^r}(x_q, n_p; \theta(n_p)) \\
J(p + (q-1)L, 2N + p) &= \frac{\partial I}{\partial x}(x_q, n_p) \, \frac{\partial \mathcal{W}_x}{\partial \theta^t}(x_q, n_p; \theta(n_p)) + \frac{\partial I}{\partial n}(x_q, n_p) \, \frac{\partial \mathcal{W}_n}{\partial \theta^t}(x_q, n_p; \theta(n_p))
\end{aligned} \quad \text{(III.30)}$$

Plus concrètement, cette matrice a $3NL$ éléments non nuls, elle réalise les produits du gradient de l'image (III.25) avec les dérivées partielles de la déformation (III.26), tout en s'adaptant à la forme du vecteur d'attitude. On peut également la comprendre comme étant composée de trois sous-matrices : ses N premières colonnes correspondent aux équations du lacet, les N suivantes au roulis, et les N dernières au tangage. La figure III.4 est une représentation schématique de la forme de la matrice J. D'après cette dernière, on comprend que chaque échantillon d'attitude $\theta(n)$ doit vérifier L équations linéaires. C'est la conséquence directe du fait que chaque ligne de L pixels est déformée par une attitude inconnue.

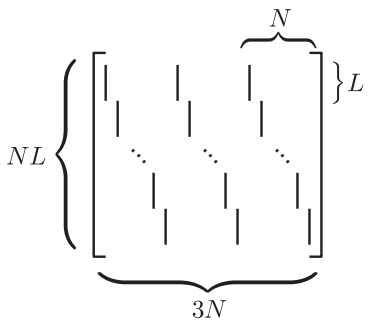


FIGURE III.4 – Illustration de la répartition des valeurs dans la matrice creuse J ; les lignes noires correspondent aux zones où la matrice n'est pas nulle.

III.4.6 Maximum de Vraisemblance (MV)

Premier cas

Nous pouvons maintenant exprimer sous forme matricielle la vraisemblance pour le premier cas de l'équation (III.21). En suivant les définitions du paragraphe précédent :

$$p(I_i, I_j|\theta) = \frac{1}{(2\pi\sigma_I^2)^{NL/2}} \exp\left(-\frac{1}{2\sigma_I^2}\left|\mathbf{i}_i + J_i\theta - \mathbf{i}_j - J_j D_{ij}\theta\right|^2\right), \quad \text{(III.31)}$$

où $|.|$ est la norme euclidienne du vecteur. Dans cette expression et pour toute la suite, nous ferons un léger abus de notation pour ne pas multiplier les indices : la matrice J_j correspond à la jacobienne de l'image $I_j^{\tau_{ij}}$ qui a été rectifiée par rapport à l'image i, alors que la matrice J_i correspond à la jacobienne de l'image I_i. La matrice D_{ij} est creuse et de taille $(3N \times 3N)$, elle décale en temps le vecteur d'attitude d'un pas de τ_{ij} ; sa forme est développée en annexe A. En posant $\mathbf{e}_{ij} = \mathbf{i}_i - \mathbf{i}_j$ le vecteur image erreur, et selon la formule de vraisemblance globale (III.23), on déduit que :

$$\begin{aligned} p(\mathcal{I}|\theta) &= \prod_{i,j;i\neq j} \frac{1}{(2\pi\sigma_I^2)^{NL/2}} \exp\left(-\frac{1}{2\sigma_I^2}\left|\mathbf{e}_{ij} + (J_i - J_j D_{ij})\theta\right|^2\right) \\ &= \mathcal{N}\Big(-\sum_{i,j;i\neq j}(J_i - J_j D_{ij})\theta, \sigma_I^2 I_d\Big), \end{aligned} \tag{III.32}$$

I_d étant la matrice identité de taille $(3N \times 3N)$. A ce stade, nous disposons d'une première relation liant les données images aux variations d'attitude. Il est possible d'estimer un $\hat{\theta}$ par la méthode du maximum de vraisemblance (MV) :

$$\hat{\theta}_{MV} = \underset{\theta}{\operatorname{argmax}}\ p(\mathcal{I}|\theta). \tag{III.33}$$

Cet estimateur ne s'appuie que sur les images pour proposer une solution, il recherche le θ qui maximise la probabilité (III.32) sans tenir compte d'information extérieure sur celui-ci. En pratique, il est plus simple de maximiser le logarithme de cette probabilité, aussi appelé log-vraisemblance. Son expression est la suivante :

$$\log(p(\mathcal{I}|\theta)) = \sum_{i,j;i\neq j} -\frac{1}{2\sigma_I^2}\left|\mathbf{e}_{ij} + (J_i - J_j D_{ij})\theta\right|^2 - M\frac{NL}{2}\log(2\pi\sigma_I^2). \tag{III.34}$$

La variable M désigne le nombre de paires d'images valides[1]. En dérivant l'équation quadratique précédente selon θ, et en recherchant la valeur pour laquelle elle s'annule, nous obtenons une solution analytique :

$$\hat{\theta}_{MV} = -\Big(\sum_{i,j;i\neq j}(J_i - J_j D_{ij})^T(J_i - J_j D_{ij})\Big)^{-1}\sum_{i,j;i\neq j}(J_i - J_j D_{ij})^T\mathbf{e}_{ij}. \tag{III.35}$$

Ce résultat est équivalent à la solution d'un problème aux moindres carrés linéaire qui suppose implicitement un bruit gaussien sur les mesures. Il est intéressant de noter que la variance du bruit n'est pas utile dans le calcul de $\hat{\theta}_{MV}$, ce qui rend la méthode simple et attractive. Le développement de la solution est disponible en annexe A.

Second cas

La vraisemblance dans le second cas, équation (III.22), peut également se développer sous forme matricielle en suivant le même raisonnement que précédemment :

$$p(I_i, I_j|\theta) = \frac{1}{(2\pi\sigma_I^2)^{NL/2}}\exp\left(-\frac{1}{2\sigma_I^2}\left|\mathbf{i}_i - \mathbf{i}_j - J_j K_{ij}\theta\right|^2\right), \tag{III.36}$$

1. en pratique $M = 12$; nous avons 4 images et le nombre total de paires que nous pouvons former pour les comparer de façon symétrique est de 4×3.

où K_{ij} est une matrice creuse de taille $(3N \times 3N)$, construite à partir du noyau k_{ij}. Celle-ci effectue la convolution de chaque angle de l'attitude du vecteur θ ; la structure de cette matrice est également décrite en annexe A. D'après la relation (III.23), la vraisemblance globale s'écrit :

$$\begin{aligned} p(\mathcal{J}|\theta) &= \prod_{i,j;i\neq j} \frac{1}{(2\pi\sigma_I^2)^{NL/2}} \exp\left(-\frac{1}{2\sigma_I^2}\left|\mathbf{e}_{ij} - J_j K_{ij}\theta\right|^2\right) \\ &= \mathcal{N}\Big(\sum_{i,j;i\neq j} J_j K_{ij}\theta, \sigma_I^2 I_d\Big) \end{aligned} \tag{III.37}$$

Une nouvelle fois, l'estimateur du maximum de vraisemblance permet d'obtenir une solution au problème d'estimation des variations d'attitude. On exprime en premier lieu la log-vraisemblance :

$$\log(p(\mathcal{J}|\theta)) = \sum_{i,j;i\neq j} -\frac{1}{2\sigma_I^2}\left|\mathbf{e}_{ij} - J_j K_{ij}\theta\right|^2 - M\frac{NL}{2}\log(2\pi\sigma_I^2), \tag{III.38}$$

pour déterminer la valeur de θ qui atteint le mode de la distribution de probabilité :

$$\hat{\theta}_{MV} = \Big(\sum_{i,j;i\neq j} K_{ij}^T J_j^T J_j K_{ij}\Big)^{-1} \sum_{i,j;i\neq j} K_{ij}^T J_j^T \mathbf{e}_{ij}. \tag{III.39}$$

Les équations (III.34) et (III.38) sont similaires aux équations obtenues par la méthode de Lucas-Kanade [Baker 04], tout en étant adaptées au contexte satellitaire et à l'acquisition d'images par des caméras *pushbroom*. On peut retrouver ce type de recalage multi-image en super résolution, où plusieurs images de basse résolution doivent être alignées avant d'être assemblées en une unique image de haute résolution [Farsiu 05].
D'un point de vue calculatoire, la première solution (III.35) implique deux fois l'approximation linéaire vue en (III.24) et nécessite le calcul de deux jacobiennes pour les deux images, alors que la solution (III.39) ne réalise qu'une approximation pour un seul calcul de jacobienne sur l'image déformée. Lorsque la minimisation est itérative et que la plupart des calculs sont répétés, la deuxième solution sera plus rapide en temps d'exécution.

III.4.7 Exemple d'estimation par MV

Nous présentons un exemple de roulis estimé par maximum de vraisemblance sur un jeu de données selon les deux formules (III.32) et (III.37). Les données traitées sont synthétiques et fournies par la société EADS Astrium ; l'acquisition *pushbroom* est simulée sur des images aériennes réelles selon un procédé très réaliste et qui nous est totalement inconnu. Nous développons ce sujet sur les données au chapitre suivant dans la partie résultats, l'avantage principal est que la variation d'attitude choisie pour la simulation est disponible et constitue une vérité terrain. Ce jeu est composé de 4 images de taille (900×1000) pixels et servira de base de comparaison pour tous les estimateurs que nous construirons par la suite. On peut relever certaines différences intéressantes dans les solutions obtenues.
Les deux figures III.5 et III.6 permettent de comparer la vraie variation de roulis en vert, à son estimation en rouge. Le résidu en gris foncé correspond à la différence entre les deux courbes. Dans le premier cas, figure III.5, le résidu présente un petit bruit haute fréquence, et une oscillation basse fréquence importante. A l'inverse, celui de la figure III.6

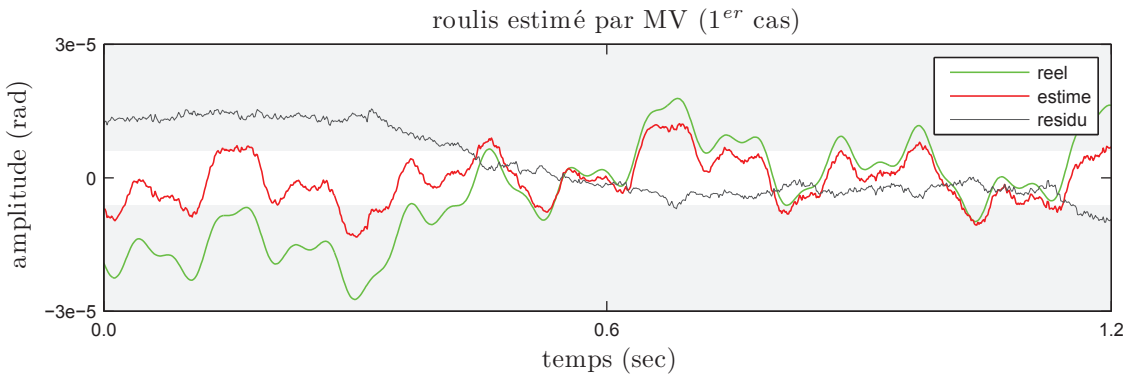


FIGURE III.5 – Exemple d'estimation du roulis par MV pour la vraisemblance (III.32) (courbe rouge), roulis réel (courbe verte), et résidu obtenu (courbe grise). La bande blanche délimite la taille d'un pixel sur l'image, la déformation sur les images est par endroit supérieure au pixel. Le résidu présente des oscillations importantes en basses fréquences.

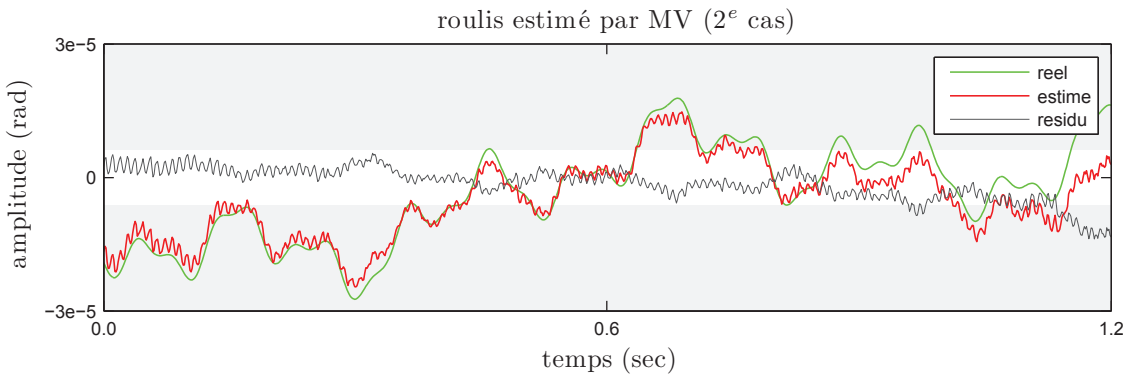


FIGURE III.6 – Exemple d'estimation du roulis par MV pour la vraisemblance (III.37) (courbe rouge), tangage réel (courbe verte), et résidu obtenu (courbe grise). La bande blanche délimite la taille d'un pixel sur l'image. Le résidu présente des oscillations importantes en hautes fréquences.

oscille plus en haute fréquence, mais moins en basse fréquence.
On peut interpréter ces observations de la façon suivante : la première estimation implique une linéarisation sur les deux images dans l'équation (III.32), cette double approximation pénalise d'autant plus la précision lorsque les déformations sont grandes dans les images, ce qui correspond visuellement à la zone grise dans les figures. La seconde estimation ne nécessite qu'une linéarisation puisqu'une des images sert de référence dans l'expression (III.37), mais elle présente le problème comme une déconvolution. Inverser les effets d'une convolution est un concept généralement mal posé car il consiste à retrouver une information perdue dans le signal par l'opération de convolution. Il en résulte assez souvent une amplification du bruit dans la zone où le signal a été coupé, en l'occurrence dans les hautes fréquences pour notre exemple en figure III.6.
Dans les deux cas, on remarque que l'approximation linéaire (III.24) est satisfaisante lorsque l'amplitude du roulis réel est contenue dans la bande blanche. Dans cette zone, la déformation géométrique sur les images ne dépasse pas un pixel. Au delà dans la partie grise, le roulis estimé peine à suivre la vraie variation. C'est la raison pour laquelle il est nécessaire d'itérer plusieurs fois la résolution des équations (III.35) ou (III.39) avec l'estimé de l'itération précédente pour affiner la solution. Il est également important que cet estimé ne soit pas parasité, sinon ce bruit peut dégrader la stabilité numérique du système et être amplifié à chaque itération.
Pour améliorer le processus d'estimation et tenir compte de certaines caractéristiques physiques des variations d'attitude, nous choisissons d'utiliser l'estimateur du MAP de l'équation (III.9) en intégrant le terme $p(\theta)$ à la vraisemblance des images. Ce dernier peut prendre différentes formes suivant les propriétés que l'on souhaite pour l'estimé $\hat{\theta}$, c'est ce que nous allons détailler dans les paragraphes suivants.

III.5 *A priori* sur l'attitude

L'utilisation d'un terme *a priori* peut être justifiée par les remarques du paragraphe précédent. Dans la plupart des problèmes, on dispose assez souvent d'une idée vague sur ce que l'on cherche à estimer, que ce soit en terme d'amplitude, de localisation ou d'évolution temporelle. Nous verrons plus loin que dans le pire des cas, l'absence d'information peut également être modélisée dans le cadre du MAP ; cela revient généralement à utiliser le MV.
Nous avons présenté au cours des chapitres précédents les causes des variations d'attitude, elles sont le résultat de phénomènes physiques relativement bien compris, et par conséquent elles observent certaines règles. Le satellite est globalement stable dans sa trajectoire [Gordon 05] ; les aléas de l'espace peuvent le faire dévier de sa trajectoire, mais cette dérive sera automatiquement détectée à un certain seuil par plusieurs capteurs, et les moteurs du satellite le recaleront immédiatement. Les oscillations du satellite autour de ses trois angles sont donc contraintes en amplitude. A l'intérieur du satellite, les moteurs vibrent et sont une autre cause des variations d'attitude. Mécaniquement, tous les moteurs vibrent selon certaines fréquences, ce phénomène est même très utilisé dans l'industrie pour diagnostiquer des défauts de fonctionnement dans les machines tournantes [Doebling 96, Carden 04]. Un moteur en « bonne santé » est supposé avoir des vibrations très stables au cours du temps, toute modification du contenu fréquentiel des vibrations peut être interprétée comme un problème. Cette propriété porte le nom de stationnarité, et peut se comprendre assez simplement dans une représentation temps-fréquence d'un

signal en figure III.7. Le signal représenté en vert contient trois fréquences à 15 Hz, 50 Hz et 100 Hz, elles sont clairement identifiables sur le spectrogramme par trois lignes droites blanches. Dans l'intervalle de temps $[2.6, 3.2]$ secondes, une irrégularité apparait dans le signal - un échelon additionné à un chirp linéaire - et se traduit dans le spectrogramme par une modification de son contenu fréquentiel. Selon notre définition, nous considérons que le signal est stationnaire dans les intervalles $[0.0, 2.6]$ secondes, et $[3.2, 3.7]$ secondes.
L'étude des phénomènes stationnaires est un domaine de recherche important, et la lit-

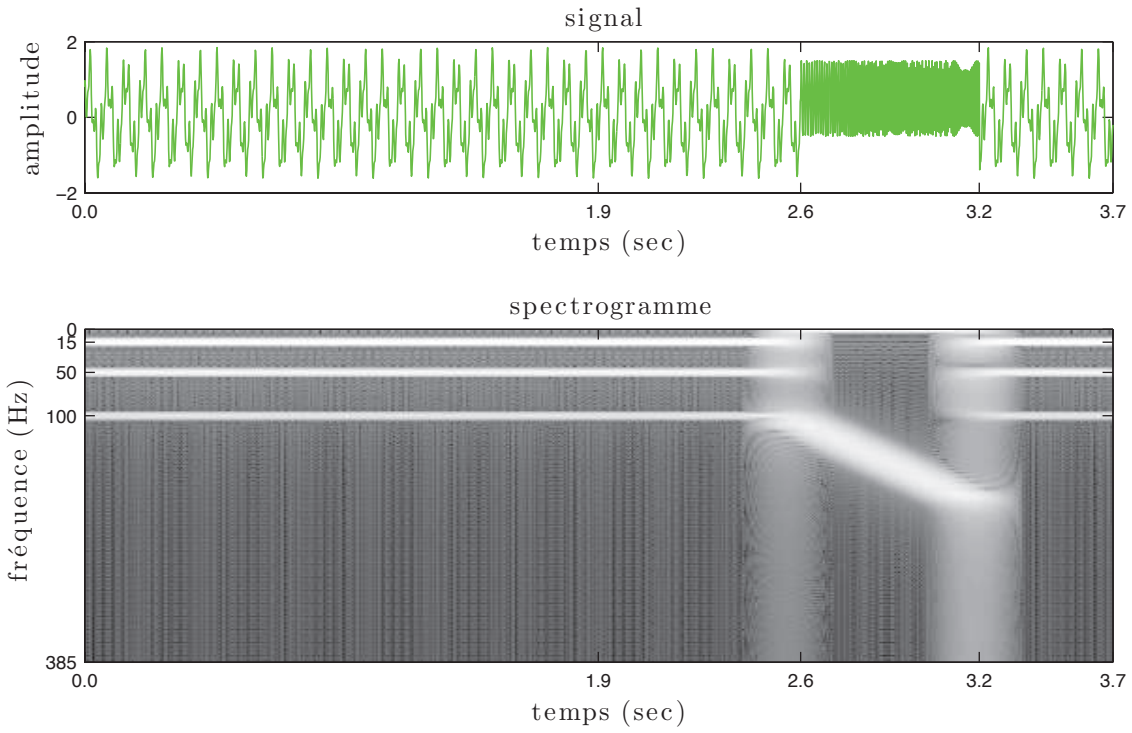


FIGURE III.7 – Le signal tracé en vert est composé de 3 fréquences : 15 Hz, 50 Hz et 100 Hz. Son spectrogramme, ou représentation temps-fréquence, au dessous permet de comprendre le concept de stationnarité : le contenu fréquentiel est identique sur les intervalles $[0.0, 2.6]$ secondes, et $[3.2, 3.7]$ secondes, c'est une zone stationnaire. Entre $[2.6, 3.2]$ secondes, un échelon additionné à un chirp linéaire apparait dans le signal, provoquant une non-stationnarité du signal.

térature est abondante sur ce sujet. Le lecteur peut se reporter à [Flandrin 99, Xiao 07] pour une analyse très intéressante de ce concept. Nous nous contenterons de remarquer que les variations d'attitude possèdent cette caractéristique de stationnarité au cours du temps à quelques exceptions près [2]. Enfin pour justifier ce propos d'un point de vue mécanique, les mouvements du satellite peuvent difficilement présenter des irrégularités, ceci à cause de l'inertie même de l'engin spatial.
Les *a priori* que nous allons développer par la suite cherchent à rendre compte de ces propriétés. La représentation spectrale d'un signal, que nous utiliserons souvent par la suite, est un outil intéressant pour mieux percevoir les effets de l'*a priori*. Nous continuerons

2. Le satellite est piloté en permanence et certaines manœuvres influencent le contenu spectral de ses moteurs. De même sa structure est soumise à des contraintes fortes dans l'espace ce qui peut modifier légèrement ses modes de résonance propres.

de tester les solutions proposées sur le même jeu de données que celui présenté au paragraphe précédent. La figure III.8 rappelle la forme du signal que nous cherchons, ainsi que son spectre de puissance.
Plusieurs basses fréquences sont présentes dans le signal, il est de même très régulier au

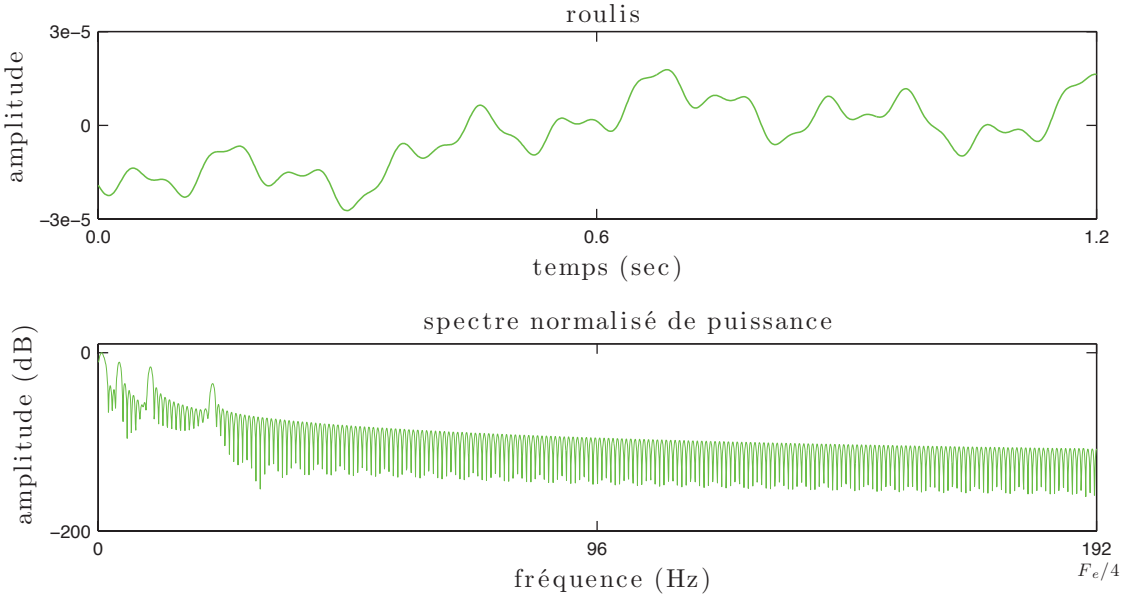


FIGURE III.8 – Représentation graphique du roulis, que nous chercherons à estimer dans les prochains paragraphes avec l'estimateur du MAP, et de son spectre associé. Le signal est régulier et composé essentiellement de basses fréquences. Trois harmoniques apparaissent clairement sur le spectre autour de 4 Hz, 10 Hz, et 22 Hz. La première est plus large, et semble être composée de plusieurs très basses fréquences.

cours du temps, ce qui est en lien avec les observations que nous avons faites précédemment.

III.5.1 Notations

Par la suite, afin de simplifier les expressions mathématiques intégrant les *a priori*, nous développerons les estimateurs du MAP pour un seul des angles de l'attitude à chaque fois. En effet, la plupart des *a priori* ont des paramètres, aussi connus sous le nom de hyperparamètres, qui prendront des valeurs différentes suivant chaque angle. Il est plus simple de prendre en compte un seul angle pour la lecture des équations, bien que le problème puisse s'exprimer selon le vecteur d'attitude complet. Nous indicerons par α les variables qui dépendent d'un des trois angles : le lacet l, le roulis r ou le tangage t. Ainsi, θ^α est le sous-vecteur de θ et de taille $(N \times 1)$, tel que :

$$\theta^\alpha = [\theta^\alpha(n_1) \ \dots \ \theta^\alpha(n_N)]^T. \tag{III.40}$$

De même, J^α devient la sous-matrice de J de taille $(NL \times N)$, et associée à l'angle d'attitude choisi ; il a été remarqué au paragraphe III.4.5 que J est composée de trois sous-matrices. Enfin, D^α et K^α sont respectivement les sous-matrices de D pour le décalage, et de K pour la convolution, elles sont de taille $(N \times N)$. On peut remarquer qu'elles sont identiques

dans leur traitement pour chaque angle, c'est à dire $K^l = K^r = K^t$ et $D^l = D^r = D^t$. Nous insistons sur le fait que les équations qui suivent sont valables pour n'importe quel angle choisi et que le problème peut toujours être exprimé sous une forme globale. Seulement nous verrons au chapitre suivant que le stockage en mémoire des matrices peut être critique et qu'il est préférable de minimiser indépendamment chaque angle.

III.5.2 Amplitude

Un premier *a priori* relativement simple consiste à supposer que les variations d'attitude sont des variables aléatoires indépendantes et identiquement distribuées (i.i.d.) suivant une loi normale centrée de variance $\sigma^2_{\theta^\alpha}$, ainsi :

$$\theta^\alpha \sim \mathcal{N}(\mathbf{0}, \sigma^2_{\theta^\alpha} I_d). \tag{III.41}$$

Afin de mieux comprendre ce qu'implique la relation (III.41), il est possible de simuler des valeurs de θ^α qui suivent cette statistique. La figure III.9 montre trois exemples de fonctions aléatoires de ce type où $\sigma_{\theta^\alpha} = 10^{-5}$. Le tracé du haut représente l'évolution temporelle, et celui du bas le spectre de puissance associé.
Intuitivement, on se rend compte que cet *a priori* contraint l'amplitude de l'attitude à évo-

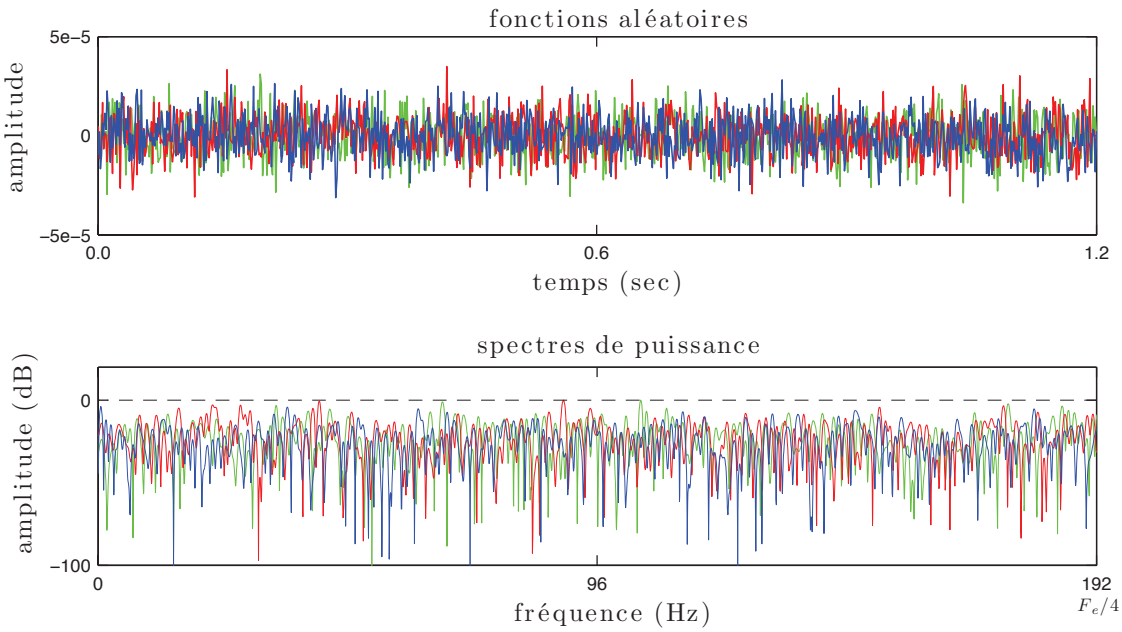


FIGURE III.9 – Le graphique du haut montre trois exemples de fonctions aléatoires échantillonnées selon l'équation (III.41), où σ_{θ^α} est fixé à 10^{-5}. Le graphique du bas représente les trois spectres de puissance associés aux fonctions ; on peut noter qu'ils sont très plats, toutes les fréquences étant présentes de façon égale (c'est la caractéristique de la réalisation d'un bruit blanc). Le spectre théorique de l'*a priori* est tracé en tirets noirs.

luer autour de zéro, et dans un intervalle fixé par la variance ; les trois fonctions évoluent presque exclusivement dans l'intervalle $[-3\sigma_{\theta^\alpha}, 3\sigma_{\theta^\alpha}]$. En revanche, aucune fréquence n'est favorisée par l'*a priori* au regard des trois spectres très plats de ces fonctions.

Formulation du MAP

On peut déduire assez simplement l'estimateur du MAP à partir des relations (III.9) (p.32), (III.41), (III.32) (p.40) et (III.37) (p.41). Associé à la première vraisemblance des images (III.32), nous obtenons le développement suivant par complétion du carré :

$$\begin{aligned} p(\theta^\alpha|\mathcal{J}) &\propto p(\mathcal{J}|\theta^\alpha)p(\theta^\alpha) \\ &\propto \prod_{i,j;i\neq j} \exp\Big(-\frac{1}{2\sigma_I^2}(\mathbf{e}_{ij} + (J_i^\alpha - J_j^\alpha D_{ij}^\alpha)\theta^\alpha)^T(\mathbf{e}_{ij} + (J_i^\alpha - J_j^\alpha D_{ij}^\alpha)\theta^\alpha)\Big) \exp\Big(-\frac{1}{2\sigma_{\theta^\alpha}^2}\theta^{\alpha T}\theta^\alpha\Big) \\ &\propto \exp\Big(-\frac{1}{2}(\theta^\alpha - \bar{\theta}^\alpha)^T\big(\sigma_I^{-2}\sum_{i,j;i\neq j}(J_i^\alpha - J_j^\alpha D_{ij}^\alpha)^T(J_i^\alpha - J_j^\alpha D_{ij}^\alpha) + \sigma_{\theta^\alpha}^{-2}I_d\big)(\theta^\alpha - \bar{\theta}^\alpha)\Big), \end{aligned} \tag{III.42}$$

avec :

$$\bar{\theta}^\alpha = -\sigma_I^{-2}(\sigma_I^{-2}\sum_{i,j;i\neq j}(J_i^\alpha - J_j^\alpha D_{ij}^\alpha)^T(J_i^\alpha - J_j^\alpha D_{ij}^\alpha) + \sigma_{\theta^\alpha}^{-2}I_d)^{-1}\sum_{i,j;i\neq j}(J_i^\alpha - J_j^\alpha D_{ij}^\alpha)^T\mathbf{e}_{ij}. \tag{III.43}$$

En posant :

$$A^\alpha = \sigma_I^{-2}\sum_{i,j;i\neq j}(J_i^\alpha - J_j^\alpha D_{ij}^\alpha)^T(J_i^\alpha - J_j^\alpha D_{ij}^\alpha) + \sigma_{\theta^\alpha}^{-2}I_d, \tag{III.44}$$

on retrouve dans cette expression la forme d'une distribution normale, on peut déduire que :

$$p(\theta^\alpha|\mathcal{J}) \propto \mathcal{N}\Big(\sigma_I^{-2}A^{\alpha-1}\sum_{i,j;i\neq j}(J_i^\alpha - J_j^\alpha D_{ij}^\alpha)^T\mathbf{e}_{ij}, A^{\alpha-1}\Big). \tag{III.45}$$

Dans le deuxième cas de vraisemblance des images donné par l'équation (III.37), l'estimateur du MAP se déduit de la même façon :

$$p(\theta^\alpha|\mathcal{J}) \propto \exp\Big(-\frac{1}{2}(\theta^\alpha - \bar{\theta}^\alpha)^T\big(\sigma_I^{-2}\sum_{i,j;i\neq j}K_{ij}^{\alpha T}J_j^{\alpha T}J_j^\alpha K_{ij}^\alpha + \sigma_{\theta^\alpha}^{-2}I_d\big)(\theta^\alpha - \bar{\theta}^\alpha)\Big), \tag{III.46}$$

avec cette fois-ci :

$$\bar{\theta}^\alpha = \sigma_I^{-2}(\sigma_I^{-2}\sum_{i,j;i\neq j}K_{ij}^{\alpha T}J_j^{\alpha T}J_j^\alpha K_{ij}^\alpha + \sigma_{\theta^\alpha}^{-2}I_d)^{-1}\sum_{i,j;i\neq j}K_{ij}^{\alpha T}J_j^{\alpha T}\mathbf{e}_{ij}. \tag{III.47}$$

Comme précédemment, on retrouve la forme d'une distribution normale dans la loi *a posteriori* :

$$p(\theta^\alpha|\mathcal{J}) \propto \mathcal{N}\Big(\sigma_I^{-2}A^{\alpha-1}\sum_{i,j;i\neq j}K_{ij}^{\alpha T}J_j^{\alpha T}\mathbf{e}_{ij}, A^{\alpha-1}\Big), \tag{III.48}$$

où :

$$A^\alpha = \sigma_I^{-2}\sum_{i,j;i\neq j}K_{ij}^{\alpha T}J_j^{\alpha T}J_j^\alpha K_{ij}^\alpha + \sigma_{\theta^\alpha}^{-2}I_d. \tag{III.49}$$

Ces calculs sont ceux d'une régression standard bayesienne, tels que très clairement décrits en page 9 de [Rasmussen 06]. Un développement plus détaillé est disponible en annexe A.

L'estimateur du MAP recherche le $\hat{\theta}^\alpha$ qui maximise la distribution *a posteriori*. Dans ce contexte gaussien, le mode de la distribution est simplement donné par sa moyenne, soit $\hat{\theta}_{MAP}^\alpha = \bar{\theta}^\alpha$ dans les deux cas. Supposons que nous connaissons les deux valeurs de σ_I

et σ_{θ^α} - nous présenterons une façon de les estimer au chapitre suivant - , nous pouvons tracer les estimations obtenues pour le même jeu de données que celui présenté au paragraphe précédent en figure III.10 et III.11.

On constate que pour les deux estimateurs, le roulis déduit est très semblable à celui ob-

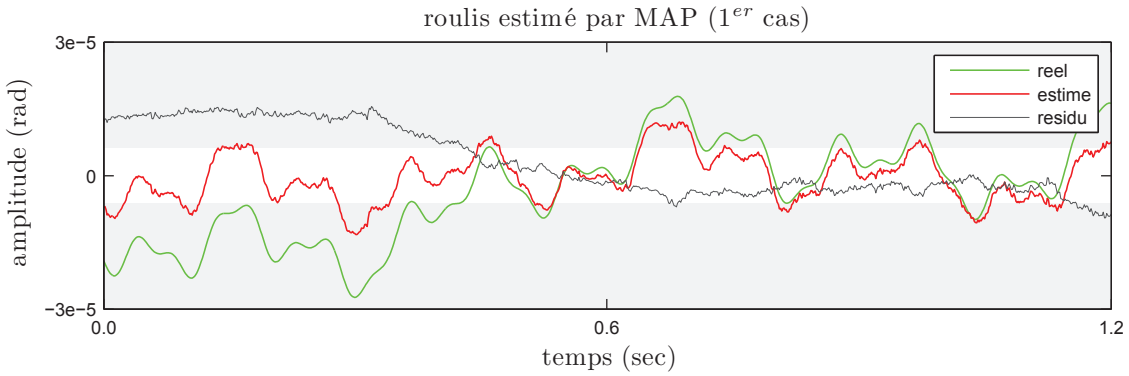


FIGURE III.10 – Exemple d'estimation du roulis par MAP selon la loi *a posteriori* enoncée en (III.42) ; $\sigma_I = 0.012$ et $\sigma_{\theta^r} = 9.2 \times 10^{-7}$. L'estimé est identique à celui obtenu en figure III.5, l'*a priori* n'apporte pas d'information supplémentaire par rapport aux données image.

tenu par maximum de vraisemblance en figures III.5 et III.6. La seule différence notable se situe dans le recentrage autour de zéro de la courbe rouge pour le second cas. Selon le noyau de convolution, le processus de déconvolution peut perdre le décalage initial - également nommé *offset* en anglais -, et dégrader les performances d'un algorithme itératif d'estimation. Ce type d'*a priori* sur la norme peut rectifier ce problème, mais il reste relativement pauvre dans l'information qu'il apporte sur les variations d'attitude. Il ne privilégie aucune fréquence, et peut même avoir tendance à sous-estimer la variable recherchée.

L'*a priori* présenté ici est souvent utilisé pour résoudre des problèmes mal conditionnés,

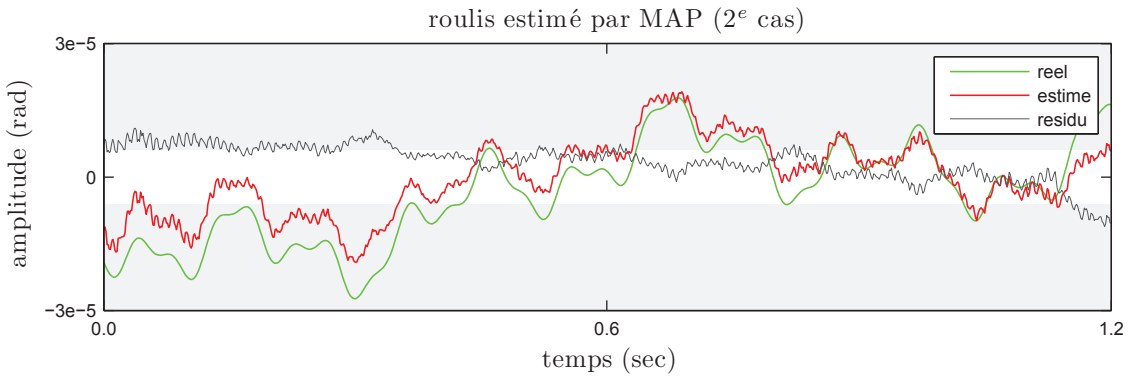


FIGURE III.11 – Exemple d'estimation du roulis par MAP selon la loi *a posteriori* calculée en (III.46) ; $\sigma_I = 0.012$ et $\sigma_{\theta^r} = 1.5 \times 10^{-6}$. En comparaison au résultat de la figure III.6, l'*a priori* a recentré la courbe du roulis en rouge autour de zéro.

pour lesquels le bruit de mesure sur les données peut faire varier de façon importante le

calcul d'une solution. Il est généralement présenté comme un terme de régularisation de type Tikhonov [Tikhonov 77], car son expression reste linéaire selon θ^α, facilitant ainsi la résolution du problème.
La similitude entre les solutions données par l'estimateur du MAP et celui du MV nous montre ici que le problème n'est pas réellement mal conditionné. Nous avons remarqué au paragraphe III.4.5 que chaque angle doit vérifier L équations ; les mesures apportées par tous les pixels étant nombreuses, elles contribuent d'autant plus à l'estimation par rapport à un faible *a priori*. En revanche, elles ne fournissent pas d'indication sur l'évolution temporelle de l'attitude, ce qui dégrade les performances de l'algorithme de recalage d'image.

III.5.3 Dérivée seconde

Une autre approche peut consister à faire des suppositions sur la dérivée temporelle des variations d'attitude. Nous avons remarqué que les estimations faites jusqu'à présent montrent des oscillations hautes fréquences importantes. Ces parasites forment localement des petits pics faisant augmenter brutalement la dérivée de θ^α. En choisissant l'hypothèse suivante :

$$\frac{\partial^2\theta^\alpha}{\partial^2 n} \sim \mathcal{N}(\mathbf{0}, \sigma^2_{\theta^\alpha} I_d), \tag{III.50}$$

et en fixant une variance $\sigma^2_{\theta^\alpha}$ suffisamment petite, l'*a priori* doit inciter les variations d'attitude à évoluer de façon lisse au cours du temps. Notons ici que nous choisissons la dérivée seconde qui pénalise d'autant plus les petites oscillations. En utilisant un opérateur de différence pour le calcul de cette dérivée, nous obtenons :

$$L^\alpha\theta^\alpha \sim \mathcal{N}(\mathbf{0}, \sigma^2_{\theta^\alpha} I_d), \tag{III.51}$$

où L^α est une matrice de taille $(N \times N)$ convoluant le vecteur d'attitude avec le noyau $[1\ -2\ 1]/n_e$, n_e étant le pas d'échantillonnage en temps ; ce vecteur est généralement connu sous le nom de laplacien numérique. En pratique, L^α réalise une opération de filtrage passe-haut de θ^α pour mieux les contraindre. L'expression de l'*a priori* se développe de la façon suivante :

$$\begin{aligned} p(L^\alpha\theta^\alpha) &= \frac{1}{(2\pi\sigma^2_{\theta^\alpha})^{N/2}} \exp\Big(-\frac{1}{2\sigma^2_{\theta^\alpha}}|L^\alpha\theta^\alpha|^2\Big) \\ &= \mathcal{N}(\mathbf{0}, \sigma^2_{\theta^\alpha}(L^{\alpha T}L^\alpha)^{-1}) \end{aligned} \tag{III.52}$$

La figure III.12 montre une nouvelle fois trois exemples de fonctions aléatoires générées selon la statistique précédente. Ces fonctions varient très lentement au cours du temps, et leurs spectres confirment la présence de très basses fréquences. Contrairement à l'*a priori* précédent, celui-ci favorise une structure spectrale dans l'estimation. Il est généralement utilisé pour renforcer le lissage en pénalisant les hautes fréquences.

Formulation du MAP

La formulation du maximum *a posteriori* est ici très similaire à la précédente dans le premier cas :

$$p(\theta^\alpha|\mathcal{J}) \propto \mathcal{N}\Big(\bar{\theta}^\alpha, A^{\alpha-1}\Big) \tag{III.53}$$

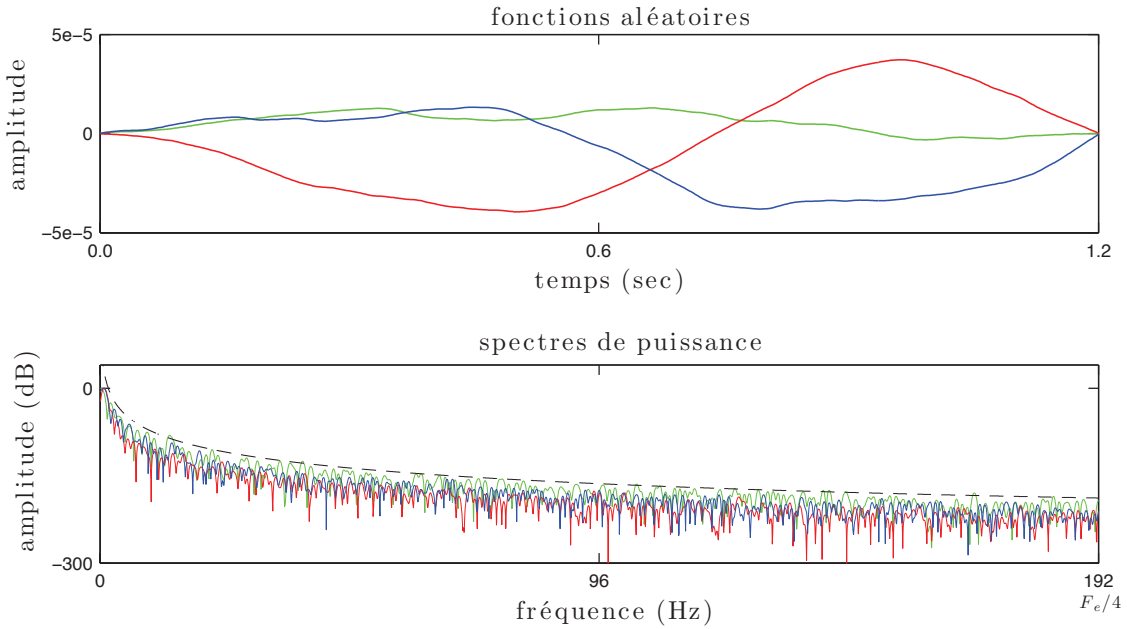


FIGURE III.12 – Le graphique du haut présente trois fonctions aléatoires tirées selon l'*a priori* défini en (III.50) avec $\sigma_{\theta^\alpha} = 3 \times 10^{-8}$; chaque courbe évolue lentement au cours du temps. La figure du dessous montre les trois spectres de puissance associés, confirmant la présence de basses fréquences. Le spectre théorique associé à l'*a priori* est tracé en tirets noirs.

avec :

$$A^\alpha = \sigma_I^{-2} \sum_{i,j;i\neq j} (J_i^\alpha - J_j^\alpha D_{ij}^\alpha)^T (J_i^\alpha - J_j^\alpha D_{ij}^\alpha) + \sigma_{\theta^\alpha}^{-2} L^{\alpha T} L^\alpha, \tag{III.54}$$

et :

$$\bar{\theta}^\alpha = -\sigma_{\theta^\alpha}^{-2} A^{\alpha -1} \sum_{i,j;i\neq j} (J_i^\alpha - J_j^\alpha D_{ij}^\alpha)^T \mathbf{e}_{ij}. \tag{III.55}$$

Dans le second cas, nous obtenons encore :

$$p(\theta^\alpha|\mathcal{J}) \propto \mathcal{N}\left(\bar{\theta}^\alpha, A^{\alpha -1}\right) \tag{III.56}$$

avec :

$$A^\alpha = \sigma_I^{-2} \sum_{i,j;i\neq j} K_{ij}^{\alpha T} J_j^{\alpha T} J_j^\alpha K_{ij}^\alpha + \sigma_{\theta^\alpha}^{-2} L^{\alpha T} L^\alpha, \tag{III.57}$$

et :

$$\bar{\theta}^\alpha = \sigma_{\theta^\alpha}^{-2} A^{\alpha -1} \sum_{i,j;i\neq j} K_{ij}^{\alpha T} J_j^{\alpha T} \mathbf{e}_{ij} \tag{III.58}$$

pour le second. Nous pouvons à nouveau traiter les données précédentes avec cet estimateur du MAP modifié, les figures III.13 et III.14 présentent les résultats obtenus pour l'estimation du roulis.

Les deux estimés sont assez nettement améliorés par rapport à ceux observés en figures III.10 et III.11, les résidus présentent moins d'oscillations parasites hautes fréquences et le contenu basse fréquence est maintenu. On peut relever un détail intéressant dans le second cas, où le décalage de la courbe rouge en amplitude par rapport à la courbe verte

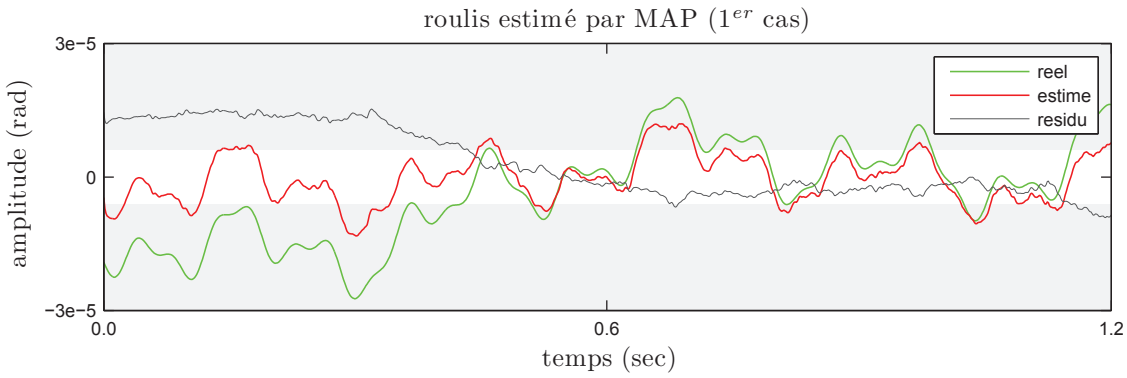


FIGURE III.13 – Exemple d'estimation du roulis par MAP selon la loi (III.53) et avec l'*a priori* (III.50) ; $\sigma_I = 0.012$ et $\sigma_{\theta^r} = 7.3 \times 10^{-8}$. Les oscillations hautes fréquences observées dans les résultats précédents sont ici en partie atténuées. On peut remarquer que l'estimation en basse fréquence n'est pas améliorée, malgré un *a priori* favorisant les basses fréquences.

est plus important que dans l'estimation par MV en figure III.6. Le problème est lié à la perte du décalage initial par convolution, comme nous l'avons observé au paragraphe précédent. Alors que l'*a priori* sur la norme encourageait la courbe à se centrer autour de zéro, l'*a priori* sur la dérivée (III.50) renforce les très basses fréquences, ce qui peut contribuer à l'augmentation du décalage.

On peut remarquer que la cohérence temporelle est introduite ici par la minimisation

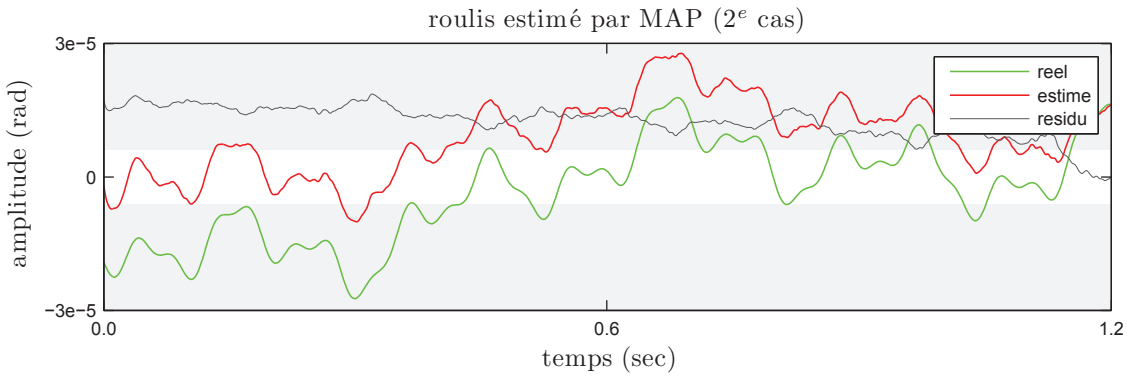


FIGURE III.14 – Exemple d'estimation du roulis par MAP selon la loi (III.56) et avec l'*a priori* (III.50) ; $\sigma_I = 0.012$ et $\sigma_{\theta^r} = 6.2 \times 10^{-8}$. Les fortes oscillations hautes fréquences observées dans les résultats précédents ont presque disparu, mais on peut noter une augmentation du décalage en amplitude de l'estimé par rapport à la courbe verte, ceci dû au renforcement des basses fréquences selon l'*a priori*.

de la dérivée temporelle. Numériquement, on peut réécrire cette contrainte de la façon suivante :

$$\theta^\alpha(n-1) - 2\theta^\alpha(n) + \theta^\alpha(n+1) = \epsilon_{\theta^\alpha}(n), \text{ avec } n \in \{2, 3, \ldots, N-1\} \text{ et } \epsilon \sim \mathcal{N}(0, \sigma^2_{\theta^\alpha}). \tag{III.59}$$

En pratique, l'algorithme cherche à minimiser une combinaison linéaire entre trois échantillons temporels d'attitude successifs. Cette relation décrit une correspondance possible entre échantillons, nous allons voir que l'on peut en trouver plusieurs autres.

III.5.4 Processus Gaussiens

Les *a priori* que nous avons développés jusqu'à maintenant considèrent soit que les variations d'attitude sont découplées au cours du temps, mais contraintes à se centrer autour de zéro, soit que leur dérivée seconde est petite, ce qui renforce le lissage au cours du temps. Supposons maintenant que chaque échantillon temporel est dépendant de ses voisins, ceci afin d'adopter l'idée de structure temporelle dans le signal à estimer. Cette dépendance peut se traduire sous la forme d'un processus gaussien comme suit :

$$\theta^{\alpha}(n) \sim \mathcal{GP}(0, k^{\alpha}(n, n')), \tag{III.60}$$

où $k^{\alpha}(n, n')$ est une fonction de covariance qui dicte les relations de l'échantillon d'attitude à l'instant n avec ses voisins aux instants n'. Une fonction de covariance couramment utilisée est la suivante :

$$k^{\alpha}(n, n') = \sigma^2_{\theta^{\alpha}} \exp\left(-\frac{1}{2l^{\alpha}}(n - n')^2\right), \tag{III.61}$$

celle-ci prend la forme d'une gaussienne comme nous pouvons le voir en figure III.15, où $l^{\alpha} \in \mathbb{R}^+$ est une variable qui élargit ou réduit le lobe de la fonction, à la façon d'un facteur d'échelle.
Si deux variables sont indépendantes alors leur covariance est nulle. Intuitivement, on comprend alors que la fonction (III.61) attribue une covariance importante à deux éléments proche en temps, en effet si $|n - n'| \to 0$ alors $k^{\alpha}(n, n') \to \sigma^2_{\theta^{\alpha}}$, inversement si $|n - n'| \to \infty$ alors $k^{\alpha}(n, n') \to 0$. Tout comme précédemment, nous pouvons tracer des fonctions aléatoires échantillonnées selon la statistique définie en (III.60) avec différentes valeurs de l^{α} ; celles-ci sont représentées en figure III.15. On peut voir que la fonction de covariance assure un lissage local proportionnel à l'élargissement de son lobe.

On remarque également que la fonction de covariance a une influence directe sur le spectre des fonctions tirées : lorsque l^{α} tend vers la période d'échantillonnage, le spectre s'étale sur toutes les fréquences à la façon d'un bruit blanc ; ceci est représenté sur la colonne gauche de la figure. Nous avons déjà observé un spectre similaire pour l'*a priori* sur l'amplitude. Inversement lorsque l^{α} tend vers la période d'observation du signal, qui est ici de 1.2 secondes, le spectre se concentre sur les très basses fréquences ; la colonne droite de la figure permet de voir ce phénomène. Enfin notons que la moyenne, supposée nulle ici, et la variance $\sigma^2_{\theta^{\alpha}}$ permettent d'ajuster l'amplitude du signal dans un intervalle. Ces observations font que cet *a priori* est nettement plus flexible que les deux précédents.
La fonction de covariance définie en (III.61) est très utilisée dans l'inférence de processus gaussiens ; elle présente l'avantage d'être simple dans son expression mathématique, et de produire des solutions analytiques lorsqu'elle est introduite dans l'estimateur du MAP. Il existe de nombreuses autres fonctions de covariances pouvant s'adapter à des non-stationnarités dans le signal - nous renvoyons le lecteur au chapitre 4 de [Rasmussen 06] -, mais celles-ci ont souvent plusieurs hyperparamètres à estimer sans formulation analytique simple. Par chance, le contexte stationnaire des variations d'attitude fait du noyau gaussien une fonction de covariance bien adaptée à l'estimation des mouvements du satellite. En effet, la fonction (III.61) dépend seulement de la différence

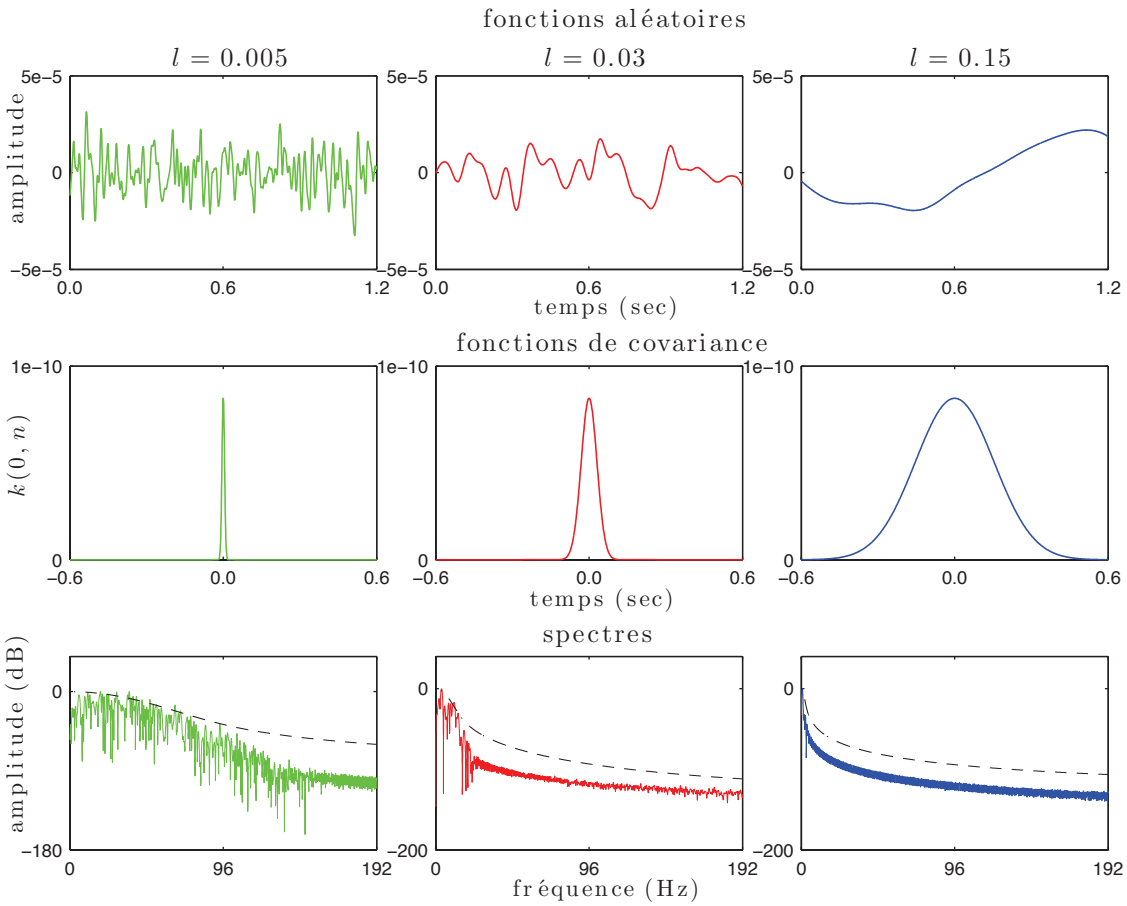


FIGURE III.15 – Exemples de fonctions aléatoires tirées selon l'*a priori* défini en (III.60), avec la fonction de covariance (III.61), et pour différentes valeurs de l^{α} (0.005, 0.03 et 0.15 de gauche à droite). $\sigma_{\theta^{\alpha}}$ a été fixé à 10^{-5}, on constate que les fonctions se maintiennent dans un intervalle autour de $[-3\sigma_{\theta^{\alpha}}, 3\sigma_{\theta^{\alpha}}]$. Le paramètre l^{α} influence le contenu spectral de la fonction *a priori* : une petite valeur de l^{α} implique des échantillons corrélés sur une petite période, la fonction de covariance est étroite et les hautes fréquences sont présentes, et inversement une grande valeur de l^{α} corrèle les échantillons sur une période plus longue, le lobe de la fonction de covariance est large et les basses fréquences prédominent dans le signal aléatoire.

de temps entre deux échantillons $(n - n')$, ce qui est une caractéristique des processus stationnaires.
Une fonction arbitraire entre deux échantillons n'est pas nécessairement une fonction de covariance valide. Pour cela, elle doit être au préalable semi-definie positive. k^α doit être continûment dérivable par rapport à n, et :

- $k^\alpha(0,0) = 0$,
- $k^\alpha(n,n') \geqslant 0$ pour tout couple $(n,n') \in \mathbb{R}^2$.

La matrice de covariance associée $\Sigma^\alpha_{ij} = k^\alpha(n_i, n_j)$, appelée matrice de Gram, est également définie semi-positive. Enfin, il est possible de combiner les fonctions de covariance pour représenter des signaux au contenu fréquentiel plus varié [Rasmussen 06], comme le montre en exemple la figure III.16.
Les trois signaux aléatoires représentés en figure III.16 sont générés par la somme de

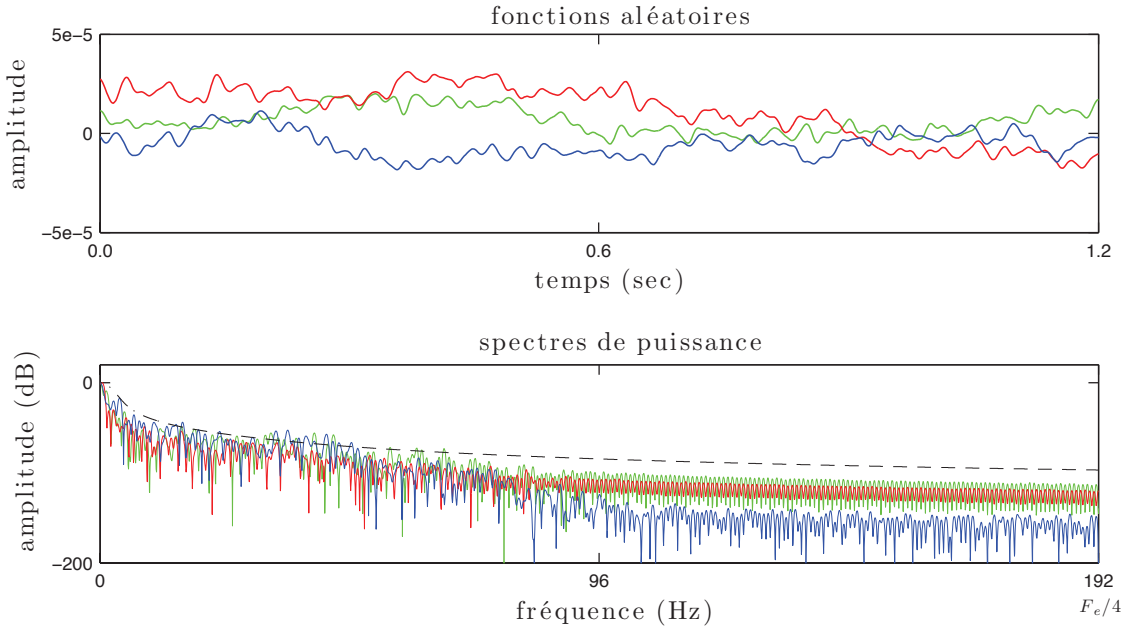


FIGURE III.16 – Le graphique du haut montre trois signaux aléatoires tirés avec l'*a priori* de l'équation (III.60) ; la fonction de covariance k^α est définie comme la somme de deux noyaux gaussiens $k^\alpha_1 + k^\alpha_2$ selon l'équation (III.61) et représentée en figure III.17. Les paramètres choisis sont $l^\alpha_1 = 0.1$, $l^\alpha_2 = 0.007$, $\sigma_{1,\theta^\alpha} = 10^{-5}$ et $\sigma_{2,\theta^\alpha} = 3 \times 10^{-6}$. Les spectres associés démontrent la forte présence des basses fréquences, et un étalement sur les moyennes fréquences.

deux fonctions de covariance tracées en figure III.17 ; ils montrent chacun des oscillations basses fréquences importantes et des moyennes fréquences plus petites. Au regard des spectres, on constate que le pic basse fréquence est bien identifié dans les trois cas, contrairement aux pics moyennes fréquences qui s'étalent sur le spectre, et varient selon le signal.

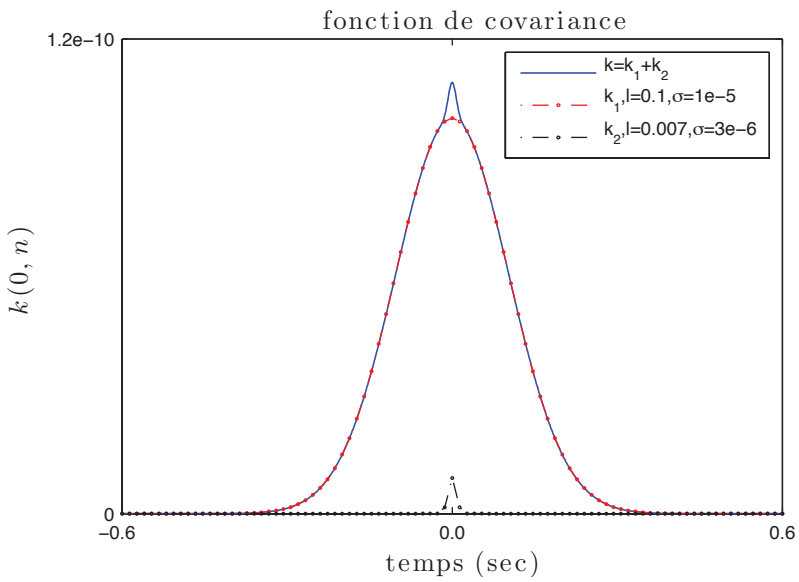


FIGURE III.17 – La fonction de covariance, tracée ici en bleu, a permis de générer les fonctions aléatoires de la figure III.16. On remarque bien par sa forme la présence de deux covariances gaussiennes de paramètres différents, tracées en pointillés rouge et vert, qui font apparaitre les basses et moyennes fréquences observées dans les simulations.

Formulation du MAP

Soit Σ^α la matrice de covariance de taille $(N \times N)$, l'*a priori* de type processus gaussien peut s'écrire comme suit :

$$\theta^\alpha \sim \mathcal{N}(\mathbf{0}, \Sigma^\alpha), \tag{III.62}$$

l'estimateur du MAP prend la même forme que l'équation (III.53) dans les deux cas, avec un $\bar{\theta}^\alpha$ respectivement identique à l'équation (III.55) pour le premier cas et (III.58) pour le second (p.50), mais une matrice A^α légèrement modifiée. Pour le premier cas nous obtenons :

$$A^\alpha = \sigma_I^{-2} \sum_{i,j;i\neq j} (J_i^\alpha - J_j^\alpha D_{ij}^\alpha)^T (J_i^\alpha - J_j^\alpha D_{ij}^\alpha) + \Sigma^{\alpha-1}, \tag{III.63}$$

et dans le deuxième cas :

$$A^\alpha = \sigma_I^{-2} \sum_{i,j;i\neq j} K_{ij}^{\alpha\,T} J_j^{\alpha\,T} J_j^\alpha K_{ij}^\alpha + \Sigma^{\alpha-1}. \tag{III.64}$$

Les résultats sur le jeu de données utilisé jusqu'à maintenant sont visibles en figures III.18 et III.19. Dans les deux cas, le bruit parasite a totalement disparu, la fonction de covariance agissant à la manière d'un filtre sur l'estimé.

En comparaison des résultats précédents, le résidu présente cependant quelques moyennes fréquences, qui ont dû être coupées par un choix de fenêtre trop large, ou favorisant trop les basses fréquences.

La figure III.15 nous permet de voir une limite de ce type d'*a priori* pour la fonction de covariance choisie. Il n'est pas possible d'être sélectif sur une bande de fréquence, dans la connaissance que l'on souhaite apporter au signal à estimer. En effet, un *a priori* haute

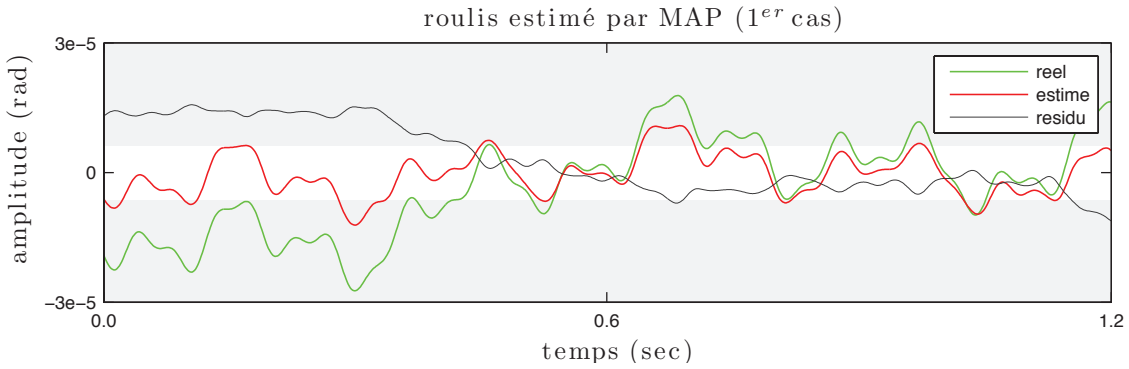


FIGURE III.18 – Exemple d'estimation du roulis par MAP avec un *a priori* de type processus gaussien ; $\sigma_I = 0.012$, $\sigma_{\theta^r} = 3 \times 10^{-2}$ et $l^r = 0.02$. Le signal rouge est entièrement lissé, on voit apparaitre quelques moyennes fréquences dans le résidu qui n'apparaissaient pas dans les résultats précédents.

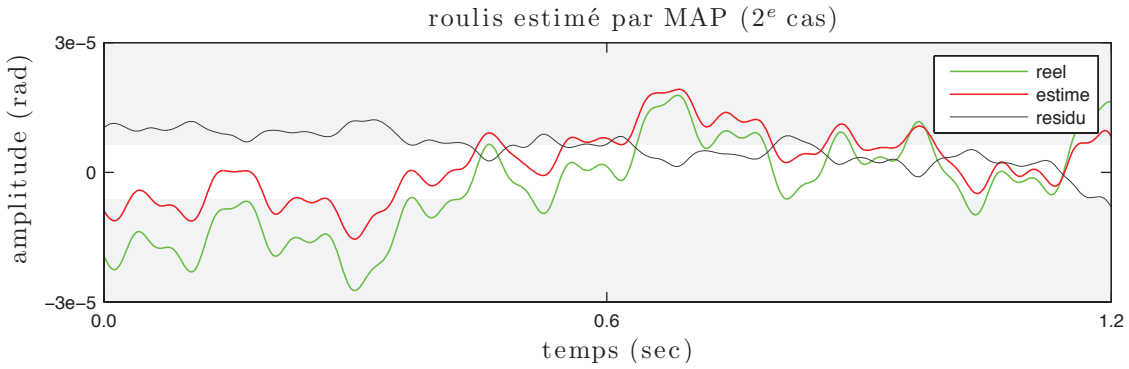


FIGURE III.19 – Exemple d'estimation du roulis par MAP avec un *a priori* de type processus gaussien ; $\sigma_I = 0.012$, $\sigma_{\theta^r} = 3 \times 10^{-2}$ et $l^r = 0.02$. Nous observons, comme un figure III.18, des moyennes fréquences dans le résidu qui ont dû être coupées par l'*a priori*.

fréquence étale le spectre jusqu'à une fréquence limite, tout en conservant les basses fréquences. Supposons que nous disposons d'informations sur la localisation de certaines raies fréquentielles, la fonction de covariance gaussienne ne permettra pas de les préciser dans l'estimation.

III.6 Processus autorégressif

Nous avons précisé en début de partie que les variations d'attitude étaient essentiellement liées aux vibrations des moteurs. L'étude et la modélisation de ces phénomènes vibratoires est un domaine de recherche assez vaste en acoustique et mécanique [Wowk 91, Rao 04]. Les méthodes d'analyse sont multiples entre les approches fréquentielles, temporelles et modales, dont un des objectifs est souvent de détecter les harmoniques du signal. Parmi les techniques temporelles, le modèle autorégressif (AR) est régulièrement cité pour son efficacité [Makhoul 75, Marple 86]. Dans son principe, le signal est modélisé à chaque instant de temps comme une combinaison linéaire de ses valeurs passées :

$$\theta^{\alpha}(n) = \sum_{p=1}^{P^{\alpha}} a_p^{\alpha}\theta^{\alpha}(n-p) + \epsilon_{\theta^{\alpha}}(n), \tag{III.65}$$

où a_p^{α} pour $1 \leqslant p \leqslant P^{\alpha}$ désignent les paramètres du modèle, et P^{α} son ordre ; $\epsilon_{\theta^{\alpha}}(n)$ est un bruit blanc centré de variance $\sigma^2_{\theta^{\alpha}}$ assimilé à une erreur de prédiction. Ce modèle peut être compris comme un type particulier de processus gaussien [Rasmussen 06], il est également très souvent utilisé pour faire du filtrage ou de la prédiction linéaire sur des séries temporelles [Kalman 60, Wiener 64, Makhoul 75]. L'ordre du modèle est un critère important dans sa précision, comme le présente la figure III.20 où plusieurs spectres de modèles AR à différents ordres sont tracés et comparés au signal modélisé. Ce dernier correspond à celui de la figure III.7 sans la portion non stationnaire. On constate assez logiquement que le modèle AR suit de mieux en mieux le spectre du signal lorsque son ordre augmente, mais il est aussi intéressant de noter qu'il commence par spécifier les hautes fréquences, comme en témoigne le pic à 100 Hz de la courbe rouge, avant de s'étendre sur les basses fréquences. Les signaux hautes fréquences étant visibles sur une courte période en temps, seuls quelques coefficients seront nécessaires pour les modéliser. En basse fréquence, un intervalle d'observation beaucoup plus long, et donc un ordre du modèle plus grand, seront utiles pour prendre en compte les évolutions du signal. En théorie, un signal aléatoire peut même être prédit sans erreur par prédiction linéaire s'il est observé sur un passé infini [Gravey 81]. Notons qu'il est possible d'être sélectif pour concentrer le modèle AR dans une bande de fréquence que l'on souhaite analyser [Makhoul 75], ces méthodes procèdent essentiellement par filtrage du signal.
D'après l'équation (III.65), on peut constater que :

$$\theta^{\alpha}(n) - \sum_{p=1}^{P^{\alpha}} a_p^{\alpha}\theta^{\alpha}(n-p) \sim \mathcal{N}(0, \sigma^2_{\theta^{\alpha}}). \tag{III.66}$$

Cette relation linéaire peut s'écrire sous forme matricielle de la façon suivante :

$$R^{\alpha}\boldsymbol{\theta}^{\alpha} \sim \mathcal{N}(\mathbf{0}, \sigma^2_{\theta^{\alpha}} I_d), \tag{III.67}$$

où R^{α} est une matrice creuse de taille $(N - P^{\alpha} \times N)$ contenant les coefficients a_p^{α} du modèle sur ses diagonales, cette matrice est décrite en annexe A. Par la suite, nous appellerons

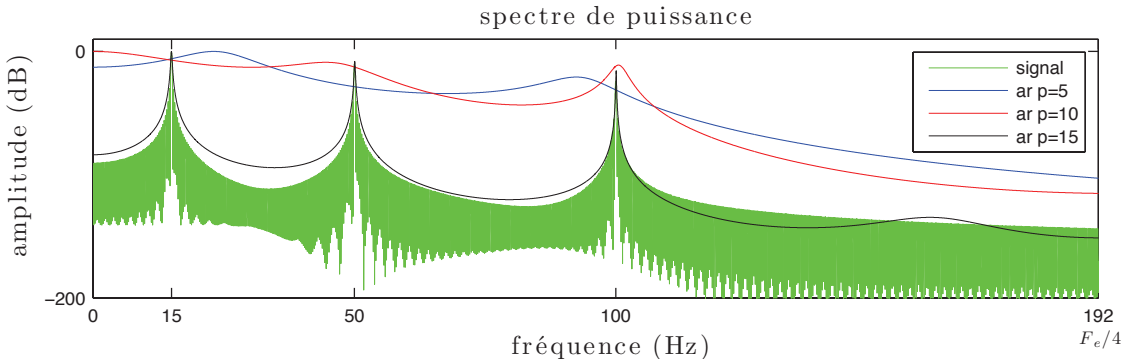


FIGURE III.20 – Spectres de modèles autorégressifs (équation (III.65)) du signal présenté en figure III.7 pour différents ordres ($P^\alpha = 5$ en bleu, $P^\alpha = 10$ en rouge et $P^\alpha = 15$ en noir). Lorsque P^α augmente, la précision du modèle autour des harmoniques du signal grandit. Le modèle favorise en premier lieu les pics hautes fréquences, comme en témoigne le pic à 100 Hz de la courbe rouge, avant de se propager aux pics basses fréquences.

$\mathbf{a}^\alpha$ le vecteur $(P^\alpha \times 1)$ des coefficients du modèle AR :

$$\mathbf{a}^\alpha = [a_1^\alpha \ \dots \ a_{P^\alpha}^\alpha]^T. \tag{III.68}$$

On peut développer l'expression de l'*a priori* de la même manière que pour l'équation (III.52) :

$$\begin{aligned} p(R^\alpha \theta^\alpha) &= \frac{1}{(2\pi\sigma_{\theta^\alpha}^2)^{N/2}} \exp\Big(-\frac{1}{2\sigma_{\theta^\alpha}^2}|R^\alpha \theta^\alpha|^2\Big) \\ &= \mathcal{N}(\mathbf{0}, \sigma_{\theta^\alpha}^2 (R^{\alpha T} R^\alpha)^{-1}). \end{aligned} \tag{III.69}$$

La figure III.21 montre trois simulations de fonctions aléatoires suivant la statistique (III.69). L'ordre du modèle AR est de 150, et les coefficients ont été estimés sur le roulis réel des données visibles en figure III.8.

De cette figure, on comprend que le modèle de l'équation (III.69) permet de générer des *a priori* beaucoup plus spécifiques que ceux que nous avons vus jusqu'à présent. En supposant que l'on connaisse certaines harmoniques du signal, cette information peut être ajoutée dans le processus d'estimation au moyen du modèle AR. Il est intéressant de remarquer une nouvelle fois dans les spectres des fonctions la précision des pics hautes fréquences. Alors que l'*a priori* du paragraphe précédent avec une fonction de covariance gaussienne ne permet pas d'encourager spécifiquement les hautes fréquences, l'*a priori* autorégressif semble être particulièrement bien dédié à cette tâche.

Formulation du MAP

Une nouvelle fois dans l'écriture du maximum *a posteriori*, seule la matrice A^α change. L'équation (III.53) (p.49) reste valable dans les deux cas avec $\bar{\theta}^\alpha$ égale à l'équation (III.55) pour le premier cas et l'équation (III.58) pour le second (p.50). Nous avons donc :

$$A^\alpha = \sigma_I^{-2} \sum_{i,j;i\neq j} (J_i^\alpha - J_j^\alpha D_{ij}^\alpha)^T (J_i^\alpha - J_j^\alpha D_{ij}^\alpha) + \sigma_{\theta^\alpha}^{-2} R^{\alpha T} R^\alpha, \tag{III.70}$$

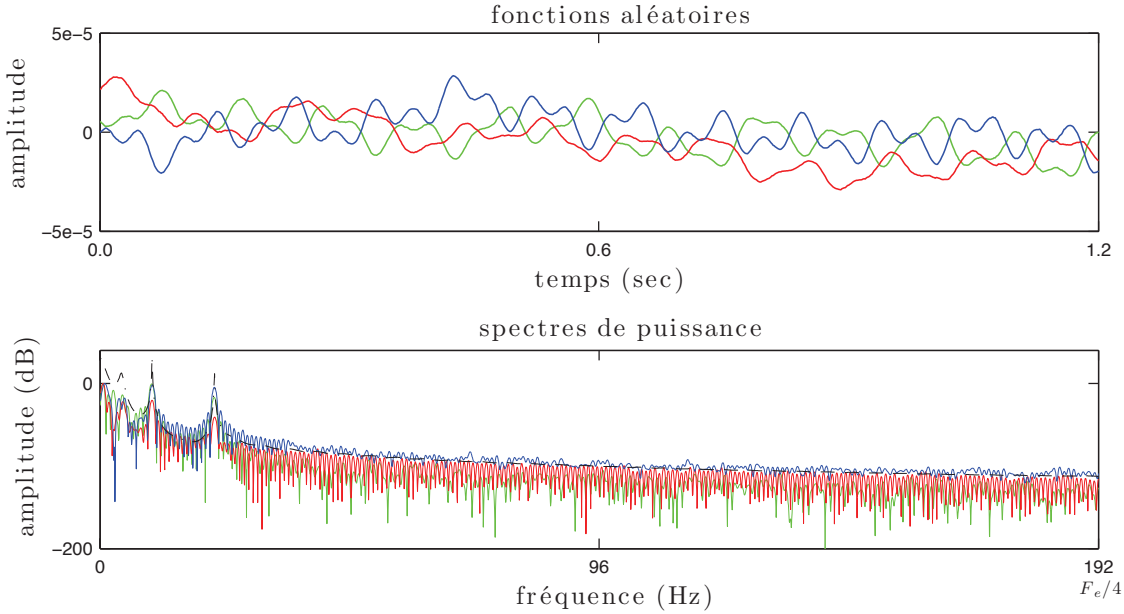


FIGURE III.21 – Représentation temporelle et spectrale de trois fonctions aléatoires tirées selon l'*a priori* autorégressif défini en (III.69). Le modèle AR contient 150 coefficients, estimés à partir du roulis réel des données (figure III.8). L'ordre du modèle a été fixé par expertise suivant le nombre de raies que contient le roulis réel. On peut noter la précision des pics hautes fréquences dans le spectre, ce modèle permet de générer des signaux plus spécifiques.

pour le premier terme de vraisemblance et :

$$A^{\alpha} = \sigma_I^{-2} \sum_{i,j;i\neq j} K_{ij}^{\alpha\,T} J_j^{\alpha\,T} J_j^{\alpha} K_{ij}^{\alpha} + \sigma_{\theta^{\alpha}}^{-2} R^{\alpha\,T} R^{\alpha}, \tag{III.71}$$

pour le second. Les résultats obtenus avec ces nouveaux estimateurs sont disponibles aux graphiques III.22 et III.23. Notons que pour cette expérience, les coefficients $\mathbf{a}^{\alpha}$ sont supposés connus, ce qui n'est souvent pas le cas en pratique. Le schéma algorithme permettant de les estimer sera présenté au chapitre suivant.

Le premier estimé en rouge de la figure III.22 ne montre plus d'oscillations parasites. On

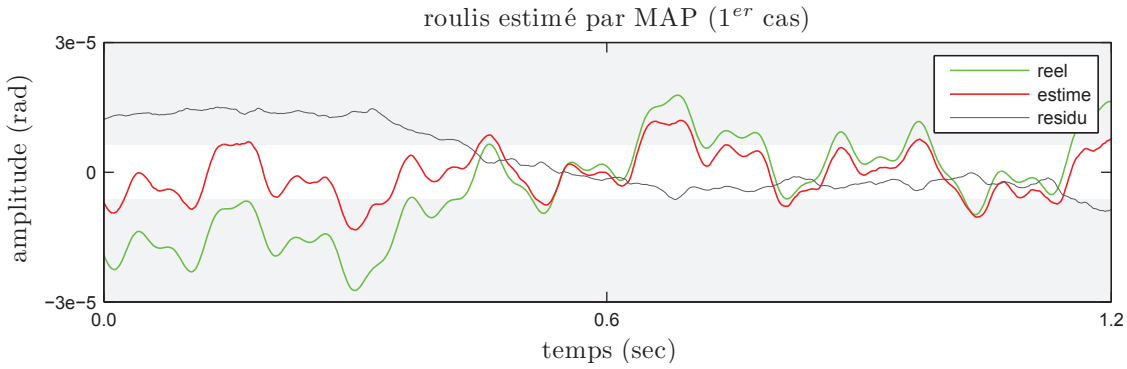


FIGURE III.22 – Exemple d'estimation du roulis par MAP avec un *a priori* autorégressif ; le modèle contient 150 coefficients, estimés à partir du roulis réel ; $\sigma_I = 0.012$ et $\sigma_{\theta^r} = 3.3 \times 10^{-8}$. Le résidu obtenu est essentiellement dû à l'approximation linéaire des termes images, il ne contient plus de fréquences parasites comme celles observées en figure III.5.

peut voir que la courbe restitue correctement une partie des fréquences du roulis réel en vert. Une oscillation très basse fréquence persiste dans le résidu, elle est sans doute liée à la linéarisation des termes images ; une discussion a été faite à ce sujet aux paragraphes III.4.5 et III.4.6.

Le résultat de la figure III.23 est encore plus intéressant. Le bruit parasite lié à la déconvolution a disparu, les basses fréquences sont mieux restituées que dans le premier cas, et le résidu ne semble être lié qu'au terme de linéarisation de l'image. Alors que l'estimateur du MV montrait des avantages et inconvénients aux deux termes de vraisemblance des images, l'utilisation d'un *a priori* adéquat met en avant la seconde solution qui restitue plus fidèlement le signal.

III.7 Discussion sur les estimateurs

L'estimateur du MAP a de réels atouts lorsqu'on dispose d'informations supplémentaires sur le processus recherché en plus des données observées. L'*a priori* choisi doit cependant être suffisamment informatif dans notre cas, car nous disposons déjà de nombreuses mesures dans les images qui peuvent facilement prendre le pas sur l'*a priori* lorsque celui-ci se révèle trop vague. Nous avons notamment expérimenté cela pour le premier *a priori* sur l'amplitude des variations d'attitude. Ce principe est un des fondements des méthodes bayesiennes, qui stipule qu'en présence d'un nombre croissant de mesures, le maximum *a posteriori* tend vers le maximum de vraisemblance [Jaynes 03,

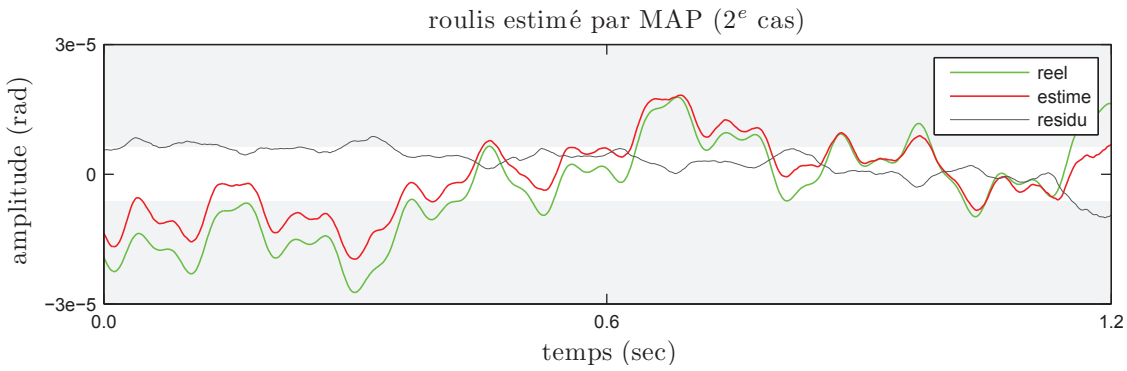


FIGURE III.23 – Exemple d'estimation du roulis par MAP avec un *a priori* autorégressif ; le modèle contient 150 coefficients, estimés à partir du roulis réel ; $\sigma_I = 0.012$ et $\sigma_{\theta^r} = 3.3 \times 10^{-8}$. L'estimé est meilleur que celui observé en figure III.22, l'*a priori* a supprimé les défauts observés en figure III.6 et liés à la déconvolution.

Mackay 02].
Chacun des *a priori* que nous venons de décrire sont liés par une même statistique qui est la loi normale. Ils font partie d'une même classe, et les processus gaussiens sont une généralisation des trois autres ; nous renvoyons le lecteur au livre [Rue 05] pour plus de détails à ce sujet, ainsi qu'à l'annexe de [Rasmussen 06] sur les processus de Markov gaussiens. Le tableau III.1 propose une synthèse des *a priori* que nous avons présentés dans cette partie.
L'*a priori* le plus informatif utilisant le modèle AR se révèle être également le plus performant dans l'estimation par MAP. Malheureusement c'est également celui qui dispose du plus grand nombre d'hyperparamètres à estimer au préalable : l'ordre P^α, le vecteur des coefficients $\mathbf{a}^\alpha$ de taille $(P^\alpha \times 1)$ et l'écart type σ_{θ^α}. L'*a priori* défini à partir de la fonction de covariance gaussienne nécessite pour sa part l'inférence d'au moins deux hyperparamètres dans sa forme la plus simple : l^α et σ_{θ^α}. Enfin les deux premiers *a priori* n'ont besoin que d'un hyperparamètre : σ_{θ^α}. Notons que nous avons omis dans les quatre cas le paramètre σ_I de bruit dans les images que nous supposerons connu. Il serait même possible d'ajouter un troisième échelon d'inférence, en plus de l'estimation des variations d'attitude et des hyperparamètres, si nous souhaitons laisser une incertitude dans le choix entre différents modèles pour θ^α. Ce principe est également au coeur des méthodes bayesiennes, comme le présente très clairement [Mackay 92, Rasmussen 06, Punska 99]. Bien que l'inférence complète soit attractive d'un point de vue théorique, elle reste difficile à appliquer en pratique étant donné le nombre d'intégrales à évaluer sans formulation analytique.
Dans notre étude, le maximum de vraisemblance présente l'avantage de ne nécessiter aucun hyperparamètre, nous avons même remarqué au paragraphe III.4.6 que la connaissance de σ_I était inutile. Précisons encore une fois que la bonne qualité relative des estimés en figures III.5 et III.6 est due à la quantité de données image dont nous disposons. On peut alors se poser la question de réduire le nombre de mesures pour gagner en mémoire sur l'ordinateur et en vitesse d'estimation. Nous tenterons de répondre à cette question au chapitre suivant ; une analyse similaire a été menée sur ce sujet dans [Baker 04] où les auteurs montrent que conserver seulement 10% des pixels dans les

a priori	Amplitude	Dérivée seconde	Processus gaussien	Autorégressif
hyperparamètres	$[\sigma_{\theta^l}, \sigma_{\theta^r}, \sigma_{\theta^t}]^T$	$[\sigma_{\theta^l}, \sigma_{\theta^r}, \sigma_{\theta^t}]^T$	$[\sigma_{\theta^l}, \sigma_{\theta^r}, \sigma_{\theta^t}, l^l, l^r, l^t]^T$	$[\sigma_{\theta^l}, \sigma_{\theta^r}, \sigma_{\theta^t}, \mathbf{a}^l, \mathbf{a}^r, \mathbf{a}^t, P^l, P^r, P^t]^T$
avantages	recentre l'attitude autour de 0	pénalise les hautes fréquences	forme du spectre *a priori* ajustable, peu d'hyperparamètres	modèle de spectre *a priori* très précis, très bien adapté aux signaux stationnaires
inconvénients	aucune information temporelle	spectre *a priori* peu informatif		beaucoup d'hyperparamètres à définir

TABLE III.1 – Récapitulatif sur les *a priori*

images n'impacte pas le processus d'estimation. Enfin, nous verrons également que l'estimateur du MV, étant donné son bruit parasite, ne peut pas être utilisé dans un algorithme itératif où chaque image est interpolée plusieurs fois.

III.8 Contraintes sur l'évolution temporelle

La discussion précédente a montré que la connaissance d'hyperparamètres est nécessaire pour pouvoir utiliser l'estimateur du MAP. Est-il possible de s'affranchir de ces hyperparamètres en choisissant un modèle pour les variations d'attitude suffisamment contraignant et évitant tout parasite dans l'estimation ? Alors que le maximum *a posteriori* « encourage » l'attitude à se comporter comme l'*a priori*, nous souhaitons lui imposer une façon d'évoluer : rester lisse au cours du temps à la manière d'un processus stationnaire basse fréquence, tout en étant suffisamment flexible pour suivre les données.

III.8.1 Modèle polynômial par morceaux

Les méthodes de régression ont souvent favorisé les polynômes pour leur simplicité et la linéarité de leurs coefficients par rapport à la fonction polynômiale. La figure III.1 est un exemple typique de régression par polynôme, le lecteur peut se reporter à l'introduction du chapitre 3 de [Bishop 07] pour une explication bien détaillée. Formellement un polynôme Φ de degré D s'écrit comme suit :

$$\Phi(n) = \sum_{d=0}^{D} b_d n^d, \tag{III.72}$$

où b_d pour $d \in \{0, 1, \ldots, D\}$ sont les coefficients. On pourrait supposer qu'à un degré suffisamment grand, le polynôme puisse approcher tout type de fonction. En réalité, au delà des problèmes de stabilité numérique que pose un degré trop important, la fonction polynômiale peut ne pas être suffisamment flexible pour suivre les mesures, même lorsque D augmente. La figure III.24 présente un exemple d'échec de régression polynomiale, où le polynôme de degré 50 ne parvient pas à faire mieux que le polynôme de degré 15.

Ce problème peut être mis en parallèle avec celui de l'interpolation, où les mesures sont non bruitées et uniques par échantillon de temps. L'objectif est de trouver une fonction passant par toutes les mesures. On peut montrer qu'il existe un unique polynôme de degré D passant par $D + 1$ points. Seulement, le polynôme estimé peut devenir très instable et osciller fortement entre chaque point, ce que caractérise le phénomène de Runge [Runge 01]. Pour ces raisons, plusieurs approches locales ont été développées pour l'interpolation telles que les polynômes de Lagrange ou de Hermite, jusqu'aux splines qui sont très connues en traitement d'image et bien exposées dans [Unser 99]. Ces méthodes ont été adaptées pour des mesures bruitées, ainsi que pour la régression, on peut citer notamment les splines de lissage [Wahba 90]. Nous souhaitons développer ce principe de polynômes locaux, et l'adapter à notre problème.

Par la suite, nous garderons les notations adoptées au paragraphe III.5.1, l'indice α pouvant désigner un des trois angles de l'attitude, et les variables non indicées par α étant communes aux trois angles. Nous voulons découper notre signal en M fenêtres de l_p échantillons au cours du temps, et estimer les coefficients d'un polynôme de degré D dans chaque fenêtre. Le modèle choisi est donc polynômial par morceaux, composé de

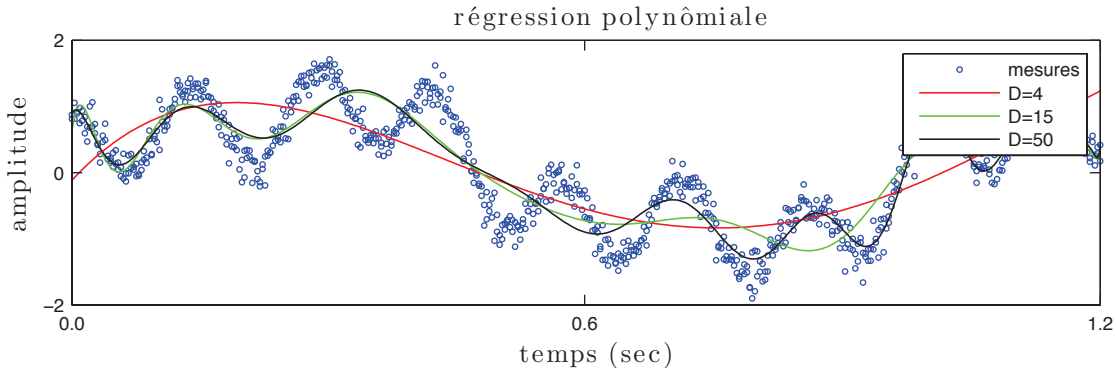


FIGURE III.24 – Régression polynômiale sur des mesures bruitées en bleu à différents ordres : $D = 4$, 15 et 50. Le polynôme de degré 4 estime les basses fréquences, les polynômes de degré 15 et 50 donnent des résultats similaires sans parvenir à suivre les oscillations rapides.

M polynômes qui sont définis entre deux instants de temps $[n_{1+l_p^m}, n_{l_p^{m+1}}]$ pour $m \in \{0, 1, \ldots, M-1\}$; les indices $l_p^m = ml_p$ appartiennent à $l_p\mathbb{Z}$, l'ensemble des entiers multiples de l_p tels que $ml_p < N$, N étant le nombre total d'échantillons de temps de l'attitude. Le modèle complet peut s'exprimer mathématiquement de la façon suivante :

$$\theta^{\alpha}(n) = \sum_{m=0}^{M-1} w_m(n)\Phi_m^{\alpha}(n), \tag{III.73}$$

où les w_m correspondent à des fonctions de pondérations telles que :

$$w_m(n) = \begin{cases} 1 & \text{pour } n \in [n_{1+l_p^m}, n_{l_p^{m+1}}] \\ 0 & \text{ailleurs.} \end{cases} \tag{III.74}$$

Enfin, nous redéfinissons les polynômes d'ordre D introduits à l'équation (III.72) de la façon suivante :

$$\Phi_m^{\alpha}(n) = \sum_{d=0}^{D} b_{d,m}^{\alpha} n^d, \tag{III.75}$$

où $b_{d,m}^{\alpha}$ sont maintenant les coefficients des polynômes suivant l'attitude et la fenêtre considérées. Tenter d'estimer le signal sous cette forme, alors que les polynômes sont indépendants, créera des discontinuités entre chaque fenêtre, or nous souhaitons que l'estimé soit lisse et stationnaire. Afin d'assurer la continuité du signal aux intersections, nous choisissons les contraintes suivantes aux points $n_{l_p^m}$ pour $m \in \{1, 2, \ldots, M-1\}$:

$$\Phi_m^{\alpha(d)}(n_{l_p^m}) = \Phi_{m+1}^{\alpha(d)}(n_{l_p^m}) \text{ pour } d \in \{0, 1, \ldots, D-1\}, \tag{III.76}$$

où l'opérateur $.^{(d)}$ désigne la dérivée temporelle à l'ordre d. La relation précédente impose à deux polynômes successifs d'être égaux en leur point de rencontre, et de dérivées égales jusqu'à l'ordre $D-1$. On comprend alors que ce système de contraintes entre polynômes voisins consolide l'ensemble du modèle en réduisant les degrés de liberté sur les

coefficients $b^{\alpha}_{d,m}$. Prenons un exemple au point $n_{l_p^m}$ pour $D = 3$, l'équation (III.76) donne les 3 relations suivantes :

$$\begin{aligned} b^{\alpha}_{0,m} + b^{\alpha}_{1,m} n_{l_p^m} + b^{\alpha}_{2,m} n^2_{l_p^m} + b^{\alpha}_{3,m} n^3_{l_p^m} &= b^{\alpha}_{0,m+1} + b^{\alpha}_{1,m+1} n_{l_p^m} + b^{\alpha}_{2,m+1} n^2_{l_p^m} + b^{\alpha}_{3,m+1} n^3_{l_p^m} \\ b^{\alpha}_{1,m} + 2b^{\alpha}_{2,m} n_{l_p^m} + 3b^{\alpha}_{3,m} n^2_{l_p^m} &= b^{\alpha}_{1,m+1} + 2b^{\alpha}_{2,m+1} n_{l_p^m} + 3b^{\alpha}_{3,m+1} n^2_{l_p^m} \\ 2b^{\alpha}_{2,m} + 6b^{\alpha}_{3,m} n_{l_p^m} &= 2b^{\alpha}_{2,m+1} + 6b^{\alpha}_{3,m+1} n_{l_p^m}. \end{aligned} \tag{III.77}$$

De façon très similaire aux splines [Unser 99], il n'y a plus qu'un seul degré de liberté par segment de temps $[n_{1+l_p^m}, n_{l_p^{m+1}}]$. Cependant, la méthode est légèrement différente pour les splines car elles définissent des points de contrôle en chaque instant de temps, contrairement à notre approche qui choisit des points espacés de l_p échantillons temporels. De fait, les splines sont en mesure de suivre des discontinuités, ce que nous avons décidé de ne pas faire étant donné le contexte. Enfin, rappelons que nous nous trouvons dans un problème de régression où plusieurs mesures bruitées servent à estimer une inconnue, l'équivalent des splines dans ce cas est nommé splines de lissage [Wahba 90]. Ce modèle ajoute une contrainte de régularisation avec une variable scalaire de balance entre les données et la fonction régularisante qu'il faut estimer, ce que nous souhaitons éviter. La solution que nous adoptons ici reste très classique en traitement du signal.
Nous pouvons à nouveau estimer une courbe de régression sur les mesures de la figure III.24 à l'aide de notre nouveau modèle. Le résultat est visible en figure III.25 pour lequel les paramètres de régression sont $D = 3$, $l_p = 10$ et $M = 40$; ce dernier est déduit de la longueur du signal et de la taille de fenêtre. Selon nos souhaits, la courbe rouge suit très bien les fluctuations des mesures, on perçoit ainsi les bénéfices d'une estimation polynômiale locale.
Le vecteur des coefficients polynômiaux $\mathbf{b}^{\alpha}$ suffit à décrire entièrement notre signal, c'est

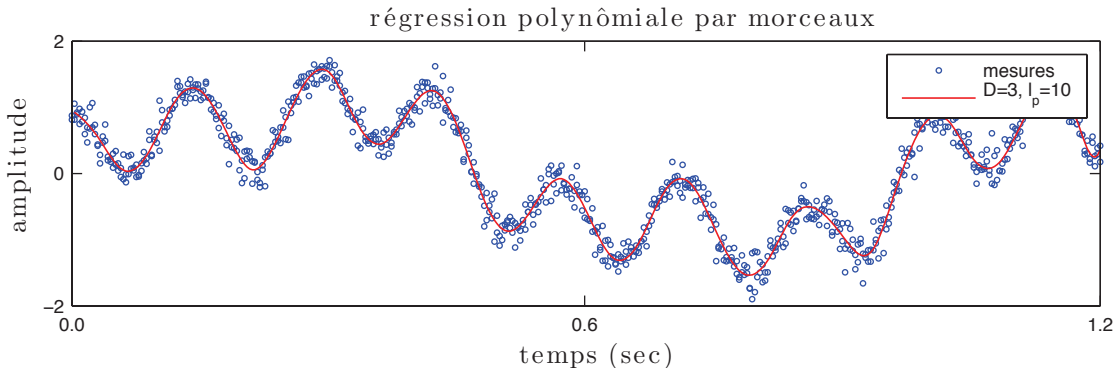


FIGURE III.25 – Régression sur le jeu de données de la figure III.24 à l'aide du modèle défini en (III.73), et avec les contraintes (III.76). Le résultat est représenté par la courbe rouge avec les paramètres suivants : $D = 3$, $l_p = 10$ et $M = 40$. Les oscillations hautes fréquences, qui ont une période de l'ordre de 50 échantillons de temps, sont très bien suivies par le modèle.

un vecteur de taille $(M(D+1) \times 1)$ tel que :

$$\mathbf{b}^{\alpha} = [b^{\alpha}_{0,0} \ \dots \ b^{\alpha}_{D,0} \ \ b^{\alpha}_{0,1} \ \dots \ b^{\alpha}_{D,1} \ \dots \ b^{\alpha}_{0,M-1} \ \dots \ b^{\alpha}_{D,M-1}]^T. \tag{III.78}$$

Le système d'équations vu en (III.77) montre que les contraintes définies en (III.76) sont linéaires par rapport aux coefficients, et peuvent s'exprimer sous forme matricielle comme

ci-dessous :

$$C^{\alpha}\mathbf{b}^{\alpha} = \mathbf{0}, \tag{III.79}$$

où C^{α} est une matrice creuse de taille $((M-1)D \times M(D+1))$; nous reportons le lecteur à l'annexe A pour la construction simple mais longue de cette matrice. Remarquons ici que C^{α} est identique pour les trois angles, elle a pour indice α en tant que sous matrice de taille correspondant à $\mathbf{b}^{\alpha}$. De même, le modèle (III.73) peut s'écrire comme le produit d'une matrice et d'un vecteur :

$$\theta^{\alpha} = W^{\alpha}\mathbf{b}^{\alpha}, \tag{III.80}$$

la matrice W^{α} intègre les fonctions w_m de (III.74), elle est creuse et de taille $(N \times M(D+1))$; une nouvelle fois nous invitons le lecteur à se porter à l'annexe A pour observer sa forme.

III.8.2 Expression de la vraisemblance

Au cours des paragraphes précédents, nous avons considéré l'inférence directe de θ en connaissance des données images. Le modèle polynômial par morceaux que nous venons de décrire peut caractériser l'attitude seulement à partir des coefficients $\mathbf{b}$; ce dernier est un vecteur $(3M(D+1) \times 1)$ tel que $\mathbf{b} = [\mathbf{b}^{lT} \;\; \mathbf{b}^{rT} \;\; \mathbf{b}^{tT}]^T$. Nous allons donc réécrire le problème d'estimation en fonction de ce nouveau vecteur, en connaissance des ensembles de contraintes $\mathcal{C}$ et d'images $\mathcal{I}$.
Définir un *a priori* sur les coefficients polynômiaux peut être hasardeux ; leurs valeurs varient au cours du temps, et ce d'autant plus que le degré auquel ils sont associés dans le polynôme augmente. Plutôt que de choisir un *a priori* gaussien d'écart-type très large - qu'il faut fixer par ailleurs - et qui n'enrichira pas l'estimateur, l'absence d'information peut être modéliser par une loi uniforme sur $\mathbb{R}$ de chaque coefficient. Ce terme de probabilité est non intégrable, mais il peut être vu comme une constante dans l'estimateur et nous amène à estimer $\mathbf{b}$ par maximum de vraisemblance :

$$\hat{\mathbf{b}}_{MV} = \underset{\mathbf{b}}{\operatorname{argmax}}\; p(\mathcal{I}|\mathbf{b}, C\mathbf{b} = \mathbf{0}), \tag{III.81}$$

où les contraintes sont exprimées sous forme matricielle, à la manière de l'équation (III.79). C est la matrice creuse de contraintes complète, elle est de taille $(3(M-1)D \times 3M(D+1))$ et a les sous-matrices C^l, C^r et C^t sur sa diagonale :

$$C = \begin{bmatrix} C^l & 0 & 0 \\ 0 & C^r & 0 \\ 0 & 0 & C^t \end{bmatrix}. \tag{III.82}$$

III.8.3 Estimateur du MV

Nous avons de nouveau deux expressions possibles pour le MV, celles-ci étant liées aux deux termes de vraisemblance que nous avons introduits au paragraphe III.4.4 (p.37).

Premier cas

D'après les relations (III.32) et (III.80), on déduit que :

$$p(\mathcal{I}|\mathbf{b}, C\mathbf{b} = \mathbf{0}) = \prod_{i,j;i\neq j} \frac{1}{(2\pi\sigma_I^2)^{NL/2}} \exp\left(-\frac{1}{2\sigma_I^2}\left|\mathbf{e}_{ij} + (J_i W - J_j D_{ij} W)\mathbf{b}\right|^2\right),$$

$$\text{tel que } C\mathbf{b} = \mathbf{0}, \quad \text{(III.83)}$$

W étant la matrice creuse de taille $(3N \times 3(M-1)D)$ qui a les sous-matrices W^l, W^r et W^t sur sa diagonale, de la même façon que C en (III.82). En s'appuyant sur le développement de la log-vraisemblance dans le premier cas, équation (III.34), on obtient :

$$\hat{\mathbf{b}}_{MV} = \underset{\mathbf{b}}{\text{argmax}} \; - \sum_{i,j;i\neq j} |\mathbf{e}_{ij} + (J_i W - J_j D_{ij} W)\mathbf{b}|^2, \text{ tel que } C\mathbf{b} = \mathbf{0}. \quad \text{(III.84)}$$

Second cas

De la même manière, d'après les relations (III.37) et (III.80), on voit que :

$$p(\mathcal{I}|\mathbf{b}, C\mathbf{b} = \mathbf{0}) = \prod_{i,j;i\neq j} \frac{1}{(2\pi\sigma_I^2)^{NL/2}} \exp\left(-\frac{1}{2\sigma_I^2}\left|\mathbf{e}_{ij} - J_j K_{ij} W\mathbf{b}\right|^2\right),$$

$$\text{tel que } C\mathbf{b} = \mathbf{0}. \quad \text{(III.85)}$$

A partir de la log-vraisemblance dans le second cas en équation (III.38), on déduit une nouvelle fois que :

$$\hat{\mathbf{b}}_{MV} = \underset{\mathbf{b}}{\text{argmax}} \; - \sum_{i,j;i\neq j} |\mathbf{e}_{ij} - J_j K_{ij} W\mathbf{b}|^2, \text{ tel que } C\mathbf{b} = \mathbf{0}. \quad \text{(III.86)}$$

Nous avons dans les deux cas un problème de minimisation aux plus grands carrés linéaire avec contrainte d'égalité linéaire, qui peut se résoudre entre autres par factorisation QR, ou à l'aide des multiplicateurs de Lagrange.

III.8.4 Résultats sur les données

Nous avons traité une nouvelle fois le jeu de données présenté au paragraphe III.4.7 à l'aide des deux estimateurs présentés en (III.84) et (III.86). La première figure III.26 montre le résultat associé au premier estimateur, pour lequel nous avons fixé $D = 3$ et $l_p = 10$. La courbe rouge rivalise très clairement avec la meilleur estimation obtenue avec *a priori* autorégressif en figure III.22. Les défauts observés dans le résidu sont les mêmes, bien que celui de la figure III.26 soit un peu plus lisse.
La seconde figure III.27 correspond aux résultats de l'estimateur (III.86), pour les mêmes choix de paramètre. Ils sont très bons en comparaison au roulis estimé de la figure III.23, et les observations peuvent rejoindre celles de la figure précédente.

III.8.5 Discussion

Nous avons cherché dans cette partie à développer une méthode qui s'affranchisse d'hyperparamètres à estimer, en référence à la précédente discussion du paragraphe III.7.

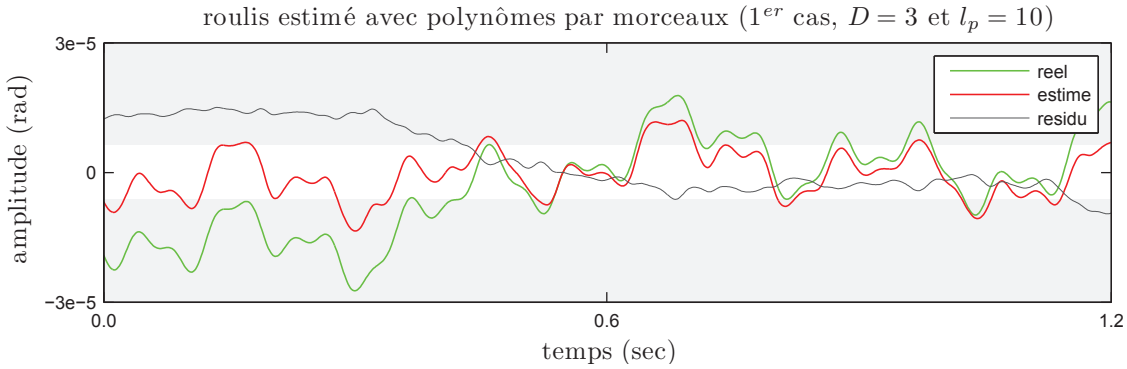


FIGURE III.26 – Exemple de roulis estimé par MV, équation (III.84), avec un modèle polynômial par morceaux tel qu'il est défini en (III.73) avec pour paramètres $\mathrm{D} = 3$ et $\mathrm{l_p} = 10$. La courbe rouge est comparable au meilleur résultat obtenu en figure III.22 avec un modèle autorégressif comme *a priori*.

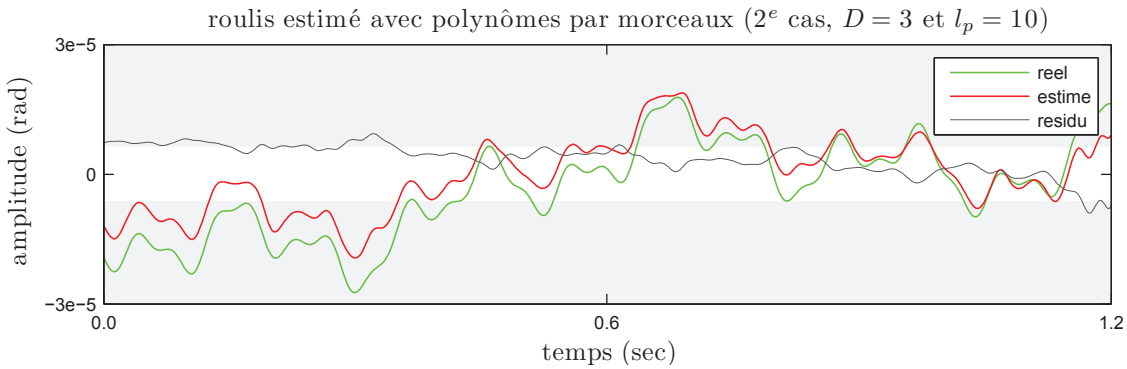


FIGURE III.27 – Exemple de roulis estimé par MV selon la relation (III.86), les paramètres du modèle polynomial par morceaux sont $\mathrm{D} = 3$ et $\mathrm{l_p} = 10$. Le résidu obtenu en gris est sensiblement le même que celui obtenu dans le meilleur des cas, figure III.23.

En réalité, il reste toujours 2 paramètres à déterminer : D et l_p. Notons que nous choisissons d'attribuer un même pas de fenêtre l_p pour tous les angles, car étant liés au même phénomène physique, il est fort probable qu'ils partagent également une même longueur d'onde. On peut remarquer que ces deux paramètres sont très liés, comme le montre la figure III.28 : un petit degré implique une petite taille de fenêtre, et inversement.
Les courbes de régression en rouge sont sensiblement les mêmes, et le nombre de coeffi-

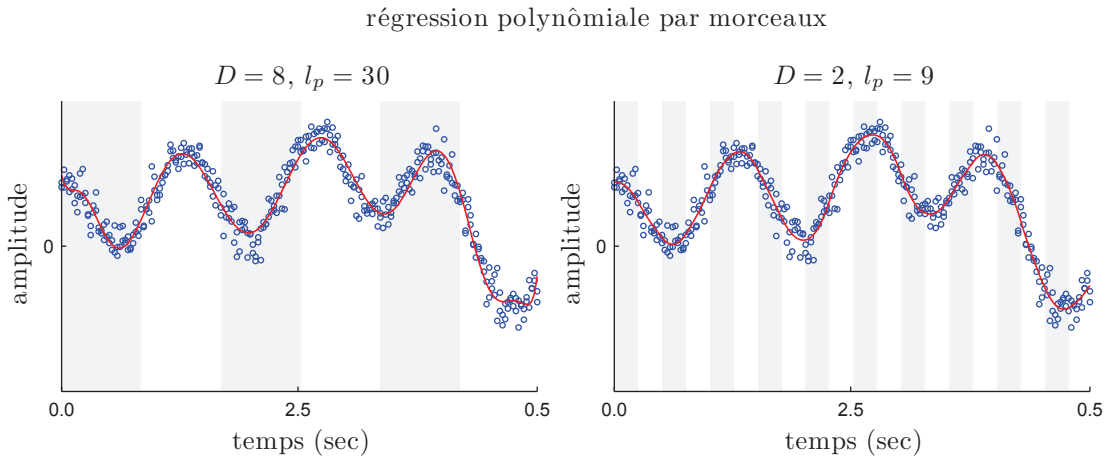


FIGURE III.28 – Exemple de régression polynômiale par morceaux dans 2 configurations différentes : (D = 8, l_p = 30) pour la courbe rouge à gauche et (D = 2, l_p = 9) pour celle de droite. Les résultats sont sensiblement identiques, ainsi que le nombre de coefficients polynômiaux estimés (54 à gauche contre 60 à droite). Les stries grises et blanches permettent d'apprécier dans chacun des cas la taille et la position des fenêtres.

cients estimés dans les deux cas est pratiquement identique : 54 pour la courbe de gauche contre 60 pour la courbe de droite. Il ne doit pas exister de relation linéaire entre D et l_p, car le polynôme est de plus en plus contraint par ses dérivées lorsque son degré augmente, ce que nous tenterons d'observer au chapitre suivant. En pratique, on constate aussi que le calcul numérique devient instable lorsque le degré du polynôme est trop grand ($D > 10$). Le choix d'une taille de fenêtre trop grande peut également nous ramener au cas de la figure III.24 où le polynôme ne suit plus les mesures malgré son ordre. On privilégiera donc plutôt des petites fenêtres avec un polynôme de degré inférieur à 5 comme paramètres d'estimation.
Si nous omettons D en le fixant à 3, comme il est couramment fait pour les splines selon un rapport entre temps de calcul et qualité d'estimation [Unser 99], il ne reste plus qu'une variable à choisir. Contrairement aux hyperparamètres vus le long de la section III.5, l_p est défini sur un domaine discret, ce qui peut faciliter la recherche d'un optimal. D'autres éléments doivent entrer en compte comme la criticité du choix d'un paramètre, ou autrement dit le risque encouru par un choix non optimal. Nous développerons ces points au cours du chapitre suivant.

III.9 Fusion de données

Nous avons vu au cours du chapitre précédent que le satellite embarquait plusieurs capteurs de positionnement pour corriger ses dérives spatiales. Ils peuvent également

servir en asservissement dans une boucle de contrôle, pour vérifier que l'engin spatial exécute une commande qui lui a été soumise. C'est autour de ces capteurs que la majorité des solutions à notre problème ont été proposées dans la littérature [Cramer 00, Poli 02], malgré les observations que nous avons faites sur leur faible résolution temporelle.
Nous nous intéressons dans cette partie au *star tracker*, qui nous donne une mesure absolue sous-échantillonnée de l'attitude. Nous rappelons que le faible échantillonnage de ce capteur est lié aux limites technologiques des appareils embarqués dans l'environnement spatial, ainsi qu'à son coût que l'on souhaite contrôler. C'est une caméra de type *pinhole* qui observe les étoiles. Tout mouvement de rotation du satellite entraine un déplacement des étoiles dans les images du *star tracker*. Ce dernier cherche à chaque pas de temps à recaler ses observations avec des atlas d'étoiles connus à l'avance pour donner une information sur l'attitude. On peut le préférer au gyroscope qui fournit une mesure relative sur le positionnement, le recalage multi-image qui a perdu sa référence bénéficiera certainement plus d'une information absolue dans un contexte de fusion. Cependant, il reste possible de combiner les deux dans le modèle probabiliste adopté, dans le cas où des mesures de gyroscope sont disponibles.

III.9.1 Extension du modèle probabiliste

Replaçons-nous dans le contexte du paragraphe III.3, où en plus des images $\mathcal{I}$, s'ajoutent les mesures *star tracker* $\mathbf{s}$. L'objectif est maintenant d'ajouter cette information supplémentaire dans le schéma d'inférence. On peut développer la loi *a posteriori* de la façon suivante :

$$\begin{aligned} p(\theta|\mathcal{I},\mathbf{s}) &= \frac{p(\mathcal{I}|\theta,\mathbf{s})p(\mathbf{s}|\theta)p(\theta)}{p(\mathbf{s})p(\mathcal{I}|\mathbf{s})} \\ &= \frac{p(\mathcal{I}|\theta)p(\mathbf{s}|\theta)p(\theta)}{p(\mathbf{s})p(\mathcal{I})}, \end{aligned} \tag{III.87}$$

où nous faisons l'hypothèse d'indépendance entre $\mathcal{I}$ et $\mathbf{s}$, étant mesurées par deux capteurs différents, on a donc $p(\mathcal{I}|\mathbf{s}) = p(\mathcal{I})$ et $p(\mathcal{I}|\theta,\mathbf{s}) = p(\mathcal{I}|\theta)$. Dans cette expression, on retrouve des éléments déjà vus auparavant : la vraisemblance des images $p(\mathcal{I}|\theta)$ ainsi que sa version marginalisée $p(\mathcal{I})$, et l'*a priori* sur les variations d'attitude $p(\theta)$. Les nouveaux termes sont : la vraisemblance des mesures *star tracker* $p(\mathbf{s}|\theta)$, et son intégration sur l'ensemble des variations d'attitude possibles $p(\mathbf{s})$.
Intéressons-nous aux termes qui ne dépendent que de θ, la loi *a posteriori* devient proportionnelle à :

$$p(\theta|\mathcal{I},\mathbf{s}) \propto p(\mathcal{I}|\theta)p(\mathbf{s}|\theta)p(\theta). \tag{III.88}$$

L'estimateur du MAP cherche à maximiser cette probabilité, nous avons :

$$\hat{\theta}_{MAP} = \underset{\theta}{\operatorname{argmax}}\ p(\theta|\mathcal{I},\mathbf{s}). \tag{III.89}$$

Dans cette équation, $p(\mathbf{s}|\theta)$ est la contribution du *star tracker* au processus d'estimation, dont nous allons détailler la forme par la suite.

III.9.2 Bruit du capteur et vraisemblance

Le *star tracker* est une mesure image dont le bruit n'est pas simple à modéliser, nous en avons discuté au début du chapitre au paragraphe III.4.1. A cela viennent s'ajouter plusieurs difficultés : étant donné le contexte spatial, le capteur image ne peut pas être aussi

bon que ceux dont nous disposons sur terre car il se doit d'être avant tout robuste, enfin l'environnement dans lequel il évolue est très complexe à caractériser. Les lentilles de la caméra peuvent se contracter ou se dilater suivant la température et le rythme diurne du satellite. Au sein des images, plusieurs sources lumineuses peuvent contaminer la mesure : réflexion du soleil, comètes, supernova, rayons cosmiques, etc. On peut aussi constater une dérive linéaire du capteur sur des très longues périodes d'observation.
La figure III.29 montre un exemple d'enregistrement du *star tracker* pour le jeu de données que nous traitons. Le rapport entre les périodes d'échantillonnage est de l'ordre de 100 ; entre deux mesures *star tracker*, une centaine de lignes ont été acquises par chaque caméra *pushbroom*.
En tant que mesure absolue de l'attitude, l'équation d'acquisition du *star tracker* au cours

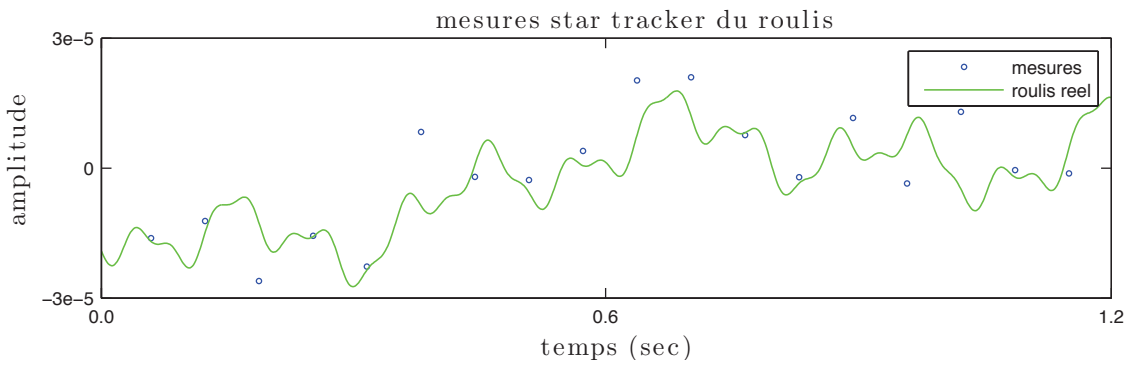


FIGURE III.29 – Représentation de l'enregistrement du *star tracker* en points bleus, et comparaison au roulis réel en vert. Les mesures sont sous échantillonnées d'un rapport de 100 en comparaison des caméras *pushbroom*.

du temps peut s'écrire comme suit :

$$\mathbf{s}(n_s) = \boldsymbol{\theta}(n_s) + \boldsymbol{\epsilon}_s(n_s), \tag{III.90}$$

où n_s sont les instants de temps tels que $n_s = f_s n$; $f_s \in \mathbb{R}^+$ est un facteur de sous-échantillonnage associé au *star tracker*. Le vecteur $\mathbf{s}(n_s)$ de taille (3×1) à l'instant n_s se construit comme le vecteur d'attitude en équation (III.12) : $\mathbf{s}(n_s) = [s^l(n_s)\ s^r(n_s)\ s^t(n_s)]^T$.
Nous avons dit que ce capteur pouvait présenter une dérive linéaire sur des très longues périodes de temps, à l'échelle de quelques secondes nous ferons l'hypothèse d'un bruit blanc sur les mesures du *star tracker* :

$$\boldsymbol{\epsilon}_s(n_s) \sim \mathcal{N}(\mathbf{0}, \sigma_s^2 I_d). \tag{III.91}$$

Nous supposerons ici que le bruit est indépendant et identiquement distribué suivant le temps, nous pouvons réécrire la relation (III.90) sous forme vectorielle :

$$\mathbf{s} = F_s \boldsymbol{\theta} + \boldsymbol{\epsilon}_s, \tag{III.92}$$

$\mathbf{s}$ est le vecteur de taille $(3N_s \times 1)$ cumulant les mesures du lacet, du roulis et du tangage, tel que :

$$\mathbf{s} = [s^l(n_{s_1})\ \dots\ s^l(n_{s_{N_s}})\ s^r(n_{s_1})\ \dots\ s^r(n_{s_{N_s}})\ s^t(n_{s_1})\ \dots\ s^t(n_{s_{N_s}})]^T, \tag{III.93}$$

F_s est la matrice de sous-échantillonnage $(3N_s \times 3N)$ qui extrait du vecteur d'attitude les échantillons correspondants aux instants d'enregistrement du *star tracker*. Des équations (III.91) et (III.92), on peut déduire assez simplement la vraisemblance recherchée :

$$p(\mathbf{s}|\theta) = \frac{1}{(2\pi\sigma_s^2)^{3N_s/2}} \exp\Big(-\frac{1}{2\sigma_s^2}|\mathbf{s} - F_s\theta|^2\Big). \tag{III.94}$$

Nous développerons une nouvelle fois les expressions suivantes en ne considérant qu'un seul angle dans l'attitude. Ainsi en suivant les notations précédentes, $\mathbf{s}^\alpha$ correspond au vecteur $(N_s \times 1)$ de mesures du *star tracker* extrait de $\mathbf{s}$ pour un angle, et F_s^α est la sous-matrice de F_s et de taille $(N_s \times N)$.

III.9.3 Expression du MAP

La loi *a posteriori* se calcule à partir des équations (III.94) et (III.88). Nous pouvons décliner son expression pour toutes les combinaisons d'*a priori* et cas de vraisemblance que nous avons vus. Nous choisissons de développer par la suite seulement certains exemples intéressants pour ne pas alourdir la lecture du document ; les calculs sont sensiblement les mêmes entre chaque solution.
Considérons dans un premier temps un *a priori* sur l'amplitude de l'attitude, tel que nous l'avons déjà vu au paragraphe III.5.2 (p.46). Nous obtenons pour la première vraisemblance des images :

$$\begin{aligned} p(\theta^\alpha|\mathcal{I}, \mathbf{s}^\alpha) \propto \prod_{i,j;i\neq j} \exp\Big(-\frac{1}{2\sigma_I^2}(\mathbf{e}_{ij} + (J_i^\alpha - J_j^\alpha D_{ij}^\alpha)\theta^\alpha)^T(\mathbf{e}_{ij} + (J_i^\alpha - J_j^\alpha D_{ij}^\alpha)\theta^\alpha)\Big) \\ \exp\Big(-\frac{1}{2\sigma_s^2}(\mathbf{s}^\alpha - F_s^\alpha\theta^\alpha)^T(\mathbf{s}^\alpha - F_s^\alpha\theta^\alpha)\Big) \exp\Big(-\frac{1}{2\sigma_{\theta^\alpha}^2}\theta^{\alpha T}\theta^\alpha\Big). \end{aligned} \tag{III.95}$$

Après quelques développements, on trouve une forme très ressemblante à l'équation (III.53) avec le résultat suivant :

$$p(\theta^\alpha|\mathcal{I}) \propto \mathcal{N}\Big(\bar{\theta}^\alpha, A^{\alpha-1}\Big), \tag{III.96}$$

où :

$$A^\alpha = \sigma_I^{-2} \sum_{i,j;i\neq j} (J_i^\alpha - J_j^\alpha D_{ij}^\alpha)^T(J_i^\alpha - J_j^\alpha D_{ij}^\alpha) + \sigma_s^{-2} F_s^{\alpha T} F_s^\alpha + \sigma_{\theta^\alpha}^{-2} I_d, \tag{III.97}$$

et :

$$\bar{\theta}^\alpha = -A^{\alpha-1}(\sigma_{\theta^\alpha}^{-2} \sum_{i,j;i\neq j} (J_i^\alpha - J_j^\alpha D_{ij}^\alpha)^T \mathbf{e}_{ij} + \sigma_s^{-2} F_s^{\alpha T} \mathbf{s}^\alpha). \tag{III.98}$$

L'expression pour la seconde vraisemblance des images est encore sensiblement la même, la matrice de variance et la moyenne changent comme suit :

$$A^\alpha = \sigma_I^{-2} \sum_{i,j;i\neq j} K_{ij}^{\alpha T} J_j^{\alpha T} J_j^\alpha K_{ij}^\alpha + \sigma_s^{-2} F_s^{\alpha T} F_s^\alpha + \sigma_{\theta^\alpha}^{-2} I_d, \tag{III.99}$$

$$\bar{\theta}^\alpha = A^{\alpha-1}(\sigma_{\theta^\alpha}^{-2} \sum_{i,j;i\neq j} K_{ij}^{\alpha T} J_j^{\alpha T} \mathbf{e}_{ij} + \sigma_s^{-2} F_s^{\alpha T} \mathbf{s}^\alpha). \tag{III.100}$$

Le mode de la distribution correspond au MAP et est encore donné par $\bar{\theta}^\alpha$ ici.

III.9.4 Résultats sur les données

Nous avons précisé que le *star tracker* fournit une mesure sous-échantillonnée de l'attitude, il ne pourra donc apporter qu'une information basse fréquence au processus d'estimation. Etant donné que le premier développement de vraisemblance d'image perd une partie basse fréquence dans le signal estimé, comme en témoigne entre autres la figure III.5 (p.42), ce cas semble très propice à la fusion. Nous nous proposons donc d'évaluer l'approche sur le jeu de données à partir de l'expression (III.98) uniquement, sachant que les résultats seront similaires dans le second cas mais légèrement moins représentatifs de l'intérêt du *star tracker* vu qu'une bonne partie des basses fréquences sont déjà estimées par ce modèle.
La figure III.30 présente le résultat de fusion obtenu à partir de l'équation (III.98). Rappelons que l'*a priori* sur l'amplitude donne des résultats très similaires au maximum de vraisemblance, on retrouve donc un roulis estimé particulièrement bruité avec des parasites hautes fréquences.
Les pics rouges estimés sur les mesures du *star tracker* sont particulièrement intéressants.

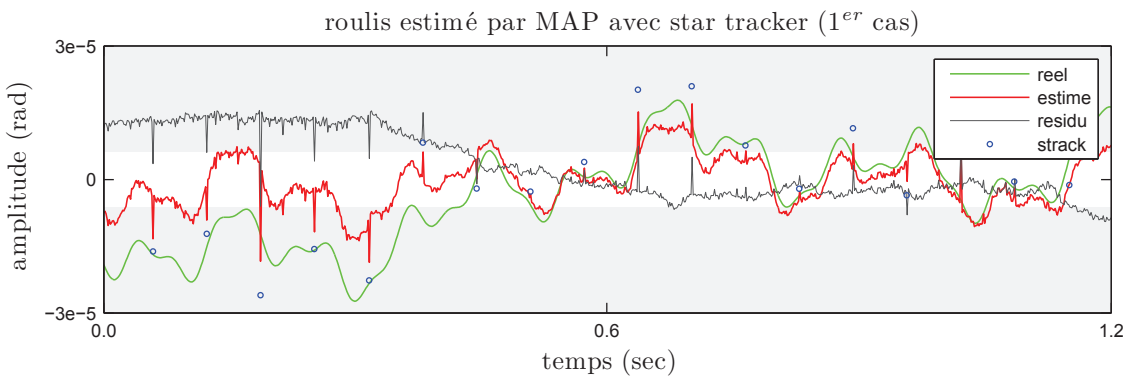


FIGURE III.30 – Exemple d'estimation de roulis par MAP avec *a priori* sur l'amplitude et données *star tracker* (représentées en rond bleu) ; $\sigma_I = 0.012$, $\sigma_{\theta^r} = 9.2 \times 10^{-7}$ et $\sigma_s = 4 \times 10^{-8}$. Le résultat est globalement le même que celui de la figure III.10, les mesures du *star tracker* contribuent ponctuellement en créant les pics rouges dans le roulis estimé. A ces endroits, l'estimation intègre les images, l'*a priori* et le *star tracker*. En dehors de ces échantillons, faute de modèle temporel, le *star tracker* ne contribue pas à l'inférence.

Sans cohérence temporelle, l'estimation du roulis est ponctuelle pour chaque instant de temps. Le *star tracker* informe sur les pas de temps où il a effectué une mesure, sans que cette information se propage aux estimations voisines.
Il est obligatoire d'ajouter une information temporelle sur les variations d'attitude pour que le *star tracker* soit plus utile dans le schéma d'inférence. Nous avons déjà vu plusieurs solutions au cours des paragraphes précédents, et le modèle autorégressif s'est révélé particulièrement efficace dans cette tâche. L'intégration de cet *a priori* est très simple à partir des équations que nous avons déjà obtenues en (III.97) et (III.98), seule la matrice de variance change et nous pouvons la réécrire comme suit :

$$A^{\alpha} = \sigma_I^{-2} \sum_{i,j;i\neq j} K_{ij}^{\alpha\,T} J_j^{\alpha\,T} J_j^{\alpha} K_{ij}^{\alpha} + \sigma_s^{-2} F_s^{\alpha\,T} F_s^{\alpha} + \sigma_{\theta^{\alpha}}^{-2} R^{\alpha\,T} R^{\alpha}. \qquad \text{(III.101)}$$

La figure III.31 montre le nouveau résultat de fusion avec un modèle temporel. La différence avec le résultat précédent est assez nette : les qualités de filtrage de l'estimateur observées en figure III.22 (p.60) apparaissent à nouveau, et l'information basse fréquence portée par le *star tracker* est réellement présente. L'oscillation basse fréquence dans le résidu, qui accompagnait le premier cas de vraisemblance des images, a en partie disparu ici.
Cependant, on peut encore noter la présence de discontinuités assez franches sur chaque

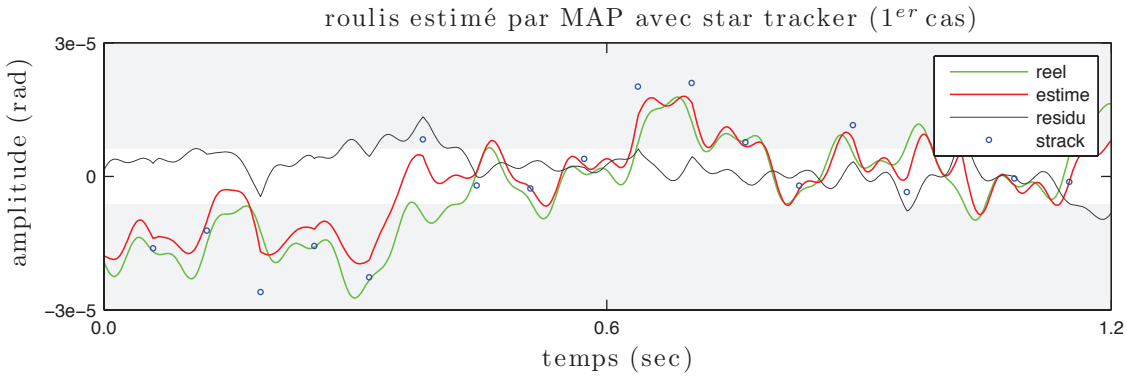


FIGURE III.31 – Exemple d'estimation par MAP avec fusion des images, du *star tracker* et de l'*a priori* autorégressif ; $\sigma_I = 0.012$, $\sigma_{\theta^r} = 8 \times 10^{-9}$ et $\sigma_s = 1.5 \times 10^{-8}$. L'influence du *star tracker* est très positive car une partie des basses fréquences est restituée dans le signal estimé. Le modèle autorégressif est toujours efficace pour le suivi des hautes fréquences. En revanche, des pics dans le résidu persistent autour des mesures bleues.

échantillon du *star tracker*, ce que nous souhaitons éviter dans le cas d'un algorithme d'estimation itératif ; les pics sont plus particulièrement visibles sur le résidu. Ces mesures ponctuelles demeurent relativement bruitées en haute fréquence, alors même que le *star tracker*, de part sa fréquence d'échantillonnage, ne devrait influencer que les très basses fréquences. Afin de forcer cette idée, nous allons utiliser un modèle temporel pour le *star tracker*.

III.9.5 Modèle polynômial basse fréquence

De façon similaire à l'introduction du paragraphe III.8.1, nous proposons d'utiliser un modèle polynomial pour expliquer les enregistrements du *star tracker*. Cela suppose qu'à l'échelle d'acquisition d'une image, l'information basse fréquence que peut fournir le *star tracker* à l'estimation des variations d'attitude est contenue dans un polynôme de degré suffisamment bas pour ne pas tomber dans les dérives observées en figure III.24 (p.64), où un ordre trop important provoque des instabilités dans le modèle estimé.
Pour un des angles de l'attitude, cela revient à écrire que :

$$s^{\alpha}(n_s) = \sum_{d=0}^{D_{s^\alpha}} c_d^{\alpha} n_s^d + \epsilon_{s^\alpha}(n_s), \tag{III.102}$$

où D_{s^α} est le degré du polynôme, c_d^{α} pour $d \in \{0, 1, \ldots, D_{s^\alpha}\}$ ses coefficients, et $\epsilon_{s^\alpha}(n_s)$ un bruit blanc de mesure centré et de variance $\sigma^2_{s^\alpha}$. L'estimation de ces coefficients est

assez simple, dès lors qu'on remarque que l'expression précédente est équivalente à :

$$\mathbf{s}^{\alpha} = W_s^{\alpha}\mathbf{c}^{\alpha} + \epsilon_{s^{\alpha}}, \tag{III.103}$$

avec W_s^{α} la matrice de taille $(N_s \times D_{s^{\alpha}} + 1)$ qui combine linéairement les coefficients polynomiaux du vecteur $\mathbf{c}^{\alpha}$ de taille $(D_{s^{\alpha}} + 1 \times 1)$ et tel que $\mathbf{c}^{\alpha} = [c_0^{\alpha} \ \dots \ c_{D_{s^{\alpha}}}^{\alpha}]^T$. La vraisemblance de ces coefficients en connaissance de $\mathbf{s}^{\alpha}$ se déduit aisément :

$$p(\mathbf{c}^{\alpha}|\mathbf{s}^{\alpha}) = \frac{1}{(2\pi\sigma_{s^{\alpha}}^2)^{N_s/2}} \exp\Big(-\frac{1}{2\sigma_{s^{\alpha}}^2}|\mathbf{s}^{\alpha} - W_s^{\alpha}\mathbf{c}^{\alpha}|^2\Big). \tag{III.104}$$

L'estimateur du MV nous donne dans ce cas-ci une solution analytique simple au problème :

$$\hat{\mathbf{c}^{\alpha}}_{MV} = (W_s^{\alpha T} W_s^{\alpha})^{-1} W_s^{\alpha T}\mathbf{s}^{\alpha} \tag{III.105}$$

Pour tenir compte de ce modèle dans la fusion, nous devons remplacer les mesures *star tracker* $\mathbf{s}$ par les coefficients polynomiaux $\mathbf{c}$ que nous venons de calculer dans l'équation (III.88) ; $\mathbf{c} = [\mathbf{c}^{lT} \ \ \mathbf{c}^{rT} \ \ \mathbf{c}^{tT}]^T$ est le vecteur complet des coefficients de taille $(D_{s^l} + D_{s^r} + D_{s^t} \times 1)$. L'estimateur du MAP devient :

$$\hat{\theta}_{MAP} = \underset{\theta}{\operatorname{argmax}} \ p(\theta|\mathcal{I}, \mathbf{c}), \tag{III.106}$$

avec :

$$p(\theta|\mathcal{I}, \mathbf{c}) \propto p(\mathcal{I}|\theta)p(\mathbf{c}|\theta)p(\theta). \tag{III.107}$$

Dans cette équation, seul $p(\mathbf{c}|\theta)$ reste à développer. En rapport aux résultats que nous avons obtenus en figures III.30 et III.31, nous souhaitons que le *star tracker* influence cette fois-ci toutes les mesures de l'attitude. Le modèle polynômial permet d'accéder à des mesures en dehors de l'échantillonnage du *star tracker*, et peut être comparé à θ^{α} en supposant que :

$$B^{\alpha}\theta^{\alpha} - W^{\alpha}\mathbf{c}^{\alpha} \sim \mathcal{N}(\mathbf{0}, \sigma_{c^{\alpha}}^2 I_d), \tag{III.108}$$

où W^{α} est la matrice $(N \times D_{s^{\alpha}})$ qui calcule les valeurs du polynôme de coefficients $\mathbf{c}^{\alpha}$ aux instants de temps associés à θ^{α}. B^{α} est une matrice $(N \times N)$ de filtrage de l'attitude par un passe bas, elle permet de s'assurer que le modèle basse fréquence est bien comparé aux basses fréquences de θ^{α}. Sans cet opérateur, le résidu de l'équation (III.108) contiendrait les moyennes et hautes fréquences de l'attitude, et l'estimateur aurait tendance à réduire ces fréquences pour rendre le modèle juste. En pratique, $B^{\alpha}\theta^{\alpha}$ est une opération de convolution avec un filtre RIF (Réponse Impulsionnelle Finie) dont nous maîtrisons la fréquence de coupure ; la matrice B^{α} est détaillée en annexe A. Une représentation du spectre de ce filtre est visible en figure III.32 en comparaison avec le spectre du roulis réel.
La vraisemblance des coefficients en connaissance de l'attitude découle normalement de l'expression (III.108) :

$$p(\mathbf{c}^{\alpha}|\theta^{\alpha}) = \frac{1}{(2\pi\sigma_{c^{\alpha}}^2)^{N/2}} \exp\Big(-\frac{1}{2\sigma_{s^{\alpha}}^2}|B^{\alpha}\theta^{\alpha} - W^{\alpha}\mathbf{c}^{\alpha}|^2\Big). \tag{III.109}$$

Considérons une nouvelle fois la loi *a posteriori* avec un *a priori* autorégressif, l'ajout du terme de probabilité précédent amène aux résultats suivants :

$$p(\theta^{\alpha}|\mathcal{I}) \propto \mathcal{N}\Big(\bar{\theta}^{\alpha}, A^{\alpha-1}\Big), \tag{III.110}$$

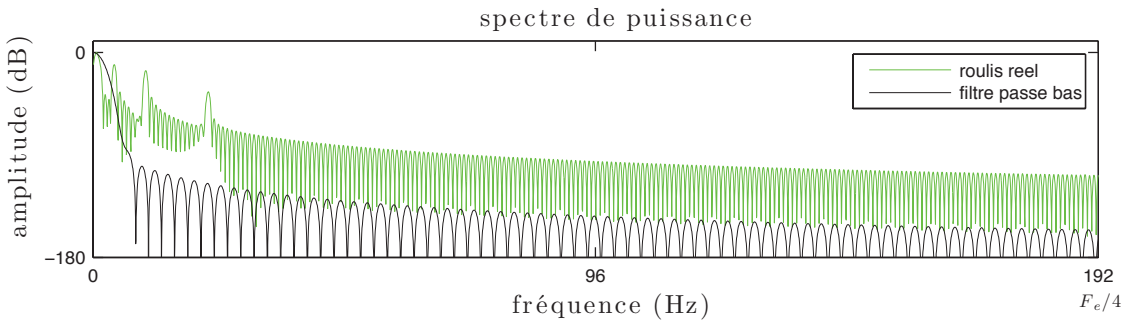


FIGURE III.32 – Représentation du spectre du filtre passe bas de l'équation (III.108) en noir et du roulis réel en vert. La fréquence de coupure à −3dB est de l'ordre de 1.5Hz, sachant que la fréquence d'échantillonnage du *star tracker* est de 16Hz ; celui-ci restituera donc au mieux les fréquences inférieures à 8Hz. On remarque que le filtre ne conserve principalement que la plus basse fréquence du signal.

avec :

$$A^{\alpha} = \sigma_I^{-2} \sum_{i,j;i\neq j} K_{ij}^{\alpha\,T} J_j^{\alpha\,T} J_j^{\alpha} K_{ij}^{\alpha} + \sigma_{\theta^{\alpha}}^{-2} R^{\alpha\,T} R^{\alpha} + \sigma_{c^{\alpha}}^{-2} B^{\alpha\,T} B^{\alpha}, \tag{III.111}$$

et :

$$\bar{\theta}^{\alpha} = -A^{\alpha\,-1}(\sigma_{\theta^{\alpha}}^{-2} \sum_{i,j;i\neq j} K_{ij}^{\alpha\,T} J_j^{\alpha\,T} \mathbf{e}_{ij} + \sigma_{c^{\alpha}}^{-2} W^{\alpha} \mathbf{c}^{\alpha}). \tag{III.112}$$

Nous pouvons à nouveau traiter le jeu de données et observer le résultat en figure III.33. Contrairement à la solution précédente, le roulis estimé ne présente pas de discontinuités. Le modèle polynômial, inféré à partir des mesures du *star tracker* et représenté en magenta sur la figure, montre bien la présence d'une oscillation très basse fréquence dans le roulis réel. Cette information est intégrée dans la courbe rouge, ainsi que la présence de fréquences plus hautes au travers de l'*a priori* autorégressif.

On constate que le résidu change de forme, les basses fréquences sont nettement atténuées. Le résultat obtenu est sensiblement meilleur que ceux obtenus précédemment avec le premier cas de vraisemblance des images, et est un peu plus juste que les meilleurs résultats observés en figures III.23 et III.27.

III.10 Conclusion

Nous venons de présenter dans ce chapitre un cadre mathématique complet permettant d'estimer les variations d'attitude du satellite lors de l'acquisition des images. Il prend en compte la géométrie du plan focal et des caméras *pushbroom* pour effectuer un recalage multi-image. Il cherche à intégrer des informations que nous avons *a priori* sur les signaux à estimer, comme la propriété de stationnarité liée à leur origine physique. Enfin il propose de fusionner de façon cohérente des mesures basses fréquences obtenues par le *star tracker* aux données images. Il est intéressant de noter que la fusion permet de compléter deux problèmes inverses : celui du recalage d'images où les basses fréquences de l'attitude sont mal estimées, et celui du *star tracker* où les hautes fréquences sont manquantes.

Toutes les solutions proposées ici n'ont pas les mêmes performances et le même degré de

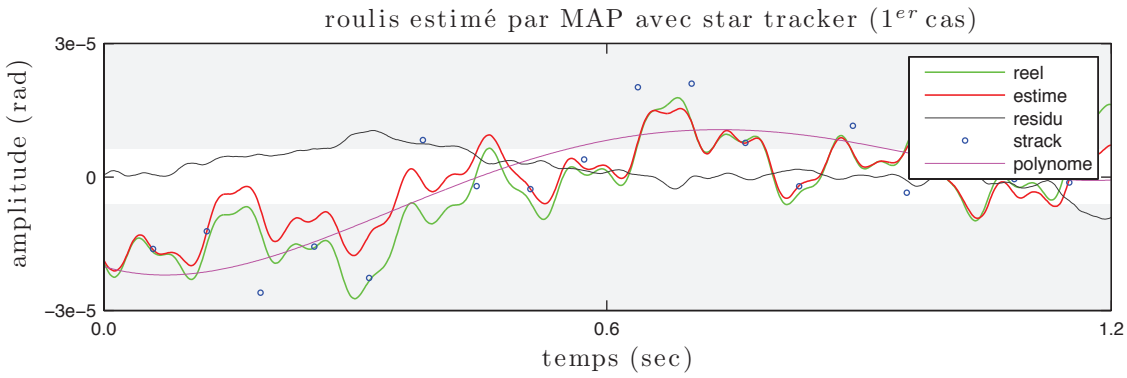


FIGURE III.33 – Exemple d'estimation du roulis par MAP suivant l'équation (III.109), avec un modèle polynômial de paramètres $D_{s^r} = 4$ et $\mathbf{c}^r = 10^{-3} \times [1.9 \quad -4.9 \quad 4.5 \quad -1.7 \quad 0.2]^T$ et le même *a priori* autorégressif que pour le résultat de la figure III.22 ; $\sigma_I = 0.012$, $\sigma_{\theta^r} = 2 \times 10^{-8}$ et $\sigma_{c^r} = 1.5 \times 10^{-7}$. La courbe rouge tente de suivre le modèle polynômial en magenta déduit à partir des mesures du *star tracker*. Les hautes fréquences sont encouragées par l'*a priori* autorégressif, le résidu montre cependant qu'une petite partie de celles-ci est filtrée par le modèle choisi.

simplicité. Elles impliquent pour la plupart l'estimation préalable de paramètres, ou hyperparamètres, pour assurer leur bon fonctionnement, étape que nous avons totalement négligée dans ce chapitre. Nous avons notamment vu que certains estimateurs très performants ont de nombreux paramètres. En pratique, l'estimation de ceux-ci est difficile, parfois même instable, et peut nécessiter un temps de calcul important. Un estimateur simple peut être préféré à une solution plus complexe pour ces raisons. Enfin nous avons comparé les différentes méthodes sur la base d'un unique jeu de données, les résultats observés sont donc trop restreints pour tirer des conclusions quant aux qualités des différents estimateurs.
Le chapitre suivant tente d'aborder le problème d'estimation de façon pragmatique ; différents algorithmes sont proposés et des résultats sur plusieurs jeux de données vont permettre de dégager les solutions les plus adaptées.

CHAPITRE IV

Algorithmes d'estimation et résultats

IV.1 Introduction

Nous avons construit au chapitre précédent le cadre formel de notre étude. Celui-ci explique comment les variations d'attitude peuvent être estimées à partir des images acquises par le satellite. Il justifie l'utilisation d'une connaissance *a priori* sur le signal recherché, et il propose de fusionner les mesures obtenues par un *star tracker* dans un schéma global d'inférence.
Les discussions précédentes ont mis en avant des points importants encore non résolus. Plusieurs hyperparamètres sont inconnus alors que ceux-ci mesurent l'importance et la forme que doivent prendre l'*a priori* vis-à-vis des données images. De même la linéarisation au premier ordre de l'image déformée limite l'estimation de déplacement à un pixel en amplitude, pourtant certaines oscillations du satellite peuvent engendrer des déformations plus grandes. Enfin nous avons supposé jusqu'à maintenant que les images étaient calibrées radiométriquement, alors qu'elles sont de différentes modalités pour recomposer des images en couleur.
Ce chapitre propose différentes solutions pour estimer les hyperparamètres : elles fonctionnent soit sur le principe de minimisation de l'erreur de prédiction, soit sur la minimisation de l'erreur photométrique. Ensuite, des méthodes rapides cherchant à simplifier la problématique de recalage multi-modal sont comparées sur la base de notre problème. Enfin l'ensemble des estimateurs est rappelé dans un tableau, l'algorithme itératif global permettant de trouver l'attitude est exposé, puis toutes les méthodes de notre étude sont évaluées et comparées sur des jeux de données satellitaires.

IV.2 Estimation des hyperparamètres

Nous avons vu jusqu'à maintenant qu'en dehors du maximum de vraisemblance, tous les estimateurs ont des hyperparamètres qui ont été supposés connus. Soit $\boldsymbol{\lambda}$ le vecteur qui regroupe tous les hyperparamètres d'un estimateur, le problème que nous avons intuitivement résolu est le suivant :

$$\hat{\theta}_{MAP} = \underset{\theta}{\operatorname{argmax}}\ p(\theta|\mathcal{I};\boldsymbol{\lambda}). \tag{IV.1}$$

Nous supposons ici qu'à l'opposé de θ, λ désigne un vecteur inconnu mais non issu d'un processus aléatoire. A ce propos on peut remarquer que les résultats d'inférence obtenus auparavant sur θ sont des distributions de probabilités, hormis celui plus spécifique du modèle polynômial par morceaux.
Ce principe marque la différence entre les approches purement bayesiennes où une distribution de probabilité est associée à chaque hyperparamètre, et les méthodes fréquentistes où on suppose qu'une valeur optimale constante existe pour chacun d'entre eux [Jaynes 03, Mackay 02]. Une inférence complète est possible dans le cadre bayesien, mais comme nous en avons brièvement discuté au chapitre précédent, elle nécessite la résolution de plusieurs échelons d'incertitude [Mackay 92, Rasmussen 06] pour lesquels les solutions simples sont rares.

IV.2.1 Erreur de prédiction

Pour prendre en compte ces nouvelles valeurs inconnues, on peut choisir de résoudre le problème suivant :

$$\hat{\theta}_{MAP}, \hat{\lambda}_{MAP} = \underset{\theta,\lambda}{\text{argmax}}\ p(\theta|\mathcal{I};\lambda), \tag{IV.2}$$

où par rapport à l'équation (IV.1), le vecteur d'hyperparamètres λ a été ajouté à la procédure de maximisation. Par la suite, nous appelerons λ^α le vecteur d'hyperparamètres associé à $\alpha \in \{l, r, t\}$ qui désigne le lacet, le roulis ou le tangage. Appliquons cette idée à l'estimateur du MAP avec *a priori* sur l'amplitude, qui a été vu au paragraphe III.5.2 et où $\lambda^\alpha = \sigma_{\theta^\alpha}$ est un scalaire. En développant le logarithme de la loi a *posteriori* en équation (III.46), on obtient :

$$\begin{aligned} \log(p(\theta^\alpha|\mathcal{I};\sigma_{\theta^\alpha})) &\propto \log(p(\mathcal{I}|\theta^\alpha)) + \log(p(\theta^\alpha;\sigma_{\theta^\alpha})) \\ &= -\sum_{i,j;i\neq j} \frac{1}{2\sigma_I^2}|\mathbf{e}_{ij} - J_j^\alpha K_{ij}^\alpha \theta^\alpha|^2 - \frac{1}{2\sigma_{\theta^\alpha}^2}|\theta^\alpha|^2. \end{aligned} \tag{IV.3}$$

Dans cette expression, les deux termes quadratiques sont nécessairement positifs donc le résultat est globalement négatif ou nul. Une solution triviale à la maximisation proposée en (IV.2) consiste à faire tendre le second terme *a priori* vers 0, on a donc $\sigma_{\theta^\alpha} \to +\infty$ pour que $\sigma_{\theta^\alpha}^{-2}|\theta^\alpha|^2 \to 0$. Cette solution ne présente aucun intérêt, puisqu'elle revient à utiliser l'estimateur du MV alors même que nous souhaitons spécifier une connaissance sur θ^α. Cette ambiguïté a été relevée dans de nombreux domaines, on peut citer [Brunet 10] pour une référence récente sur la problématique de recalage d'image.
Le problème doit être considéré autrement pour aboutir à une solution, reposons le comme suit [Brunet 10] :

$$\hat{\theta}_{MAP} = \underset{\theta}{\text{argmax}}\ p(\theta|\mathcal{I};\underset{\lambda}{\text{argmin}}\ \mathcal{E}(\lambda)), \tag{IV.4}$$

où $\mathcal{E}$ est une fonction d'erreur dépendante des hyperparamètres et que nous souhaitons cette fois-ci minimiser. On comprend que l'inférence est découpée en deux étapes : une recherche des hyperparamètres optimaux dans un premier temps et consécutivement une estimation de l'attitude. Une solution simple peut consister à choisir en premier des valeurs pour λ^α, calculer le vecteur $\hat{\theta}_{\lambda^\alpha}^\alpha$ associé par l'estimateur du MAP - nous appelons $\hat{\theta}_{\lambda^\alpha}^\alpha$ le vecteur d'attitude pour l'angle α qui a été estimé à partir de λ^α comme vecteur

d'hyperparamètres - et évaluer cette estimation par rapport aux données dont nous disposons grâce à la vraisemblance des images $p(\mathcal{I}|\theta^{\alpha}_{\lambda^{\alpha}})$:

$$p(\mathcal{I}|\theta^{\alpha}_{\lambda^{\alpha}}) = \prod_{i,j;i\neq j} \frac{1}{(2\pi\sigma_I^2)^{NL/2}} \exp\left(-\frac{1}{2\sigma_I^2}\left|\mathbf{e}_{ij} - J^{\alpha}_j K^{\alpha}_{ij}\theta^{\alpha}_{\lambda^{\alpha}}\right|^2\right). \tag{IV.5}$$

Nous choisissons de comparer $\theta^{\alpha}_{\lambda^{\alpha}}$ aux images uniquement à l'aide de la fonction de vraisemblance pour ne pas retomber dans le cas de l'équation (IV.3). Alors, minimiser la fonction d'erreur empirique suivante :

$$\mathcal{E}(\boldsymbol{\lambda}^{\alpha}) = \sum_{i,j;i\neq j} \left|\mathbf{e}_{ij} - J^{\alpha}_j K^{\alpha}_{ij}\hat{\theta}^{\alpha}_{\lambda^{\alpha}}\right|^2, \tag{IV.6}$$

est équivalent à choisir le $\boldsymbol{\lambda}^{\alpha}$ pour lequel $\theta^{\alpha}_{\lambda^{\alpha}}$ est le plus vraisemblable à la vue des images. En procédant ainsi pour plusieurs $\boldsymbol{\lambda}^{\alpha}$, on construit une fonction d'erreur échantillonnée aux valeurs choisies. Mais on comprend également que l'hyperparamètre optimal sera une nouvelle fois tel que $\hat{\theta}^{\alpha}_{\lambda^{\alpha}} = \theta^{\hat{\alpha}}_{MV}$, ou autrement dit celui qui annule la contribution de l'*a priori*.
Ce problème peut être vu de façon plus éloquente en sélection de modèle, où la minimisation de l'erreur empirique aboutit systématiquement au choix du modèle le plus complexe. En réduisant sa marge d'erreur aux données - ce qui est équivalent à minimiser l'erreur empirique - , le modèle est capable de suivre parfaitement toutes les fluctuations des mesures. Pourtant il est peu performant car une donnée supplémentaire dans l'apprentissage de ce modèle peut radicalement changer sa forme. Ce phénomène intéressant a été observé et étudié à de nombreuses reprises dans différents contextes, nous reportons le lecteur à [Andrew 09] pour une introduction didactique à celui-ci, et [Mackay 92] pour son interprétation bayésienne.
Plutôt que de chercher le modèle qui à un $\boldsymbol{\lambda}^{\alpha}$ fixé correspond le mieux aux mesures, il est plus judicieux de rechercher celui qui minimise l'erreur de prédiction sur l'ensemble des données. Cette idée peut s'exprimer comme ci-dessous :

$$\mathcal{E}(\boldsymbol{\lambda}^{\alpha}) = \mathbb{E}\left[\sum_{i,j;i\neq j} \left|\mathbf{e}_{ij} - J^{\alpha}_j K^{\alpha}_{ij}\hat{\theta}^{\alpha}_{\lambda^{\alpha}}\right|^2\right], \tag{IV.7}$$

où $\mathbb{E}$ désigne l'espérance sur les variables aléatoires que sont le vecteur de différences $\mathbf{e}$ et la matrice jacobienne J^{α} construites à partir des images. La formule (IV.7) est également connue sous le nom d'erreur de généralisation [Andrew 09, Rasmussen 06]. Etant donné que nous ne disposons que d'une quantité limitée d'images, le calcul de cette espérance ne peut qu'être approché par des méthodes statistiques.

IV.2.2 Validation Croisée (VC)

Il existe de nombreuses méthodes permettant la sélection de modèles [Akaike 74, Schwarz 78] ou l'estimation d'hyperparamètres [Hansen 00]. Elles se distinguent généralement par leur robustesse ou leur rapidité de temps de calcul. La validation croisée est une technique approchant l'erreur de prédiction qui, mis à part son temps de calcul prohibitif, fait souvent consensus pour la qualité des résultats obtenus [Wahba 90, Bartoli 09]. Ce critère est très générique et assez simple à appliquer, ce qui explique son utilisation massive dans de nombreux domaines. Il existe plusieurs variantes à la VC [Andrew 09, Brunet 10], on peut citer notamment la *k-fold cross validation* et son cas particulier la *leave-one-out cross validation*. Nous allons détailler ces méthodes par la suite.

Principe

Le principe général de la validation croisée consiste à séparer les données en deux catégories : une d'apprentissage servant à estimer les paramètres d'un modèle, et l'autre de test permettant d'évaluer l'erreur de prédiction du modèle inféré précédemment. En répétant cette opération plusieurs fois, le modèle qui aura l'erreur de prédiction la plus faible sera élu comme le plus performant.
Associé à notre contexte, les données d'entrée correspondent aux images et plus particulièrement aux NL pixels qui composent chacune d'entre elles. Pour un estimateur choisi, les modèles comparés correspondent à ce même estimateur avec différentes valeurs de $\boldsymbol{\lambda}^{\alpha}$; soit $\boldsymbol{\lambda}^{\alpha}_{m}$ pour $m \in \{1, 2, \ldots, M\}$ les vecteurs des hyperparamètres échantillonnés choisis. On considère la séparation des pixels en une partie apprentissage indicée par $.^{(a)}$ et une partie test indicée par $.^{(t)}$. Les ensembles d'images $\mathcal{J}^{(a)}$ et $\mathcal{J}^{(t)}$, les matrices $J^{\alpha(a)}$ et $J^{\alpha(t)}$, ainsi que les vecteurs $\mathbf{e}^{(a)}$ et $\mathbf{e}^{(t)}$ sont respectivement construits selon ces deux groupes. La validation croisée se résume alors au schéma IV.1.
Cet algorithme procède par recherche exhaustive, ce qui explique sa relative lenteur.

Pour m de 1 jusqu'à M, faire :

1. rechercher l'attitude sur le jeu d'apprentissage :

$$\hat{\theta}^{\alpha}_{\lambda^{\alpha}} = \underset{\theta^{\alpha}}{\operatorname{argmax}}\ p(\theta^{\alpha}|\mathcal{J}^{(a)}; \boldsymbol{\lambda}^{\alpha}_{m}),$$

2. évaluer la qualité de la prédiction sur le jeu de test :

$$\mathcal{E}(\boldsymbol{\lambda}^{\alpha}_{m}) = \sum_{i,j;i\neq j} \left| \mathbf{e}^{(t)}_{ij} - J^{\alpha(t)}_{j} K^{\alpha}_{ij} \hat{\theta}^{\alpha}_{\lambda^{\alpha}} \right|^2.$$

Déduire le vecteur de paramètre optimal :

$$\hat{\boldsymbol{\lambda}}^{\alpha} = \underset{\boldsymbol{\lambda}^{\alpha}}{\operatorname{argmin}}\ \mathcal{E}(\boldsymbol{\lambda}^{\alpha}).$$

FIGURE IV.1 – Algorithme de validation croisée standard.

La qualité d'estimation de $\hat{\boldsymbol{\lambda}}^{\alpha}$ dépend du nombre de valeurs auxquelles on souhaite évaluer la fonction d'erreur de prédiction. En choisissant un intervalle raisonnable autour de notre croyance sur les hyperparamètres, et avec un échantillonnage correct dans cet intervalle, on obtient assez souvent des résultats probants. Notons que ce schéma algorithmique est présenté pour l'estimation des hyperparamètres selon un angle, mais on peut également l'appliquer pour rechercher globalement ceux pour tous les angles. Seulement M croît de façon exponentielle avec l'ajout de nouveaux hyperparamètres à déterminer, ce qui augmente grandement le temps de calcul. Il est préférable de fixer les hyperparamètres indépendamment selon les angles, ce qui est sous-optimal mais déjà très satisfaisant comme nous le verrons dans les résultats. Sachant que les variations d'attitude sont stationnaires, il est intéressant de remarquer qu'on peut se contenter de portions d'images pour effectuer la validation croisée. Enfin concernant la répartition des pixels entre apprentissage et test, elle se fait de façon aléatoire et il est commun d'attribuer 70% de leur nombre au premier et les 30% restant au second.

Dans sa version standard, la validation croisée est souvent critiquée pour être un estimateur de variance trop importante, spécialement lorsque la quantité de données est faible. En effet, l'attribution aléatoire d'une mesure entre le groupe apprentissage et le groupe test fait que plusieurs répétitions de l'algorithme ne montrent pas toujours les même résultats. On peut aussi constater que cette technique « perd » 30% de données pour l'apprentissage, ce qui peut biaiser l'estimation en supposant que ces mesures non prises en compte contenaient des informations essentielles.
Les figures IV.2 et IV.3 montrent plusieurs résultats de validation croisée sur un même jeu de données - celui que nous avons utilisé au cours du chapitre précédent et présenté au paragraphe III.4.7 - où nous avons sélectionné une portion de (400×500) pixels sur les images pour déduire les hyperparamètres. La première figure considère l'estimateur du MAP décrit en équation (III.46) (p.47) avec *a priori* sur l'amplitude, et la seconde celui proposé en équation (III.56) (p.50) avec *a priori* sur la dérivée seconde. L'hyperparamètre σ_{θ^r} est recherché dans les deux cas ; nous présentons des résultats où nous utilisons 100% des pixels divisés entre apprentissage et test - soit 200000 -, et 25% de ceux-ci pour observer la variance de l'estimateur.
La comparaison des deux figures est intéressante, alors que l'hyperparamètre dans le

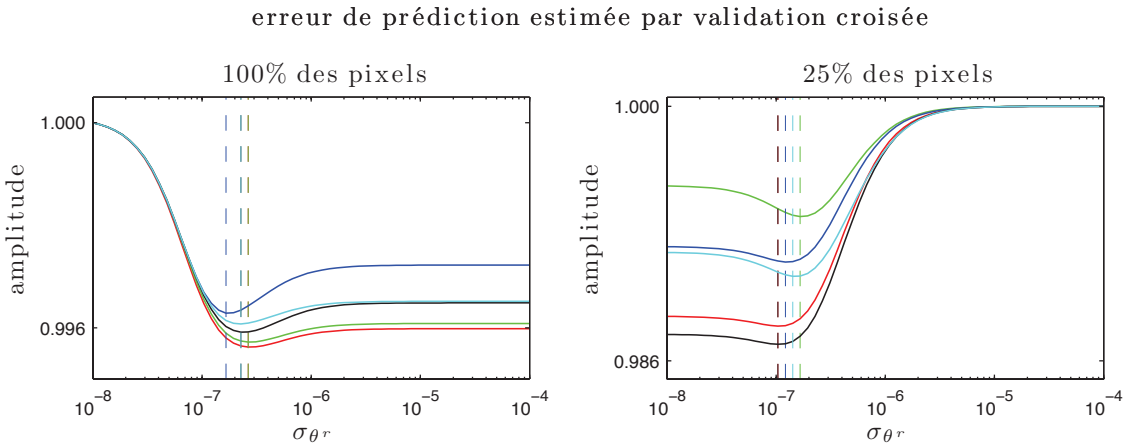


FIGURE IV.2 – Tracé de l'erreur de prédiction déduite par validation croisée de l'estimateur équation (III.46) (p.47) ; l'hyperparamètre recherché est σ_{θ^r}. L'algorithme est exécuté 5 fois de suite, correspondant aux 5 courbes de différentes couleurs. Dans la figure de gauche, 100% des pixels des images sont utilisés pour l'inférence, alors que la figure de droite n'utilise que 25% de ceux-ci (200000 contre 50000 pixels respectivement). A l'échelle de l'intervalle de recherche (60 valeurs échantillonnées entre $[10^{-8}, 10^{-4}]$), la validation croisée montre des résultats relativement stables entre 10^{-7} et 2×10^{-7} pour les valeurs inférées de σ_{θ^r}. Les minimums de chaque courbe sont assez bien marqués et repérés par les lignes en tirets (ces dernières se superposent dans certains cas). On remarque que la diminution du nombre de pixels influence peu le résultat.

premier cas en figure IV.2 semble assez bien déterminé par validation croisée, même avec des données réduites, le second cas de la figure IV.3 montre des courbes avec un bassin de convergence plus large, et on voit notamment apparaître le problème de variance dans l'estimateur lorsqu'on n'utilise que 25%, avec une valeur minimum de σ_{θ^r} autour de 2×10^{-9} et une valeur maximum de presque 7×10^{-7}.

On constate qu'au-delà du nombre de mesures, la validation croisée n'a pas la même

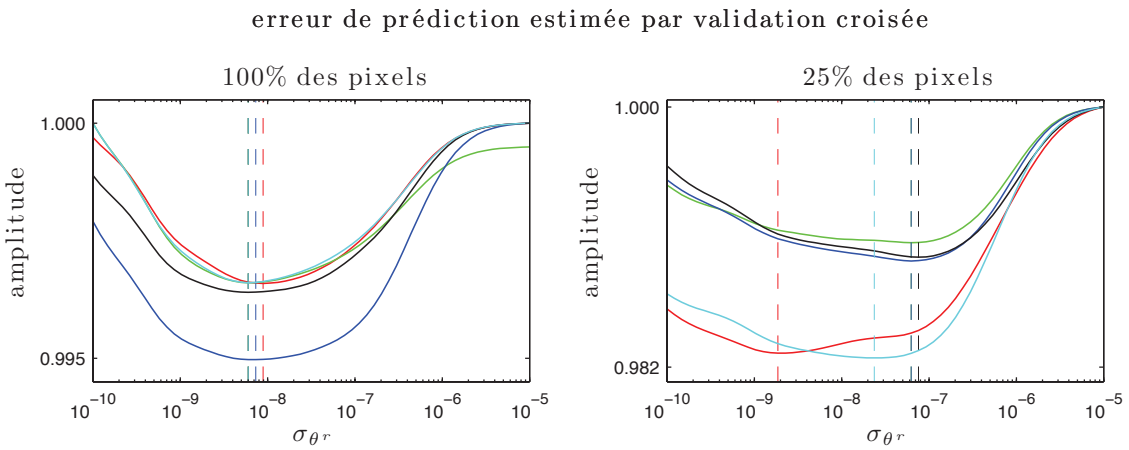


FIGURE IV.3 – Tracé de l'erreur de prédiction pour l'estimateur du MAP équation (III.56) (p.50) où l'hyperparamètre recherché est σ_{θ^r}. Les 5 courbes ont été estimées par validation croisée dans les mêmes conditions que la figure IV.2. Pour l'*a priori* sur la dérivée seconde, les bassins de convergence des courbes sont plus larges que ceux de la figure précédente. Si l'estimateur concentre assez bien les valeurs de σ_{θ^r} autour de 9×10^{-9} sur la figure de gauche, on observe une variance plus importante des résultats lorsque la quantité diminue ; c'est une caractéristique de la validation croisée.

sélectivité suivant les estimateurs du MAP choisis, alors que ceux-ci ont une formulation mathématique très proche. Il est possible que cela soit lié à la qualité de l'*a priori*, comme la quantité d'information qu'il apporte sur le processus à estimer.

k-fold cross validation

La validation croisée de type *k-fold* cherche à réduire la variance de l'algorithme standard en exploitant toutes les données et en calculant une espérance sur plusieurs validations croisées. Pour cela, les données sont séparées aléatoirement en plusieurs paquets ; soit C leur nombre. A tour de rôle, un paquet est retiré pour servir de jeu de test, les autres sont regroupés pour former le jeu d'apprentissage. L'algorithme standard est appliqué tant que tous les paquets n'ont pas été testés et finalement la moyenne des C validations croisées forme l'erreur de prédiction. Soit $.^{(c^+)}$ et $.^{(c^-)}$ les indices des variables qui ont été créées respectivement en ne prenant en compte que les mesures du paquet c, et en considérant toutes les mesures sauf celles du paquet c. Le schéma de l'algorithme est présenté en figure IV.4.

Nous présentons en figures IV.5 et IV.6 les résultats associés à la *k-fold cross validation* pour les mêmes données que celles utilisées auparavant, où le nombre de paquets C a été fixé à 7. Cette nouvelle version de la validation croisée ne semble pas améliorer l'inférence de σ_{θ^r} pour l'*a priori* sur l'amplitude comme le montre la comparaison entre les figures IV.2 et IV.5. Les valeurs résultantes sont toujours bien localisées autour de 1.5×10^{-7} mais leur variance n'est pas vraiment améliorée. En revanche l'*a priori* sur les dérivées secondes semblent mieux profiter de cet algorithme comme en témoignent les figures IV.3 et IV.6. Le premier graphique réalisé avec les données complètes montre que les minima de toutes les courbes sont concentrés en un même point en abscisse. Le second présente

Pour c de 1 jusqu'à C, faire :
  Pour m de 1 jusqu'à M, faire :

1. rechercher l'attitude sur le jeu d'apprentissage :

$$\hat{\theta}^{\alpha}_{\lambda^{\alpha}} = \underset{\theta^{\alpha}}{\operatorname{argmax}}\ p(\theta^{\alpha}|\mathcal{J}^{(c^-)};\lambda^{\alpha}_{m}),$$

2. évaluer la qualité de la prédiction sur le jeu de test :

$$\mathcal{E}_c(\lambda^{\alpha}_{m}) = \sum_{i,j;i\neq j} \left| e^{(c^+)}_{ij} - J^{\alpha(c^+)}_{j} K^{\alpha}_{ij} \hat{\theta}^{\alpha}_{\lambda^{\alpha}} \right|^2 .$$

Calculer la moyenne des erreurs de prédiction :

$$\mathcal{E}(\lambda^{\alpha}) = \mathbb{E}(\mathcal{E}_c(\lambda^{\alpha})),$$

Déduire le vecteur de paramètres optimal :

$$\hat{\lambda}^{\alpha} = \underset{\lambda^{\alpha}}{\operatorname{argmin}}\ \mathcal{E}(\lambda^{\alpha}).$$

FIGURE IV.4 – Algorithme de validation croisée type *k-fold*.

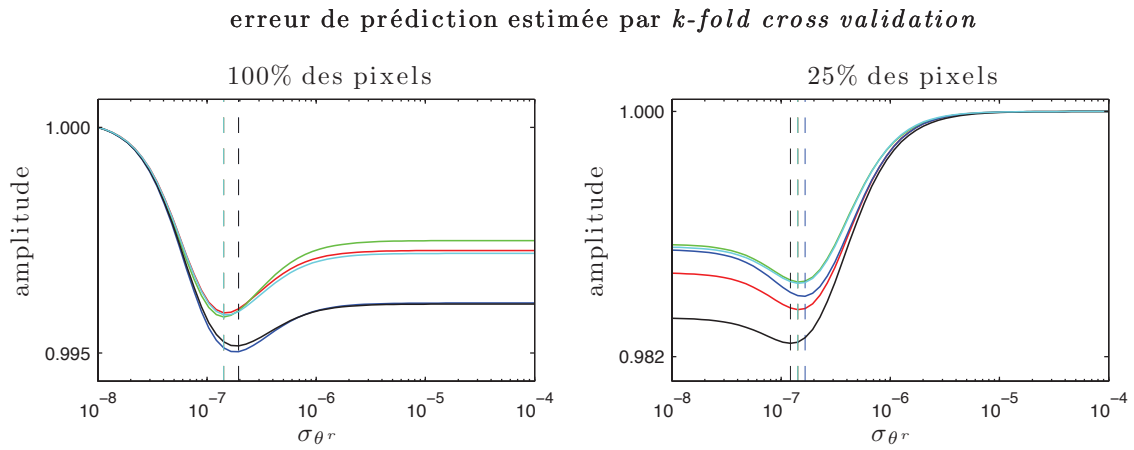


FIGURE IV.5 – Estimation de l'erreur de prédiction par *k-fold cross validation* avec $C = 7$. L'expérience menée est comparable à la figure IV.2. Contrairement au cas de la figure IV.6, cette version de la validation croisée n'apporte pas d'amélioration importante dans la variance des résultats.

des bassins de convergence plus réguliers, et des valeurs inférées de σ_{θ^r} plus rapprochées.
Bien que cet algorithme soit un peu plus performant dans le second cas, son coût algo-

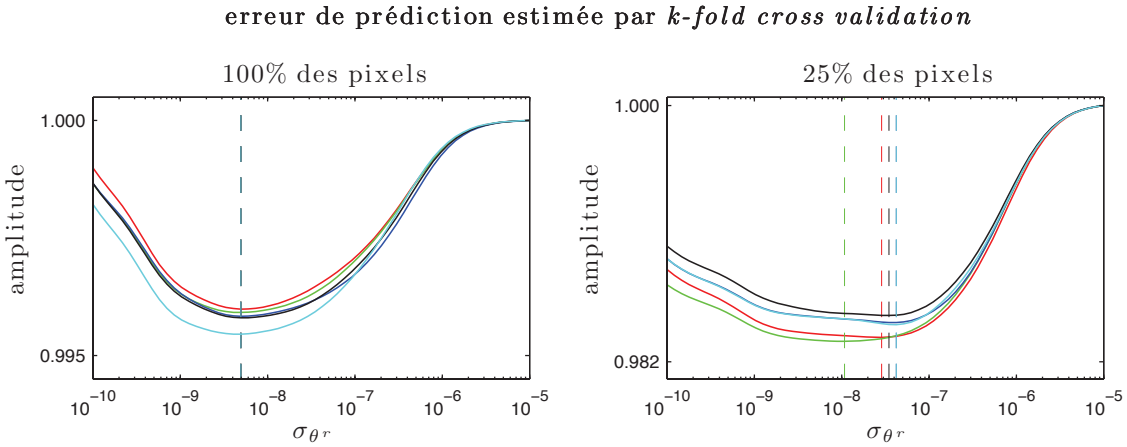


FIGURE IV.6 – Estimation de l'erreur de prédiction par *k-fold cross validation* avec $C = 7$. Le contexte est le même que celui de la figure IV.3 où σ_{θ^r} est l'hyperparamètre recherché. Les 5 courbes adoptent des formes presque identiques, avec dans le cas des mesures complètes une même valeur optimale. La variance dans le second est réduite, les valeurs se sont recentrées sur un intervalle de $[10^{-8}, 4 \times 10^{-8}]$.

rithmique est nettement plus conséquent que le précédent puisque la validation croisée dans sa version standard est ici répétée C fois. Le choix d'une des méthodes peut se faire en fonction de l'hyperparamètre recherché, mais aussi en fonction de la quantité de données que nous utilisons pour l'estimer. La réduire peut accélérer les calculs mais également augmenter les risques de mauvaise décision sur les hyperparamètres. Dans notre cas en pratique, nous disposons d'un nombre largement suffisant de mesures dans les images pour nous contenter de la validation croisée classique. D'autant plus que comme nous le verrons dans la dernière partie des expérimentations, le choix d'une valeur pour un hyperparamètre n'est pas toujours critique pour l'estimation consécutive de l'attitude.

leave-one-out cross validation

La validation croisée de type *leave-one-out* est un cas particulier de la version précédente. Plutôt que de grouper les données par paquet, on considère que chaque mesure est un paquet. Cette approche consiste à supprimer une mesure pour l'apprentissage, et à évaluer l'erreur de prédiction sur celle-ci uniquement. En répétant l'opération pour toutes les mesures, on construit l'erreur de prédiction. Le schéma algorithmique est le donc le même que celui présenté en figure IV.4, avec cette fois-ci $M = NL$.
Cette approche est prohibitive en temps de calcul, et n'est utilisée que dans des cas particuliers où très peu de mesures sont disponibles, ce qui n'est pas le cas pour les données images. En revanche, nous sommes confrontés dans notre étude à ce problème pour estimer le degré du modèle polynômial sur les acquisitions du *star tracker*. Nous présentons un exemple d'estimation du degré d'un polynôme en figure IV.7.
L'algorithme devient déterministe dans ce cas, il n'y a plus de notion d'aléa car chaque mesure est évaluée ; le résultat est donc unique pour un même jeu de données. Cette ver-

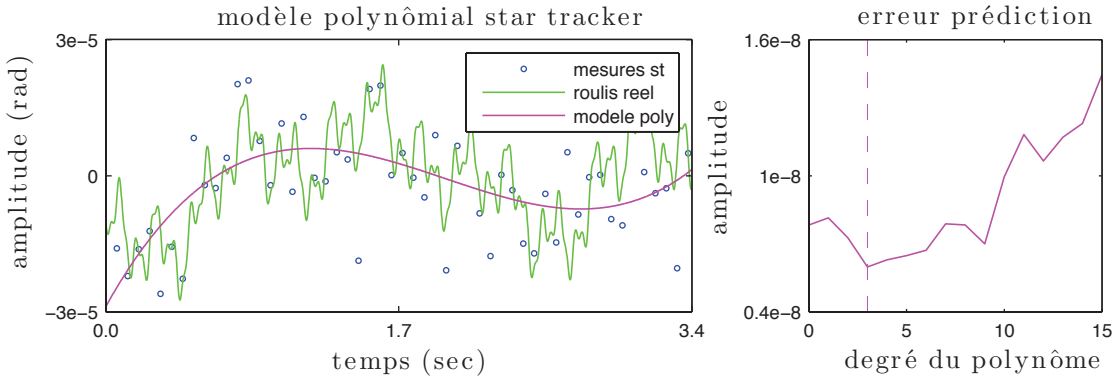


FIGURE IV.7 – Modèle polynômial (courbe en magenta) estimé sur 53 mesures de *star tracker* en bleu. Le degré du polynôme ($D_s = 3$) est déduit par la méthode de validation croisée de type *leave-one-out*, l'erreur de prédiction est représentée à droite. Malgré la grande disparité dans les mesures, le polynôme estimé semble être une bonne approximation des basses fréquences du roulis réel (courbe verte).

sion de la validation croisée est également référencée sous le nom de *PREdictive Sum of Squares* (PRESS) dans la littérature [Rasmussen 06, Bartoli 08]. Il existe des implémentations rapides de cette méthode pour l'estimation de plusieurs hyperparamètres comme proposé par [Bartoli 09], et il serait intéressant de les envisager pour compléter l'étude sur l'estimation des hyperparamètres à partir des images.

IV.2.3 Erreur photométrique

La validation croisée est un critère générique, elle ne s'appuie pas sur la nature des données, et ne cherche pas à les comprendre. Malgré ses performances intéressantes pour l'estimation des hyperparamètres, ils existent certains cas où elle ne peut pas être appliquée correctement. Prenons l'exemple de la fusion au paragraphe III.9.5 (p.74), où l'estimateur du MAP est le suivant :

$$\hat{\boldsymbol{\theta}}_{MAP} = \underset{\boldsymbol{\theta}}{\operatorname{argmax}}\ p(\boldsymbol{\theta}|\mathcal{I}, \mathbf{c}; \boldsymbol{\lambda}), \tag{IV.8}$$

avec $\boldsymbol{\lambda} = [\boldsymbol{\sigma}_\theta \ \ \boldsymbol{\sigma}_c]^T$ le vecteur des hyperparamètres pour l'attitude et :

$$p(\boldsymbol{\theta}|\mathcal{I}, \mathbf{c}) \propto p(\mathcal{I}|\boldsymbol{\theta}) p(\mathbf{c}|\boldsymbol{\theta}; \boldsymbol{\sigma}_c) p(\boldsymbol{\theta}; \boldsymbol{\sigma}_\theta). \tag{IV.9}$$

Nous rappelons ici que $\mathbf{c}$ est le vecteur des coefficients du modèle polynomial basse fréquence de l'attitude. Si nous cherchons à appliquer la validation croisée pour estimer dans le cas du roulis les deux hyperparamètres σ_{θ^r} et σ_{c^r}, on observe le résultat en figure IV.8. Pour ce test, l'estimateur du MAP utilise le premier modèle de vraisemblance des images, telle qu'il est décrit au paragraphe III.4.1 avec un *a priori* autorégressif. Rappelons que les basses fréquences sont mal estimées par ce modèle.
On constate sur le graphique de gauche que le roulis estimé ne profite pas des mesures *star tracker*, alors même que le modèle polynômial donne une assez bonne information sur les basses fréquences du roulis réel. Le tracé des iso-contours de la fonction $\mathcal{E}(\sigma_{c^r}, \sigma_{\theta^r})$ sur le graphique de droite montre une convergence selon l'axe σ_{θ^r}, et des lignes fuyantes

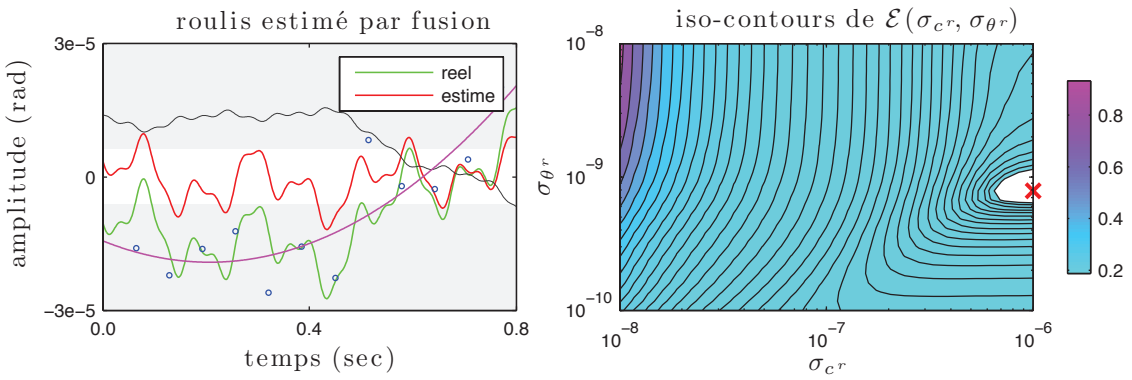


FIGURE IV.8 – Exemple d'estimation par validation croisée des hyperparamètres σ_{θ^r} et σ_{c^r} pour la fusion image et *star tracker*. L'estimateur du MAP est présenté au chapitre précédent, paragraphe III.9.5. Les images utilisées sont de taille (600×500), et 11 mesures *star tracker* (en rond bleu sur le graphique de gauche) ont permis d'estimer le modèle polynômial (en magenta) de degré $D_{s^r} = 2$. Le résidu (en noir) est le même que celui sans fusion, l'estimateur du MAP considère les informations du *star tracker* comme fausses. La figure de droite montre les iso-contours de la fonction d'erreur de prédiction évaluée en 900 points (30 selon chaque axe), le minimum est marqué par la croix rouge. La convergence selon l'hyperparamètre σ_{θ^r} est nette sur une valeur de 7.8×10^{-10}, en revanche $\hat{\sigma}_{c^r}$ atteint la borne supérieure de son intervalle.

sans convergence selon l'axe σ_{c^r}. Le minimum déduit se situe sur la croix rouge, avec $\hat{\boldsymbol{\lambda}}^r = [\hat{\sigma}_{c^r} \;\; \hat{\sigma}_{\theta^r}]^T = [7.8 \times 10^{-10} \;\; 10^{-6}]^T$. Le problème se situe au niveau du critère de test de l'équation (IV.7), qui pour la première vraisemblance d'image (III.32) prend la forme suivante :

$$\mathcal{E}(\boldsymbol{\lambda}^\alpha) = \mathbb{E}\left[\sum_{i,j;i\neq j} \left|\mathbf{e}_{ij} + (J_i^\alpha - J_j^\alpha D_{ij}^\alpha)\theta_{\lambda^\alpha}^\alpha\right|^2\right]. \tag{IV.10}$$

La vraisemblance des images ayant été linéarisée - comme nous l'avons décrit au paragraphe III.4.5 -, la déformation n'est pas supposée dépasser un certain seuil en amplitude. Les mesures du *star tracker* sont donc considérées comme aberrantes par l'erreur de prédiction (IV.10), et seul l'*a priori* autorégressif prend de l'importance. En pratique, la validation croisée fait tendre σ_{c^r} vers l'infini, ce qui contribue à écarter le modèle polynomial de l'inférence ; c'est ce que nous observons à la figure IV.8 où le bassin de convergence se creuse sur la borne supérieure de l'intervalle des valeurs échantillonnées pour σ_{c^r}.
Si nous remontons un peu plus haut dans le chapitre précédent jusqu'au paragraphe III.4.1, on se rappelle que la vraisemblance originale des images est la suivante pour le premier cas :

$$I_i\big(\mathcal{W}(x, n; \theta(n))\big) - I_j^{\tau_{ij}}\big(\mathcal{W}(x, n; \theta(n + \tau_{ij}))\big) \sim \mathcal{N}\big(0, \sigma_I^2\big), \tag{IV.11}$$

où le terme de gauche peut être compris comme l'erreur photométrique entre les deux images I_i et I_j. Ainsi, la véritable façon d'évaluer la qualité d'une attitude estimée $\hat{\theta}$ pour un vecteur d'hyperparamètres donné $\hat{\boldsymbol{\lambda}}$ serait de calculer cette erreur photométrique, plutôt que de considérer l'erreur de prédiction en équation (IV.10). Pour tenir compte de cette idée, nous adaptons l'algorithme de recherche des hyperparamètres IV.1 comme présenté

en figure IV.9, où M désigne une nouvelle fois le nombre d'échantillons auxquels on souhaite calculer l'erreur.
Plusieurs remarques peuvent être faites sur cet algorithme. Notons déjà qu'il n'y a plus

Pour m de 1 jusqu'à M, faire :

1. rechercher l'attitude :

$$\hat{\theta}^{\alpha}_{\lambda^{\alpha}} = \underset{\theta^{\alpha}}{\operatorname{argmax}}\ p(\theta^{\alpha}|\mathcal{I}, \mathbf{c}^{\alpha}; \boldsymbol{\lambda}^{\alpha}_{m}),$$

2. évaluer la qualité de l'estimation avec l'erreur photométrique :

$$\mathcal{E}(\boldsymbol{\lambda}^{\alpha}_{m}) = \sum_{i,j;i\neq j} \sum_{n=1}^{N} \sum_{x=1}^{L} \left| I_i\left(\mathcal{W}(x, n; \theta^{\alpha}_{\lambda^{\alpha}}(n))\right) - I_j^{\tau_{ij}}\left(\mathcal{W}(x, n; \theta^{\alpha}_{\lambda^{\alpha}}(n + \tau_{ij}))\right)\right|^2.$$

Déduire le vecteur de paramètres optimal :

$$\hat{\boldsymbol{\lambda}}^{\alpha} = \underset{\boldsymbol{\lambda}^{\alpha}}{\operatorname{argmin}}\ \mathcal{E}(\boldsymbol{\lambda}^{\alpha}).$$

FIGURE IV.9 – Algorithme de validation photométrique.

de jeux d'apprentissage et de test, toutes les données sont considérées en même temps dans la première étape d'estimation de l'attitude ; cette méthode d'estimation est donc déterministe. Le temps de calcul semble être similaire à celui de la validation croisée. En pratique l'étape 2 de calcul de l'erreur photométrique est plus coûteuse car elle prend en compte toutes les données, alors que la validation croisée n'en considère qu'une fraction, et elle nécessite l'interpolation des images pour l'attitude $\hat{\theta}$. Enfin cette méthode est également exhaustive dans sa recherche des hyperparamètres, le résultat dépend des intervalles considérés et du nombre d'échantillons auxquels l'erreur photométrique est mesurée.
La figure IV.10 est le résultat du même test que celui conduit en figure IV.8, en utilisant cette fois-ci la validation photométrique. A la vue du graphique de gauche, on remarque que le roulis estimé est meilleur, essentiellement parce qu'il prend mieux en compte les mesures du *star tracker*. Le bassin de convergence de la fonction d'erreur photométrique est très net, avec un minimum marqué par la croix rouge en $[\hat{\sigma}_{c^r} \quad \sigma_{\hat{\theta}^r}]^T = [3.2 \times 10^{-9} \quad 1.2 \times 10^{-7}]^T$. En comparaison du résultat précédent, la contribution de l'*a priori* a été un peu réduite au profit du modèle polynômial. Selon la forme des courbes autour du minimum, l'hyperparamètre σ_{c^r} apparait même être mieux déterminé que σ_{θ^r}.
L'estimation de paramètres de déformation par critère photométrique est historique dans les problèmes de vision par ordinateur [Szeliski 06], en revanche son utilisation dans l'inférence d'hyperparamètres est moins commune ou peu explicitée. En comparaison à la validation croisée, l'article [Brunet 10] qui décrit clairement son utilisation montre aussi que ce critère photométrique est plus fiable dans le cadre d'un recalage d'image non rigide avec points d'intérêts. Le principal défaut de cette approche est aussi son temps de calcul dû à une recherche exhaustive, mais il est important de comprendre que dans ce contexte de fusion, la validation croisée ne peut pas être appliquée et que l'utilisation des mesures photométriques est une manière robuste et simple de contourner ce problème.

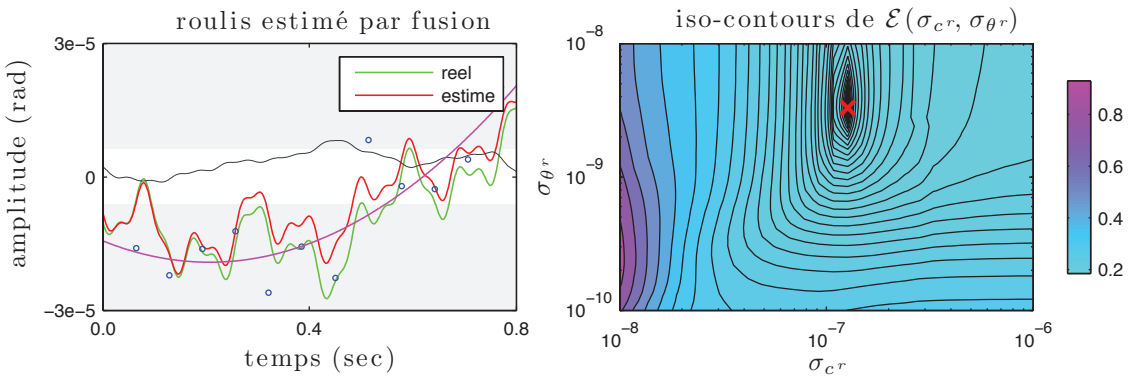


FIGURE IV.10 – Exemple d'estimation par validation photométrique des hyperparamètres σ_{θ^r} et σ_{c^r} pour la fusion image et *star tracker*. Les conditions de l'expérience sont identiques à celles de la figure IV.8. En comparaison, le roulis estimé suit mieux les basses fréquences du vrai roulis, le résidu est également plus faible. Le bassin de convergence de la fonction d'erreur photométrique est très net, avec un minimum en $[\hat{\sigma}_{c^r} \;\; \hat{\sigma}_{\theta^r}]^T = [3.2 \times 10^{-9} \;\; 1.2 \times 10^{-7}]^T$.

En prenant un peu plus de recul sur cette solution, on voit qu'elle propose de rééchantillonner la fonction d'origine de vraisemblance des images comme la plupart des méthodes MCMC pourraient le faire, à la différence près que nous choisissons un intervalle fixe de recherche des hyperparamètres comme pour la validation croisée.
La validation par photométrie est également suffisamment générique pour estimer tout type d'hyperparamètres, la recherche du degré du polynôme associé à la taille de fenêtre optimale en est un exemple pour le modèle polynômial sous contraintes que nous avons décrit au paragraphe III.8.1. La figure IV.11 est un nouvel exemple d'inférence dans le cas où les valeurs recherchées sont discrètes. Ici, la fonction d'erreur est plus difficile à interpréter car le minimum à l'emplacement de la croix rouge se distingue moins bien que pour les cas précédents. Globalement, elle privilégie des valeurs petites pour le degré du polynôme et la taille de fenêtre. Le tracé des minimums locaux en tirets rouges tend à confirmer qu'il n'y a pas de relation linéaire entre D et l_p, comme nous l'avons supposé au chapitre précédent.
Il pourrait être intéressant de corréler la recherche des paramètres optimaux avec une information sur le temps de calcul de l'algorithme pour l'inférence complète du signal. En effet, la taille de fenêtre implique plus ou moins de coefficients polynômiaux à estimer ; de même le degré joue beaucoup sur le temps de calcul malgré les contraintes énoncées entre chaque polynôme. En favorisant les solutions rapides, il est possible que la convergence sur un minimum global soit meilleure. Les deux critères de validation - croisée et photométrique - que nous venons d'analyser seront comparés dans la partie expérimentations sur les données. Nous proposons également en fin de chapitre un tableau récapitulatif de leur utilisation en fonction des estimateurs.

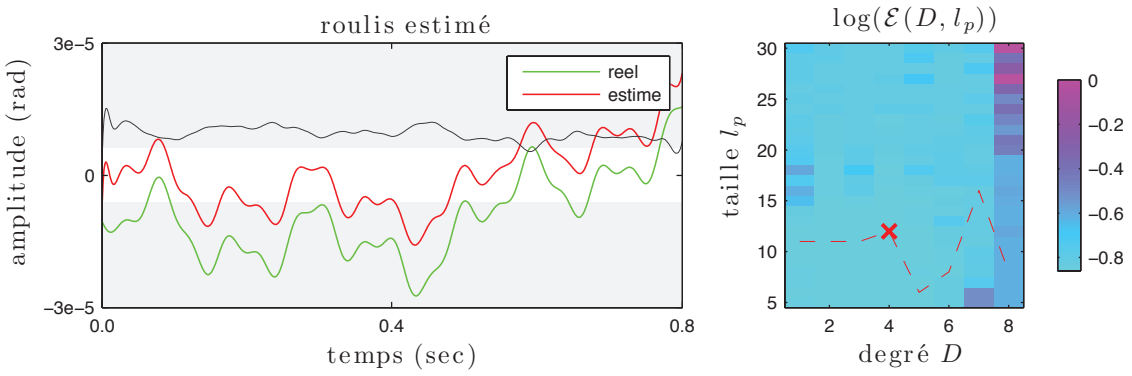


FIGURE IV.11 – Exemple d'inférence par critère photométrique du degré et de la taille de fenêtre pour le modèle polynômial avec contraintes. La fonction d'erreur est plus complexe à interpréter ici car le minimum en $\hat{D} = 4$ et $\hat{l_p} = 12$ ne se dégage pas clairement de son voisinage. Elle privilégie globalement des petites tailles de fenêtre et des polynômes de degré inférieur à 6. Le parcours des minimums locaux selon chaque degré tracé en tiret rouge ne montre pas de relation linéaire entre taille de l_p et D. Il aurait plutôt tendance à confirmer l'utilisation de petites valeurs pour les hyperparamètres.

IV.2.4 Vraisemblance marginale

La vraisemblance marginale a été introduite au début du chapitre précédent, paragraphe III.3. Nous rappelons ici son expression :

$$p(\mathcal{I}) = \int p(\mathcal{I}|\theta)p(\theta)d\theta, \tag{IV.12}$$

elle se situe au dénominateur de la loi *a posteriori* et correspond à une marginalisation des données fournies par les images sur l'ensemble des variations d'attitude admissibles. Ce terme est souvent négligé dans les problèmes de vision, car il ne dépend pas de θ. Son évaluation nécessite généralement l'utilisation de méthodes stochastiques qui peuvent être coûteuses en temps, ou des techniques d'approximations analytiques valables dans des cas restreints. Cette omission est d'autant plus justifiée que, comme nous l'avons remarqué, ce terme est indépendant de la variable θ que nous recherchons. Notons que quelques travaux originaux l'utilisent pour estimer le paramètre de régularisation dans un problème de déconvolution d'image satellitaire [Jalobeanu 01, Jalobeanu 02].
Pourtant, ce terme a une grande importance dans les problèmes d'inférence bayesienne, il permet notamment de comprendre à quel point un modèle peut expliquer les données que nous observons. Un modèle complexe est en mesure de comprendre et d'expliquer des classes de données très variées, mais il risque de proposer une solution complexe en considérant par exemple que le bruit fait partie du signal. Ce phénomène est généralement appelé surapprentissage - ou *overfitting* dans la littérature anglaise -, le modèle ne fait plus la distinction entre la structure du signal et celle du bruit. A l'inverse, un modèle trop simple ne cherche pas à comprendre les fluctuations du signal, qu'il considère majoritairement comme du bruit. La propriété de la vraisemblance marginale est d'incorporer automatiquement cet arbitrage entre modèle simple et complexe, pour prendre une décision optimale à la vue des données. Ce concept a été décrit très clairement dans [Mackay 92] où plusieurs modèles pour l'interpolation de mesures sont dissociés par la

vraisemblance marginale, et très bien exploité dans le cadre des processus gaussiens dans [Rasmussen 06].
Nous faisons référence à cette méthode comme étant une piste de recherche intéressante. De la façon dont nous avons construit nos modèles avec des fonctions de probabilités gaussiennes, la vraisemblance marginale peut avoir une solution analytique. En optimisant cette fonction par un algorithme de type descente de gradient, la recherche des hyperparamètres optimaux peut être bien plus rapide que par les critères de validation croisée ou photométrique. Nous avons conduit quelques tests dans ce sens, mais les résultats sont encore trop prématurés pour être exposés.

IV.2.5 Coefficients du modèle polynômial basse fréquence

Dans notre étude au chapitre précédent, deux estimateurs du MAP se distinguent des autres par leur formulation. En premier, celui de l'équation (III.106) (p.75) qui correspond au modèle de la fusion image et *star tracker* et dont nous rappelons l'expression ci-dessous :

$$\hat{\theta}_{MAP}, \hat{\mathbf{c}}_{MAP} = \underset{\theta,\mathbf{c}}{\text{argmax}}\ p(\theta|\mathcal{I}, \mathbf{c}), \tag{IV.13}$$

avec :

$$p(\theta|\mathcal{I}, \mathbf{c}) \propto p(\mathcal{I}|\theta)p(\mathbf{c}|\theta)p(\theta). \tag{IV.14}$$

Les coefficients $\mathbf{c}$ du modèle polynômial basse fréquence peuvent aussi être compris comme des hyperparamètres estimés à partir des mesures *star tracker* $\mathbf{s}$. La solution proposée au paragraphe III.9.5 est en réalité une approximation de :

$$\hat{\theta}_{MAP} = \underset{\theta,\mathbf{c}}{\text{argmax}}\ p(\theta, \mathbf{c}|\mathcal{I}, \mathbf{s}), \tag{IV.15}$$

par :

$$\hat{\theta}_{MAP} = \underset{\theta}{\text{argmax}}\ p(\theta|\mathcal{I}, \underset{\mathbf{c}}{\text{argmax}}\ p(\mathbf{c}|\mathbf{s})), \tag{IV.16}$$

où plusieurs échelons d'incertitude sont omis comme celui lié au modèle polynômial, ou à l'estimation des coefficients $\mathbf{c}$. Il serait important pour compléter l'étude de comparer les résultats obtenus selon les deux approches ; jusqu'à maintenant notre intuition nous fait penser que la première solution sera bien plus longue et complexe à calculer par rapport à la seconde pour un gain de précision probablement minime.

IV.2.6 Coefficients du modèle autorégressif

Le second estimateur est celui du MAP selon l'*a priori* autorégressif que nous réécrivons ici à partir des équations (III.9) (p.32) et (III.69) (p.58) :

$$\hat{\theta}_{MAP} = \underset{\theta}{\text{argmax}}\ p(\theta|\mathcal{I}, \mathbf{a}), \tag{IV.17}$$

avec :

$$p(\theta|\mathcal{I}, \mathbf{a}) = p(\theta|\mathcal{I})p(\theta|\mathbf{a}). \tag{IV.18}$$

Ces équations sont vraies si nous supposons que $\mathcal{I}$ est indépendant du vecteur des coefficients AR $\mathbf{a}$ - ce que nous admettrons - , et que nous connaissons $\mathbf{a}$ avec certitude. Nous avons expliqué au cours du chapitre précédent que les modes de vibrations d'une structure mécanique peuvent être bien connus. Cependant suivant les conditions de fonctionnement du satellite, ces modes peuvent dévier au cours du temps. Il peut être intéressant

de les estimer pour affiner le modèle, nous les avons donc présumés inconnus dans notre étude. A ce stade, le problème à résoudre devient le suivant :

$$\hat{\theta}_{MAP}, \hat{\mathbf{a}}_{MAP} = \underset{\theta,\mathbf{a}}{\text{argmax}}\ p(\theta, \mathbf{a}|\mathcal{I}). \tag{IV.19}$$

Une inférence complète par des méthodes de type MCMC serait complexe ici : les coefficients sont nombreux, il est difficile d'appliquer un *a priori* sur ceux-ci en dehors du fait qu'ils doivent suivre des conditions de stabilité [Makhoul 75], et on peut ajouter que l'ordre du modèle est indéterminé. Il est plus simple d'utiliser un $\hat{\theta}$ pour pouvoir les estimer par une des différentes méthodes présentées dans [Makhoul 75].
Nous avons vu que l'estimateur du MV montrait déjà des résultats intéressants bien que bruités en haute fréquence, de plus il ne nécessite aucun hyperparamètre. Nous proposons donc d'utiliser $\hat{\theta}_{MV}$ pour estimer le modèle AR, tout en étant sélectif sur la bande de fréquence du signal modélisé pour éviter que le bruit haute fréquence soit appris par le modèle. En pratique, la sélection par bande de fréquence se fait par préfiltrage du signal [Makhoul 75], bien que certaines solutions récentes proposent de s'affranchir de cette étape [Zheng 00]. Dans notre cas nous souhaitons concentrer l'apprentissage des coefficients sur les basses fréquences. Le problème peut se reformuler de la façon suivante :

$$\hat{\theta}_{MAP} = \underset{\theta}{\text{argmax}}\ p(\theta|\mathcal{I}, \underset{\mathbf{a}}{\text{argmax}}\ p(\mathbf{a}|\check{\theta}_{MV})), \tag{IV.20}$$

où $\check{\theta}_{MV}$ est l'attitude estimée par maximum de vraisemblance et filtrée selon ses basses fréquences. On distingue donc clairement trois étapes dans cette approche : l'estimation de l'attitude par maximum de vraisemblance, la détermination des coefficients AR sur $\check{\theta}_{MV}$, puis l'inférence de l'attitude par le MAP. Il existe plusieurs méthodes pour calculer les valeurs de $\mathbf{a}$ à partir du signal, nous choisissons de les estimer à partir des équations de Yule-Walker construites d'après $\check{\theta}_{MV}$, comme étant une solution très couramment utilisée [Makhoul 75].
La figure IV.12 montre les résultats associés à cette technique avec l'estimateur du MAP de l'équation (III.71) (p.60). Le graphique du haut explique bien le contexte d'estimation du modèle AR. Le roulis estimé par maximum de vraisemblance est filtré par un passe bas de fréquence de coupure à -3dB de 25Hz, pour concentrer le modèle autorégressif sur les basses fréquences. On peut remarquer sur le spectre du roulis en gris clair un pic de fréquence autour de 115Hz, qui correspond exactement aux oscillations parasites en hautes fréquences que l'on voit sur la représentation temporelle du roulis par MV au second graphique. C'est une conséquence de la déconvolution, qui amplifie le bruit dans les zones d'ombre où le signal a été perdu par convolution ; nous proposons en annexe une analyse simple de ce phénomène. Le modèle AR en magenta que nous estimons prend assez bien la forme du spectre du roulis en basse fréquence, tout en réduisant son apprentissage des hautes fréquences du signal. En comparaison, le spectre du modèle autorégressif oracle en vert - estimé à partir du roulis réel - reconnait parfaitement les basses fréquences, tout en restant très lisse sur les hautes fréquences puisqu'il n'y a pas de bruit. Le résultat en bas montre que l'estimateur du MAP, avec modèle autorégressif appris sur l'estimé par MV, donne des résultats très satisfaisants. En comparaison de l'estimé par oracle, le résidu est presque nul à un petit décalage près en amplitude. Ajoutons que l'ordre du modèle AR $P^r = 212$ a été estimé sur $\check{\theta}^r_{MV}$ avec le critère d'Akaike [Akaike 74]. Il existe de nombreuses approches pour estimer P comme le précise [Peyrard 01], c'est un

problème complexe qui n'a pas encore de solution idéale. Nous avons porté notre choix sur la méthode d'Akaike pour sa simplicité d'utilisation et sa rapidité d'exécution, elle est assez standard et bien reconnu pour les modèles AR [Makhoul 75], bien qu'elle ait tendance à surestimer l'ordre.

Quelques questions simples peuvent naturellement se poser comme : pourquoi ne pas utiliser directement l'estimateur du MV filtré comme solution au problème, étant donné qu'on doit l'estimer au préalable pour apprendre le modèle AR ? Une partie de la réponse se trouve d'abord sur le graphique du milieu, où $\check{\theta}_{MV}$ est comparé à $\hat{\theta}_{MAP}$. On peut distinguer par instant des petites irrégularités sur le signal bleu $\check{\theta}^{r}_{MV}$, par exemple autour de 0.9 secondes, alors que la courbe rouge reste très régulière comme le doivent être les signaux stationnaires. Il est important de voir que le filtre basse fréquence est relativement lâche dans son contenu spectral, il ne coupe pas complètement le pic de bruit, qui est même légèrement relevé par le modèle AR. Son principe est de maintenir le signal, tout en supprimant ce qui est clairement invraisemblable, comme des oscillations supérieures à $F_e/4$; F_e étant la fréquence d'échantillonnage des caméras *pushbroom*. Il faut également comprendre que le cadre bayesien mélange les images et l'*a priori* de façon optimale en une étape, permettant aux données de prendre le dessus sur l'*a priori* si celui-ci se révèle mauvais et vise versa. Le MV filtré pour sa part est construit en deux étapes qui ne cherchent pas à savoir à un moment si l'une est plus importante que l'autre. En soi, $\check{\theta}_{MV}$ se rapproche plus de l'estimateur du MAP avec *a priori* sur la dérivée seconde, mais dans une version sous optimale et également plus rapide à calculer ; il n'y a pas d'hyperparamètre à rechercher dans ce cas.

L'autre partie de la réponse se comprend dans le concept même de l'apprentissage, où une fraction des données est utilisée pour déduire et apprendre un modèle ou des paramètres rapidement, l'inférence complète étant réalisée sur toutes les données. Suivant cette idée, $\check{\theta}_{MV}$ est seulement estimé sur une portion des données. Malheureusement le modèle AR nécessite une certaine période d'observation pour apprendre les basses fréquences, ce qui peut être contradictoire avec l'idée d'apprentissage ; dans l'exemple 0.9 secondes ont été prises pour estimer le modèle AR. En pratique, le nombre de pixels en hauteur qui correspond à la taille du capteur *pushbroom* peut être limité et il n'est pas utile d'apprendre sur une trop longue période car la restitution de très basses fréquences par recalage d'image est difficile. Dans ce cas extrême, le *star tracker* complètera mieux le schéma d'estimation que l'*a priori*. Enfin d'après les hypothèses sur les variations d'attitude, le modèle autorégressif estimé à un moment sera valable pour de très nombreuses images par la suite.

Une autre question peut concerner la fréquence de coupure du filtre, qui a été choisie de façon empirique ici. Par véritable expertise, soit dans la meilleure des situations on connait les fréquences de résonance, auquel cas notre schéma d'estimation ne servirait qu'à confirmer notre croyance, sinon il est toujours possible d'avoir une vague idée de la bande de fréquence d'intérêt. Enfin ajoutons que les pics de bruit peuvent être isolés de façon précise sur le spectre étant donné que nous connaissons les noyaux de convolution k_{ij}.

IV.3 Recalage multimodal

Nous avons jusqu'ici considéré que les images sont enregistrées sur une même bande spectrale de la lumière, alors que le satellite utilise des caméras de différentes sensibilités

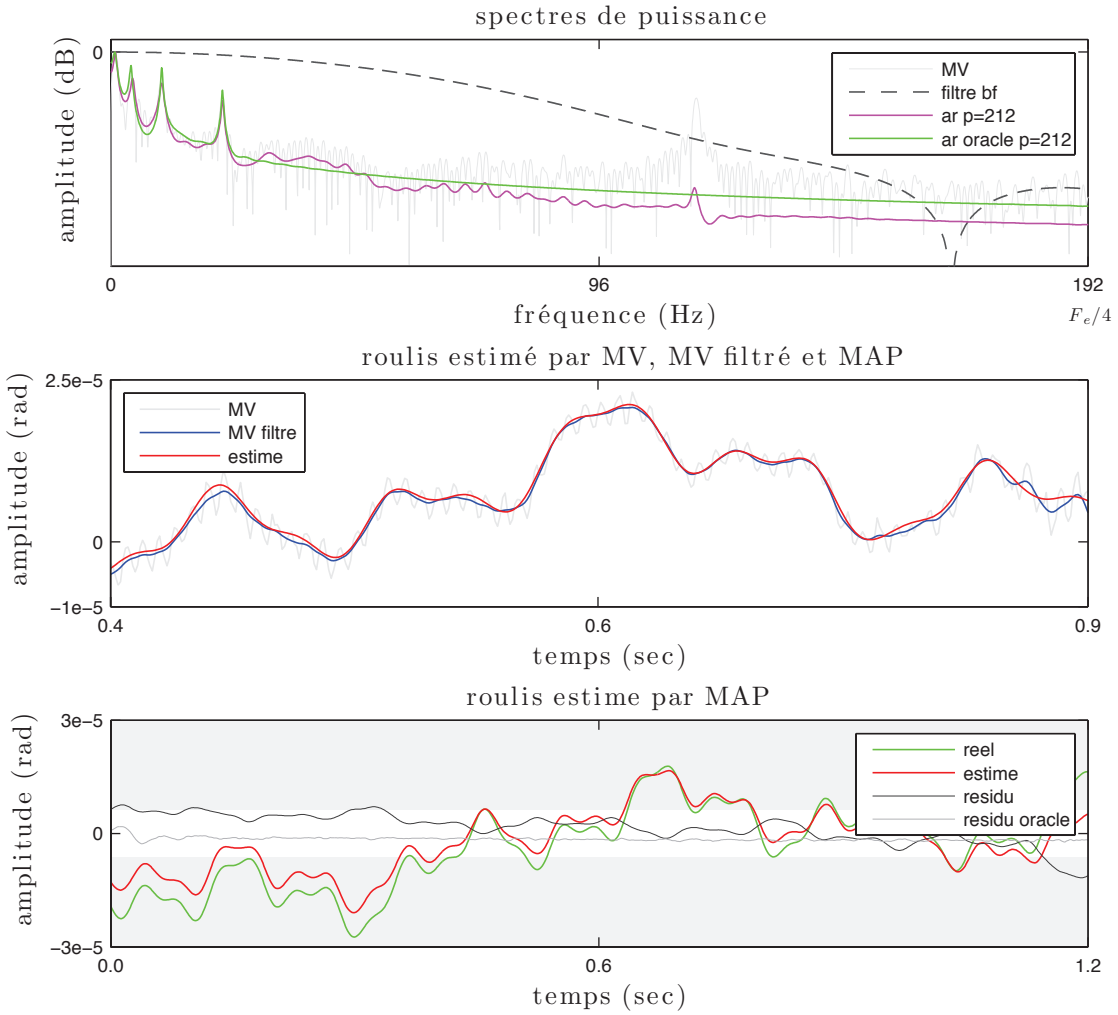


FIGURE IV.12 – Exemple d'estimation du roulis avec l'estimateur du MAP de l'équation (IV.20) (p.93), le second terme de vraisemblance des images (III.37) (p.41) est ici utilisé. Le modèle autorégressif est appris sur une période de 0.9 secondes, d'après une portion des images de taille (300×700), les hyperparamètres sont déduits du même jeu d'apprentissage avec $\sigma_I = 0.012$ et $\sigma_{\theta^r} = 5.7 \times 10^{-9}$. La figure du haut montre les spectres de $\hat{\theta}^r_{MV}$ en gris clair, du filtre basse fréquence en noir et des modèles AR estimés sur les données et par oracle. L'ordre $P_r = 212$ est déterminé par le critère de Akaike. Le second graphique permet de comparer au cours du temps l'estimateur du MV, son filtrage et l'estimateur du MAP ; il est intéressant de noter que les petites oscillations du MV sont directement liées au pic autour de 115Hz visible dans son spectre. La troisième figure montre l'estimation du MAP avec le modèle AR inféré plus haut, le résultat est sensiblement identique au MAP avec modèle AR appris sur le roulis réel (oracle).

au spectre lumineux pour construire des images en couleur. La première hypothèse du paragraphe III.4.1 selon laquelle le bruit entre deux images est gaussien et centré devient fragilisée. Pour s'en convaincre, la figure IV.13 propose d'observer la différence entre les différents canaux de couleur d'une image prise par un appareil photo classique type *pinhole*. Si on néglige les effets du dématriçage lié au filtre du capteur photosensible - ce filtre est souvent de type Bayer [Bayer 76] -, on peut considérer que les trois images rouge, verte et bleue sont parfaitement recalées.
Les images de différence à la troisième ligne du tableau montrent que contrairement à notre hypothèse, le bruit caractéristique n'est pas blanc mais plutôt très structuré. La plupart des zones homogènes ressortent et font par la même occasion apparaitre les contours des objets. Les histogrammes des erreurs observées à la dernière ligne du tableau ne sont pas simples à interpréter dans leur forme, mais on peut dire que la tentative de faire correspondre une distribution gaussienne à ces données n'est pas réussie.
On peut tout de même relever quelques points positifs. Les trois images sont globalement très ressemblantes, il est même assez difficile à première vue de faire une distinction entre chacune. Les effets d'ombre sont plus marqués sur le canal bleu, et le grand écran semble très lumineux dans cette même image. Certains objets ont une intensité lumineuse perçue qui est inversée entre les canaux, comme la tasse orange en haut à droite, ou le boitier à cd en bas à gauche. Cependant une grande majorité des objets est présente sur les trois images ; les contours sont identiques et seule l'intensité lumineuse à laquelle ils sont perçus varie. Cette impression se confirme en observant plus attentivement les différences entre canaux, où les objets transparaissent plus par leur différence d'intensité dans les zones homogènes que par des contours qui seraient non existants dans une des deux images [1]. On peut soulever un dernier détail intéressant sur la forme des distributions, où la différence entre canaux rouge et bleu se distingue des deux autres. Ceci peut être lié au filtre Bayer qui possède deux fois plus de photosites dans la bande spectrale verte que dans les deux autres bandes. Ainsi le rapport signal sur bruit de l'image verte est meilleur, ce qui aurait tendance à réduire la variance des distributions auxquelles elle est impliquée. Ce point n'est *a priori* pas transposable à notre cas, chaque caméra *pushbroom* possède un nombre identique de photosites.
Les remarques concernant la figure IV.13 peuvent s'expliquer par le contexte physique de l'acquisition : les modalités observées sont très rapprochées dans le spectre de la lumière visible, les informations communes sont donc nombreuses. Le problème auquel nous sommes confronté est très différent des phénomènes multimodaux complexes régulièrement observés en imagerie médicale lorsque des données obtenues par différents appareils doivent être mises en correspondance [Viola 97, Roche 98]. L'imagerie satellitaire propose aussi des solutions pour des cas de recalage difficiles avec variations de luminosité [Foroosh 02] et utilisation de la bande infrarouge. Beaucoup de solutions ont été proposées et elles dérivent pour la plupart de deux concepts : l'information mutuelle [Viola 97] et la corrélation croisée [Zitov 03]. Le premier est très coûteux en temps de calcul, et doit être privilégié pour des cas de mise en correspondance de données radicalement opposées par leur nature ; par exemple un modèle 3-D avec une image 2-D. Le second a montré à plusieurs reprises de très bons résultats expérimentaux tout en étant très rapide à calculer - bien que cet estimateur n'apparaisse pas comme le plus précis se-

1. Au delà de notre problème et de ce simple constat, la dualité entre contour et région est à la base de nombreux algorithmes de segmentation d'image, nous reportons le lecteur à [Cocquerez 95] pour une bonne introduction à ces problèmes.

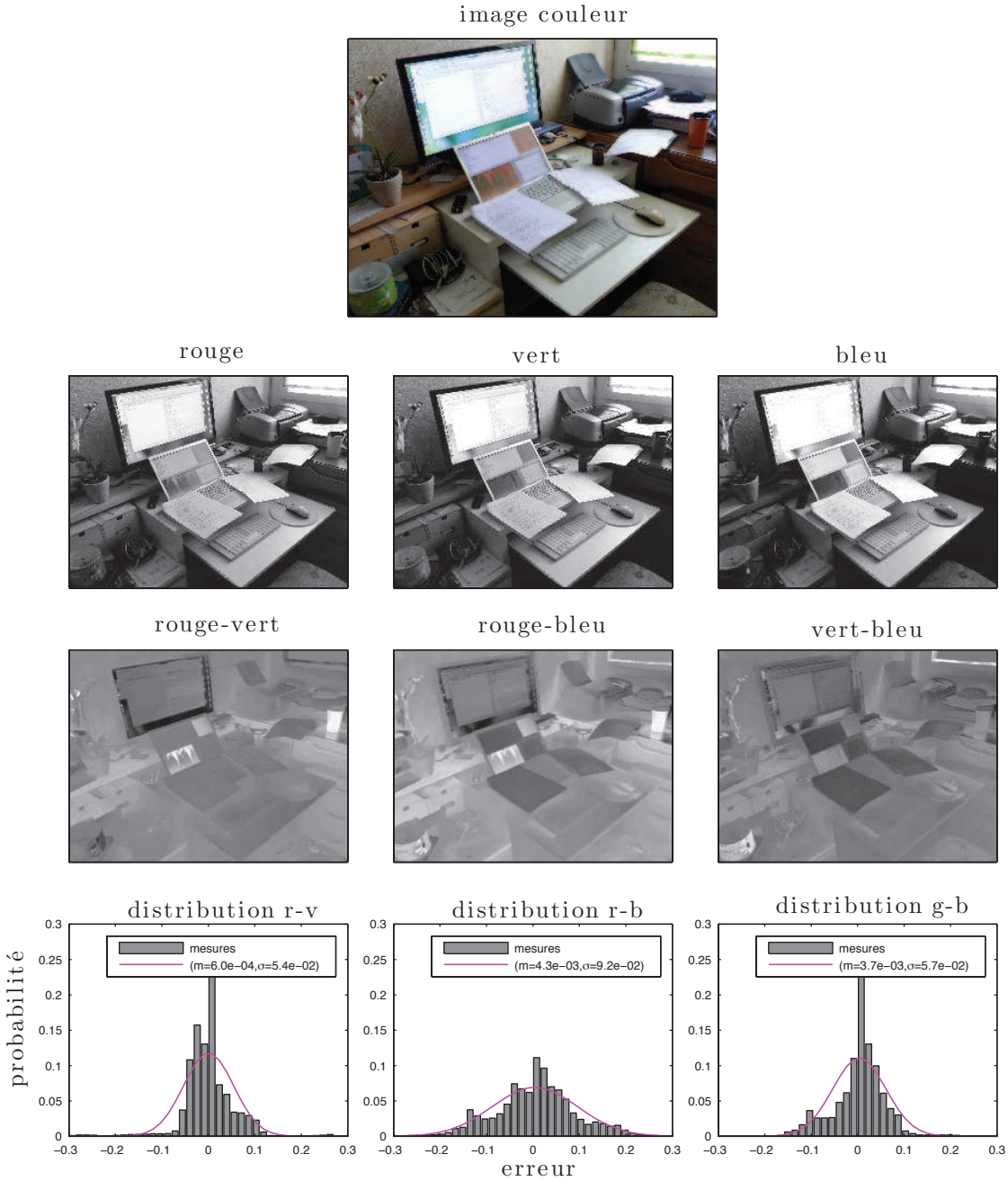


FIGURE IV.13 – Exemple d'image couleur de taille (1000×1300) décomposée selon ses trois canaux rouge, vert et bleu (seconde ligne du tableau). Les différences de luminosité entre chaque paire de modalité sont représentées en troisième ligne. Les densités de probabilité des erreurs sont tracées à la dernière ligne. La dynamique des images originalement de 0 à 255 a été recadrée sur l'intervalle $[0, 1]$. Les histogrammes sont construits par décompte du nombre d'échantillons appartenant aux 64 intervalles entre $[-0.5, 0.5]$. Le modèle théorique de loi normale en magenta est estimé par la fonction `normfit.m` de Matlab.

lon cette étude théorique intéressante [Robinson 04] - , mais le choisir reviendrait à considérer notre étude de la même façon que ce que nous avons exposé au chapitre 2 au sujet des points d'intérêts. Plusieurs solutions simples ont été proposées pour l'algorithme du Lucas Kanade [Baker 04, Irani 98], qui est le fondement de notre approche.

IV.3.1 Images brutes

Nous venons de voir que les modalités enregistrées par chaque caméra étaient très proches, et contenaient des informations presque identiques. La première idée consiste donc à traiter les images brutes sans tenir compte de leurs différences. Le jeu de données que nous avons utilisé jusqu'à maintenant était calibré radiométriquement - même modalité pour chaque caméra -, à chaque fois le recalage s'est fait en considérant les images brutes. Nous choisissons donc un nouveau jeu de données possédant quatre images multispectrales de taille (900×1000) : panchromatique, rouge, verte et bleue. La figure IV.14 montre le résultat obtenu pour l'estimation du roulis avec un *a priori* de type processus gaussien, l'estimateur du MAP utilise ici la seconde vraisemblance des images.
Le résultat obtenu est plutôt satisfaisant, la courbe de résidu est régulière ce qui fait pen-

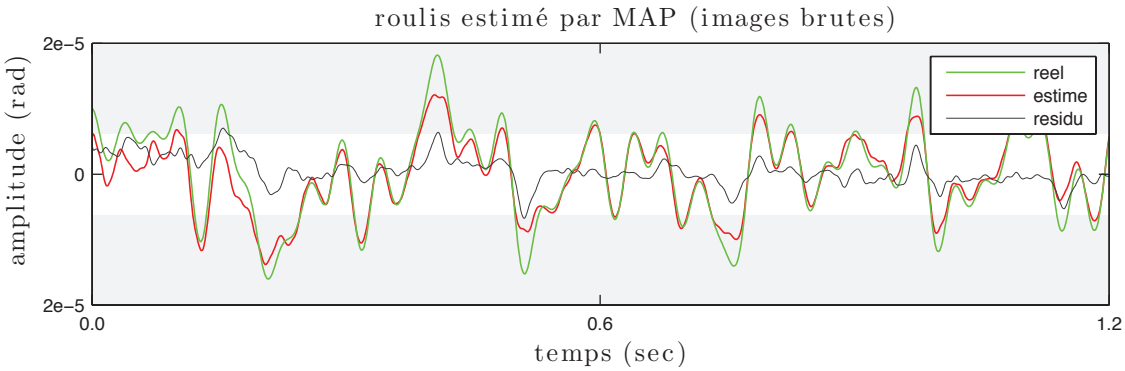


FIGURE IV.14 – Exemple d'estimation du roulis par MAP et *a priori* de processus gaussien sur des images multispectrales brutes ; $\sigma_I = 0.01$, $\sigma_{\theta^r} = 1.128$ et $l^r = 0.008$. L'écart-type du résidu est de 2.1×10^{-6}, globalement et malgré le caractère multimodal des données, l'estimateur se comporte assez bien avec une courbe de résidu sans artéfact particulier.

ser que certaines inversions de contraste importantes ne biaisent pas réellement l'estimé. Les limitations semblent être dues uniquement à la linéarisation du terme image.

IV.3.2 Images de contours

Une partie des méthodes de recalage multimodal propose d'utiliser des éléments invariants dans les images qui ne sont pas sujets aux changements de luminosité ou aux inversions de contraste. Les contours entre autres appartiennent à cette famille d'éléments, et ont été exploités à plusieurs reprises dans le contexte multimodal [Irani 98] ou pour tenter d'améliorer les performances du recalage [Baker 04]. Afin d'extraire ces informations saillantes et supposées communes dans les deux images, on peut procéder par un filtrage passe haut de celles-ci avant de les mettre en correspondance. Outre la rapidité qu'implique cette opération de filtrage - c'est une convolution -, on peut voir que l'information basse fréquence est éliminée dans les deux images, et par la même occasion les

inversions de contraste. La figure IV.15 permet de s'en apercevoir au niveau de l'image de la tasse ; alors que le contraste est important entre les deux images, les filtres font disparaitre cette information contradictoire.
Il existe plusieurs façons d'extraire les hautes fréquences d'une image, parmi les plus simples on peut relever les méthodes utilisant le premier ordre de la dérivée de l'image :

$$I_c(x,n) = \sqrt{(\mathbf{d}_x * G * I)(x,n)^2 + (\mathbf{d}_n * G * I)(x,n)^2}. \tag{IV.21}$$

Cette équation est l'expression de l'opérateur de Canny [Canny 86, Cocquerez 95], qui propose de calculer la norme des gradients selon x et n de l'image filtrée par G. Les vecteurs $\mathbf{d}_x = [1\ 0\ -1]^T$ et $\mathbf{d}_n = [1\ 0\ -1]$ sont les opérateurs différentiels qui approchent les dérivées partielles de l'image numériquement. La matrice G contient les coefficients d'une gaussienne en deux dimensions, elle est généralement de taille inférieure à (11×11) et réalise le filtrage basse fréquence de l'image I pour améliorer la robustesse du détecteur face au bruit. Il est possible d'adopter d'autres opérateurs pour le calcul du gradient ou du filtrage, mais l'équation de base (IV.21) reste la même.
Le détecteur de Marr-Hildreth [Marr 80] utilise pour sa part la dérivée seconde de l'image :

$$I_c(x,n) = |(D_2 * G * I)(x,n)|, \tag{IV.22}$$

où D_2 est une matrice (3×3) calculant le laplacien numérique de l'image par convolution, et G est à nouveau un filtre de lissage gaussien. Notons qu'ici nous souhaitons conserver les images énergétiques en sortie de ces équations, nous ne recherchons pas de carte de contours par seuillage, ce qui est généralement la seconde étape dans la détection de contours [Cocquerez 95] et a parfois été envisagée pour le recalage d'images.
En analysant de près les résultats de l'expérience en figure IV.15, on peut voir certaines différences dans les deux méthodes. Globalement, les informations communes entre les deux images ressortent bien dans les deux cas, cependant les contours par Canny sont plus nombreux et plus nets que par Marr-Hildreth. C'est une des raisons pour laquelle le premier est préféré au second dans le contexte général de détection de contours. Mais au regard des erreurs entre les images filtrées par chaque méthode, le résidu produit selon Marr-Hildreth semble plus s'approcher d'un bruit blanc que celui obtenu par Canny. En réalité, ce dernier est plus sensible à tout type de contour car son filtre passe haut est moins sélectif que celui construit selon la dérivée seconde. Il détecte notamment les effets spéculaires sur la droite de la tasse, qui sont eux-mêmes plus ou moins marqués selon la modalité observée. Ainsi les contours d'un même objet peuvent être légèrement décalés par des effets de lumière, ce que montre clairement l'image d'erreur après filtrage de Canny. Le détecteur de Marr-Hildreth se contente de faire ressortir les contours les plus marqués, mais par conséquent il propose moins d'information pour le recalage. Une nouvelle fois, les distributions de probabilité ne prennent pas la forme d'une distribution normale, par contre assez curieusement nous retrouvons une distribution typique des gradients d'une image naturelle alors que nous traitons ici de l'erreur entre deux images. Cette distribution est de type queue lourde où une grande partie des valeurs sont concentrées autour de zéro, tout en attribuant une probabilité plus importante aux grandes valeurs par rapport à une distribution normale. Cette statistique a été observée dans [Field 94] et élégamment exploitée dans [Fergus 06] comme *a priori* pour la déconvolution aveugle d'image.
La figure IV.16 présente les résultats obtenus par ces deux approches sur le même jeu de données que celui utilisé au paragraphe précédent. Nous avons choisi une taille de (9×9)

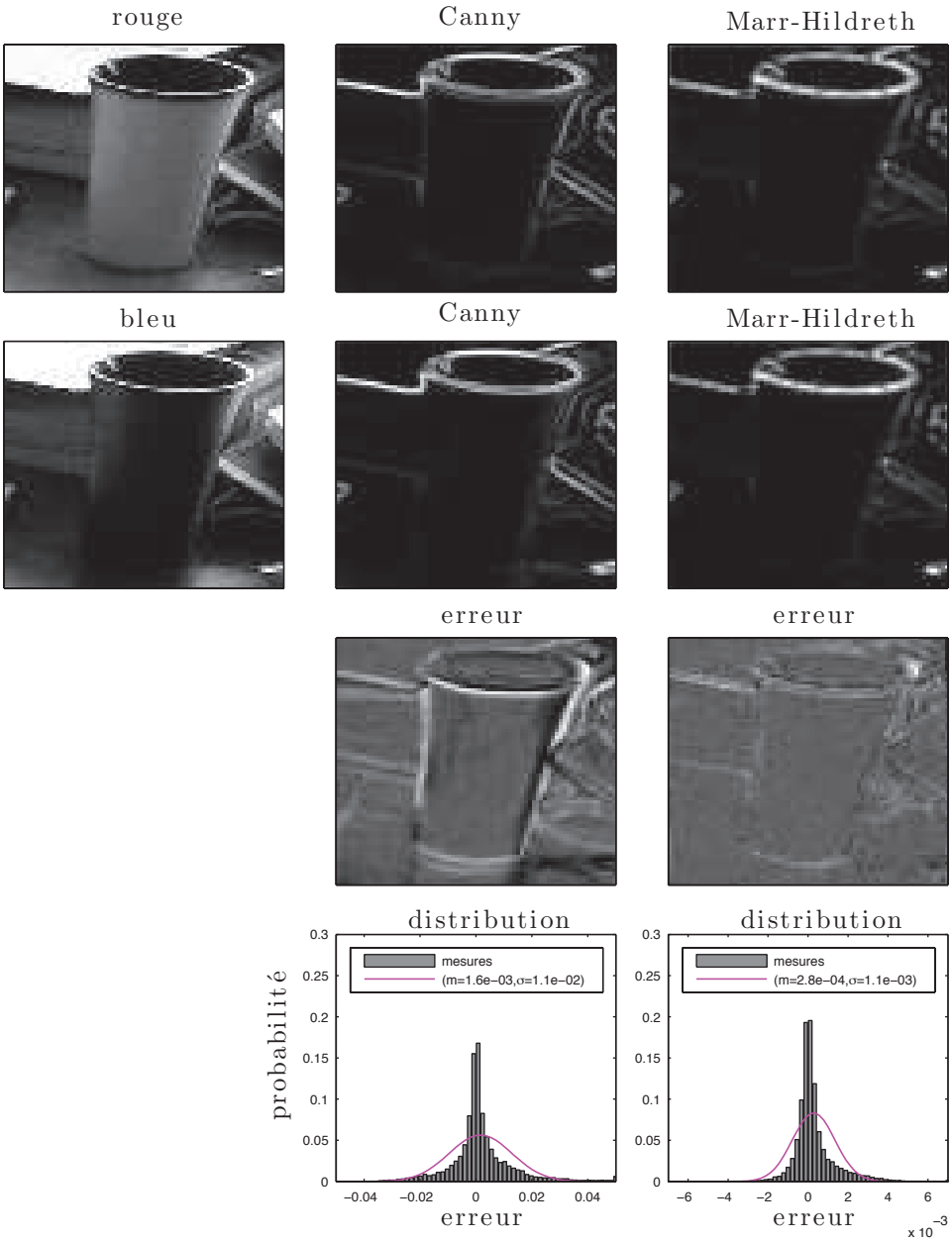


FIGURE IV.15 – Exemple de comparaison des modalités rouge et bleu sur une image (150×150) par extraction des contours. La première colonne correspond aux images brutes dont les intensités s'étendent entre $[0, 1]$ (elles sont comprises entre $[0, 255]$ à l'origine), la seconde au traitement fait par la méthode de Canny [Canny 86] et la troisième par la méthode de Marr-Hildreth [Marr 80]. Les histogrammes sont calculés sur 64 intervalles entre $[-0.05, 0.05]$ et $[-0.007, 0.007]$ respectivement pour les deux cas. Les modèles de loi normale en magenta appliqués sur les distributions sont estimés à partir des images d'erreur par l'algorithme `normfit.m` de Matlab.

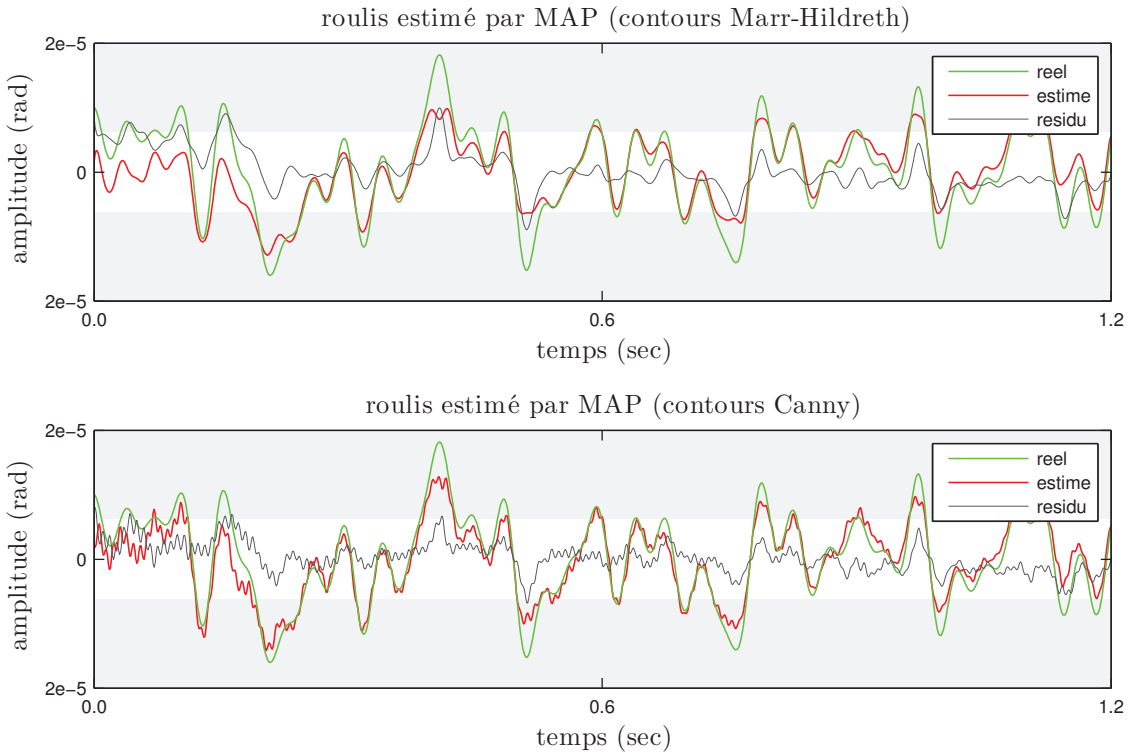


FIGURE IV.16 – Exemple d'estimation du roulis par MAP et *a priori* de processus gaussien sur des images multispectrales filtrées. La figure du haut montre le résultat d'estimation après extraction des contours par Marr-Hildreth, les hyperparamètres estimés sont $\sigma_I = 0.0003$, $\sigma_{\theta^r} = 0.088$ et $l^r = 0.014$. Celle du bas correspond à l'estimation avec les contours de Canny où les hyperparamètres sont cette fois-ci $\sigma_I = 0.0017$, $\sigma_{\theta^r} = 0.014$ et $l^r = 0.0043$. La seconde méthode est sensiblement meilleure que la première avec un écart-type des résidus de 2.3×10^{-6} contre 3.1×10^{-6} pour la première, mais les résultats restent légèrement en deçà de ceux de la figure IV.14.

pour le filtre basse fréquence, elle donne le meilleur résultat sur l'écart-type du résidu. En comparaison à la figure IV.14, les résultats sont un peu moins bons visuellement, et ceci est confirmé par l'écart-type des résidus qui est de 0.18 et 0.25 en unité de taille pixel pour les contours de Marr-Hildreth et Canny respectivement, contre 0.17 pour le traitement direct des images brutes. Selon ces valeurs, la méthode de Canny semble préférable à celle de Marr-Hildreth, alors même que nos observations de la figure IV.15 tendaient à affirmer le contraire. Cependant, la forme du résidu dans le graphique du bas fait apparaitre un bruit haute fréquence qui n'est pas souhaitable dans notre cas.
Ces résultats peuvent s'expliquer de différentes façons. Les deux approches procèdent par filtrage pour révéler les hautes fréquences, il en résulte assez souvent une amplification du bruit sur cette gamme de fréquence, qui est ensuite rectifiée par un lissage de l'image selon un filtre passe bas. Les contours ont alors tendance à s'épaissir en perdant un peu en localisation, certains détails peuvent être perdus par le lissage et le bruit n'est pas strictement filtré car il est réparti sur les pixels voisins. Nous avons également vu sur la figure IV.15 que les contours peuvent ne pas correspondre complètement suivant les effets d'ombre et de spécularité. Ces éléments peuvent induire en erreur le recalage et dégrader les performances d'un estimateur.

IV.3.3 Correction radiométrique

La solution basée sur la correction radiométrique part du principe qu'il existe une fonction liant les intensités perçues entre deux images. Soit $f : \mathbb{R} \to \mathbb{R}$ cette fonction, nous pouvons écrire la relation suivante pour la paire d'images i et j :

$$I_i(x, n) = f(I_j(x, n)). \tag{IV.23}$$

Il est difficile d'imaginer la forme que devrait prendre une telle fonction, pour cette raison des modèles assez simples de type polynômial ont été utilisés pour rectifier des variations de luminosité [Baker 04, Szeliski 06], ou pour le recalage multimodal en imagerie médicale [Guimond 01]. Les coefficients peuvent être estimés par régression linéaire entre les deux images avant recalage, en supposant qu'à l'origine les déformations ne sont pas d'une grande amplitude. Notre contexte est donc favorable à cette solution, soit r_d les coefficients du polynôme pour $d \in [0 : D]$, l'équation (IV.23) peut se réécrire :

$$I_i(x, n) = \sum_{d=0}^{D} r_d I_j(x, n)^d. \tag{IV.24}$$

La figure IV.17 propose une nouvelle fois d'observer cette solution sur l'image de la tasse, où les intensités de l'image rouge sont mises en correspondance avec celles de l'image bleue par un polynôme de degré trois. Il est intéressant de voir que la couleur de la tasse a bien été rattrapée par le modèle de correction radiométrique, en revanche la luminosité de la scène à l'arrière plan est fortement assombrie, alors même qu'elle semblait bien correspondre initialement entre les deux modalités. L'erreur entre l'image rouge rectifiée et sa correspondante bleue confirme que le modèle n'est pas en mesure de suivre toutes les variations d'intensité. La distribution de probabilité des erreurs prend une forme particulière, en suivant presque la loi normale estimée en magenta tout en présentant un pic aberrant proche de 0.2. On touche ici la limite de cette méthode qui n'est pas suffisamment locale pour expliquer certaines variations d'intensité, mais elle est simple et permet de maintenir un maximum d'informations dans les images, notamment dans celles dont

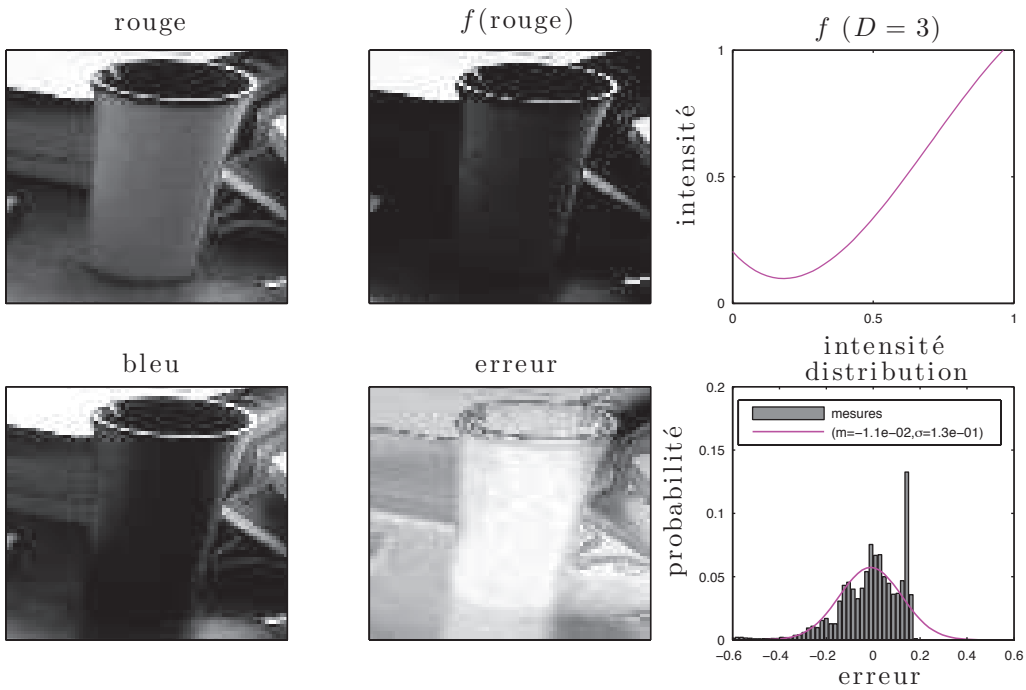


FIGURE IV.17 – Exemple de correction radiométrique sur l'image rouge à partir de la modalité bleue f(rouge). Le modèle du graphique en haut à droite est un polynôme de degré 3 selon l'équation (IV.24), qui associe à une intensité rouge entre $[0, 1]$ sa nouvelle valeur. L'erreur associée compare l'image bleue avec f(rouge) sur la figure au milieu en bas. L'histogramme des distributions est calculé sur 64 intervalles entre $[-0.6, 0.6]$, et le modèle normal est estimé par la méthode `normfit.m` de matlab.

la radiométrie n'est pas rectifiée.
Les résultats associés au jeu de données multispectral sont présentés en figure IV.18.

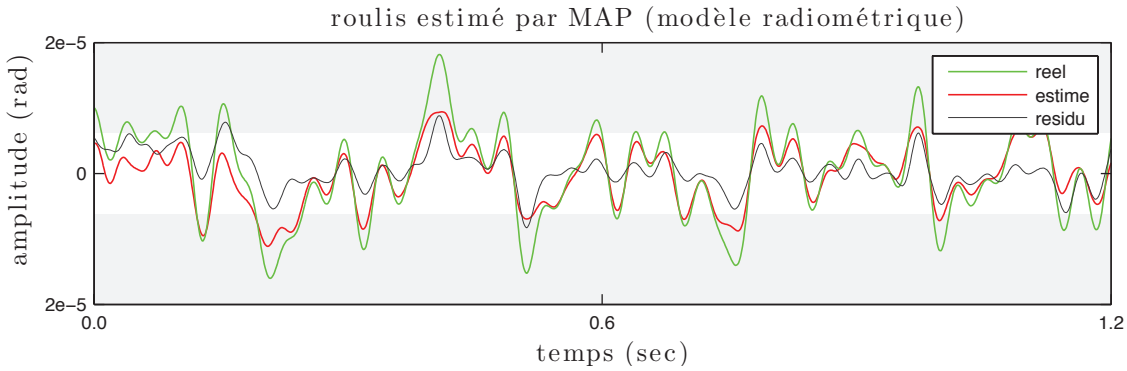


FIGURE IV.18 – Exemple d'estimation du roulis par MAP et *a priori* de processus gaussien, sur des images multispectrales de radiométrie rectifiées par un polynôme selon l'équation (IV.24). Les hyperparamètres sont $\sigma_I = 0.013$, $\sigma_{\theta^r} = 0.014$ et $l^r = 0.009$. L'écart-type du résidu est de 2.8×10^{-6}, ce qui positionne cette solution entre les deux proposées à la figure IV.16. Cependant la forme du résidu est meilleure car elle est moins parasitée par des fréquences du bruit.

Afin de rendre un peu plus locale la correction radiométrique, nous avons estimé des polynômes de degré 1 suivant l'équation (IV.24) sur des fenêtres successives selon l'axe n de taille $(L \times 10)$. Ce choix est relativement empirique, mais nous souhaitions conserver une vitesse de calcul comparable aux méthodes précédentes avec un modèle relativement simple. Une nouvelle fois cette solution est moins performante que le traitement direct des images brutes, avec un écart-type de 0.23 rapporté à la taille d'un pixel, mais l'estimé est un peu meilleur que ceux proposés en figure IV.16, notamment parce qu'il n'amplifie pas de fréquence parasite.

IV.3.4 Discussion

Les résultats que nous obtenons confirment nos observations du début du paragraphe IV.3, selon lesquelles les images spectrales sont très ressemblantes et que notre problème de recalage multimodal n'est pas aussi important que ceux vus dans d'autres disciplines telles que l'imagerie médicale. Encore une fois, notre contexte où les images sont prises à des écarts de temps très court rend peu probable des changements de luminosité importants dans la scène photographiée. La meilleure solution semble être la plus simple : conserver les images brutes.
Globalement, les expérimentations que nous avons menées ont tendance à confirmer ces résultats, mais ils ne sont que préliminaires et nécessitent plus de recul avec des traitements sur de nombreuses données. Il est intéressant de noter également que contrairement à une majorité d'algorithmes de recalage qui utilisent une image de référence pour rectifier une image déformée, notre approche est symétrique dans la mise en correspondance de par le contexte. Chaque image joue alternativement le rôle de référence à laquelle les autres sont comparées, et il est possible qu'un biais provoqué dans un sens soit rectifié lorsque le rôle des images est inversé. Les tests que nous avons faits jusqu'à maintenant confirment aussi l'importance de prendre en compte chaque paire d'images,

mais ils ne sont pas encore suffisants pour valider cette intuition. Cette observation est peut-être à mettre en lien avec une petite classe d'algorithmes de recalage qui propose d'utiliser les deux images alternativement comme des références pour la mise en correspondance [Authesserre 09], l'auteur montre que cette approche est particulièrement robuste pour des images très bruitées. Enfin, cette idée peut être mise en relation à la solution de [Farsiu 05] au problème de super résolution ; chaque paire d'images est contrainte dans sa déformation pour mieux se fusionner en une unique image super résolue.
On peut déjà affirmer assez logiquement que le recalage multispectral est un peu moins bon que le monomodal, mais dans notre cas l'estimation des hyperparamètres est plus spécifiquement instable comme en témoigne la figure IV.16. La valeur du facteur d'échelle l^r estimé pour les contours de Canny est trop petite, ce qui provoque les oscillations hautes fréquences dans la courbe rouge du roulis. Un meilleur apprentissage des paramètres de l'*a priori* nous aurait certainement conduit à des résultats similaires à la figure IV.14. La méthode de correction radiométrique est moins touchée par ce problème, et le résultat est assez satisfaisant si on considère que nous sommes restés à un niveau assez simple d'exploitation de ce modèle. Une étude plus prolongée permettrait de choisir un degré de polynôme et une taille de fenêtre adaptée pour que le modèle soit valide localement. Tout ceci doit se faire en considérant les temps de calcul qu'implique une telle opération. Enfin, nous n'avons pas parlé de l'image panchromatique dans l'introduction de cette section, bien qu'elle ait été prise en compte dans tous les traitements que nous avons faits. Dans son principe, la caméra panchromatique est sensible à un spectre du visible large qui inclut les modalités rouge, verte et bleue. Une piste de recherche originale serait d'observer si cette image peut être vue comme une combinaison linéaire des trois autres, et de voir si le recalage multi-image est amélioré par ce principe.

IV.4 Estimation itérative

Nous avons vu au cours du chapitre précédent que la linéarisation du terme de l'image déformée était valable pour des déplacements inférieurs à la taille du pixel. Jusqu'à maintenant, tous les exemples de résultats obtenus confirment cette limitation. Pour la repousser, les méthodes fondées sur l'algorithme de Lucas-Kanade procèdent par estimation itérative où les images sont rectifiées avec un estimé courant à chaque itération, le résultat final étant la somme de tous les estimés. Ce procédé est très répandu pour de nombreux problèmes même en dehors de la vision par ordinateur. Il fait partie de la classe des algorithmes d'optimisation selon le gradient, et l'algorithme de Lucas-Kanade utilise plus particulièrement la méthode de Gauss-Newton [Baker 04].
Le schéma général de l'algorithme que nous adoptons est celui de la figure IV.19, où t est le compteur des itérations et $\mathcal{M}$ l'estimateur qui a été choisi parmi le MV, le MAP avec ses différents *a priori*, le modèle polynômial par morceaux et la fusion avec les données *star tracker*. Cette fonction prend en entrée l'ensemble des données disponibles $\mathcal{I}$ pour les images et $\mathbf{s}$ pour le *star tracker* avec ses hyperparamètres associés $\boldsymbol{\lambda}$, pour déterminer l'attitude :

$$\hat{\theta} = \mathcal{M}(\mathcal{I}, \mathbf{s}, \hat{\boldsymbol{\lambda}}). \tag{IV.25}$$

Nous conservons $\mathbf{s}$ en entrée de tous les modèles pour assurer le côté générique de la formule, bien que seule la fusion exploite ces mesures. De même, nous réaffirmons que les coefficients $\mathbf{a}$ pour le modèle autorégressif et $\mathbf{c}$ pour le modèle polynômial de la fusion sont des hyperparamètres associés à leur estimateur. Malgré le choix d'une méthode

d'inférence qui diffère de celles des paragraphes IV.2.2 et IV.2.3, ils appartiennent bien au vecteur $\boldsymbol{\lambda}$. Nous résumons au tableau IV.1 l'ensemble des estimateurs proposés avec les hyperparamètres qui les composent, ainsi que les équations ou méthodes d'inférence suggérées pour les utiliser.

Dans l'algorithme de la figure IV.19, $\mathcal{I}^w$ correspond à l'ensemble des images filtrées -

- ◇ Initialisation des variables :
  $\boldsymbol{\theta} = \mathbf{0}$, $\Delta\boldsymbol{\theta} = \mathbf{0}$, $t = 0$

- ◇ Pour $\alpha \in \{r, t, l\}$,
  - ○ estimer les hyperparamètres $\boldsymbol{\lambda}^\alpha$ associés à $\mathcal{M}$ (tableau IV.1),
  - ○ faire :
    1. déformer les images $\mathcal{I}$ avec l'estimé courant $\boldsymbol{\theta}$ et corriger leur radiométrie (filtrage ou modèle) pour donner $\mathcal{I}^w$,
    2. estimer la mise à jour de l'attitude : $\Delta\boldsymbol{\theta}^\alpha = \mathcal{M}(\mathcal{I}^w, \mathbf{s}^\alpha, \boldsymbol{\lambda}^\alpha)$,
    3. mettre à jour l'attitude : $\boldsymbol{\theta}^\alpha = \boldsymbol{\theta}^\alpha + \Delta\boldsymbol{\theta}^\alpha$
  - ○ tant que $|\Delta\boldsymbol{\theta}^\alpha| > \delta$.

FIGURE IV.19 – Algorithme d'estimation itératif.

seulement s'il est décidé de ne pas utiliser les données brutes directement - et interpolées par une méthode bicubique aux nouvelles positions de pixel données par $\boldsymbol{\theta}$; nous précisons bien ici $\boldsymbol{\theta}$ plutôt que $\boldsymbol{\theta}^\alpha$ afin que le tangage puis le lacet profitent des estimations précédentes. Nous choisissons d'adopter une optimisation indépendante selon chaque angle pour plusieurs raisons. Les matrices prennent une place mémoire importante et au delà d'une certaine limite, la mémoire vive sature et force l'ordinateur à basculer sur la mémoire de masse, ce qui ralentit considérablement les calculs. De même, l'estimation de tous les hyperparamètres peut devenir impraticable à des dimensions trop importantes. Enfin nous avons observé par expérimentation que l'estimation seule du roulis était pratiquement identique à celle faite en considérant tous les angles à la fois. Une fois les images rectifiées du roulis, l'estimation du tangage et de ses hyperparamètres deviennent mieux conditionnés, et il en va de même du lacet. Il est possible d'ajouter une dernière étape d'estimation globale, mais celle-ci n'améliore pas de façon significative le résultat. L'algorithme prend globalement la forme d'une descente de gradient incrémentielle où les paramètres sont optimisés indépendamment, avec une méthode de Gauss-Newton dans sa boucle interne. Enfin, le δ est une condition d'arrêt sur la norme de la mise à jour de l'attitude que nous fixons à 10^{-4}. Pour toutes les expériences qui suivent, nous avons limité l'algorithme à 7 itérations pour chaque angle si le critère d'arrêt n'est pas atteint pendant la convergence.

IV.5 Résultats sur des données

Nous présentons ici des résultats sur des données satellitaires que la société EADS Astrium nous a fait parvenir pour mener nos expériences. Elles se composent de quatre

Données	Estimateur	*A priori*	$\boldsymbol{\lambda}$	Estimation $\boldsymbol{\lambda}$	$p(\mathcal{I}\|\boldsymbol{\theta})$	Estimation $\boldsymbol{\theta}$
images $\mathcal{I}$	Maximum de vraisemblance				éq. (III.32), §III.4.6	éq. (III.35), §III.4.6
					éq. (III.37), §III.4.6	éq. (III.39), §III.4.6
	Polynômial par morceaux		$[l_p, D]^T$	Validation Croisée (VC) §IV.2.2 ou Validation Photométrique (VP) §IV.2.3	éq. (III.32), §III.4.6	éq. (III.84), §III.8.3
					éq. (III.37), §III.4.6	éq. (III.86), §III.8.3
	Maximum *a posteriori*	Amplitude, éq. (III.41), §III.5.2	$[\sigma_{\theta^l}, \sigma_{\theta^r}, \sigma_{\theta^t}]^T$		éq. (III.32), §III.4.6	éq. (III.42), §III.5.2
					éq. (III.37), §III.4.6	éq. (III.46), §III.5.2
		Dérivée seconde, éq. (III.50), §III.5.3			éq. (III.32), §III.4.6	éq. (III.53), §III.5.3
					éq. (III.37), §III.4.6	éq. (III.56), §III.5.3
		Processus gaussien, éq. (III.60), §III.5.4	$[\sigma_{\theta^l}, \sigma_{\theta^r}, \sigma_{\theta^t}, l^l, l^r, l^t]^T$		éq. (III.32), §III.4.6	éq. (III.63), §III.5.4
					éq. (III.37), §III.4.6	éq. (III.64), §III.5.4
		autorégressif, éq. (III.67), §III.6	$[\sigma_{\theta^l}, \sigma_{\theta^r}, \sigma_{\theta^t}, \mathbf{a}^{lT}, \ldots$ $\mathbf{a}^{rT}, \mathbf{a}^{tT}, P^l, P^r, P^t]^T$	VC §IV.2.2 ou VP §IV.2.3 pour σ, §IV.2.6 pour $\mathbf{a}$ et P	éq. (III.32), §III.4.6	éq. (III.70), §III.6
					éq. (III.37), §III.4.6	éq. (III.71), §III.6
images $\mathcal{I}$, *star tracker* $\mathbf{s}$	Fusion	tous ceux du MAP	$+[\mathbf{c}^{lT}, \mathbf{c}^{rT}, \mathbf{c}^{tT}, D_{s^l}, \ldots$ $D_{s^r}, D_{s^t}, \sigma_{c^l}, \sigma_{c^r}, \sigma_{c^t}]^T$	VP §IV.2.3 pour σ, VC §IV.2.2 pour D et §III.9.5 pour $\mathbf{c}$	éq. (III.32), §III.4.6	éq. (III.110), §III.9.5
					éq. (III.37), §III.4.6	éq. (III.110), §III.9.5

TABLE IV.1 – Tableau des estimateurs.

jeux aux caractéristiques différentes qui sont détaillés avant chaque test. Ces données sont toutes synthétiques, les déformations liées aux variations d'attitude sont simulées dans les images selon un procédé qui nous est totalement inconnu. En revanche cette simulation se veut la plus réaliste possible en prenant en compte le bruit d'acquisition des capteurs photométrique et sensoriel, les distorsions radiométriques et géométriques, les effets de terrain tels que l'élévation, et la structure du satellite pour lier les phénomènes oscillatoires à sa mécanique. L'avantage principal de ces données synthétiques est de disposer de la vérité terrain sur les variations d'attitude pour la confronter à nos estimations, ce que ne permettent pas les données réelles. Nous disposons également des images non déformées, c'est à dire telles que nous sommes supposés les avoir après recalage. Selon Astrium, la difficulté à traiter ces simulations est comparable aux cas réels. Ces expériences doivent être vues comme une première étape de validation de l'approche proposée, avant de la tester dans le futur sur des images réelles.
Aux données s'ajoutent les matrices des dérivées partielles de la fonction de transformation $\mathcal{W}$ par rapport à x et n pour le lacet, le roulis et le tangage en chaque point de l'image ; nous rappelons que x est orienté suivant l'axe de la caméra *pushbroom* et n selon le temps. Elles sont calculées numériquement par différences finies et couplées avec des modèles d'élévation numérique du terrain ; la méthode est présentée dans le brevet [Crombez 08]. A une position du satellite donnée, il est possible de projeter une ligne de visée pour chaque pixel de la caméra sur le modèle surfacique du terrain et de comparer son intersection à celle obtenue par une même projection où les angles d'observation du satellite ont varié légèrement. Nous en avons parlé au cours du chapitre précédent, le procédé inverse consistant à calculer le lieu d'intersection d'un point de la surface sur l'image acquise par le satellite au cours du temps est nettement plus compliqué à modéliser [Gupta 95, Gupta 97]. L'utilisation de ces matrices permettent de s'affranchir d'un modèle analytique pour la fonction $\mathcal{W}$, tout en relâchant l'hypothèse stricte d'une terre plane puisqu'elles intègrent une information d'élévation sur le terrain. Elles sont utilisées pour le calcul de J à l'équation (III.30) (p.39).
Les variations selon les trois angles ont été simulées pour toutes les données, mais nous ne montrons ici que des résultats d'estimation sur le roulis et le tangage. Le lacet est une rotation dans le plan de l'image et engendre des déformations essentiellement sur ses bords, elles sont bien moins visibles que les effets liés au roulis ou au tangage. La simulation a considéré des variations d'attitude d'amplitude similaire pour les trois rotations, provoquant des décalages supérieurs à la taille d'un pixel pour le roulis et le tangage, et de l'ordre du centième de pixel pour le lacet, ce qui est bien inférieur à notre précision de restitution de l'attitude. Précisons tout de même que l'effet du lacet sera croissant avec la longueur de la caméra *pushbroom*, la qualité de son estimation sera comparable aux résultats que nous présentons sur les deux autres angles pour des capteurs suffisamment longs et des déformations proches du demi pixel.
Quelles que soient les données, nous avons maintenu des intervalles [2] constants pour l'estimation des hyperparamètres, ainsi que des filtres identiques pour l'inférence du modèle AR et du modèle polynômial basse fréquence. Les hyperparamètres sont estimés sur une portion de (300×800) pixels sélectionnée aléatoirement dans les images. Nous souhaitons laisser l'ordinateur prendre des décisions à la vue des images, bien qu'une expertise humaine puisse être plus efficace sur plusieurs points. Les temps de calcul sont don-

2. 30 valeurs échantillonnées entre $[10^{-9.5}, 10^{-6.5}]$ pour la majorité des σ_{θ^α} et entre $[10^{-8}, 10^{-6}]$ pour les σ_{c^α}, ceux du processus gaussien sont estimés sur 10 échantillons entre $[0.01, 1]$ pour σ_{θ^α} et $[0.01, 0.1]$ pour l^α

nés à titre indicatif pour comparer les différentes méthodes, ils ne reflètent pas ce qu'un algorithme optimisé est capable de faire mais doivent plutôt être pris pour une borne supérieure du temps de traitement possible. En effet, certains compromis ont été faits pour pouvoir tester toutes les méthodes dans un même programme générique sans qu'une approche soit privilégiée aux dépens d'une autre. L'algorithme fonctionne sous Matlab et la machine de test est dotée d'un processeur Core2 duo cadencé à 3GHz avec 3.8Go de mémoire vive.

IV.5.1 Jeu de données ASTRO

Données

Ce premier jeu contient 4 images monospectrales de taille (900×2564), ce qui signifie que les caméras sont sensibles aux mêmes composantes spectrales de la lumière. Cette étape de validation par des images de même modalité est intéressante car elle permet de s'affranchir du problème non trivial de recalage multimodal. La fréquence d'échantillonnage d'acquisition des images 1-D par les capteurs *pushbroom* est de 770Hz, ce qui correspond à une période d'acquisition de 3.3 secondes pour les images 2-D complètes. Nous disposons également de données *star tracker* échantillonnées à 16Hz, soit 54 mesures réparties dans l'intervalle de temps. Une partie de ces données a notamment permis d'apprécier les performances d'une majorité des estimateurs, en référence aux tests que nous avons menés jusqu'à maintenant.
Les quatre caméras sont respectivement espacées de $\tau^{12} = 33.5$, $\tau^{23} = 40$ et $\tau^{34} = 20$ en échantillons de temps - ou nombre d'images 1D - si on se réfère à la figure II.1 (p.20) du chapitre 2. En conséquence, le premier noyau de convolution tel que nous l'avons introduit au paragraphe III.4.3 du chapitre précédent, équation (III.16), sera :

$$k_{12} = \left[0.5 \;\; 0.5 \;\; \underbrace{0 \; \ldots \; 0}_{\text{32 zeros}} \;\; -1\right]^T, \tag{IV.26}$$

la forme de celui du centre est encore plus simple :

$$k_{23} = \left[1 \;\; \underbrace{0 \; \ldots \; 0}_{\text{39 zeros}} \;\; -1\right]^T, \tag{IV.27}$$

tout comme le dernier :

$$k_{34} = \left[1 \;\; \underbrace{0 \; \ldots \; 0}_{\text{19 zeros}} \;\; -1\right]^T. \tag{IV.28}$$

En remarquant que $k_{ij} = -k_{ji}$, l'ensemble des noyaux se construit sans difficulté.

Attitude

Les caractéristiques du satellite simulé font que les variations d'attitude sont basses fréquences. La figure IV.20 permet d'observer la vérité terrain pour les deux angles que nous souhaitons estimer - le roulis et le tangage - avec leur spectre de fréquence associé.
Les spectres sont sensiblement les mêmes avec trois harmoniques autour de 22Hz, 10Hz et 4Hz, les très basses fréquences en dessous de 3Hz varient légèrement entre les trois angles. On note que le décalage à zéro selon l'axe des ordonnées est important pour le tangage, il est à l'origine du grand pic à zéro sur son spectre. Ce décalage ne pourra être estimé que par les mesures *star tracker*. La bande blanche délimite une nouvelle fois la taille d'un pixel, ce qui nous fait comprendre l'importance de la déformation.

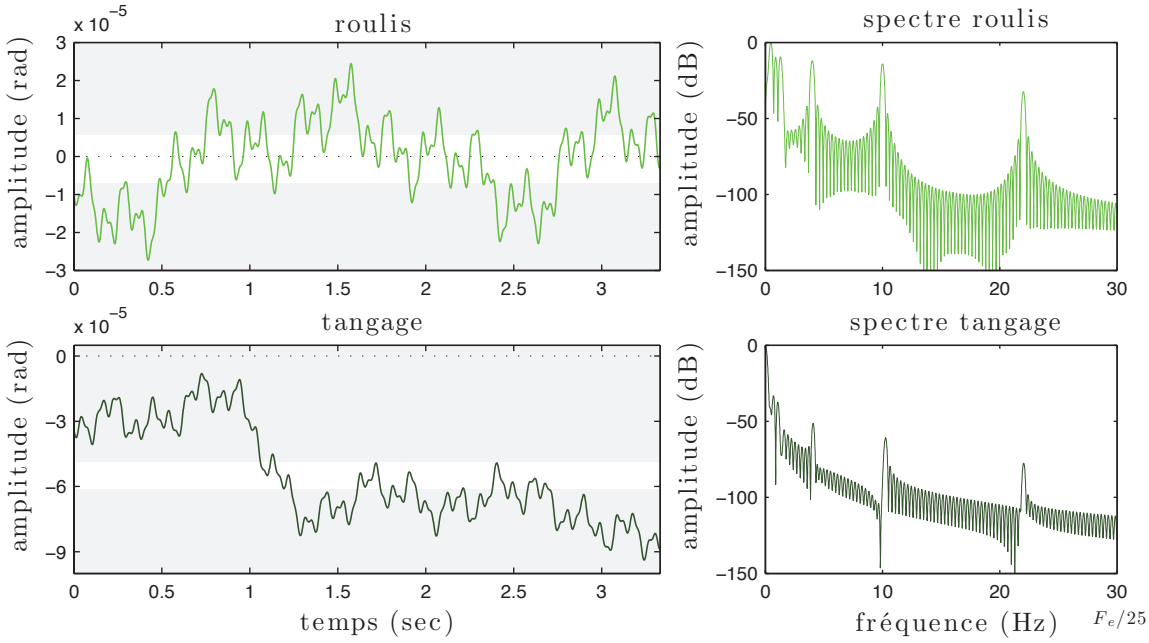


FIGURE IV.20 – Représentation temporelle et spectrale de l'attitude réelle pour le jeu de données ASTRO. La bande blanche représente la taille d'un pixel comparée à l'amplitude de la déformation, la ligne pointillée permet de se rendre compte du décalage des courbes selon l'axe des ordonnées.

Star tracker

Les mesures *star tracker* nous permettent d'estimer un polynôme qui modélise les basses fréquences du signal. Cette idée a été présentée au paragraphe III.9.5, et la sélection des degrés D_{s^r} et D_{s^t} se fait par validation croisée de type *leave-one-out* comme suggéré au paragraphe IV.2.2. La figure IV.21 montre les résultats de sélection de modèle pour les deux angles.

Nous avons arbitrairement choisi de limiter le degré du polynôme basse fréquence à 15, pour éviter la recherche des oscillations hautes fréquences dans le signal sur cette période de temps. Dans l'ensemble les choix de modèle sont plutôt justes vis à vis de la vérité terrain, alors que certaines mesures du *star tracker* sont très bruitées comme celles du roulis à -2×10^{-5} radian autour de 1.5 et 2 secondes.

Modèle autorégressif

L'estimateur du MAP avec *a priori* autorégressif nécessite au préalable l'inférence des coefficients $\mathbf{a}^r$ et $\mathbf{a}^t$ du modèle selon l'approche proposée au paragraphe IV.2.6. La figure IV.22 présente les résultats de cette détermination des coefficients autorégressifs pour les deux types de vraisemblance images, où le critère d'Akaike est utilisé pour trouver les ordres P^r et P^t. Nous avons justifié au paragraphe IV.2.6 l'utilisation de ce critère parmi les nombreuses solutions existantes [Peyrard 01].

On peut voir immédiatement que le modèle AR du roulis pour le premier cas de vraisemblance est victime de surapprentissage, l'ordre du modèle a été estimé à 349 contre 178 pour le second cas de vraisemblance. Une grande partie des pics du bruit sont appris par

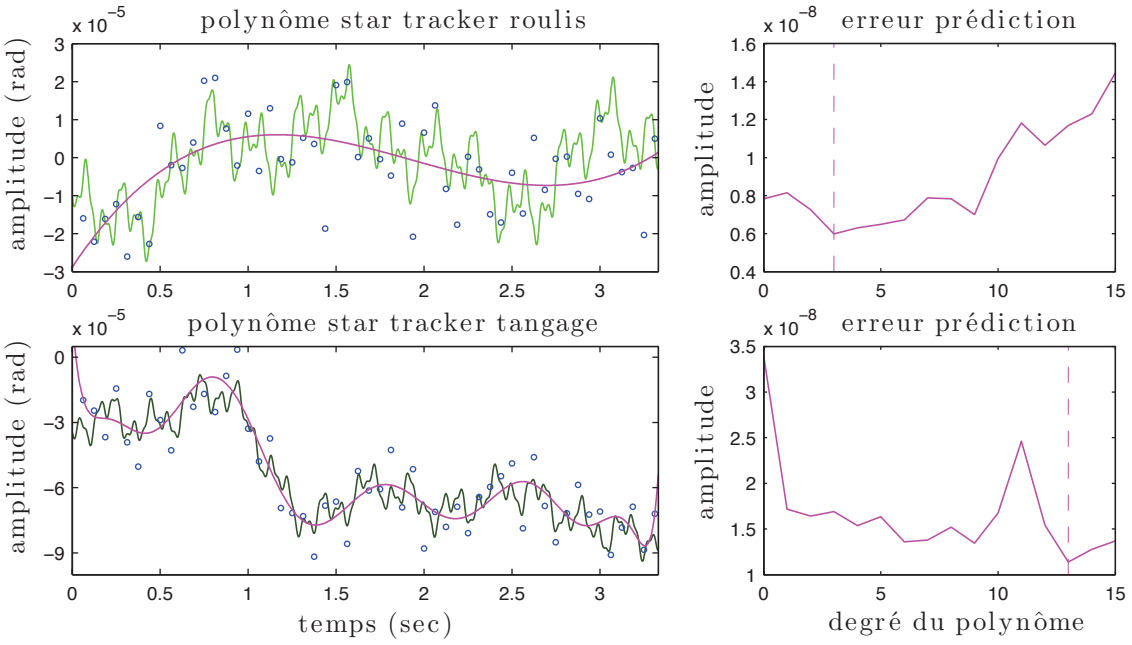


FIGURE IV.21 – Résultat sur la modélisation basse fréquence de l'attitude à partir des mesures *star tracker*. Les degrés estimés par validation croisée de type *leave-one-out* sont $D_{s^r} = 3$ et $D_{s^t} = 13$.

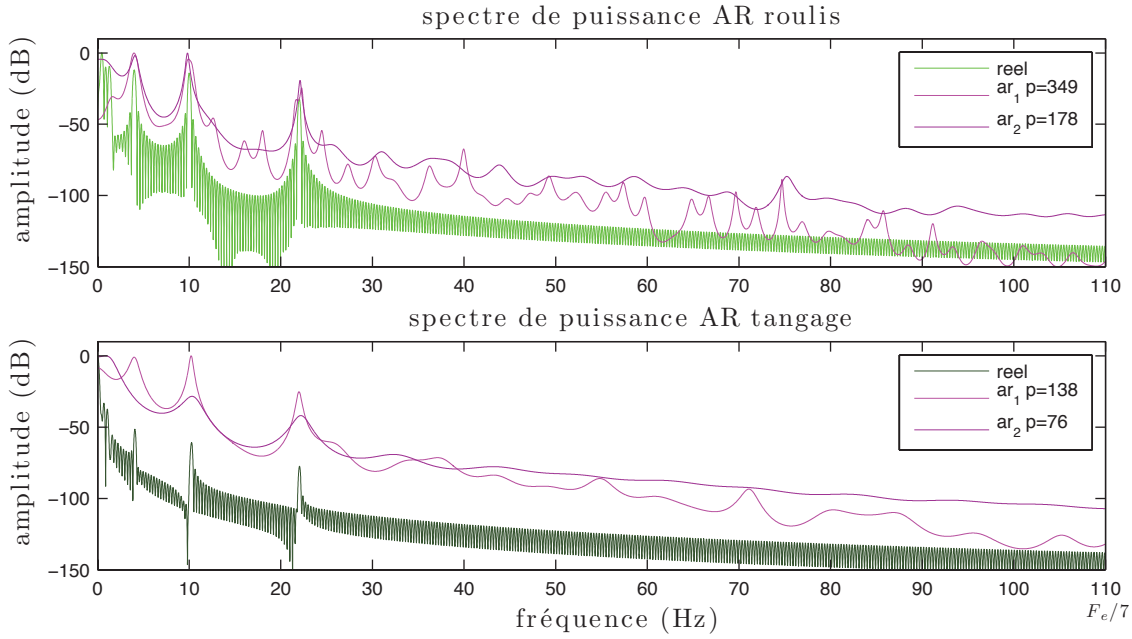


FIGURE IV.22 – Résultat sur l'estimation du modèle autorégressif pour le roulis en haut et le tangage en bas ; les deux vraisemblances pour les images proposées au chapitre précédent sont indicées par 1 et 2 respectivement pour le premier et second cas.

le modèle AR, au dépend des fréquences plus basses du signal ; le phénomène apparait également à un moindre niveau sur le spectre du tangage. Cette observation est d'autant plus étonnante que nous avions observé dès le début du chapitre précédent que l'estimateur du MV était légèrement moins bruité pour le premier cas de vraisemblance des images ; rappelons ici que le modèle AR est estimé à partir d'une version filtrée du maximum de vraisemblance de l'attitude. En réalité, nous avions également vu que les basses fréquences du signal n'étaient pas bien restituées pour le premier cas, ceci est confirmé par l'absence d'apprentissage de basse fréquence par le modèle AR malgré son ordre élevé. Les pics de fréquence du signal utile sont donc rapprochés du bruit moyenne fréquence et malgré le filtrage, le critère d'Akaike les considère comme une partie du signal. Ce critère est rapide, mais n'est pas suffisamment performant pour distinguer signal et bruit dans cette situation. La vraisemblance des images par convolution est avantagée par un spectre basse fréquence plus riche du signal estimé par MV.

Résultats

Nous présentons ici des résultats quantitatifs de comparaison de l'attitude estimée par rapport à la vérité terrain, en parallèle à un calcul de score à partir des images recalées selon notre estimation. Ce dernier correspond à une moyenne de corrélation croisée normalisée (CCN) [Szeliski 06], évaluée sur 500 fenêtres de taille (9×9) et dont les positions sont tirées aléatoirement dans les images. Sur chacune d'entre elles, les vignettes sont extraites à un endroit donné et l'évaluation est faite pour chaque paire d'image. La version normalisée ajoute une légère robustesse au caractère multimodal des images - cette observation est en rapport aux prochains jeux de données -, bien que des solutions plus performantes existent [Roche 99]. La valeur de la CCN est comprise entre $[-1, 1]$, où 1 correspond à une corrélation parfaite et inversement pour -1. Cette évaluation à partir des images est utile car dans un contexte réel, il faut pouvoir utiliser une métrique pour évaluer nos résultats.
Le tableau IV.2 montre l'écart-type de l'erreur [3] entre attitude réelle et attitude estimée en unité de pixel pour l'estimateur du MAP avec différents *a priori* ; les scores de CCN et le temps d'estimation de l'algorithme en secondes complètent les résultats. Nous adoptons la charte graphique suivante pour tous les tableaux qui suivent : dans chaque ligne le meilleur score est en vert et le plus mauvais en rouge, ils seront de plus écrits en gras lorqu'ils représentent le meilleur et le pire dans l'absolu parmi tous les tableaux pour un même jeu de données. Les valeurs identiques ont été départagées par le troisième chiffre décimal pour plus de clarté dans le tableau, bien que dans notre cas une précision au millième de pixel ait peu de sens. Nous avons utilisé la validation croisée et la validation photométrique pour estimer les hyperparamètres dans la plupart des cas - en dehors du σ_c du modèle polynômial de la fusion et du couple (l_p, D) pour le modèle polynomial par morceaux où seule la photométrie est utilisée -, le tableau présente seulement le meilleur résultat obtenu entre les deux méthodes pour chaque estimateur. Une case grisée indique que les hyperparamètres ont été estimés par validation photométrique et inversement une case blanche montre qu'ils ont été estimés par validation croisée.
Le tableau IV.3 affiche les résultats de la fusion de données *star tracker* et images pour le meilleur *a priori* dans les deux cas de vraisemblance. Les résultats obtenus avec les autres *a priori* sont soit sensiblement moins bons - c'est surtout le cas de celui sur l'amplitude -,

3. Ce critère est utilisé par Astrium comme mesure de la performance de restitution des variations d'attitude.

soit presque identiques. Si l'hyperparamètre σ_c est obligatoirement estimé par photométrie, les autres peuvent néanmoins être trouvés par validation croisée auparavant. C'est la raison pour laquelle nous avons à nouveau l'alternance de cellules blanches et grises suivant la méthode qui a obtenu le meilleur résultat.
Le tableau IV.4 correspond aux résultats du modèle polynômial par morceaux, qui cette fois-ci n'utilise que la validation photométrique.

		a priori MAP				MV
		amplitude	dérivée 2^{de}	ps gaussien	autorégressif	
vraisemblance 1	roulis	0.68	0.61	0.58	0.48	0.70
	tangage	1.54	1.39	1.44	1.40	1.49
	CCN	0.87	0.9	0.96	0.91	0.89
	temps	**1510**	1037	952	758	746
vraisemblance 2	roulis	0.25	0.22	0.22	0.23	0.34
	tangage	0.66	0.28	0.25	**0.25**	1.31
	CCN	0.97	0.97	**0.98**	0.97	0.97
	temps	285	260	491	268	**175**

TABLE IV.2 – Résultats du MAP et du MV sur le jeu de données ASTRO. Les mesures pour le roulis et le tangage sont les écarts-types de l'erreur entre attitude réelle et attitude estimée (le meilleur résultat tend vers zéro). La mesure CCN est un score de corrélation croisée normalisée (le meilleur résultat tend vers un). Le temps correspond au nombre de secondes d'exécution de l'algorithme complet (le meilleur est le plus petit nombre). Le meilleur score et le plus mauvais score pour chaque ligne sont respectivement écrits en vert et en rouge. Une case grisée indique que le meilleur résultat a été obtenu par sélection photométrique des hyperparamètres, sinon une case blanche correspond à une estimation de ceux-ci par validation croisée.

Discussion

Le calcul de CCN avant recalage est égal à 0.77 ; hormis le premier cas de modèle polynômial par morceaux, tous les estimateurs améliorent ce critère de qualité image. Si on s'attache à l'écart-type de l'attitude qui est à l'origine de 0.84 pixel pour le roulis et de 1.74 pixel pour le tangage avant recalage, on constate une nouvelle fois que cette valeur avec l'attitude estimée est réduite, à la même exception près. Les résultats sont particulièrement flagrant pour le deuxième cas de vraisemblance des images qui procède par convolution : ils sont à la fois plus précis, plus rapides et meilleurs selon le critère de CCN. Globalement, les estimateurs du MAP avec *a priori* de type processus gaussien ou autorégressif se démarquent des autres méthodes.
Le modèle polynômial par morceaux est particulièrement mauvais, mais on peut trouver certaines explications sur cette observation. Premièrement, la sélection automatique des hyperparamètres l_p et D n'est pas optimale. Si l'exemple de la figure IV.11 montre assez justement qu'il faut privilégier une petite fenêtre avec un petit degré de polynôme,

	MAP avec fusion	
	dérivée 2^{de}	autorégressif
	vraisemblance 1	vraisemblance 2
roulis	0.66	**0.20**
tangage	1.40	0.25
CCN	0.90	0.97
temps	984	522

TABLE IV.3 – Résultats de la fusion sur le jeu de données ASTRO.

	polynômial par morceaux	
	vraisemblance 1	vraisemblance 2
roulis	**0.86**	0.27
tangage	**1.81**	0.86
CCN	**0.75**	0.94
temps	1426	908

TABLE IV.4 – Résultats du modèle polynômial par morceaux sur le jeu de données ASTRO.

elle ne présente pas de minimum aussi marqué que pour l'estimation d'autres hyperparamètres. Nous avons observé expérimentalement que la sélection tendait à favoriser des polynômes de degré 4 sur des fenêtres de temps inférieures à 10. Le système à résoudre devient très contraint avec de nombreux coefficients à déterminer, ce qui provoque des temps de calcul très longs lorsque l'on cherche à estimer le signal complet. Des polynômes de degrés 2 ou 3 avec des tailles de fenêtre comprises entre 10 et 15 sont souvent suffisants. Secondement, le temps associé à l'attitude croit linéairement et les valeurs prises par les coefficients polynômiaux varient énormément entre le début et la fin du signal, ce qui provoque des instabilités dans la résolution numérique du système. La méthode est donc doublement pénalisée par le choix d'un polynôme de degré trop élevé : le temps de calcul augmente et le système numérique devient instable. Notons que la période d'apprentissage de 1.04 seconde, identique pour toutes les méthodes, est trop courte pour voir apparaitre ce problème ; elle ne pénalise donc pas ces solutions instables.

La figure IV.23 résume assez bien ces deux observations : le modèle polynômial par

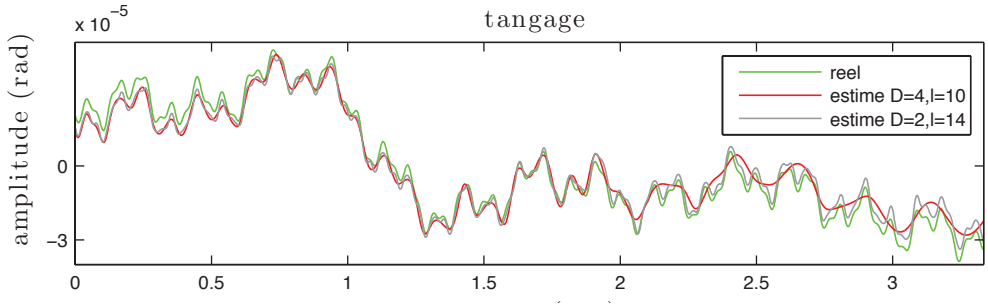


FIGURE IV.23 – . Estimation du tangage sur le jeu de données ASTRO ; les hyperparamètres du modèle en gris ont été choisis à la main, contrairement à ceux de la courbe rouge.

morceaux de degré 2 pour une fenêtre de taille supérieure suit mieux les variations du tangage du début jusqu'à la fin. A l'inverse celui qui a été choisi automatiquement, avec $D = 4$ et $l_p = 10$, dérive de plus en plus au fil du temps. C'est assez flagrant autour de trois secondes où les oscillations hautes fréquences ne sont plus suivies ; l'écart de valeur entre le coefficient du degré 4 de la première à la dernière fenêtre est important, ce qui déstabilise la résolution numérique du problème. Pour contrer cette limite, une solution consisterait à normaliser le temps entre $[0, 1]$ sur chaque fenêtre d'observation des polynômes ; tous les polynômes sont ainsi évalués sur une même base de temps.

Si en apparence la validation croisée semble donner des meilleurs résultats, les valeurs d'écart-type et de CCN relevées sont particulièrement resserrées entre les deux méthodes. En dehors d'un temps de calcul de 10% à 30% supérieur pour la validation photométrique, il est difficile ici de privilégier une méthode d'estimation des hyperparamètres par rapport à l'autre.

On peut remarquer que les scores de CCN sont très proches entre chaque *a priori* pour la seconde vraisemblance, alors que les précisions données par les écarts-types sont assez disparates entre le MV et le MAP avec modèle autorégressif. Ce critère met tout de même en avant une des meilleurs méthodes avec l'*a priori* de type processus gaussien.

Les temps de calcul doivent être lus ici avec un certain recul, ils prennent en compte l'es-

timation des hyperparamètres et chaque itération jusqu'à la convergence de l'algorithme. La première vraisemblance nécessite plus d'itérations que la seconde et le calcul supplémentaire d'une jacobienne sur une image, ce qui explique ces résultats. Le temps de calcul du MV peut donner une indication sur la vitesse de l'algorithme hors estimation des hyperparamètres, mais en pratique on a constaté que la vitesse de résolution des équations matricielles sur une itération avec Matlab dépendait beaucoup de l'*a priori* choisi et du conditionnement du problème. Dans le schéma itératif choisi où les estimations précédentes servent aux suivantes, l'amplification du bruit peut parfois être critique au point de rendre le système très long à résoudre, c'est notamment ce qui s'est produit pour le MAP avec *a priori* sur l'amplitude et selon la première vraisemblance.

Recalage des images

Nous avons sélectionné la méthode ayant donné les meilleurs résultats sur la moyenne des écarts-types de l'erreur entre le roulis et le tangage. Celle-ci est le MAP avec fusion selon la seconde vraisemblance, *a priori* autorégressif et estimation des hyperparamètres par validation photométrique ; $\sigma_I = 0.014$, $\sigma_{\theta^r} = 1.1 \times 10^{-8}$, $\sigma_{\theta^t} = 1.8 \times 10^{-8}$, $\sigma_{c^r} = 8.5 \times 10^{-7}$, $\sigma_{c^t} = 3.2 \times 10^{-7}$. La figure IV.24 montre le résultat du recalage sur l'image panchromatique avec les variations d'attitude estimées.
L'erreur de recalage est globalement réduite : la plupart des contours disparaissent, et d'autres persistent parce que la déformation n'a pas été suffisamment bien estimée. Au regard du résidu des courbes, on remarque une lente dérive linéaire pour le tangage qui semble être une conséquence de la déconvolution avec des effets de bords, et une oscillation très basse fréquence sur le roulis. Sur certaines portions la variance du résidu est inférieure au dixième de pixel ce qui constitue une bonne précision dans ce contexte.

IV.5.2 Jeu de données MLF

Données

Quatre images multispectrales de taille (900×2564) composent ce second jeu ; en l'occurrence une de modalité panchromatique, une rouge, une verte et une bleue. Les conditions sont identiques aux données ASTRO pour ce qui est de la fréquence d'échantillonnage et de la configuration du plan focal.

Attitude

Les oscillations simulées pour ce satellite sont moyennes et hautes fréquences. La figure IV.25 montre le tracé temporel de l'attitude que nous recherchons, ainsi que les spectres associés au roulis et au tangage. On distingue assez clairement onze harmoniques communes aux deux signaux entre 2Hz et 24Hz. L'amplitude est modérée mais suffisante pour nécessiter plusieurs itérations.

Star tracker

Les mesures *star tracker* nous informent sur les basses fréquences de l'attitude. Conformément à notre méthode, nous recherchons un polynôme sur ces mesures où le degré est estimé par validation croisée de type *leave-one-out*. La figure IV.26 montre le résultat de cette approche.
La sélection du modèle polynômial indique que le signal ne contient pas de basse fré-

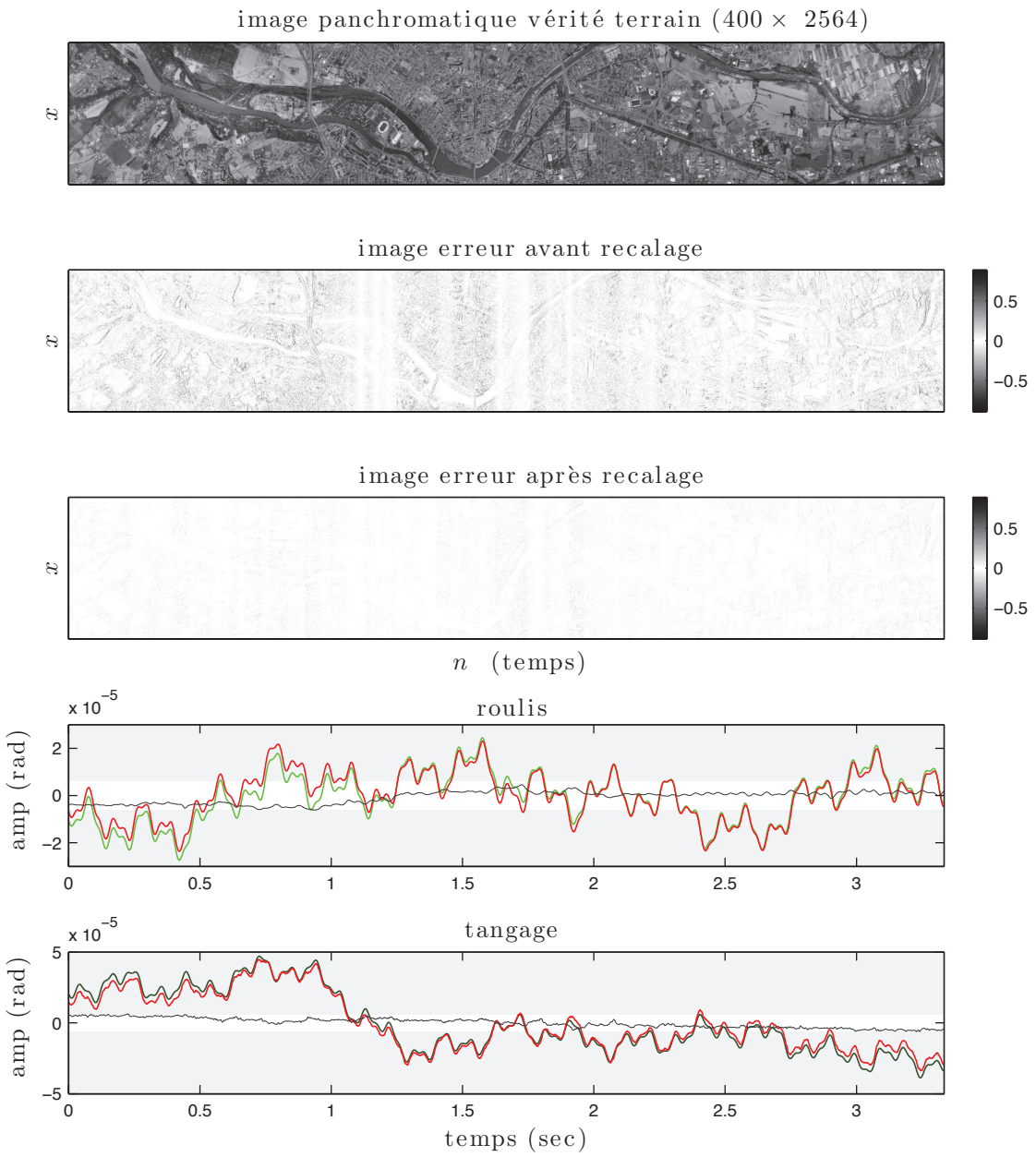


FIGURE IV.24 – . Résultat du recalage d'image sur le jeu de données ASTRO ; le résultat est obtenu par MAP avec fusion selon la seconde vraisemblance et avec *a priori* autorégressif. La figure du haut est un extrait (400 × 2564) de l'image panchromatique non déformée, en dessous les différences d'intensités entre les images vérité terrain / image déformée et vérité terrain / image rectifiée (soit avant et après recalage). Les deux graphiques suivant représentent l'attitude réelle en vert, l'estimée en rouge et le résidu en noir ; ce dernier se maintient dans la zone blanche délimitant la taille d'un pixel. Le décalage en zéro a été supprimé sur les courbes pour plus de clarté dans les résultats.

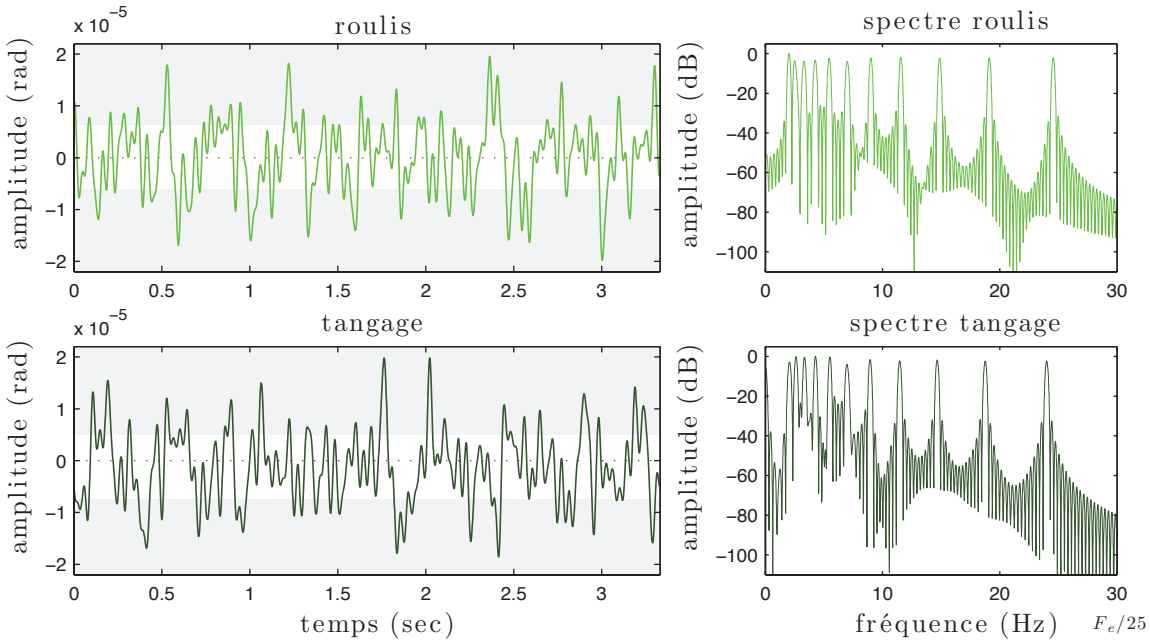


FIGURE IV.25 – Tracé temporel et spectral de l'attitude recherchée sur le jeu de données MLF.

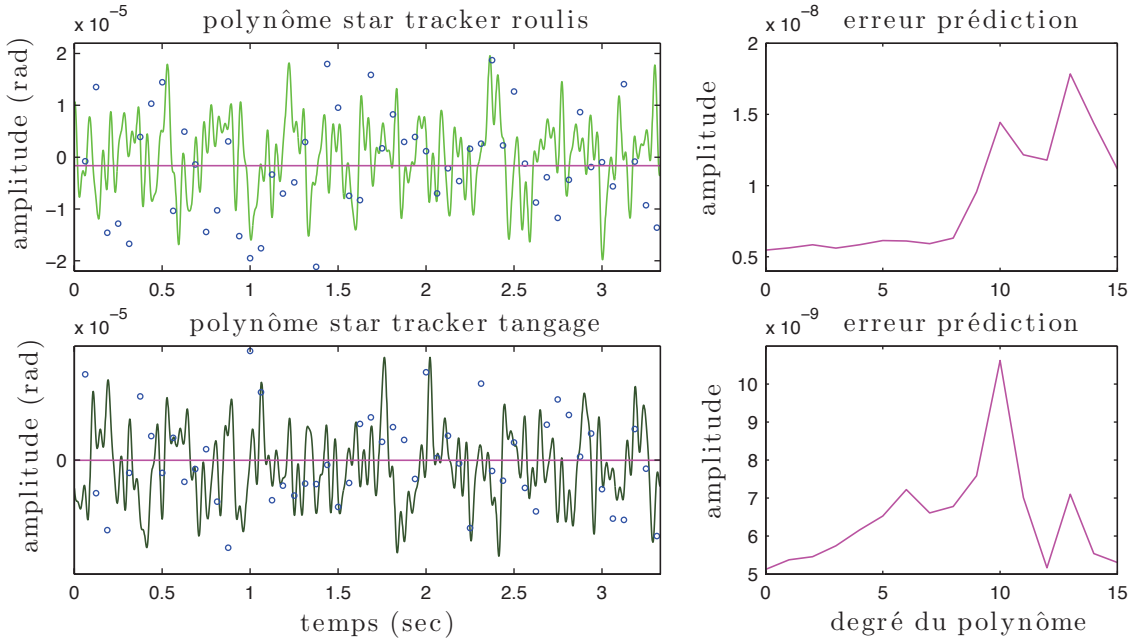


FIGURE IV.26 – Modélisation des mesures *star tracker* de MLF, $D_{s^r} = 0$ et $D_{s^t} = 0$.

quence - un polynôme de degré nul a été trouvé dans les deux cas -, ce qui est plutôt en accord avec nos remarques du paragraphe précédent sur l'attitude recherchée. Si certaines mesures en cercle bleu paraissent bonnes en comparaison de la vérité terrain, il est important de comprendre que l'échantillonnage du *star tracker* à 16Hz limite sa capacité de restitution des fréquences à 8Hz au maximum. A cette remarque doit s'ajouter le bruit particulièrement important du *star tracker*, et le fait qu'évaluer la vraisemblance de chacune de ces mesures indépendamment est assez difficile ; on peut dire que l'algorithme a fait un choix de modèle assez judicieux.

Modèle autorégressif

Le modèle autorégressif estimé à partir des données est présenté en figure IV.27. Dans le cas du roulis, la première vraisemblance favorise un modèle très simple à l'ordre $P^r = 6$, indiquant que le signal est composé de basses et moyennes fréquences. Le modèle estimé avec la seconde vraisemblance est plus précis avec un ordre égal à 295, il isole assez nettement quatre pics entre 10Hz et 25Hz. Le bruit est aussi légèrement appris par le modèle au delà de 45Hz, seulement il reste assez minoritaire par rapport aux harmoniques du signal ; rappelons que notre représentation de la figure IV.27 est en décibels.
Assez curieusement, l'effet inverse se produit pour l'apprentissage du modèle sur le tangage : la première vraisemblance isole assez bien les harmoniques du signal, contrairement au second cas de vraisemblance qui reste assez vague sur l'information qu'il apporte sur les moyennes fréquences. On peut remarquer au tableau IV.5 que le score du modèle autorégressif pour la première vraisemblance est relativement bon pour l'estimation du tangage, à comparer de son score pour le roulis.

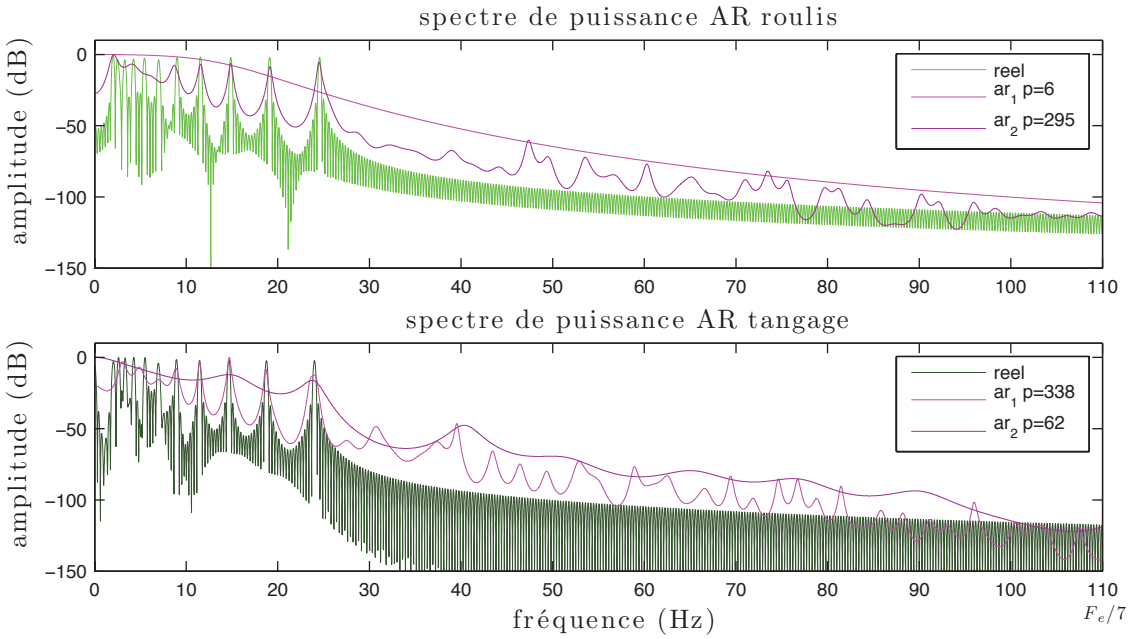


FIGURE IV.27 – Modèles autorégressifs estimés pour la première et seconde vraisemblance, d'indices respectifs 1 et 2, sur les données MLF.

Résultats

Nous présentons les résultats de la même façon qu'à l'expérience précédente ; le tableau IV.5 correspond aux résultats obtenus avec l'estimateur du MAP, le tableau IV.6 décrit les performances de la fusion avec le meilleur *a priori* et le IV.7 est associé au modèle polynômial par morceaux. Le score de CCN avant recalage est de 0.76 et les écarts-types du roulis et du tangage réels sont respectivement de 0.55 et 0.56 pixel.

		a priori MAP				MV
		amplitude	dérivée 2^{de}	ps gaussien	autorégressif	
vraisemblance 1	roulis	0.46	0.52	0.19	0.52	0.48
	tangage	1.72	0.43	0.17	0.28	0.73
	CCN	0.73	0.83	0.94	0.86	0.81
	temps	1515	1509	818	1184	747
vraisemblance 2	roulis	0.11	0.10	**0.08**	0.08	0.11
	tangage	0.19	0.17	0.19	0.18	**0.16**
	CCN	0.95	0.95	0.95	**0.96**	0.95
	temps	205	259	333	254	146

TABLE IV.5 – Résultats du MAP et du MV sur le jeu de données MLF. Les mesures pour le roulis et le tangage sont les écarts-types de l'erreur entre attitude réelle et attitude estimée (le meilleur résultat tend vers zéro). La mesure CCN est un score de corrélation croisée normalisée (le meilleur résultat tend vers un). Le temps correspond au nombre de secondes d'exécution de l'algorithme complet (le meilleur est le plus petit nombre). Le meilleur score et le plus mauvais score pour chaque ligne sont respectivement écrits en vert et en rouge. Une case grisée indique que le meilleur résultat a été obtenu par sélection photométrique des hyperparamètres, sinon une case blanche correspond à une estimation de ceux-ci par validation croisée.

Discussion

Ces résultats confirment les observations que nous avons faites auparavant, l'estimateur du MAP donne les meilleurs résultats avec la seconde vraisemblance des images, et les *a priori* privilégiés sont la dérivée seconde, le processus gaussien et le modèle autorégressif. Les performances du modèle polynômial par morceaux sont légèrement meilleurs pour estimer ces oscillations moyennes fréquences, mais elles restent un cran en dessous des autres pour les raisons précitées. On peut être assez surpris de voir que le meilleur écart-type pour le tangage est obtenu avec le maximum de vraisemblance. La figure IV.28 permet de mieux comprendre ce résultat ; l'écart-type du résidu est effectivement petit - il est de l'ordre de 0.16 pixel sur cette portion seule -, mais le signal estimé contient un bruit parasite haute fréquence assez important. L'amplitude de la déformation étant faible, l'algorithme a convergé suffisamment rapidement pour éviter une amplification trop importante de ce bruit. Le résultat graphique n'est pas satisfaisant,

	MAP autorégressif avec fusion	
	vraisemblance 1	vraisemblance 2
roulis	0.49	0.10
tangage	0.23	0.17
CCN	0.86	0.95
temps	1055	495

TABLE IV.6 – Résultats de la fusion sur le jeu de données MLF.

	polynômial par morceaux	
	vraisemblance 1	vraisemblance 2
roulis	**0.85**	0.32
tangage	**1.72**	0.34
CCN	**0.49**	0.89
temps	**3215**	1064

TABLE IV.7 – Résultats du modèle polynômial par morceaux sur le jeu de données MLF.

mais il obtient un score supérieur à ceux du MAP selon ce critère d'écart-type ; pourtant la courbe de tangage estimée en figure IV.29 selon la meilleure méthode semble plus juste.

Ces observations nous font penser que l'écart-type, qui est souvent utilisé comme un outil générique d'évaluation des résultats, n'est peut être pas bien adapté dans notre problème. Il permet de distinguer assez souvent les meilleures solutions, mais ne les isole pas complètement de solutions plus erronées. Il serait pertinent d'ajouter un critère de régularité du signal, d'autant plus que celui-ci est stationnaire et plutôt basse fréquence.

Recalage des images

De la même façon que pour les données précédentes, nous avons sélectionné la meilleure technique en terme d'écart-type de l'erreur. Il s'agit cette fois-ci du MAP avec la seconde vraisemblance, *a priori* autorégressif et sélection des hyperparamètres par validation croisée ; $\sigma_I = 0.014$, $\sigma_{\theta^r} = 2.1 \times 10^{-9}$, $\sigma_{\theta^p} = 2.3 \times 10^{-8}$. La figure IV.29 permet d'apprécier le recalage des images à partir de l'attitude estimée.

L'image d'erreur obtenue après recalage est intéressante car peu de contours sont apparents ; elle donne l'impression par endroit de n'être constituée que d'un bruit blanc, ce qui est la première hypothèse de notre étude au chapitre précédent. Les résidus observés sur l'estimation de l'attitude sont bons, surtout pour le roulis qui est inférieur au dixième de pixel. C'est une performance appréciable dans un contexte multimodal tel que le nôtre, où toutes les images sont déformées. On note encore une fois un effet de bord sur le tangage estimé après trois secondes.

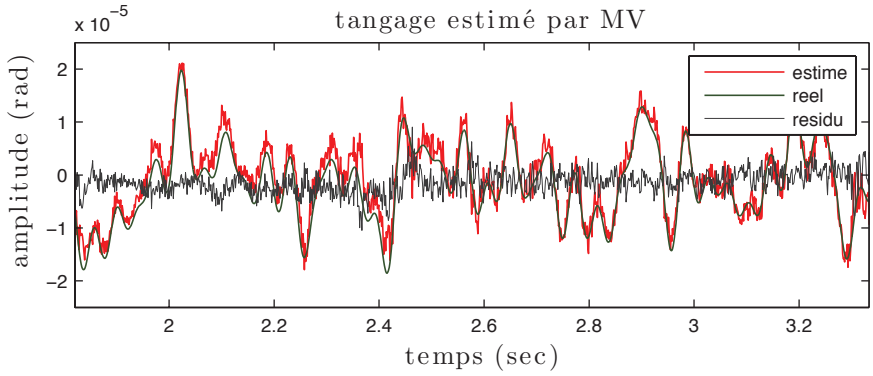


FIGURE IV.28 – Résultat d'estimation du tangage par maximum de vraisemblance, l'écart-type du résidu mesuré sur cette portion est de 0.16 pixel.

L'avantage de ce jeu multispectral est de pouvoir composer l'image couleur en sortie de l'algorithme ; c'est une des applications directes de notre étude. Un extrait de celle-ci est visible en figure IV.30, les nombreuses bavures visibles sur les contours de l'image avant recalage n'apparaissent plus sur l'image de droite. Cette dernière est un peu plus nette et surtout elle ne présente pas de tache de couleur aberrante, ce qui atteste du bon fonctionnement de la méthode.

IV.5.3 Jeu de données Reunion

Données

Pour ce jeu, nous disposons une nouvelle fois de quatre images multispectrales dans une configuration identique que les précédentes données.

Attitude

La particularité de cette simulation est d'être très basse fréquence avec une amplitude de variation importante, tout en contenant des petites oscillations hautes fréquences. C'est ce que nous observons en figure IV.31. Ce type d'oscillations est le plus difficile qui nous est proposé d'estimer ; le contenu fréquentiel est riche à la fois en basse et en haute fréquence, il diffère de plus légèrement entre le roulis et le tangage. L'amplitude de la déformation est largement supérieure à la taille d'un pixel, et les très basses fréquences seront difficiles à restituer en s'appuyant sur les images seulement. On peut noter le décalage à zéro en amplitude qui provoque ici le grand pic au début du spectre des deux angles.

Star tracker

La figure IV.32 montre le résultat de la sélection du polynôme sur les mesures *star tracker* du roulis et du tangage. Les basses fréquences sont plutôt bien estimées ici, avec un polynôme de degré 13 pour le roulis et 5 pour le tangage. On peut noter les dérives aux bords du modèle associé au roulis, ce qui peut être préjudiciable pour l'estimation de l'attitude lors de la fusion.

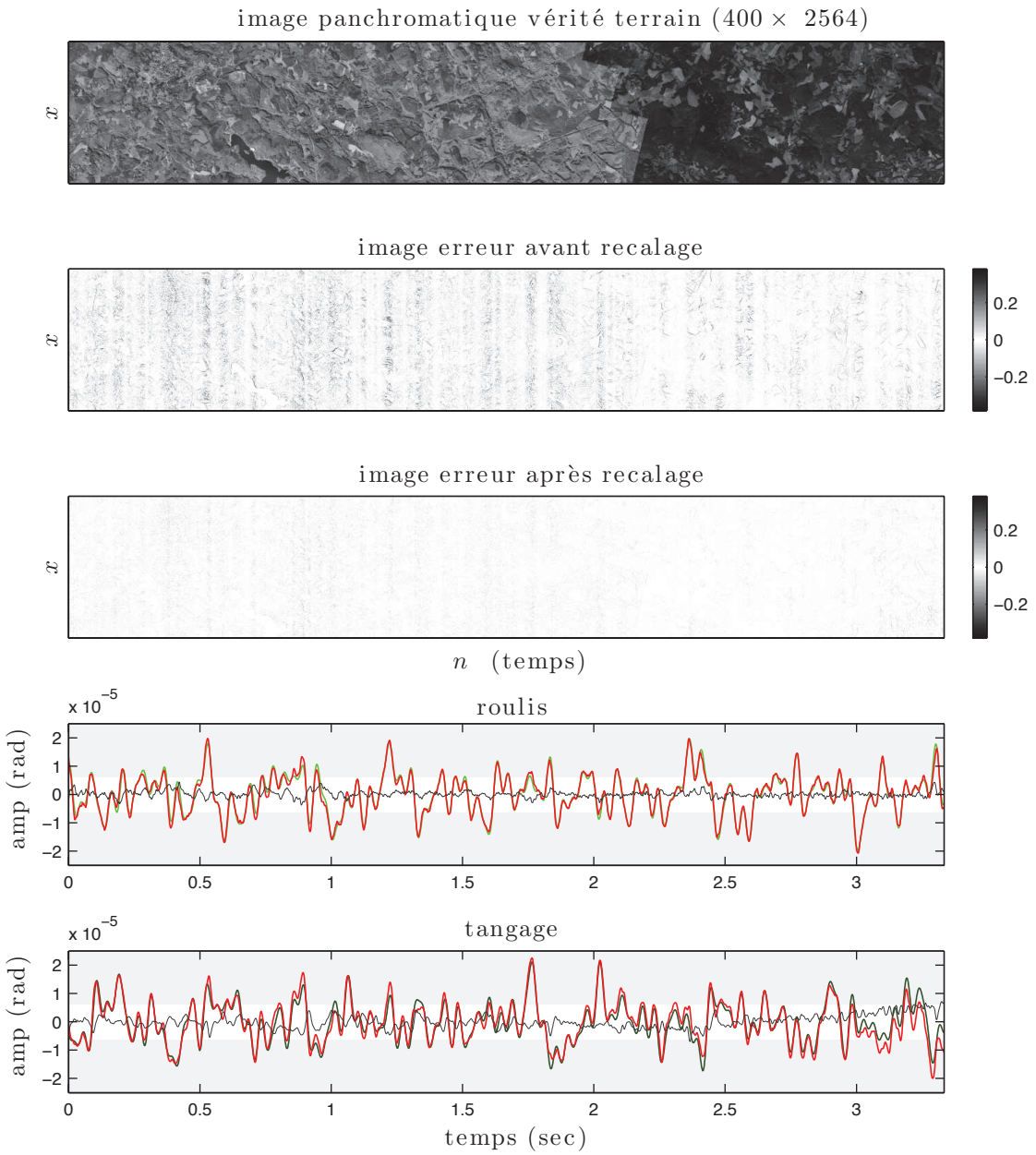


FIGURE IV.29 – . Résultat du recalage d'image sur le jeu de données MLF ; la méthode utilisée est le MAP avec *a priori* autorégressif pour le second cas de vraisemblance. La figure du haut est un extrait (400 × 2564) de l'image panchromatique non déformée, en dessous les différences d'intensités entre les images vérité terrain / image déformée et vérité terrain / image rectifiée (soit avant et après recalage). Les deux graphiques suivants représentent l'attitude réelle en vert, l'estimée en rouge et le résidu en noir ; la zone blanche délimite la taille d'un pixel.

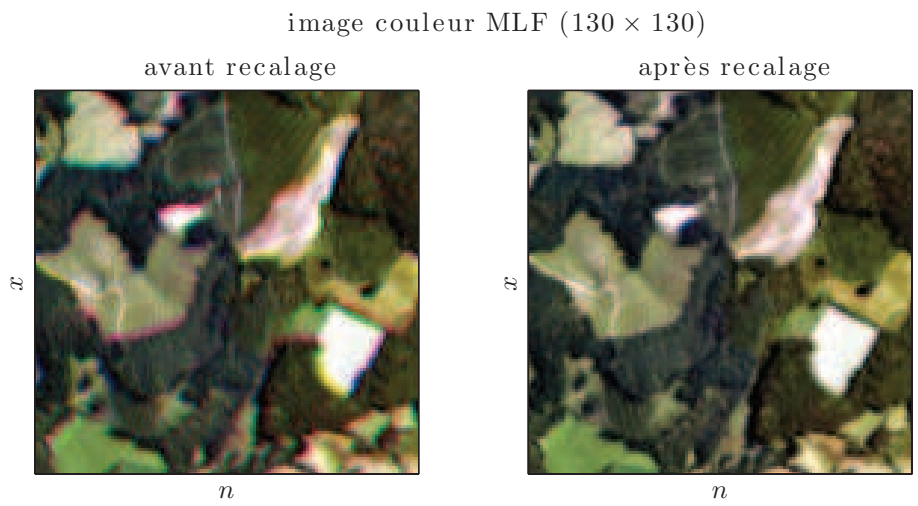


FIGURE IV.30 – Résultat du recalage sur un extrait de l'image couleur des données MLF de taille (130 × 130) pixels. Des bavures rouges, vertes et jaunes sont bien apparentes au niveau des contours sur l'image avant rectification. En comparaison elles ont presque entièrement disparu sur l'image de droite.

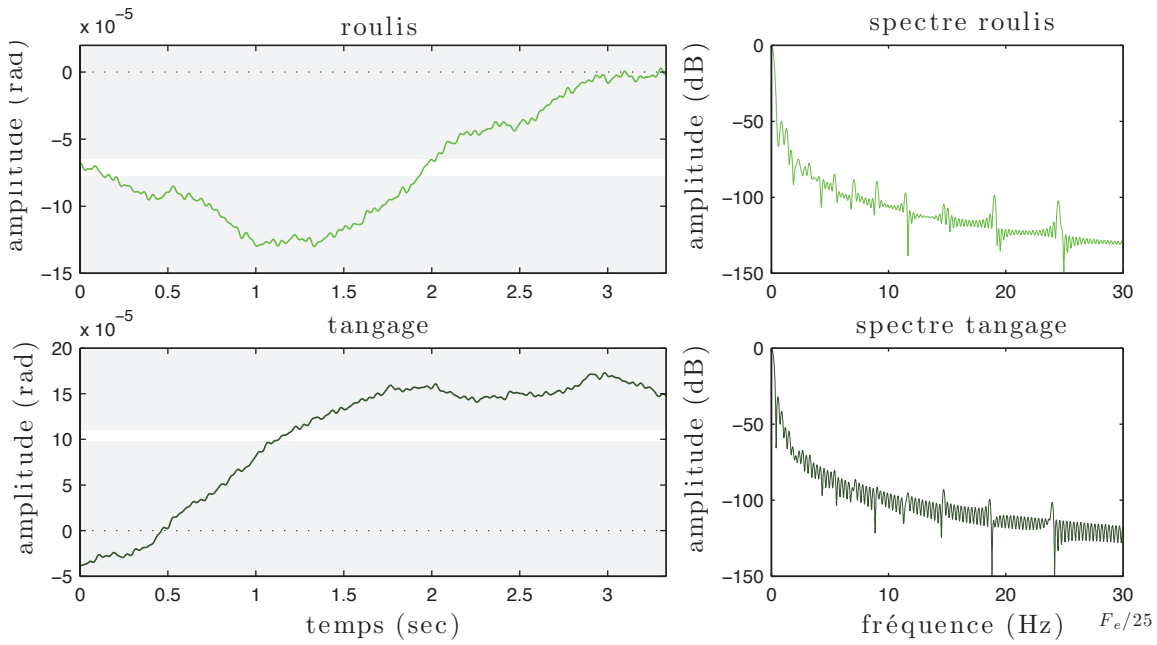


FIGURE IV.31 – Variations d'attitude simulées pour le jeu de données Reunion.

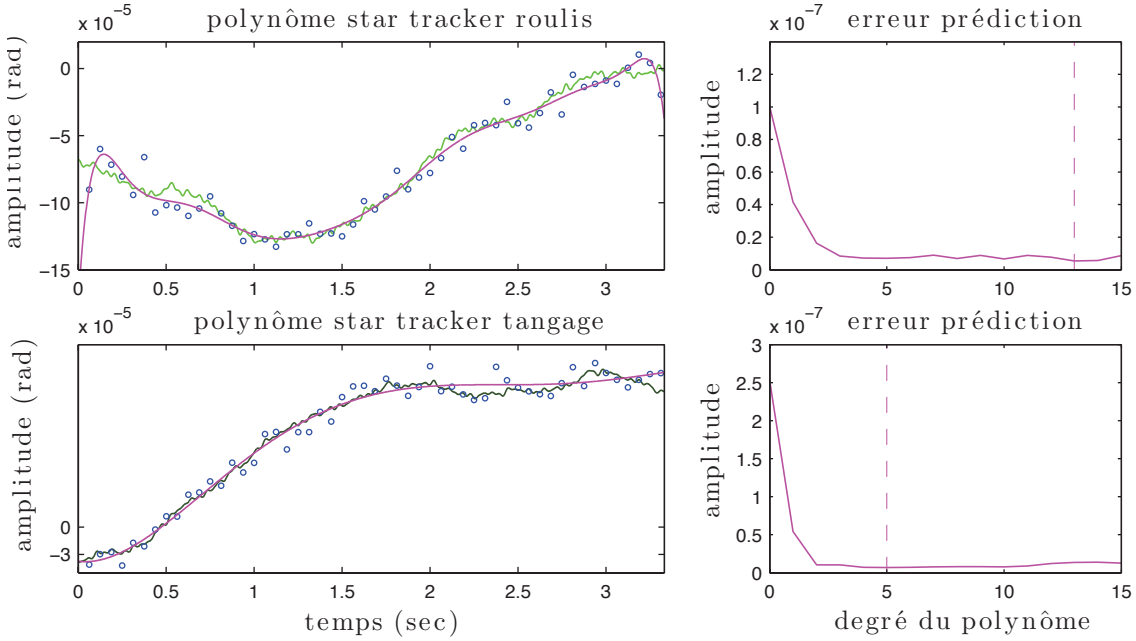


FIGURE IV.32 – Résultat sur le modèle polynômial basse fréquence estimé à partir des mesures *star tracker* du jeu Reunion ; $D_{s^r} = 13$ et $D_{s^t} = 5$.

Modèle autorégressif

La figure IV.33 montre les modèles autorégressifs estimés pour le roulis et le tangage. Ils enregistrent tous globalement la forme basse fréquence du spectre, en distinguant certaines harmoniques du signal comme celles autour de 18Hz et 25Hz. On voit également que le bruit contamine le modèle de façon plus importante que précédemment, de même le pic à 0Hz dû au décalage en amplitude du signal monopolise une partie du modèle alors que celui-ci pourrait mieux se concentrer sur les fréquences utiles. Pour contourner simplement ce problème par la suite, il faudrait supprimer le décalage de $\check{\theta}_{MV}$ avant la phase d'apprentissage.

Résultats

Les performances de chaque méthode sont résumées dans les tableaux IV.8 pour le MAP, IV.9 pour la fusion et IV.10 pour le modèle polynômial par morceaux. Le score de CCN avant mise en correspondance des images est de 0.86. et les écarts-types sont de 3.36 et 5.34 pixel respectivement pour le roulis et le tangage réels.

Discussion

Les résultats obtenus sont assez intéressants ; tout d'abord on constate que le meilleur score d'estimation du tangage est réalisé par le modèle polynômial par morceaux, ce qui tend à prouver que malgré les problèmes rencontrés, celui-ci constitue une alternative crédible. Ensuite, bien que les résultats soient assez mauvais, la fusion apporte une légère amélioration aux performances du premier cas de vraisemblance en précision et en temps de calcul. Le *star tracker* a un niveau de bruit important, mais il est bénéfique au processus

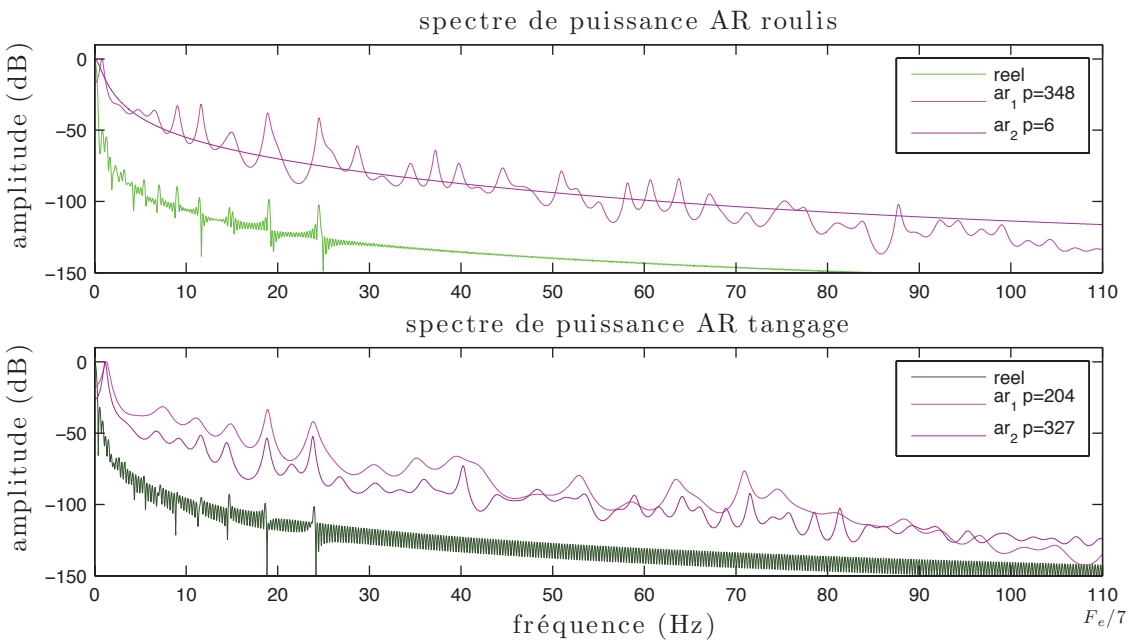


FIGURE IV.33 – Modèles autorégressifs estimés pour la première et seconde vraisemblance sur les données Reunion.

		a priori MAP				MV
		amplitude	dérivée 2^{de}	ps gaussien	autorégressif	
vraisemblance 1	roulis	3.12	2.92	3.20	2.99	3.26
	tangage	4.55	4.34	4.86	4.78	5.06
	CCN	0.86	0.88	0.90	0.88	0.85
	temps	1512	1158	1036	995	743
vraisemblance 2	roulis	1.17	0.47	0.45	0.42	2.83
	tangage	1.28	0.66	0.65	0.66	4.30
	CCN	0.87	0.96	0.96	0.96	0.91
	temps	634	286	513	**281**	304

TABLE IV.8 – Résultats du MAP et du MV sur le jeu de données Reunion. Les mesures pour le roulis et le tangage sont les écarts-types de l'erreur entre attitude réelle et attitude estimée (le meilleur résultat tend vers zéro). La mesure CCN est un score de corrélation croisée normalisée (le meilleur résultat tend vers un). Le temps correspond au nombre de secondes d'exécution de l'algorithme complet (le meilleur est le plus petit nombre). Le meilleur score et le plus mauvais score pour chaque ligne sont respectivement écrits en vert et en rouge. Une case grisée indique que le meilleur résultat a été obtenu par sélection photométrique des hyperparamètres, sinon une case blanche correspond à une estimation de ceux-ci par validation croisée.

	MAP avec fusion	
	dérivée 2de vraisemblance 1	autorégressif vraisemblance 2
roulis	2.75	**0.40**
tangage	4.41	0.64
CCN	0.89	**0.96**
temps	1057	565

TABLE IV.9 – Résultats de la fusion sur le jeu de données Reunion.

	polynômial par morceaux	
	vraisemblance 1	vraisemblance 2
roulis	**3.29**	0.44
tangage	**5.31**	**0.64**
CCN	**0.81**	0.95
temps	1189	**1695**

TABLE IV.10 – Résultats du modèle polynômial par morceaux sur le jeu de données Reunion.

d'estimation dans ces situations de signaux très basses fréquences.

Recalage des images

La meilleure méthode retenue est la même que pour le jeu de données ASTRO : MAP avec fusion selon le deuxième cas de vraisemblance et *a priori* autorégressif, en revanche les hyperparamètres ont été sélectionnés par validation croisée ; $\sigma_I = 0.012$, $\sigma_{\theta^r} = 2.9 \times 10^{-8}$, $\sigma_{\theta^t} = 5.9 \times 10^{-8}$, $\sigma_{c^r} = 2.3 \times 10^{-7}$, $\sigma_{c^t} = 1 \times 10^{-6}$. L'amplitude de la déformation est importante comme en témoigne l'image d'erreur avant recalage où tous les contours sont apparents. L'image d'erreur après rectification de l'attitude est satisfaisante, particulièrement à son début. Le résidu de l'attitude oscille très lentement et presque entièrement dans la bande blanche. On remarque encore une fois que l'estimation est assez précise au centre du signal et un peu moins bonne sur les bords.
Nous pouvons à nouveau construire l'image couleur en sortie de notre algorithme, le résultat est en figure IV.35. Au premier coup d'oeil, on peut être surpris de ne pas observer aussi bien les bénéfices du recalage qu'en figure IV.30, alors que l'amplitude de la déformation est nettement supérieure sur les données Reunion. En réalité, les variations très basses fréquences sont généralement imperceptibles à l'oeil, par chance la configuration du plan focal du satellite fait que les scènes sont observées à des écarts de temps suffisamment courts par chaque caméra. La variation de l'attitude du satellite est trop faible sur ce laps de temps, et les images rouge, verte et bleue rassemblées paraissent bien recalées, alors que sur une échelle de temps suffisamment importante la déformation est bien présente. Ceci est à mettre en correspondance avec le bon score de CCN que l'on obtient avant même de recaler les images. Les améliorations sur la figure IV.35 peuvent tout de même être perçues sur les contours, notamment au niveau de la tache blanche en bas à droite, où deux bavures rouge et verte sont bien présentes dans l'image avant recalage et disparaissent sur l'image de droite. Sur certains contours horizontaux de végétation, on peut aussi voir des reflets bleus et verts qui sont rectifiés après recalage.

IV.5.4 Jeu de données ALMF

Données

Ce dernier jeu de données est aussi composé de quatre images multispectrales, avec une configuration de plan focal et un échantillonnage identiques à ceux vus précédemment. De plus, l'attitude simulée et les mesures *star tracker* associées sont identiques au cas du jeu Reunion, seule la scène observée est différente. Par conséquent nous ne développerons pas les paragraphes attitude et *star tracker* qui sont les mêmes que pour le jeu précédent.

Modèle autorégressif

Les modèles autorégressifs déduits des données sont très simples cette fois-ci, comme le montre la figure IV.36. Quelque soit le type de vraisemblance image, l'ordre du modèle est relativement bas. Ce résultat n'est pas nécessairement mauvais, car si le signal est grossièrement supposé basse fréquence, aucune fréquence parasite n'est apprise en comparaison de la figure IV.33. Le modèle autorégressif se restreint ici à agir à la façon d'un *a priori* sur la dérivée seconde comme nous l'avons vu au chapitre précédent

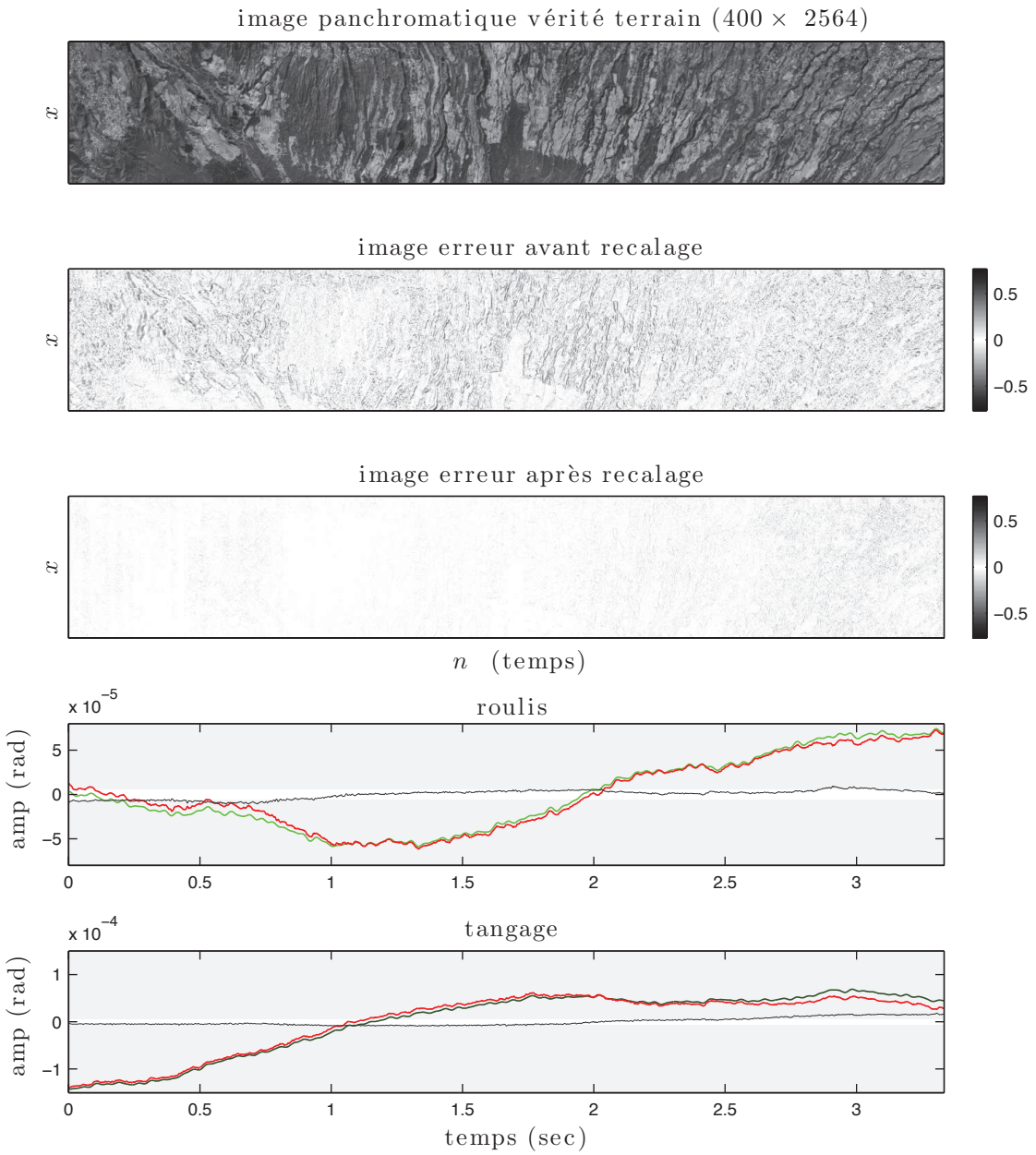


FIGURE IV.34 – Résultat du recalage d'image sur le jeu de données Reunion ; l'estimateur choisi est celui du MAP avec fusion, *a priori* autorégressif et pour la seconde vraisemblance. La figure du haut est un extrait (400 × 2564) de l'image panchromatique non déformée, en dessous les différences d'intensités entre les images vérité terrain / image déformée et vérité terrain / image rectifiée (soit avant et après recalage). Les deux graphiques suivants représentent l'attitude réelle en vert, l'estimée en rouge et le résidu en noir ; la zone blanche délimite la taille d'un pixel.

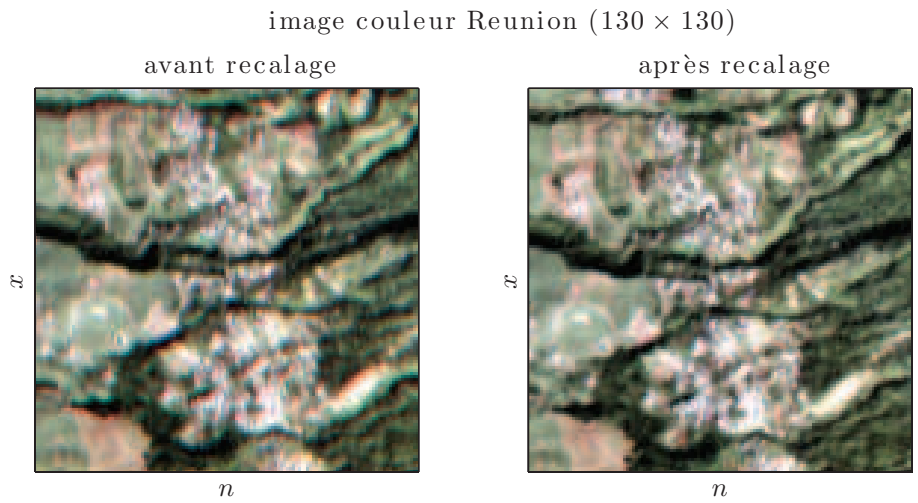


FIGURE IV.35 – Résultat de la construction des images couleurs sur le jeu de données Reunion. Les artéfacts sur l'image avant recalage sont difficiles à repérer à l'oeil nu à cause des variations très basses fréquences de l'attitude. La comparaison de la tache blanche dans le quart en bas à droite des deux images permet tout de même d'apprécier le recalage.

paragraphe III.5.3 (p.49) ; celui-ci a vraiment l'avantage de pouvoir s'accorder à de nombreuses formes spectrales.

Résultats

Comme pour les expériences précédentes, tous les résultats associés à ce jeu de données sont affichés dans les tableaux IV.11 pour l'estimateur du MAP, IV.12 pour la fusion et IV.13 pour le modèle polynômial par morceaux. Le score de CCN est de 0.85 avant alignement des images, et les écarts-types de l'attitude réelle sont logiquement identiques à ceux du jeu Reunion.

Discussion

Cette expérience est intéressante puisqu'elle propose d'estimer le même signal que pour les données Reunion, mais avec des images différentes. En comparaison, on retrouve des résultats très similaires et ce jusqu'aux temps de calcul. Le modèle polynômial confirme ses bons résultats, mais surtout les scores affichés par le MAP avec la seconde vraisemblance sont plus impressionnants : la précision est ici inférieure au quart de pixel, alors que celle observée auparavant pour les images Reunion était de l'ordre du demi pixel. Hors du contexte de fusion, le processus gaussien et le modèle autorégressif - malgré les remarques que nous avons pu faire concernant la simplicité du modèle inféré - sont particulièrement performants. La fusion tire encore son épingle du jeu en présentant une précision de l'ordre du dixième de pixel, mais cette fois-ci avec un *a priori* sur la dérivée seconde ; ce que nous allons tenter de comprendre ci-dessous. Nous n'avons pas de réelle explication sur l'écart en performance des deux expériences, il est possible que la mise en correspondance de certaines modalités sur le jeu Reunion ait été moins bonne, seulement visuellement nous n'avons pas aperçu d'inversion de contraste plus marquée

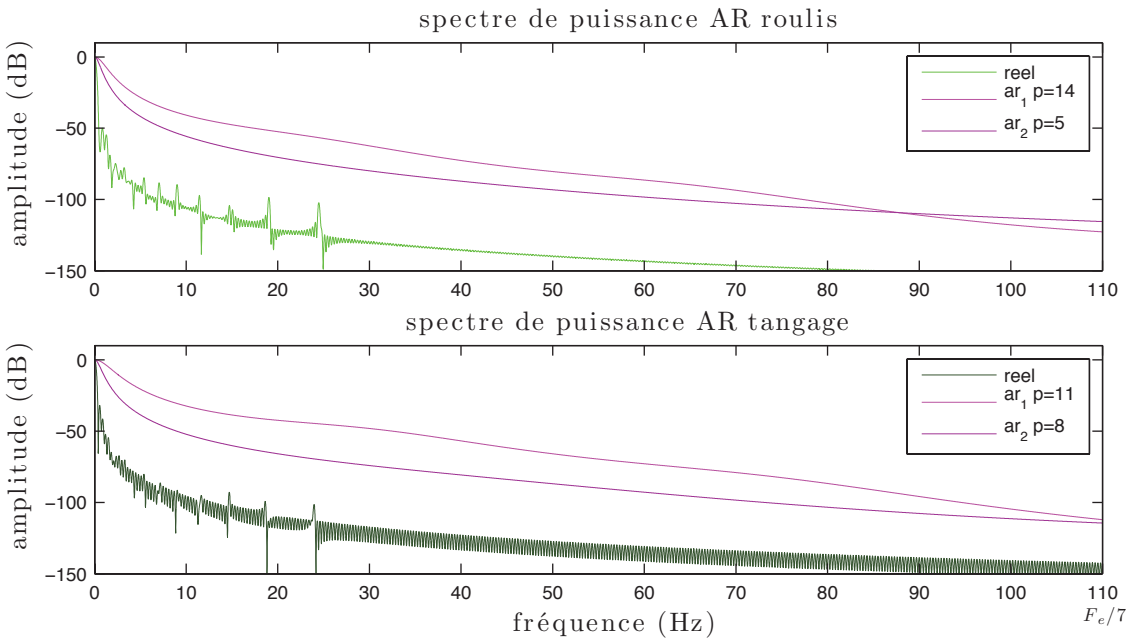


FIGURE IV.36 – Modèles autorégressifs estimés pour la première et seconde vraisemblance sur les données ALMF.

		a priori MAP				MV
		amplitude	dérivée 2^{de}	ps gaussien	autorégressif	
vraisemblance 1	roulis	**4.58**	2.90	2.99	2.90	3.18
	tangage	3.00	4.78	4.95	4.60	5.06
	CCN	0.83	0.88	0.89	0.87	0.84
	temps	1510	998	1024	1360	744
vraisemblance 2	roulis	0.64	0.17	0.19	0.18	2.61
	tangage	1.42	0.14	0.10	**0.10**	4.44
	CCN	**0.79**	0.94	**0.94**	0.94	0.90
	temps	659	260	359	**236**	306

TABLE IV.11 – Résultats du MAP et du MV sur le jeu de données ALMF. Les mesures pour le roulis et le tangage sont les écarts-types de l'erreur entre attitude réelle et attitude estimée (le meilleur résultat tend vers zéro). La mesure CCN est un score de corrélation croisée normalisée (le meilleur résultat tend vers un). Le temps correspond au nombre de secondes d'exécution de l'algorithme complet (le meilleur est le plus petit nombre). Le meilleur score et le plus mauvais score pour chaque ligne sont respectivement écrits en vert et en rouge. Une case grisée indique que le meilleur résultat a été obtenu par sélection photométrique des hyperparamètres, sinon une case blanche correspond à une estimation de ceux-ci par validation croisée.

	MAP dérivée 2de avec fusion	
	vraisemblance 1	vraisemblance 2
roulis	2.81	0.11
tangage	4.51	**0.12**
CCN	0.87	0.94
temps	1055	509

TABLE IV.12 – Résultats de la fusion sur le jeu de données ALMF.

	polynômial par morceaux	
	vraisemblance 1	vraisemblance 2
roulis	3.35	0.28
tangage	**5.20**	0.16
CCN	0.85	0.93
temps	1111	985

TABLE IV.13 – Résultats du modèle polynômial par morceaux sur le jeu de données ALMF.

entre les images des deux jeux. On remarque tout de même qu'au sein des images de vérité terrain des figures IV.34 et IV.37, l'orientation des contours dans le jeu ALMF est plus variée dans toutes les directions. Dans le jeu Reunion, les contours verticaux sont plus marqués ce qui peut biaiser l'estimation.

Recalage des images

Nous avons une nouvelle fois sélectionné l'estimateur réalisant la meilleur performance en terme d'écart-type de l'erreur sur l'attitude estimée : celui-ci est le MAP avec fusion, second cas de vraisemblance, *a priori* sur la dérivée seconde et sélection des paramètres par validation croisée ; $\sigma_I = 0.013$, $\sigma_{\theta^r} = 1.4 \times 10^{-8}$, $\sigma_{\theta^t} = 1.2 \times 10^{-7}$, $\sigma_{c^r} = 2.8 \times 10^{-7}$, $\sigma_{c^t} = 8.5 \times 10^{-7}$. La figure IV.37 présente les résultats en image sur les données ALMF.

L'image d'erreur avant recalage témoigne une nouvelle fois de l'amplitude des déformations, en dessous l'erreur après recalage se comporte un peu comme celle observée en figure IV.29 : les contours disparaissent au profit du bruit photométrique. Le résidu obtenu sur les courbes de l'attitude est légèrement bruité en haute fréquence, surtout pour le tangage. Cette observation rejoint en parti celle faite sur l'écart-type comme critère de qualité, qui a fait que cet estimateur a été préféré à d'autres. Nous y reviendrons plus loin dans la conclusion, cependant la précision obtenue reste très bonne.

Les images de couleur IV.38 permettent de se rendre compte de l'utilité du recalage. Encore une fois, les variations sont basses fréquences ce qui rend les bavures difficiles à observer comme nous l'avons remarqué au jeu de données précédent. Certains contours sur la droite du centre de l'image non recalée présentent cependant quelques artéfacts corrigés sur l'image de gauche.

IV.6 Expériences complémentaires

Dans ce paragraphe nous effectuons certains tests pour compléter notre analyse.

IV.6.1 Pourcentage de pixels dans l'estimation

Nous avons terminé le chapitre précédent en suggérant d'utiliser moins de pixels dans les images pour estimer l'attitude. Une analyse de ce type a déjà été conduite dans [Baker 04], qui tend à dire que seul 10% des pixels sont suffisants pour un recalage géométrique des images. Nous nous contentons ici de mesurer la moyenne de l'écart-type du résidu pour le roulis et le tangage sur le jeu de données ASTRO et MLF, en utilisant différents pourcentages de pixels. L'estimateur est celui du MAP avec *a priori* autorégressif dans les deux cas, où les hyperparamètres sont ceux qui ont été déterminés lors des expérimentations.

La figure IV.39 propose le résultat de ce test, en mesurant également le temps de calcul. Les deux courbes à gauche montrent clairement qu'au delà d'un certain seuil, augmenter le pourcentage de pixels n'améliore pas la précision, on aurait par contre tendance à dire qu'il stabilise les résultats en assurant une précision donnée à partir de 30%. Une extension logique de cette observation serait de calculer les niveaux d'incertitudes associés à l'estimation de l'attitude suivant le nombre de pixels utilisé. Il est important de rappeler que contrairement à [Baker 04], l'inférence se fait à l'aide d'un *a priori*, ce qui doit compenser un nombre de mesures trop faible. Les courbes d'évolution du temps de calcul

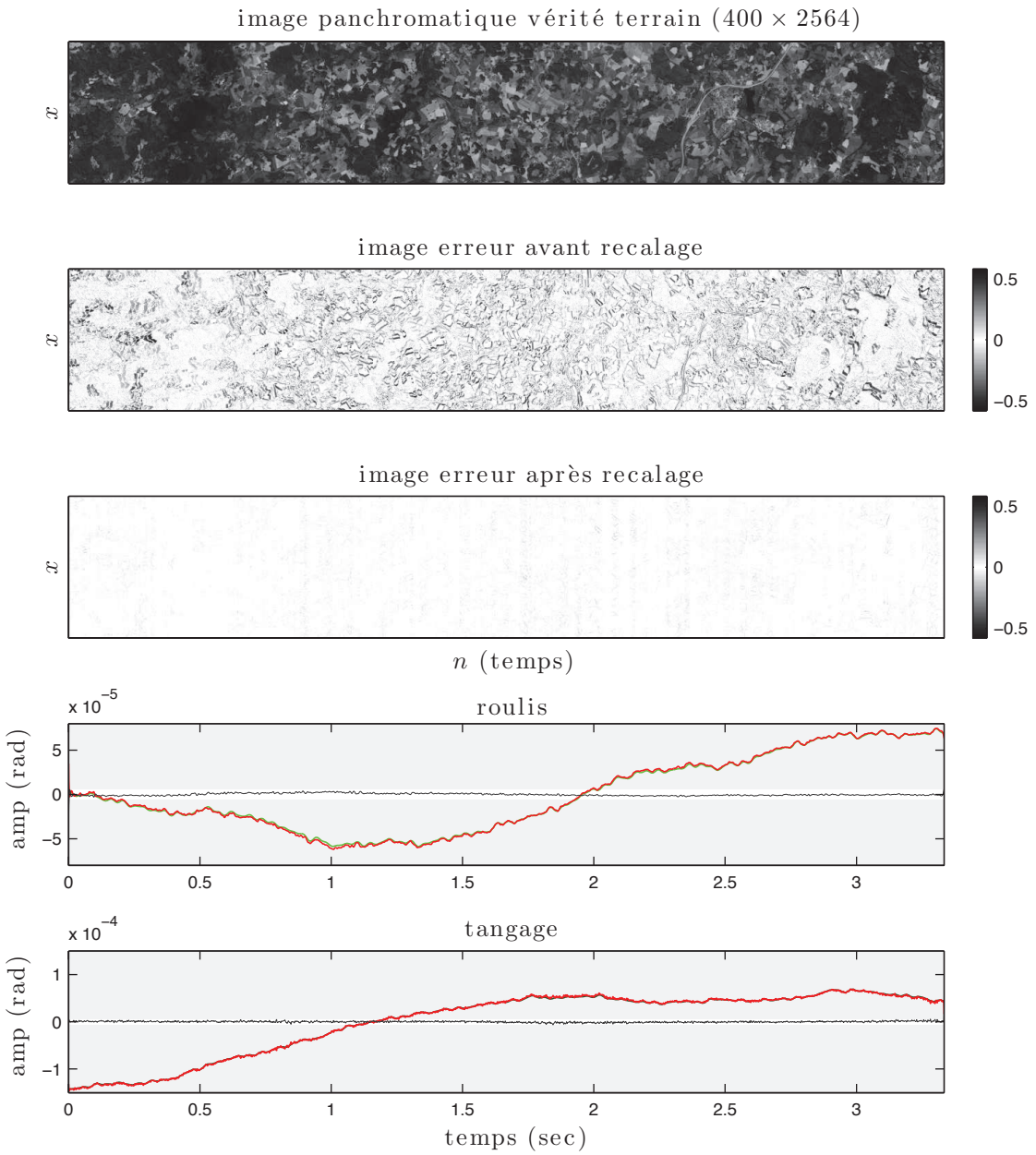


FIGURE IV.37 – Résultat du recalage d'image sur le jeu de données ALMF ; la méthode sélectionnée est le MAP avec fusion, second cas de vraisemblance et *a priori* sur la dérivée seconde. La figure du haut est un extrait (400 × 2564) de l'image panchromatique non déformée, en dessous les différences d'intensités entre les images vérité terrain / image déformée et vérité terrain / image rectifiée (soit avant et après recalage). Les deux graphiques suivants représentent l'attitude réelle en vert, l'estimée en rouge et le résidu en noir ; la zone blanche délimite la taille d'un pixel.

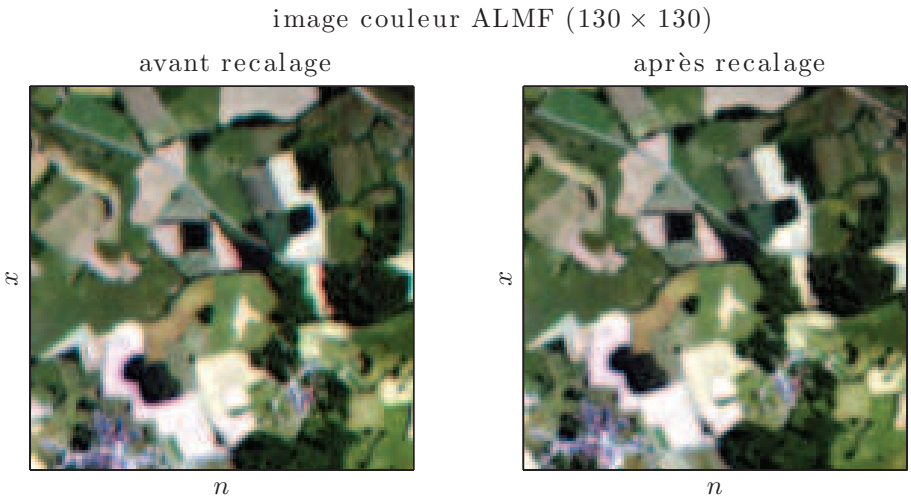


FIGURE IV.38 – Image couleur (130 × 130) construite avant et après recalage sur le jeu de données ALMF. Les artéfacts sont difficiles à observer à cause du type d'oscillation très basse fréquence des variations d'attitude. On peut en apercevoir légèrement à droite du centre de l'image, sur une bande verticale blanche où deux bavures rouge et verte apparaissent. Encore plus à droite à la limite du cadre, le contour de végétation est excessivement vert, ce qui est corrigé dans l'image de droite.

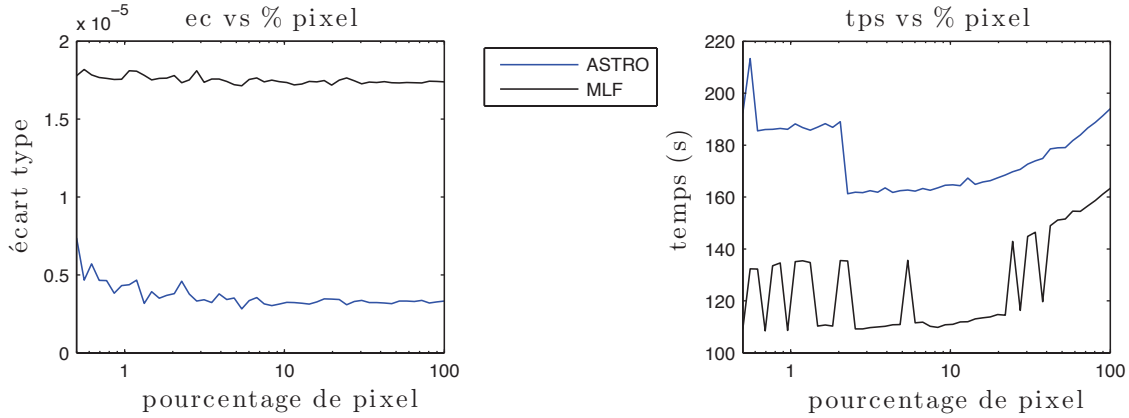


FIGURE IV.39 – Evolution de l'écart-type du résidu de l'erreur d'estimation de l'attitude suivant le pourcentage de pixels utilisés sur le graphique de gauche, évolution du temps de calcul à droite.

sont symptomatiques de ce que nous avons déjà observé dans nos expériences. Ce temps doit être lié au conditionnement du système numérique à résoudre, qui est moins bon lorsque les pixels sont trop peu nombreux.
Utiliser moins de pixels ne permet pas forcément de gagner en temps vu que l'algorithme doit toujours à la fin inverser une matrice de taille $(N \times N)$, N correspondant au nombre d'échantillons temporels de l'attitude que nous recherchons. Par contre, cela permet de gagner de la place en mémoire en évitant le stockage de matrices trop grandes - notamment la jacobienne J - et certains produits matricielles sont plus rapides à faire. Enfin il faut garder à l'esprit que les images réelles de satellite sont largement plus grandes que celles traitées dans ce rapport, cette observation sur le nombre de pixels utiles prend alors tout son sens.

IV.6.2 Influence du choix de l'hyperparamètre

Dans ce paragraphe, nous cherchons à mettre en évidence l'influence de l'hyperparamètre σ_θ sur l'écart-type de l'erreur d'estimation. Nous utilisons le même estimateur qu'au test précédent sur le jeu de données MLF, en recherchant uniquement le roulis.
La courbe noire obtenue à gauche sur la figure IV.40, qui représente l'évolution de l'écart-

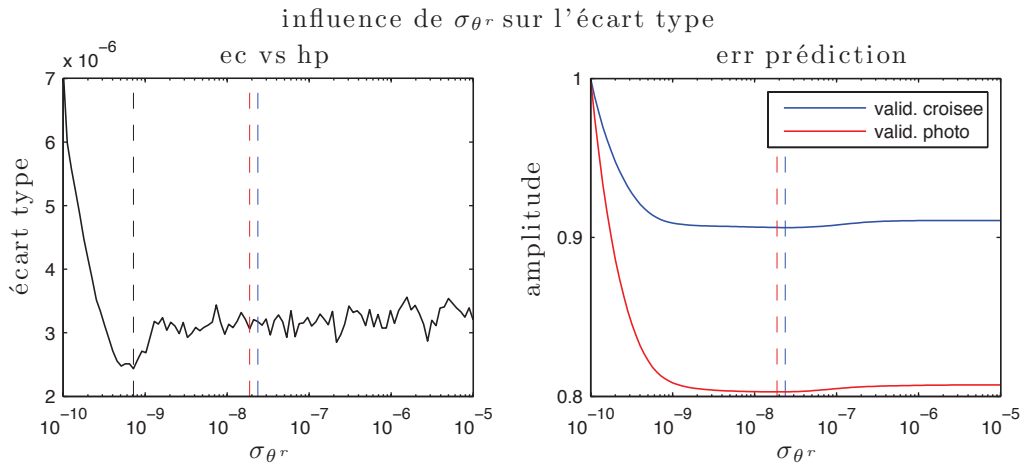


FIGURE IV.40 – Influence de l'hyperparamètre σ_θ sur l'estimation de l'attitude. La courbe noire à gauche trace l'évolution de l'écart-type du résidu entre le roulis réel et le roulis estimé, suivant la valeur de σ_{θ^r}, les courbes à droite montrent les choix qui ont été faits sur cet hyperparamètre par validation croisée et photométrique.

type du résidu suivant la valeur de σ_{θ^r}, est assez clair à interpréter : il est préférable de sur-estimer l'hyperparamètre que de le sous-évaluer, l'erreur croissant fortement entre 10^{-10} et 10^{-9}. A droite, les erreurs de prédiction sur la sélection de σ_{θ^r} par validation croisée et photométrique montrent que ces estimateurs ont tendance à sur-estimer les hyperparamètres dans notre contexte ; nous avons pu l'observer à plusieurs reprises. On peut notamment voir qu'ils n'atteignent pas la valeur idéale autour de 6×10^{-9} que suggère la courbe noire sur la figure de gauche. Ceci laisse à penser que notre approche pour déterminer les hyperparamètres peut encore être améliorée.
De façon générale, l'observation que nous venons de faire se répète sur la plupart des σ_θ que nous recherchons pour les *a priori*. Il faut cependant penser que sur-estimer l'hyper-

paramètre peut être plus critique suivant la qualité de l'*a priori* ou lorsqu'on utilise moins de pixels pour l'inférence. Enfin nous suggérons une nouvelle fois que l'écart-type n'est peut-être pas la meilleure façon d'évaluer l'erreur dans notre problème.

IV.6.3 Estimation du bruit image

Nous avons supposé dans notre travail que l'écart-type du bruit dans les images σ_I était connu, nous l'avons par ailleurs fixé égal à 1.5% de la valeur d'intensité maximum dans l'image. Cette variable pourrait également être incluse dans le processus d'estimation comme un hyperparamètre, c'est notamment ce que propose [Rasmussen 06] dans le contexte des processus gaussiens. Seulement dans cette situation, les problèmes d'unicité sont importants [Jalobeanu 02, Rasmussen 06] : les données peuvent être interprétées par un modèle différent selon chaque niveau de bruit. La figure IV.41 illustre le phénomène, où nous avons cherché par validation croisée à estimer σ_I et σ_{θ^r}. La fonction d'erreur de prédiction montre une succession de minimums locaux le long d'une ligne droite.
L'estimateur utilisé est celui du MAP selon la seconde vraisemblance avec *a priori* sur

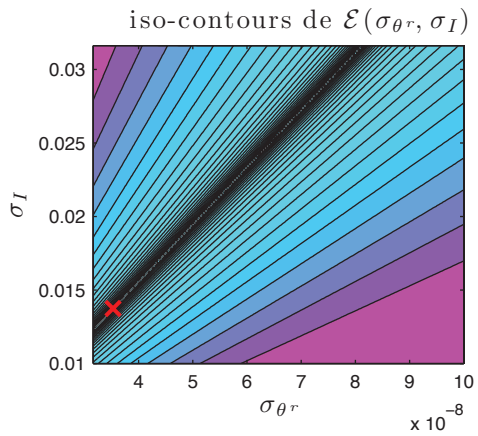


FIGURE IV.41 – Représentation de la fonction d'erreur de prédiction selon σ_I et σ_{θ^r} avec l'estimateur du MAP ; mise en évidence du problème d'unicité pour l'estimation de l'écart-type du bruit σ_I et du paramètre du modèle σ_{θ^r}.

la dérivée seconde et a été testé sur les données MLF. On peut tout de même apprécier le fait que parmi les nombreux minima locaux, les hyperparamètres choisis sont $(\sigma_I, \sigma_{\theta^r}) = [0.013, 3.5 \times 10^{-8}]$, ce qui est très proche du niveau de bruit déterminé par notre critère. Malgré ce point positif, cette recherche de σ_I ajoute trop de degré de liberté à notre système alors que dans les problématiques industrielles, il est souvent possible par expertise d'avoir précisément le niveau de bruit du capteur image. Deux approches se distinguent alors : les méthodes purement bayesiennes qui utilisent des hyper *a priori* sur les hyperparamètres pour apporter une connaissance sur ceux-ci, ceci est clairement justifié au regard de la figure IV.41, l'autre approche consiste à fixer le niveau de bruit à la main ou à le déterminer par des techniques statistiques dans les images.

IV.6.4 Signal non stationnaire

Pour ce dernier test, nous avons mélangé les jeux MLF et ALMF en introduisant des portions d'image du premier dans le second. Nous créons ainsi une discontinuité dans les images et un changement des caractéristiques stationnaires de l'attitude. Nous avons choisi d'expérimenter ce jeu synthétique avec les estimateurs du MAP pour le deuxième cas de vraisemblance et les trois *a priori* : dérivée seconde, processus gaussien et autorégressif. Les résultats sont visibles en figure IV.42.
En dehors des deux échelons marquant la transition entre les attitudes des deux jeux de

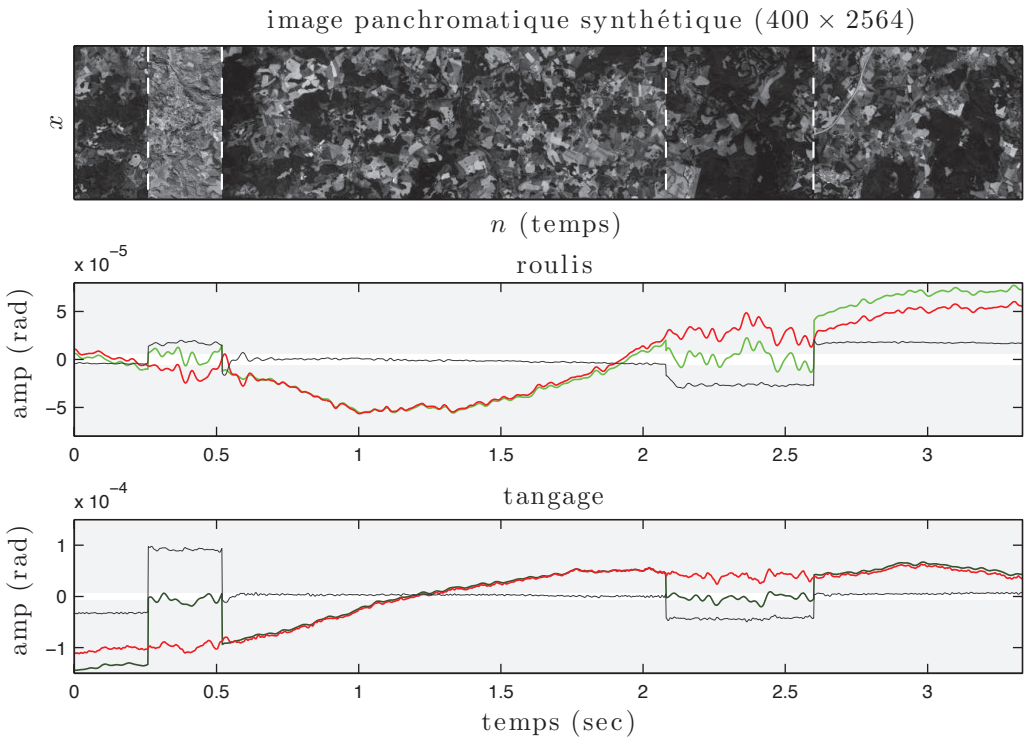


FIGURE IV.42 – Résultat du recalage d'image sur le jeu de données synthétique avec le MAP pour le second cas de vraisemblance et *a priori* de processus gaussien. Les lignes en tirets blancs délimitent les zones d'insertion des portions d'image de MLF dans les données ALMF.

données - qui assez logiquement ne sont pas estimés - , chaque estimateur a su s'adapter correctement aux différents types d'oscillations. Les résultats sont très similaires suivant les estimateurs, avec un très léger avantage pour l'*a priori* sur la dérivée seconde qui par construction émet moins d'hypothèse sur la stationnarité du signal. On aperçoit qu'aux transitions très franches, comme celle sur le roulis après 0.5 seconde, l'estimateur force le lissage ce qui provoque ces petites oscillations sur le résidu. On remarque également au regard du résidu que le modèle appris a l'air d'être un peu plus performant sur l'estimation de ALMF, ce qui semble normal étant donné que MLF est minoritaire sur ce jeu synthétique.

IV.7 Conclusion

Dans ce chapitre, nous avons présenté et testé des méthodes concrètes pour déterminer sans exception toutes les variables du chapitre précédent. L'ensemble des estimateurs a été passé en revue et éprouvé sur des données simulées, qui ont pour caractéristique d'être de difficulté de traitement comparable à des données réelles. Nous avons justifié certaines décisions, notamment sur les aspects de recalage multi-modal, ainsi que les approximations faites pour estimer certains hyperparamètres comme les coefficients autorégressifs ou polynômiaux du modèle basse fréquence. Des problèmes ont été soulevés, comme celui du modèle polynômial par morceaux qui en pratique devrait donner de meilleurs résultats sur des signaux longs, ou le critère de validation des résultats qui ne distingue pas assez les signaux réguliers ; le choix de l'écart-type s'est fait sur la base d'études industrielles faites avant nos travaux. Des pistes d'amélioration ont été proposées dans la plupart des cas.
Il est important de comprendre que les travaux exposés dans ce chapitre ne cherchent pas à obtenir les meilleurs résultats possibles en optimisant localement chaque calcul, ce qui est plus du ressort de l'ingénierie. Ils présentent plutôt les solutions les plus cohérentes vis à vis du problème, au niveau de la méthode de recalage d'images sélectionnée et de l'*a priori* utilisé. Nous avons pris le parti de guider l'estimation, en laissant prendre à l'ordinateur plusieurs décisions. Pour cette raison les degrés de liberté sont importants et les filtres peu sélectifs, mais les meilleurs estimateurs ont tout de même su s'adapter aux différents jeux de données. Rappelons que le phénomène est stationnaire, l'apprentissage des hyperparamètres pourrait être unique, se faire sur des longues périodes d'observations et être légèrement ajusté au besoin. Des paramètres peuvent être figés par expertise, notamment les coefficients autorégressifs ou la taille de la fenêtre de covariance pour les processus gaussiens. Les marges de manœuvre sont importantes à plusieurs niveaux.

CHAPITRE V

Conclusion

Nous avons présenté dans cette thèse un cadre formel et expérimental pour estimer les variations d'attitude d'un satellite à partir des images qu'il photographie. Nous commencerons par citer les avantages de l'approche proposée :

- Le schéma bayesien adopté dans le chapitre 3 est très flexible puisqu'il permet à tout moment d'apporter une information supplémentaire sur le processus à estimer, que ce soit au travers d'un capteur inertiel ou d'un nouvel *a priori*. Le *star tracker* est un exemple de capteur parmi d'autres embarqués dans le satellite et qui peut compléter la mise en correspondance des images. Chaque équation associée aux images et au *star tracker* découle du bruit du capteur qui a engendré ces mesures ; la caractérisation de ce bruit en premier lieu est une façon cohérente d'aborder le problème puisque c'est souvent l'unique information dont on dispose au début de la modélisation. De plus les valeurs des hyperparamètres sont mieux liées à la physique du système et plus facilement interprétables.

- Le recalage est multi-image, son expression permet d'estimer directement l'information haute fréquence sur l'attitude absolue, sans procéder par plusieurs étapes de traitement comme dans [de Lussy 08]. Il tient notamment compte de la géométrie du plan focal et de la linéarité du capteur en associant à tous les pixels d'une même ligne les trois valeurs d'angles de l'attitude. Enfin le choix de l'algorithme de Lucas-Kanade dans ce contexte est judicieux : les déformations sont petites, la précision de recalage est sous-pixellique et la recherche de l'attitude est orientée par la descente de gradient.

- Tous les hyperparamètres sont appris par différents algorithmes vus au chapitre 4. De fait, l'approche est très flexible et les résultats obtenus montrent que les décisions prises par l'ordinateur à la suite de l'apprentissage sont cohérentes.

- La méthode proposée tente d'apporter une réponse simple à plusieurs questions comme la mise en correspondance multimodale ou l'estimation des coefficients autorégressifs. Comme nous le verrons un peu plus loin, les pistes d'amélioration sont nombreuses et l'algorithme peut être spécifié par expertise pour augmenter la précision d'estimation.

Dans ce travail, nous n'avons pas cherché à développer de modèle géométrique complet de l'acquisition *puhsbroom* à la manière de [Gupta 97] et qui tienne compte des variations d'attitude. Notre méthode utilise des matrices de dérivées numériques de la déformation pour s'affranchir du problème d'élévation du terrain. C'est une façon simple de contourner le problème mais dont nous n'avons pas évalué la précision, les matrices étant fournies par EADS Astrium.
Les résultats que nous présentons proviennent de données synthétiques générées par le simulateur d'EADS Astrium et sur lequel nous n'avions pas d'emprise. Il est certain que nous manquons de recul sur ce que notre approche produirait sur des données réelles. En cela la solution que nous proposons est prometteuse mais sa validation, pour être concluante, nécessiterait des tests sur des données réelles.
Enfin certaines situations particulières liées aux données réelles n'ont pas été anticipées, c'est le cas par exemple de la présence de nuages dans les images. Ceux-ci se déplacent et peuvent provoquer des problèmes de parallaxe dus à leur altitude, ils doivent donc être considérés comme des données aberrantes et être supprimés du processus d'estimation. Nous avons vu au chapitre précédent que l'utilisation de moins de 25% des pixels permettait d'obtenir déjà une bonne précision. Une solution peut consister à segmenter au préalable les images en s'assurant grossièrement de choisir des pixels en dehors des nuages ; ces derniers peuvent être assez bien caractérisés par leur couleur et leur homogénéité. Ceci peut aussi être mis en lien avec l'utilisation d'estimateurs robustes pour la segmentation de mouvement dans une scène [Odobez 95] où le mouvement des nuages serait isolé du reste de la scène.

Ce dernier point sur les nuages est une idée de travail de recherche, nous en avons relevé plusieurs que nous résumons ici :

- Le recalage multimodal peut être vu d'une nouvelle façon si l'on suppose que l'image panchromatique est une combinaison linéaire des trois autres. On peut rappeler que la bande panchromatique est physiquement sensible aux trois autres bandes spectrales - rouge, verte et bleue - à la fois. Le filtrage ne semble pas être une solution intéressante car elle supprime de l'information et élève le bruit dans les images. En revanche la correction radiométrique devrait permettre de corriger certains biais et inversions de contraste si son traitement est suffisamment local tout en étant rapide.

- Le modèle polynomial basse fréquence pour la fusion peut être remplacé par un processus gaussien qui varie lentement dans le temps. Ainsi le problème aurait l'avantage de se repenser uniquement en terme de deux processus gaussiens : un basse fréquence estimé sur les mesures *star tracker* et un haute fréquence calculé à partir des images, ceci à la manière des figures III.16 et III.17 (p.54). Les processus gaussiens sont très flexibles vis-à-vis des données tout en assurant un lissage local important lorsque les hyperparamètres sont bien sélectionnés.

- La figure IV.40 (p.136) nous montre qu'un choix judicieux des hyperparamètres permettrait d'observer des résultats plus précis ; les estimateurs que nous avons utilisés ont tendance à surévaluer leur valeur. Il devrait être possible d'intégrer un

critère de régularité pour le signal estimé si l'on se réfère à sa stationnarité, afin d'encourager des solutions plus lisses.

– Nous avons décidé de faire une impasse sur les techniques d'interpolation des images dans cette thèse [1]. Pourtant elles peuvent influencer la précision de l'algorithme de Lucas-Kanade, le temps de calcul - dans notre cas l'interpolation représente de 50% à 80% du temps à chaque itération - , ainsi que le rendu final de l'image couleur. Dans notre contexte, les images sont irrégulièrement échantillonnées et notre attention s'est portée sur les méthodes de régression de noyaux telles que [Katkovnik 06] et [T. 07]. Elles ont montré de très bons résultats dans des applications de dématriçage et d'interpolation sur des grilles irrégulières pour la préservation de contours dans les images. En approchant localement le contenu des images par des polynômes, elles permettent au premier degré d'estimer le gradient selon x et n de manière assez précise. Le lecteur peut se référer au premier chapitre de [Katkovnik 06] pour une analyse détaillée de cette méthode. Nous avons mené quelques tests en utilisant le gradient estimé par regression polynômiale locale pour mettre en correspondance des images par l'algorithme de Lucas-Kanade. Le nombre d'itérations jusqu'à la convergence était sensiblement inférieur au cas d'utilisation d'un gradient plus approximatif calculé par différences finies. Ce travail est préliminaire, mais il représente une perspective intéressante.

– La complexité algorithmique et un temps de calcul précis méritent d'être analysés. Si l'estimation des hyperparamètres est mise à part [2], les étapes clés de l'agorithme sont : le calcul de la jacobienne des images, l'inversion d'une matrice $(N \times N)$ - résultat du produit de toutes matrices - qui estime un angle de l'attitude de taille $(N \times 1)$ et l'interpolation des images. Le premier point est surtout critique du point de vue stockage en mémoire mais nous pouvons utiliser une quantité réduite de pixels pour la créer. Pour le second, il existe plusieurs solutions pour l'inversion d'une matrice dont le coût en temps de calcul varie, il faudrait les comparer en terme de précision [Golub 96]. Enfin pour le dernier sujet, les méthodes d'interpolation sont nombreuses et leur temps de calcul varie souvent de façon inversement proportionnelle avec leur précision, il peut être intéressant d'utiliser un interpolateur rapide au début de la convergence de l'algorithme et plus précis à la fin pour trouver un bon compromis entre vitesse et précision.

1. Nous avons tout de même précisé que nous utilisons l'interpolateur bicubique dans le schéma de l'algorithme au paragraphe IV.4 (p.105) ; les performances de celui-ci ne sont pas optimales dans le contexte du recalage d'images.

2. Nous présumons que plusieurs hyperparamètres seront presque invariants pour un satellite donné, ceci à cause de la nature stationnaire du signal à estimer et de la structure physique du satellite qui reste la même.

ANNEXE A

Matrices et développements MV/MAP

A.1 Matrice de décalage

La matrice de décalage est utilisée pour la première fois dans l'équation III.31 (p.39), elle permet de décaler les échantillons du vecteur d'attitude dans le temps. Soit $\tau = 2$, alors pour $\alpha \in \{l, r, t\}$:

$$D^{\alpha} = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & 1 & 0 & \dots & 0 \\ \vdots & & & \ddots & & & \vdots \\ 0 & & & \dots & & 0 & 1 \\ 0 & & & \dots & & & 0 \\ 0 & & & \dots & & & 0 \end{bmatrix} \text{ et } D = \begin{bmatrix} D^{l} & 0 & 0 \\ 0 & D^{r} & 0 \\ 0 & 0 & D^{t} \end{bmatrix}. \tag{A.1}$$

Rappelons que D^{α} est de taille $(N \times N)$, les deux lignes finales de zéros permettent de compléter le vecteur d'attitude décalé par des échantillons nuls aux endroits où ils ne peuvent pas être connus. Nous avons de plus la relation $D^{\alpha}_{ij} = D^{\alpha T}_{ji}$, autrement dit cela signifie que la matrice D^{α} pour τ est la transposée de la matrice D^{α} pour $-\tau$. Une conséquence est que les lignes de zéros se retrouve cette fois-ci en haut.

Dans le cas où τ n'est pas entier, par exemple pour $\tau = 0.5$, alors :

$$D^{\alpha} = \begin{bmatrix} 0.5 & 0.5 & 0 & 0 & \dots & 0 \\ 0 & 0.5 & 0.5 & 0 & \dots & 0 \\ \vdots & & & \ddots & & \vdots \\ 0 & & \dots & 0 & 0.5 & 0.5 \\ 0 & & \dots & & & 0 \end{bmatrix}. \tag{A.2}$$

Notons que le décalage de l'attitude pour une valeur non entière de τ correspond ici à une interpolation linéaire des échantillons, ce qui a pour effet de filtrer le contenu

haute fréquence du signal. Etant donné la nature basse fréquence θ^α, la multiplication par l'opérateur D^α aura peu d'influence ici.

A.2 Matrice de convolution

La matrice de convolution est appelée à plusieurs endroits comme dans l'équation III.36 (p.40). Elle se construit à partir d'un noyau comme celui décrit en équation III.18 (p. 36). Soit $k = [k_1 \;\; k_2 \;\; k_3]^T$ un vecteur (3×1) réel, si $k_1 > 0$ - ceci signifie que la première caméra est prise comme référence de temps pour le signal à estimer -, alors la matrice de convolution par le noyau k est la suivante pour $\alpha \in \{l, r, t\}$:

$$K^\alpha = \begin{bmatrix} k_1 & k_2 & k_3 & 0 & 0 & \dots & 0 \\ 0 & k_1 & k_2 & k_3 & 0 & \dots & 0 \\ \vdots & & & \ddots & & & \vdots \\ 0 & & \dots & 0 & k_1 & k_2 & k_3 \\ 0 & & & \dots & & & 0 \\ 0 & & & \dots & & & 0 \end{bmatrix} \text{ et } K = \begin{bmatrix} K^l & 0 & 0 \\ 0 & K^r & 0 \\ 0 & 0 & K^t \end{bmatrix}. \tag{A.3}$$

Une nouvelle fois la matrice K^α est de taille $(N \times N)$, les deux lignes de zéros complètent le signal convolué au bord. Si $k_1 < 0$ - dans ce cas la seconde caméra est prise comme référence -, alors :

$$K^\alpha = \begin{bmatrix} 0 & & & \dots & & & 0 \\ 0 & & & \dots & & & 0 \\ k_1 & k_2 & k_3 & 0 & 0 & \dots & 0 \\ 0 & k_1 & k_2 & k_3 & 0 & \dots & 0 \\ \vdots & & & \ddots & & & \vdots \\ 0 & & \dots & 0 & k_1 & k_2 & k_3 \end{bmatrix} \tag{A.4}$$

A.3 Matrice d'*a priori* sur la dérivée seconde

La matrice d'*a priori* sur la dérivée seconde est introduite en équation III.50 (p. 49), elle est de taille $(3N \times 3N)$ et sa construction est la suivante :

$$L^\alpha = \begin{bmatrix} -1 & 1 & 0 & 0 & 0 & \dots & 0 \\ 1 & -2 & 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & -2 & 1 & 0 & \dots & 0 \\ \vdots & & & \ddots & & & \vdots \\ 0 & & \dots & 0 & 1 & -2 & 1 \\ 0 & & & \dots & & 1 & -1 \end{bmatrix} \text{ et } L = \begin{bmatrix} L^l & 0 & 0 \\ 0 & K^r & 0 \\ 0 & 0 & K^t \end{bmatrix}. \tag{A.5}$$

A.4 Matrice d'*a priori* autoregressif

L'équation III.67 (p. 57) exprime pour la première fois la matrice d'*a priori* autoregressif, elle est de taille $(N - P^{\alpha} \times N)$ pour chaque angle, et sa forme est comme suit :

$$R^{\alpha} = \begin{bmatrix} a_{P^{\alpha}} & a_{P^{\alpha}-1} & \dots & a_1 & -1 & 0 & 0 & \dots & 0 \\ 0 & a_{P^{\alpha}} & a_{P^{\alpha}-1} & \dots & a_1 & -1 & 0 & \dots & 0 \\ \vdots & & & & \ddots & & & & \vdots \\ 0 & \dots & & 0 & a_{P^{\alpha}} & a_{P^{\alpha}-1} & \dots & a_1 & -1 \end{bmatrix} \text{ et } R = \begin{bmatrix} R^l & 0 & 0 \\ 0 & R^r & 0 \\ 0 & 0 & R^t \end{bmatrix}. \tag{A.6}$$

A.5 Matrice de temps du modèle polynômial par morceaux

La matrice de temps du modèle polynômial par morceaux est appelée à l'équation III.80 (p.66), elle permet de reconstruire le vecteur d'attitude à partir des coefficients polynômiaux. Elle est de taille $(N \times M(D+1))$ selon chaque angle ; pour $i \in [1 : l_p]$, $m \in [1 : M]$ et $d \in [0 : D]$ chacun de ses termes est égal à :

$$W^{\alpha}\big(i + (m-1)l_p, d + 1 + (m-1)(D+1)\big) = n^{d}_{i+(m-1)l_p} \tag{A.7}$$

où n_j pour $j \in [1 : N]$ sont les instants de temps de tous les échantillons. Dans ses colonnes, elle contient les vecteurs de temps associés à chaque fenêtre et à différents degrés. Encore une fois on a :

$$W = \begin{bmatrix} W^l & 0 & 0 \\ 0 & W^r & 0 \\ 0 & 0 & W^t \end{bmatrix}. \tag{A.8}$$

A.6 Matrice de contraintes du modèle polynômial par morceaux

La matrice de contraintes du modèle polynômial par morceaux est utilisée à l'équation III.79 (p.66), nous reprenons l'exemple de l'équation III.77 pour exprimer la matrice sur un point de contrôle $n_{l_p^m}$ pour $D = 3$, nous rappelons ci-dessous les égalités obtenues :

$$\begin{aligned} b^{\alpha}_{0,m} + b^{\alpha}_{1,m} n_{l_p^m} + b^{\alpha}_{2,m} n^2_{l_p^m} + b^{\alpha}_{3,m} n^3_{l_p^m} &= b^{\alpha}_{0,m+1} + b^{\alpha}_{1,m+1} n_{l_p^m} + b^{\alpha}_{2,m+1} n^2_{l_p^m} + b^{\alpha}_{3,m+1} n^3_{l_p^m} \\ b^{\alpha}_{1,m} + 2b^{\alpha}_{2,m} n_{l_p^m} + 3b^{\alpha}_{3,m} n^2_{l_p^m} &= b^{\alpha}_{1,m+1} + 2b^{\alpha}_{2,m+1} n_{l_p^m} + 3b^{\alpha}_{3,m+1} n^2_{l_p^m} \\ 2b^{\alpha}_{2,m} + 6b^{\alpha}_{3,m} n_{l_p^m} &= 2b^{\alpha}_{2,m+1} + 6b^{\alpha}_{3,m+1} n_{l_p^m}. \end{aligned} \tag{A.9}$$

L'extrait de matrice C^{α} pour ce point de contrôle entre les polynômes des fenêtres m et $m+1$ est le suivant :

$$C^{\alpha} = \begin{bmatrix} 1 & n_{l_p^m} & n^2_{l_p^m} & n^3_{l_p^m} & -1 & -n_{l_p^m} & -n^2_{l_p^m} & -n^{-3}_{l_p^m} \\ 0 & 1 & 2n_{l_p^m} & 3n^2_{l_p^m} & 0 & -1 & -2n_{l_p^m} & -3n^2_{l_p^m} \\ 0 & 0 & 2 & 6n_{l_p^m} & 0 & 0 & -2 & -6n_{l_p^m} \end{bmatrix}. \tag{A.10}$$

Cette forme est répétée successivement sur les $M-1$ points de contrôle. Et nous avons encore :

$$C = \begin{bmatrix} C^l & 0 & 0 \\ 0 & C^r & 0 \\ 0 & 0 & C^t \end{bmatrix}. \tag{A.11}$$

A.7 Matrice de filtrage basse fréquence

La matrice de filtrage basse fréquence réalise le filtrage basse fréquence de l'attitude et est utilisée en équation III.108 (p.75). Soit $b = [b_1 \ b_2 \ b_3 \ b_4 \ b_5]^T$ le vecteur des coefficients du filtre, la matrice est la suivante :

$$B^\alpha = \begin{bmatrix} b_3 & b_4 & b_5 & 0 & 0 & 0 & \dots & 0 \\ b_2 & b_3 & b_4 & b_5 & 0 & 0 & \dots & 0 \\ b_1 & b_2 & b_3 & b_4 & b_5 & 0 & \dots & 0 \\ 0 & b_1 & b_2 & b_3 & b_4 & b_5 & \dots & 0 \\ \vdots & & & \ddots & & & & \vdots \\ 0 & \dots & 0 & b_1 & b_2 & b_3 & b_4 & b_5 \\ 0 & \dots & 0 & 0 & b_1 & b_2 & b_3 & b_4 \\ 0 & \dots & 0 & 0 & 0 & b_1 & b_2 & b_3 \end{bmatrix} \text{ et } B = \begin{bmatrix} B^l & 0 & 0 \\ 0 & B^r & 0 \\ 0 & 0 & B^t \end{bmatrix}. \tag{A.12}$$

A.8 Solution du maximum de vraisemblance

Nous repartons de l'expression III.34 (p.40) :

$$\log(p(\mathcal{I}|\theta)) = \sum_{i,j;i\neq j} -\frac{1}{2\sigma_I^2} \left|\mathbf{e}_{ij} + (J_i - J_j D_{ij})\theta\right|^2 - M\frac{NL}{2}\log(2\pi\sigma_I^2). \tag{A.13}$$

En la dérivant par rapport à θ on obtient :

$$\log(p(\mathcal{I}|\theta)) = \sum_{i,j;i\neq j} -\frac{1}{\sigma_I^2}(J_i - J_j D_{ij})^T(\mathbf{e}_{ij} + (J_i - J_j D_{ij})\theta), \tag{A.14}$$

puis en recherchant la valeur pour laquelle elle s'annule nous avons :

$$-\sum_{i,j;i\neq j}(J_i - J_j D_{ij})^T(J_i - J_j D_{ij})\theta = \sum_{i,j;i\neq j}(J_i - J_j D_{ij})^T\mathbf{e}_{ij}, \tag{A.15}$$

et :

$$\hat{\theta}_{MV} = -\Big(\sum_{i,j;i\neq j}(J_i - J_j D_{ij})^T(J_i - J_j D_{ij})\Big)^{-1}\sum_{i,j;i\neq j}(J_i - J_j D_{ij})^T\mathbf{e}_{ij}. \tag{A.16}$$

A.9 Développement du maximum *a posteriori*

Nous redéveloppons ici l'équation III.42 (p.III.42).

$$p(\theta^\alpha|\mathcal{J}) \propto p(\mathcal{J}|\theta^\alpha)p(\theta^\alpha)$$

$$\begin{aligned}
\propto & \prod_{i,j;i\neq j} \exp\Big(-\frac{1}{2\sigma_I^2}(\mathbf{e}_{ij} + (J_i^\alpha - J_j^\alpha D_{ij}^\alpha)\theta^\alpha)^T(\mathbf{e}_{ij} + (J_i^\alpha - J_j^\alpha D_{ij}^\alpha)\theta^\alpha)\Big) \exp\Big(-\frac{1}{2\sigma_{\theta^\alpha}^2}\theta^{\alpha T}\theta^\alpha\Big) \\
= & \prod_{i,j;i\neq j} \exp\Big(-\frac{1}{2}\big(\sigma_I^{-2}(\mathbf{e}_{ij} + (J_i^\alpha - J_j^\alpha D_{ij}^\alpha)\theta^\alpha)^T(\mathbf{e}_{ij} + (J_i^\alpha - J_j^\alpha D_{ij}^\alpha)\theta^\alpha) + \sigma_{\theta^\alpha}^{-2}\theta^{\alpha T}\theta^\alpha\big)\Big) \\
= & \prod_{i,j;i\neq j} \exp\Big(-\frac{1}{2}\big(\theta^{\alpha T}\big(\sigma_I^{-2}(J_i^\alpha - J_j^\alpha D_{ij}^\alpha)^T(J_i^\alpha - J_j^\alpha D_{ij}^\alpha) + \sigma_{\theta^\alpha}^{-2}I_d\big)\theta^\alpha \ldots \\
& \ldots + 2\sigma_I^{-2}\theta^{\alpha T}(J_i^\alpha - J_j^\alpha D_{ij}^\alpha)^T\mathbf{e}_{ij} + \sigma_I^{-2}\mathbf{e}_{ij}^T\mathbf{e}_{ij}\big)\Big) \\
\propto & \prod_{i,j;i\neq j} \exp\Big(-\frac{1}{2}\big(\theta^{\alpha T}\big(\sigma_I^{-2}(J_i^\alpha - J_j^\alpha D_{ij}^\alpha)^T(J_i^\alpha - J_j^\alpha D_{ij}^\alpha) + \sigma_{\theta^\alpha}^{-2}I_d\big)\theta^\alpha + 2\sigma_I^{-2}\theta^{\alpha T}(J_i^\alpha - J_j^\alpha D_{ij}^\alpha)^T\mathbf{e}_{ij}\Big).
\end{aligned}$$

En posant :

$$\bar{\theta}^\alpha = -\sigma_I^{-2}(\sigma_I^{-2}\sum_{i,j;i\neq j}(J_i^\alpha - J_j^\alpha D_{ij}^\alpha)^T(J_i^\alpha - J_j^\alpha D_{ij}^\alpha) + \sigma_{\theta^\alpha}^{-2}I_d)^{-1}\sum_{i,j;i\neq j}(J_i^\alpha - J_j^\alpha D_{ij}^\alpha)^T\mathbf{e}_{ij},$$

on peut réécrire le dernier développement comme :

$$\propto \exp\Big(-\frac{1}{2}(\theta^\alpha - \bar{\theta}^\alpha)^T\big(\sigma_I^{-2}\sum_{i,j;i\neq j}(J_i^\alpha - J_j^\alpha D_{ij}^\alpha)^T(J_i^\alpha - J_j^\alpha D_{ij}^\alpha) + \sigma_{\theta^\alpha}^{-2}I_d\big)(\theta^\alpha - \bar{\theta}^\alpha)\Big).$$

Pour mieux s'en rendre compte, il suffit d'injecter l'expression de $\bar{\theta}^\alpha$ dans cette dernière relation et de vérifier l'égalité au terme constant $\bar{\theta}^{\alpha T}\bar{\theta}^\alpha$ près. Le calcul procède par complétion du carré.

Publications de la thèse

Les travaux de thèse ont donné lieu aux publications suivantes :

- R. Perrier, E. Arnaud, P. Sturm & M. Ortner. *Estimating satellite attitude from pushbroom sensors*. In Proceedings of the conference Computer Vision and Pattern Recognition, 2010.

- R. Perrier, E. Arnaud, P. Sturm & M. Ortner. *Image-based satellite attitude estimation*. In Proceedings of the International Geoscience and Remote Sensing Symposium, 2010.

- R. Perrier, E. Arnaud, P. Sturm & M. Ortner. *Satellite image registration for attitude estimation with a constrained polynomial model*. In Proceedings of the International Conference on Image Processing, 2010.

- R. Perrier, E. Arnaud, P. Sturm & M. Ortner. *Estimation de l'attitude d'un satellite par recalage d'images*. CORESA - COmpression et REprésentation des Signaux Audiovisuels, 2010.

- R. Perrier, E. Arnaud, P. Sturm & M. Ortner. *Sensor measurements and image registration fusion to retrieve variations of satellite attitude*. In Proceedings of the Asian Conference on Computer Vision, 2010.

Ils sont couverts par le brevet suivant :

- M. Ortner, R. Perrier, E. Arnaud & P. Sturm. *Procédé d'estimation du mouvement d'un instrument d'observation à défilement survolant un corps céleste*. Numéro de dépôt : WO/2011/061123, date de publication : 26/05/2011.

Bibliographie

[Akaike 74] H. Akaike. *A new look at the statistical model identification*. IEEE Transactions on Automatic Control, vol. 19, no. 6, pages 716–723, 1974.

[Andrew 09] N. Andrew. *Learning Theory, Regularization and Model Selection*. Lecture notes on Machine Learning, `http://www.stanford.edu/class/cs229/`, 2009.

[Authesserre 09] J. B. Authesserre, R. Mégret & Y. Berthoumieu. *Formulation bidirectionnelle pour la conception d'algorithmes d'alignement d'images*. In ORASIS actes du Congrès des jeunes chercheurs en vision par ordinateur, 2009.

[Baker 04] S. Baker & I. Matthews. *Lucas-Kanade 20 Years On : A Unifying Framework*. International Journal of Computer Vision, vol. 56, no. 1, pages 221–255, 2004.

[BallAerospace 07] BallAerospace. `http://www.ballaerospace.com/file/media/D150320HAST2001_09_loRes.pdf`, 2007.

[Bartoli 08] A. Bartoli. *Maximizing the Predictivity of Smooth Deformable Image Warps through Cross-Validation*. Journal of Mathematical Imaging and Vision, vol. 31, pages 133–145, 2008.

[Bartoli 09] A. Bartoli. *On Computing the Prediction Sum of Squares Statistic in Linear Least Squares Problems with Multiple Parameter or Measurement Sets*. International Journal of Computer Vision, vol. 85, pages 133–142, 2009.

[Bayer 76] B. E. Bayer. *Color Imaging Array*, 1976.

[Bishop 07] C. M. Bishop. Pattern recognition and machine learning. Springer, 2007.

[Black 96] M. J. Black & P. Anandan. *The robust estimation of multiple motions : parametric and piecewise-smooth flow fields*. Computer Vision and Image Understanding, vol. 63, no. 1, pages 75–104, 1996.

[Brunet 10] F. Brunet, A. Bartoli, N. Navab & R. Malgouyres. *Utilisation de l'information photométrique pour la sélection des hyperparamètres en recalage géométrique d'images*. In Actes des journées COmpression et REprésentation des Signaux Audiovisuels, September 2010.

[Buades 05] A. Buades, B. Coll & J. M. Morel. *A review of image denoising algorithms, with a new one*. Multiscale Modeling and Simulation, vol. 4, no. 2, pages 490–530, 2005.

[Canny 86] F. J. Canny. *A Computational Approach to Edge Detection*. IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 8, no. 6, pages 679–698, 1986.

[Carden 04] E. P. Carden & P. Fanning. *Vibration Based Condition Monitoring : A Review*. Structural Health Monitoring, vol. 3, no. 4, pages 355–377, 2004.

[Chang 00] S. G. Chang, Bin Y. & M. Vetterli. *Adaptive wavelet thresholding for image denoising and compression*. IEEE Transactions on Image Processing, vol. 9, no. 9, pages 1532–1546, 2000.

[Cocquerez 95] J. P. Cocquerez & S. Philipp. Analyse d'images : filtrage et segmentation. Dunod, 1995.

[Cramer 00] M. Cramer, D. Stallmann & N. Haala. *Direct Georeferencing Using GPS/inertial Exterior Orientations For Photogrammetric Applications*. In International Archives of Photogrammetry and Remote Sensing, pages 198–205, 2000.

[Crombez 08] V. Crombez, U. Polverini & G. Flandin. *Method of restoring movements of the line of sight of an optical instrument*, 2008.

[d'Angelo 10] P. d'Angelo. *Hugin - Panorama photo stitcher*. `http://hugin.sourceforge.net/`, 2010.

[de Lussy 08] F. de Lussy, D. Greslou & L. Gross Colzy. *Process Line for Geometrical Image Correction of Disruptive Microvibrations*. In International Society for Photogrammetry and Remote Sensing, pages 27–35, 2008.

[Doebling 96] S.W. Doebling, C.R. Farrar, M.B. Prime & D.W. Shevitz. *Damage identification and health monitoring of structural and mechanical systems from changes in their vibration characteristics : A literature review*. Rapport technique, Los Alamos National Lab., 1996.

[Drareni 08] J. Drareni, P. Sturm & S. Roy. *Plane-Based Calibration for Linear Cameras*. In The 8th Workshop on Omnidirectional Vision, Camera Networks and Non-classical Cameras - OMNIVIS, 2008.

[Farsiu 05] S. Farsiu, M. Elad & P. Milanfar. *Constrained, Globally Optimal, Multi-Frame Motion Estimation*. In Proceedings of the Workshop on Statistical Signal Processing, pages 1396–1401, 2005.

[Fergus 06] R. Fergus, B. Singh, A. Hertzmann, S. T. Roweis & W. T. Freeman. *Removing camera shake from a single photograph*. In ACM SIGGRAPH, pages 787–794. ACM, 2006.

[Field 94] D. J. Field. *What is the goal of sensory coding ?* Neural Computation, vol. 6, pages 559–601, 1994.

[Flandrin 99] P. Flandrin. Time-frequency/time-scale analysis, volume 10 of *Wavelet Analysis and its Applications*. Academic Press Inc., San Diego, CA, 1999.

[Foroosh 02] H. Foroosh, J. Zerubia & M. Berthod. *Extension of phase correlation to subpixel registration*. IEEE Transactions on Image Processing, pages 188–200, 2002.

[GeoEye 10] GeoEye. `http://www.geoeye.com/CorpSite/assets/docs/technical-papers/2010/ASPRS_GeoEye-1_Presentation_2010.pdf`, 2010.

[Golub 96] G. H. Golub & C. F. Van Loan. Matrix computations. Johns Hopkins University Press, Baltimore, MD, USA, 1996.

[Gordon 05] P. Gordon. *Airborne Pushbroom Line Scanners : An Alternative to Digital Frame Scanners*. Geoinformatics, vol. 8, no. 1, pages 50–57, 2005.

[Gravey 81] A. Gravey & B. Picinbono. *Déterminisme et indéterminisme de certains signaux aléatoires*. In Colloque GRETSI sur le Traitement du Signal et des Images, pages 147–152, 1981.

[Gruen 03] A. Gruen & L. Zhang. *Sensor Modeling for Aerial Triangulation with Three-Line Scanner (TLS) Imagery*. Photogrammetrie, Fernerkundung, Geoinformation (PFG), no. 2, pages 85–98, 2003.

[Guimond 01] A. Guimond, A. Roche, N. Ayache & J. Meunier. *Multimodal Brain Warping Using the Demons Algorithm and Adaptative Intensity Corrections*. IEEE Transactions on Medical Imaging, vol. 20, no. 1, pages 58–69, 2001.

[Gupta 95] R. Gupta & R. I. Hartley. *Camera Estimation for Orbiting Pushbrooms*. In Proceedings of the Asian Conference on Computer Vision, 1995.

[Gupta 97] R. Gupta & R. I. Hartley. *Linear Pushbroom Cameras*. IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 19, no. 9, pages 963–975, 1997.

[Hansen 00] P. C. Hansen. *The L-Curve and its Use in the Numerical Treatment of Inverse Problems*. In Computational Inverse Problems in Electrocardiology, pages 119–142. WIT Press, 2000.

[Harris 88] C. Harris & M. Stephens. *A Combined Corner and Edge Detector*. In The Fourth Alvey Vision Conference, pages 147–151, 1988.

[Hasinoff 10] S. W. Hasinoff, F. Durand & W. T. Freeman. *Noise-optimal capture for high dynamic range photography*. In Proceedings of the conference Computer Vision and Pattern Recognition, pages 553–560, 2010.

[Hirschmüller 05] H. Hirschmüller, F. Scholten & G. Hirzinger. *Stereo Vision Based Reconstruction of Huge Urban Areas from an Airborne Pushbroom Camera (HRSC)*. In Pattern Recognition, volume 3663, pages 58–66. Springer Berlin Heidelberg, 2005.

[Inglada 04] Jordi Inglada & Giros Alain. *On the possibility of automatic multisensor image registration*. IEEE Transactions on Geoscience and Remote Sensing, vol. 42, no. 10, 2004.

[Irani 98] M. Irani & I. P. Anandan. *Robust Multi-Sensor Image Alignment*. In Proceedings of International Conference on Computer Vision, pages 959–966, 1998.

[Irani 99] M. Irani & P. Anandan. *About Direct Methods*. In Proceedings of the International Workshop on Vision Algorithms, pages 267–277, 1999.

[Jalobeanu 01] A. Jalobeanu. *Modèles, estimation bayésienne et algorithmes pour la déconvolution d'images satellitaires et aériennes*. PhD thesis, University of Nice Sophia Antipolis, 2001.

[Jalobeanu 02] A. Jalobeanu, L. Blanc-Féraud & J. Zerubia. *Hyperparameter estimation for satellite image restoration using a MCMC Maximum Likelihood method*. Pattern Recognition, vol. 35, no. 2, 2002.

[Jaynes 03] E. T. Jaynes. Probability Theory - The Logic of Science. Cambridge University Press, 2003.

[Kalman 60] R. E. Kalman. *A New Approach to Linear Filtering and Prediction Problems*. Transactions of the ASME Journal of Basic Engineering, vol. 82, no. Series D, pages 35–45, 1960.

[Katkovnik 06] V. Katkovnik, K. Egiazarian & J. Astola. Local approximation techniques in signal and image processing. SPIE Press, 2006.

[Liu 08] C. Liu, R. Szeliski, S. B. Kang, C. L. Zitnick & W. T. Freeman. *Automatic Estimation and Removal of Noise from a Single Image*. IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 30, no. 2, pages 299–314, 2008.

[Lowe 04] D. G. Lowe. *Distinctive Image Features from Scale-Invariant Keypoints*. International Journal of Computer Vision, vol. 60, pages 91–110, 2004.

[Lucas 81] B. D. Lucas & T. Kanade. *An Iterative Image Registration Technique with an Application to Stereo Vision*. In Proceedings of the International Joint Conference on Artificial Intelligence, pages 674–679, 1981.

[Luisier 11] F. Luisier, T. Blu & M. Unser. *Image Denoising in Mixed Poisson-Gaussian Noise*. IEEE Transactions on Image Processing, vol. 20, no. 3, pages 696–708, 2011.

[Mackay 92] D. J. C. Mackay. *Bayesian Interpolation*. Neural Computation, vol. 4, no. 3, pages 415–447, 1992.

[Mackay 02] D. J. C. Mackay. Information Theory, Inference and Learning Algorithms. Cambridge University Press, 2002.

[Makhoul 75] J. Makhoul. *Linear prediction : A tutorial review*. Proceedings of the IEEE, vol. 63, no. 4, pages 561–580, 1975.

[Marple 86] S L. Marple. Digital spectral analysis : with applications. Prentice-Hall, Inc., 1986.

[Marr 80] D. Marr & E. Hildreth. *Theory of Edge Detection*. Proceedings of the Royal Society of London, vol. 207, no. 1167, pages 187–217, 1980.

[McHugh 05] S. McHugh. `http://www.cambridgeincolour.com/tutorials/image-noise.htm`, 2005.

[Microsoft 10] Microsoft. `http://www.microsoft.com/ultracam/en-us/UltraCamEagle.aspx`, 2010.

[MOMS 83] MOMS. `http://science.nasa.gov/missions/moms`, 1983.

[Odobez 95] J. M. Odobez & P. Bouthemy. *Estimation robuste multiéchelle de modèles paramétrés de mouvement sur des scènes complexes*. Traitement du Signal, vol. 12, no. 2, pages 113–128, 1995.

[Peyrard 01] N. Peyrard. *Approximations de type champ moyen des modèles de champ de Markov pour la segmentation de données spatiales*. PhD thesis, Université Joseph Fourier, 2001.

[Poli 02] D. Poli. *General model for airborne and spaceborne linear array sensors*. In International Archives of Photogrammetry and Remote Sensing, volume 34, 2002.

[Punska 99] O. Punska. *Bayesian approaches to multi-sensor data fusion*. Master of Philosophy dissertation, Cambridge University, 1999.

[QuickBird 01] QuickBird. `http://www.ballaerospace.com/page.jsp?page=77`, 2001.

[Rademacher 98] P. Rademacher & G. Bishop. *Multiple-center-of-projection images*. In Proceedings of SIGGRAPH, pages 199–206, New York, NY, USA, 1998. ACM.

[Rao 04] S. S. Rao. Mechanical vibrations. Prentice Hall, 2004.

[Rasmussen 06] C. E. Rasmussen & C. Williams. Gaussian Processes for Machine Learning. MIT Press, 2006.

[Robinson 04] D. Robinson & P. Milanfar. *Fundamental performance limits in image registration*. IEEE Transactions on Image Processing, vol. 13, no. 9, pages 1185 –1199, 2004.

[Roche 98] A. Roche, G. Malandain, X. Pennec & N. Ayache. *The Correlation Ratio as a New Similarity Measure for Multimodal Image Registration*. In Proceedings of the Conference on Medical Image Computing and Computer Assisted Intervention, pages 1115–1124, London, UK, 1998. Springer-Verlag.

[Roche 99] A. Roche, G. Malandain, N. Ayache & S. Prima. *Towards a Better Comprehension of Similarity Measures Used in Medical Image Registration*. In Medical Image Computing and Computer Assisted Intervention, volume 1679, pages 555–566. Springer Berlin Heidelberg, 1999.

[Rougé 97] B. Rougé. Théorie de la chaîne image et restauration d'image optique à bruit final fixé. Habilitation à Diriger des Recherches, 1997.

[Rousseeuw 87] P. J. Rousseeuw & A. M. Leroy. Robust regression and outlier detection. John Wiley & Sons, Inc., New York, NY, USA, 1987.

[Rue 05] H. Rue & L. Held. Gaussian Markov random fields : Theory and applications, volume 104 of *Monographs on Statistics and Applied Probability*. Chapman & Hall, London, 2005.

[Runge 01] C. Runge. *Über empirische Funktionen und die Interpolation zwischen äquidistanten Ordinaten, Zeitschrift für Mathematik und Physik 46*, 1901.

[Schwarz 78] G. Schwarz. *Estimating the Dimension of a Model*. The Annals of Statistics, vol. 6, no. 2, pages 461–464, 1978.

[SPOT 86] SPOT. `http://www.cnes.fr/web/CNES-fr/301-de-spot-1-a-spot-5.php`, 1986.

[Starck 02] J. L. Starck, E. J. Candes & D. L. Donoho. *The curvelet transform for image denoising*. IEEE Transactions on Image Processing, vol. 11, no. 6, pages 670–684, 2002.

[Szeliski 06] R. Szeliski. *Image alignment and stitching : a tutorial*. Foundations and Trends in Computer Graphics and Vision, vol. 2, no. 1, pages 1–104, 2006.

[T. 07] Hiroyuki T., Sina F. & Peyman M. *Kernel regression for image processing and reconstruction*. IEEE Transactions on Image Processing, vol. 16, pages 349–366, 2007.

[Tikhonov 77] A. N. Tikhonov & V. Y. Arsenin. Solutions of Ill-Posed Problems. V. H. Winston & Sons, Washington, D.C. : John Wiley & Sons, New York,, 1977.

[Unser 99] M. Unser. *Splines : A Perfect Fit for Signal and Image Processing*. IEEE Transactions on Signal Processing, vol. 16, no. 6, pages 22–38, 1999.

[Viola 97] P. Viola & William M. Wells III. *Alignment by Maximization of Mutual Information*. International Journal of Computer Vision, vol. 24, pages 137–154, 1997.

[Wahba 90] G. Wahba. Spline models for observational data, volume 59. Society for Industrial and Applied Mathematics, 1990.

[Wiener 64] N. Wiener. Extrapolation, interpolation, and smoothing of stationary time series. The MIT Press, 1964.

[Wong 03] M. Wong. `http://www.photoxels.com/tutorial_noise.html`, 2003.

[Wowk 91] V. Wowk. Machinery vibration : measurement and analysis. McGraw-Hill, 1991.

[Xiao 07] J. Xiao, P. Borgnat & P. Flandrin. *Sur un test temps-fréquence de stationnarité*. In Colloque GRETSI sur le Traitement du Signal et des Images, pages 957–960, 09 2007.

[Zheng 00] W. X. Zheng. *Autoregressive parameter estimation from noisy data*. IEEE Transactions on Circuits and Systems II : Analog and Digital Signal Processing, vol. 47, no. 1, pages 71–75, 2000.

[Zitov 03] B. Zitov & J. Flusser. *Image registration methods : a survey*. Image and Vision Computing, vol. 21, no. 11, pages 977–1000, 2003.