Université Paris Cité
École doctorale Sciences Mathématiques de Paris Centre (ED 386)
Laboratoire MAP5, CNRS UMR 8145

Modèles pour la correction d'images thermiques issues de caméras non-refroidies sans obturateur mécanique

par **Laurent BIDAULT**

Présentée et soutenue publiquement le 18 décembre 2025

Thèse de doctorat de mathématiques appliquées

Dirigée par **Lionel moisan**
Co-encadrée par **Rémy Abergel**
Co-encadrée par **Olivier Harant**

Devant un jury composé de :

Lionel Moisan	*Professeur, Université Paris Cité*	Directeur de thèse
Jérôme Gilles	*Professeur, San Diego State University*	Rapporteur
Mauro Dalla Mura	*Maître de conférence, Institut national polytechnique de Grenoble*	Rapporteur
Julie Delon	*Professeure, École normale supérieure de Paris*	Examinatrice
Gabriele Facciolo	*Professeur, École normale supérieure de Paris-Saclay*	Examinateur
Rémy Abergel	*Ingénieur de recherche CNRS, Université Paris Cité*	Membre invité et co-encadrant
Olivier Harant	*Ingénieur, Lynred*	Membre invité et co-encadrant

Ce qui rend les choses visibles n'est pas visible.

Jacques Derrida

École doctorale 386
(Sciences mathématiques
de Paris Centre)
4 place Jussieu
75252 Paris cedex 05

Laboratoire MAP5
(Université de Paris)
45 rue des Saints-Pères
75006 Paris

LYNRED
364 route de Valence
38113 Veurey-Voroize

Remerciements

Je souhaite tout d'abord exprimer ma profonde gratitude à mon directeur de thèse, Lionel Moisan. Son investissement constant, sa disponibilité et la justesse de ses remarques ont été essentiels tout au long de ce travail. Son sens de l'explication, toujours limpide et précis, sa rigueur scientifique exemplaire et sa vision d'ensemble ont offert un cadre de recherche d'une richesse inestimable. Sa bienveillance et son humanité ont fait de ce parcours un apprentissage stimulant et épanouissant.

Je tiens également à remercier chaleureusement mon co-directeur de thèse, Rémy Abergel, pour son enthousiasme communicatif, sa créativité et la générosité de ses échanges. Sa capacité à allier rigueur scientifique et accompagnement bienveillant a rendu chaque étape de cette thèse plus claire et plus motivante. Sa disponibilité et son engagement ont largement contribué à la profondeur et à la cohérence de ce travail.

Je souhaite adresser une reconnaissance toute particulière à mon tuteur scientifique en entreprise, Olivier Harant. Son ingéniosité et sa polyvalence ont permis de trouver des solutions pertinentes face aux nombreux défis rencontrés dans le projet, qu'il s'agisse de la manipulation des caméras, de l'organisation des expérimentations ou de l'analyse des données. Au-delà de ses compétences, c'est son humanité, sa sympathie, sa disponibilité et son écoute attentive qui m'ont profondément marqué. Son enthousiasme et son investissement constants ont été pour moi une source d'énergie et de confiance, et son accompagnement a joué un rôle décisif dans la réussite de cette thèse.

Je remercie l'ensemble du laboratoire MAP5 de l'Université Paris Cité pour la qualité scientifique et humaine de son environnement de recherche. Je souhaite exprimer ma reconnaissance envers les enseignants et permanents qui ont contribué à ma formation, ainsi qu'à l'équipe administrative pour son efficacité et sa disponibilité. Les échanges scientifiques et les discussions informelles au sein du laboratoire ont été particulièrement stimulants et ont nourri ma réflexion tout au long de ces années, dans une atmosphère conviviale, marquée par l'entraide et le partage.

Je souhaite également remercier mes collègues de l'entreprise *LYNRED*, à Veurey-Voroize, pour leur expertise, leur créativité et leur esprit collaboratif, qui ont rendu ce travail plus épanouissant.

Je tiens à remercier les membres du jury, Julie Delon, Mauro Dalla Mura, Gabriele Facciolo et Jérôme Gilles, pour l'honneur qu'ils me font en acceptant d'évaluer ce travail. Je souhaite également remercier Jocelyn Chanussot et Julie Delon, membres de mon comité de suivi, pour leurs remarques et leur accompagnement au cours de ces années.

Enfin, je remercie chaleureusement mes proches et amis, et en particulier mes parents

et mon frère, pour leur soutien indéfectible et leur présence constante. Leur confiance, leur patience et leurs encouragements m'ont accompagné à chaque étape de ce parcours. Je souhaite également exprimer ma gratitude envers toutes les personnes qui, de près ou de loin, ont contribué à rendre cette thèse possible, par leurs gestes de soutien, leurs paroles d'encouragement ou simplement leur présence amicale. C'est grâce à cette humanité partagée, à cette bienveillance et à cet enthousiasme collectif que ce travail a pu voir le jour.

Table des notations

Abréviations et sigles

FPA : *Focal Plane Array* (matrice plan focal)
TFPA : Température du FPA
ROIC : *ReadOut Integrated Circuit* (circuit intégré de lecture)
NUC : *Non-uniformity Correction* (correction des non-uniformités)
FPN : *Fixed Pattern Noise* (bruit de non-uniformité spatiale)
ACP : Analyse par Composante Principale
SVD : *Singular Value Decomposition* (décomposition en valeurs singulières)
DL : *Digital Level*
i.i.d. : Indépendantes et Identiquement Distribuées
i.e. : *Id Est* (c'est-à-dire)
e.g. : *Exampli Gratia* (par exemple)
cf. : *Confer* (reportez-vous à)
vs. : *Versus* (contre)

Notations mathématiques et opérateurs

N_l : Nombre de lignes d'une image
N_c : Nombre de colonnes d'une image
$\Omega = \{1, \ldots, N_c\} \times \{1, \ldots, N_l\}$: Domaine discret de l'image
(x, y) : Coordonnées d'un pixel ($1 \leq x \leq N_c$, $1 \leq y \leq N_l$)
S : Irradiance de scène
T : Température de scène
θ : Environnement de la caméra
τ : Mesure de la TFPA
$\mathcal{D}$: Image du décalage
$\mathcal{R}$: Image de la responsivité

$\mathcal{G}$: Image du gain
$\mathcal{G}_0$: Image du gain unique
$\mathcal{O}$: Image de l'offset
B : Image du biais
$\mathscr{B}$: Famille de biais
$\mathscr{T}$: Famille des TFPA associée à $\mathscr{B}$
r : Image brute
$r^{(S)}$: Image brute d'une scène uniforme d'irradiance S
$r^{(T)}$: Image brute d'une scène uniforme d'un corps noir de température T
C : Image dont les éléments de chaque colonne sont constants
$v = \mathcal{G}_0 r$: Image pseudo-brute
N_T : Nombre de corps noirs utilisés dans $\mathscr{D}$
N_e : Nombre d'états utilisés dans $\mathscr{D}$
n_s : Nombre de simulations (pour le test d'overfitting)
N_{bits} : Nombre de bits (quantification numérique)
$\mathbf{C}$: Carte des degrés
$\mathbf{M}$: Masque des pixels non défectueux
M : Cardinal du masque
$\mathbf{N}$: Ensemble de situations (différentes scènes et TFPA)
∇_x : Gradient horizontal
∇_{xx} : Dérivée seconde horizontale
∇_y : Gradient vertical
∇ : Gradient spatial
$\mathcal{N}(\mu, \sigma^2)$: Loi normale de moyenne μ et d'écart-type σ
$\rho(M)$: Rayon spectral de la matrice symétrique réelle M
$\mathbb{R}^\Omega$: Ensemble des images discrètes réelles de domaine Ω
$\mathbb{R}_d[X]$: Ensemble des polynômes réels de degré au plus d
$\mathscr{M}_{m,n}(\mathbb{R})$: Ensemble des matrices $m \times n$ à coefficients réels
$\mathscr{M}_n(\mathbb{R})$: Ensemble des matrices carrées $n \times n$
M^* : Matrice adjointe ou transposée
$\mathbf{O}(n)$: Ensemble des matrices orthogonales réelles d'ordre n
$\mathbf{D}(n,p)$: Matrices $n \times p$ avec éléments diagonaux positifs ordonnés décroissants, le reste nul
$[\![n, p]\!]$: Ensemble des entiers dans l'intervalle $[n, p]$

Caméras utilisées

Caméra	Type	Résolution
Smart IR	Non refroidie	320x240
ATTO 640	Non refroidie	640x480
ATI 640a	Non refroidie	640x480
ATI 640b	Non refroidie	640x480
ATI 640c	Non refroidie	640x480

Jeux de données et indices spécifiques

Chaque jeu de données de calibration est noté $\mathscr{D}_k$, où k identifie la caméra utilisée pour l'acquisition de données calibrées :
$\mathscr{D}_1$: données de calibration pour la caméra ATTO 640
$\mathscr{D}_2$: données de calibration pour la caméra ATI 640a

Chaque jeu $\mathscr{D}_k$ est divisé en sous-ensembles correspondant aux différents corps noirs utilisés :
$\mathscr{D}_k^f$: Sous-ensemble issu du corps noir de température la plus froide
$\mathscr{D}_k^c$: Sous-ensemble issu du corps noir de température la plus chaude
$\tilde{\mathscr{D}}_k$: Sous-ensemble restreint aux corps noirs chauds et froids
$\mathscr{D}_k^t$: Sous-ensemble relatif à un ou plusieurs corps noirs de « test », de température intermédiaire.

Résumé

Titre : Modèles pour la correction d'images thermiques issues de caméras non-refroidies sans obturateur mécanique

Résumé : Les caméras thermiques non-refroidies mesurent le rayonnement infrarouge émis par les objets en fonction de leur température et de leur émissivité. L'image brute obtenue est cependant dégradée par un bruit structuré spatialement non uniforme, appelé *fixed pattern noise* (FPN). La correction de ce bruit repose généralement sur un modèle affine dont les composantes dépendent de l'environnement de la caméra, en particulier de la température du *focal plane array* (TFPA).

Deux grandes familles de méthodes permettent d'estimer ces composantes. Les approches « *calibration-based* », considérées comme les plus fiables, s'appuient sur des données acquises à différentes TFPA pour déterminer, pour chaque pixel, les composantes du modèle affine. La méthode *shutterless polynomial*, largement utilisée dans cette catégorie, modélise ces composantes par des polynômes en fonction de la TFPA. Les approches « *scene-based* », quant à elles, exploitent les variations de la scène observée pour estimer le FPN. Elles nécessitent toutefois des variations temporelles suffisantes et deviennent sensibles à leur absence, ce qui peut engendrer des artefacts de type « *ghosting* ».

Dans le cadre des modèles « *calibration-based* », nous proposons une modélisation polynomiale des images brutes en fonction de la TFPA, à partir de laquelle les composantes du modèle affine sont ensuite estimées. Cette approche se distingue de la méthode de référence *shutterless polynomial*, qui procède dans l'ordre inverse en estimant d'abord les composantes à partir des images brutes avant de les modéliser par des polynômes. La modélisation polynomiale des images brutes offre des performances de correction comparables, voire légèrement supérieures. La calibration est réalisée dans un environnement thermique strictement contrôlé afin de limiter l'influence de variations d'autres paramètres environnementaux que la TFPA sur la réponse du capteur. L'analyse des erreurs de fit révèle des anomalies, présentant une tendance commune au sein des pixels, que nous interprétons dans un premier temps comme pouvant résulter d'imprécisions de mesure de la TFPA lors de la calibration. Une extension, appelée « correction delta », permet d'estimer conjointement les polynômes des images brutes et les imprécisions de la TFPA. Bien qu'elle réduise la moyenne des carrés des erreurs de fit, cette extension n'améliore pas la qualité de correction, ce qui invalide l'hypothèse selon laquelle seules des imprécisions de mesure expliqueraient ces anomalies et met en évidence l'influence de flux thermiques parasites, malgré une phase de calibration conçue pour réduire leur impact. Ce constat souligne les limites des modèles paramétriques fondés uniquement sur la TFPA.

L'analyse statistique de l'espace des biais, dans le cadre du modèle affine à gain constant et biais variable, conduit à proposer un modèle de correction dit « mixte », fondé sur une famille de biais issus des données de calibration. Ce modèle reconstruit le biais correcteur comme une combinaison linéaire dans cet espace vectoriel, optimisée selon un critère de qualité d'image appliqué à l'image reconstruite. Il combine ainsi les avantages des approches « *calibration-based* » et « *scene-based* », tout en restant non paramétrique, ce qui lui confère une meilleure adaptabilité à la complexité thermique réelle. Les résultats expérimentaux montrent que ce modèle surpasse la méthode *shutterless polynomial* en robustesse et en qualité de correction, notamment dans des environnements variés.

Après correction par un modèle reposant sur des données de calibration, certains artefacts colonnaires peuvent subsister en raison de dérives temporelles du circuit électronique de lecture de la caméra. Deux méthodes sont proposées. La première, TVx, impose que la médiane du gradient horizontal de chaque colonne de l'image reconstruite soit nulle, réduisant efficacement le bruit colonnaire mais pouvant générer des discontinuités horizontales en présence de structures verticales. La seconde détecte, dans chaque colonne du gradient horizontal de l'image bruitée, des segments dits « homogènes », supposés correspondre à des zones uniformes, et sélectionne les plus représentatifs via un problème variationnel favorisant un gradient horizontal lisse. Cette dernière méthode réduit le bruit colonnaire, laissant seulement de légers artefacts basse fréquence horizontaux.

Mots-clés : Imagerie thermique non refroidie, correction des non-uniformités, modèle affine non paramétrique, méthodes variationnelles, débruitage colonnaire

Abstract

Title : Models for thermal image correction from shutterless uncooled cameras

Abstract : Uncooled thermal infrared cameras measure the radiance of objects as a function of their temperature and emissivity. The raw images are however degraded by spatially structured noise, commonly referred to as fixed pattern noise (FPN). This defect is generally corrected using an affine model whose parameters depend on the camera environment, in particular on the focal plane array temperature (TFPA).

Two main categories of methods exist for estimating these parameters. « Calibration-based » approaches, regarded as the most reliable, rely on acquisitions at different TFPA values and model each pixel's parameters with polynomials of the TFPA. The widely used *shutterless polynomial* method is the standard reference in this category. « Scene-based » approaches, in contrast, exploit variations in the observed scene. They require sufficient temporal motion to function correctly, and in the absence of enough movement, ghosting artifacts may appear.

In this work, we first propose a polynomial modeling of raw images with respect to the TFPA, from which the affine model parameters are subsequently derived. Compared with the *shutterless polynomial* method, which estimates parameters before fitting them as TFPA polynomials, our approach yields comparable or superior correction quality. A further extension, called delta correction, was designed to compensate for potential inaccuracies in TFPA measurements. While it successfully reduces fitting errors, it does not enhance the quality of the corrected images, thereby invalidating the initial hypothesis of TFPA measurement errors. Instead, the findings suggest that unmodeled physical phenomena, such as parasitic thermal fluxes, significantly affect the acquisition of raw images.

Statistical analysis of the bias space in the affine model with constant gain and variable bias shows that calibration-derived bias terms form a vector space. Based on this insight, we propose a « mixed » correction model within the framework of the affine model with constant gain and bias, where the corrective bias is searched for within this vector space. The bias is reconstructed as a linear combination optimized according to an image quality criterion applied to the reconstructed image, ensuring that the corrected image exhibits intrinsic spatial regularity. By doing so, the model combines the advantages of both « calibration-based » and « scene-based » approaches while remaining non-parametric, providing greater adaptability to the actual thermal complexity of the scene. Experimental results demonstrate that this « mixed » model outperforms the « shutterless

polynomial » method in terms of robustness and correction quality, particularly in diverse environmental conditions.

After « calibration-based » correction, column-wise artifacts may persist due to measurement drifts in the camera's readout circuit. We propose two methods to address this issue. The first, TVx, enforces that the median of the horizontal gradient in each column of the reconstructed image is zero, which reduces column-wise noise but may introduce horizontal discontinuities in the presence of vertical structures in the scene. The second method uses a variational approach to detect, in each column of the horizontal gradient of the noisy image, segments called « homogeneous », assumed to correspond to uniform regions, and selects the most representative ones in order to promote a smooth horizontal gradient in the corrected image. This approach improves correction while occasionally leaving slight low-frequency horizontal artifacts.

Keywords : Non-uniformity correction, uncooled thermal imaging, affine non-parametric model, variational problem, destriping

Sommaire

Résumé **vii**

Abstract **ix**

1 Introduction **1**
1 Imagerie thermique, principes et contexte industriel 1
2 État de l'art des modèles de correction existants 5
3 Problématique et objectifs de la thèse 11
4 Résumé des travaux et contributions 12
5 Organisation de la thèse 28

2 Principes physiques, modèles de correction existants **31**
1 Principes physiques et électroniques de l'imagerie thermique 31
2 Modélisation du capteur et état de l'art des modèles de correction 45
3 Conclusion 63

3 Modélisation du capteur, étude des données de calibration **65**
1 Modèles de correction fondé sur une modélisation polynomiale des images brutes 67
2 Analyse statistique des données de calibration 100
3 Conclusion 112

4 Modèles de correction, débruitage colonnaire **113**
1 Correction d'une image à partir d'une famille de biais 113
2 Débruitage colonnaire en imagerie thermique 149

Conclusion **177**

Annexe **181**

Bibliographie **185**

Table des matières **190**

Liste des figures **196**

Chapitre 1

Introduction

L'imagerie thermique s'est largement développée dans des domaines variés tels que les drones, la surveillance, l'imagerie médicale ou l'automobile. On distingue deux technologies principales, les caméras thermiques refroidies et les caméras thermiques non-refroidies. Nous nous intéressons ici aux capteurs non-refroidis, dont la technologie actuelle repose sur des microbolomètres. Ces caméras occupent une place croissante grâce à leur compacité et à leur coût réduit. Elles produisent toutefois des images brutes fortement dégradées par un bruit spatialement structuré, ce qui rend l'interprétation de la scène difficile. Pour pallier ces limitations, un modèle de correction affine est généralement utilisé, et l'estimation précise des paramètres de ce modèle constitue le cœur de la correction. Deux grandes familles d'approches sont alors distinguées. Les méthodes *calibration-based* reposent sur des données acquises en usine dans des conditions contrôlées, elles sont efficaces mais coûteuses en raison de la phase de calibration et sont sensibles au vieillissement du capteur. Les méthodes *scene-based* exploitent les variations de la scène observée, mais nécessitent des changements temporels suffisants pour fonctionner correctement.

Ce chapitre introductif présente d'abord le contexte industriel et scientifique de l'imagerie thermique non-refroidie et les enjeux liés à la correction des images brutes. Il décrit ensuite les principaux modèles de correction, en distinguant les approches *calibration-based* et *scene-based*, avant de préciser la problématique et le positionnement des travaux. Enfin, il résume de manière synthétique les contributions principales de la thèse.

1 Imagerie thermique, principes et contexte industriel

Pour comprendre les défis liés à la correction des images thermiques, il est nécessaire de revenir d'abord aux principes physiques qui régissent la formation de ces images, en particulier aux mécanismes du rayonnement thermique et à leur détection par des capteurs non-refroidis.

1.1 Rayonnement thermique émis par un objet

Le rayonnement thermique est l'énergie électromagnétique émise par un corps en raison de sa température absolue et de son émissivité. La quantité d'énergie émise à chaque

longueur d'onde est caractérisée par la radiance spectrale, définie comme la puissance rayonnée par unité de surface, par unité d'angle solide et par unité de longueur d'onde.

Cette radiance dépend essentiellement de la température T du corps et de son émissivité ε, une grandeur sans dimension comprise entre 0 et 1 qui décrit la capacité du matériau à émettre du rayonnement par rapport à un corps noir idéal. Un corps noir absorbe totalement toute radiation incidente et possède une émissivité égale à 1.

Pour un corps d'émissivité ε à la température T, la radiance spectrale est donnée par la loi de Planck généralisée

$$L_\lambda^{(\varepsilon)}(T) = \varepsilon \times \frac{2hc^2}{\lambda^5 \left(\exp\left(\frac{hc}{\lambda kT}\right) - 1\right)}, \tag{1.1}$$

où λ est la longueur d'onde, h est la constante de Planck, c la vitesse de la lumière, et k la constante de Boltzmann.

La radiance augmente avec la température et reste proportionnelle à l'émissivité. Cette relation constitue le fondement physique permettant, après conversion et traitement du signal, d'estimer la température des objets à partir de la radiance mesurée, à condition que l'émissivité du matériau soit connue. Dans des conditions de la vie courante, pour des températures comprises entre -50°C et 100°C, la majeure partie du rayonnement est émise dans l'infrarouge, en particulier dans la bande 3–15 μm, justifiant l'usage des caméras infrarouges pour la détection thermique.

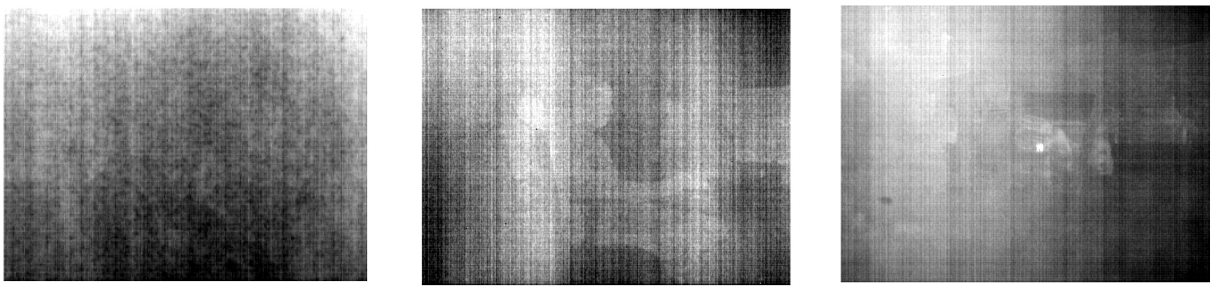
1.2 Image brute issue d'une caméra thermique non-refroidie

Les caméras thermiques se répartissent en deux grandes catégories, refroidies et non-refroidies. Elles se distinguent principalement par leur sensibilité thermique (*noise equivalent temperature difference*, NETD) et leur temps de réponse, qui définissent leur capacité à détecter des différences de température et à réagir aux variations rapides. Le système de refroidissement est composé d'un vase cryogénique dans lequel le capteur est placé pour assurer une isolation thermique efficace. Les caméras refroidies utilisent une détection photonique dans la bande infrarouge moyenne (3 µm–5 µm), offrant une grande sensibilité, environ 20 mK, et un temps de réponse très court, environ 50 µs. Elles restent cependant plus volumineuses, coûteuses et présentent une durée de vie limitée.

Ce manuscrit se concentre sur les caméras non-refroidies, dont le capteur principal est une matrice de microbolomètres. Le rayonnement thermique de la scène traverse l'optique et chauffe chaque microbolomètre, sensible aux longueurs d'onde comprises entre 7 et 15 μm. Cette variation de température modifie la résistance électrique des microbolomètres, laquelle est convertie en signal numérique par l'électronique associée. L'ensemble des signaux forme l'image brute délivrée par la caméra.

Le signal mesuré par chaque pixel ne dépend pas uniquement de la scène observée. Tout flux thermique provenant du boîtier, des éléments électroniques ou de l'environnement ambiant contribue également et dépasse souvent la puissance du rayonnement émis par la scène. On peut comparer cette situation à une caméra visible dont le boîtier serait partiellement translucide et émettrait une lumière interne. Le signal final contient alors à la fois la contribution de scène et des contributions parasites, ce qui dégrade la lisibilité de la scène dans l'image.

Ces contributions parasites, assimilables à un bruit thermique appelé *fixed pattern noise* (FPN), dominent l'image brute et rendent indispensable la correction. Le FPN présente une structure spatiale non homogène avec des motifs en colonnes et des composantes de basse fréquence, comme illustré à la figure 1.1.



(a) *Smart IR320* (b) *ATI-640b* (c) *ATI-640c*

FIGURE 1.1 – **Image brute issue de caméras thermiques non-refroidies.** La figure présente des images brutes obtenues avec deux technologies de caméras différentes. En (a) l'image acquise avec la technologie *Smart IR320* est fortement dégradée par le *fixed pattern noise* (FPN), rendant la scène illisible. En (b) et (c) les images obtenues avec la technologie *ATI-640* montrent un FPN moins prononcé, permettant de mieux distinguer les éléments de la scène. Le FPN est spécifique à chaque caméra d'une même technologie. On en conclut que le FPN combine une intensité propre à la technologie et un motif spécifique à chaque exemplaire.

Les caméras non-refroidies offrent une sensibilité thermique typique de l'ordre de 40 mK et un temps de réponse d'environ 40 ms. Elles présentent un compromis adapté aux applications où la compacité, la faible consommation et le coût de fabrication sont des critères majeurs. Ces capteurs restent toutefois sensibles aux variations thermiques de l'environnement et aux dérives électroniques, ce qui se traduit par des images brutes dégradées nécessitant des méthodes de correction adaptées. La section suivante présente le contexte industriel et les principaux domaines d'application de ces caméras.

1.3 Contexte industriel et domaines d'application

Le développement des caméras thermiques non-refroidies s'oriente vers des capteurs toujours plus compacts et à plus haute résolution, ce qui conduit à la miniaturisation des microbolomètres. En 2025, la taille des pixels atteint environ 8 µm, permettant la réalisation de capteurs offrant des résolutions accrues. La majorité des modèles actuels affichent une définition de 640×480 ou 320×240 pixels, tandis que des résolutions supérieures, déjà disponibles, restent encore peu répandues mais tendent à se diffuser progressivement, en lien avec les progrès de miniaturisation et les besoins des applications industrielles ciblées.

On distingue deux types principaux de caméras thermiques non-refroidies selon la présence ou non d'un obturateur mécanique. Les modèles équipés d'un obturateur (*shutter*) permettent de mettre à jour les paramètres du modèle de correction à partir d'images de référence générées par l'obturateur, mais cette approche interrompt le flux vidéo et augmente la taille de la caméra. Les modèles dits *shutterless* évitent l'interruption du flux et favorisent la miniaturisation, mais nécessitent actuellement une calibration usine plus

longue et coûteuse pour paramétrer précisément les modèles de correction. Le présent travail de thèse s'inscrit dans le contexte des caméras *shutterless*, technologies privilégiées pour la production industrielle chez *Lynred*.

Dans le processus de cette production industrielle, chaque caméra est calibrée en usine dans un environnement thermique contrôlé. La température environnante est progressivement modifiée, par exemple dans une étuve dont la température varie de -40 à 80°C par paliers, avec stabilisation à chaque étape. À chaque palier, la caméra acquiert des scènes uniformes, souvent à l'aide de corps noirs maintenus à des températures précises. Les données recueillies servent à paramétrer les modèles de correction, dont le principal paramètre est la température du capteur situé au plan focal (*temperature focal plane array*, TFPA). Les autres contributions thermiques, incluant les flux parasites provenant du boîtier, des composants électroniques et de l'air ambiant, forment un flux de chaleur complexe autour du capteur. Elles doivent rester stables d'une acquisition à l'autre, que leur répartition spatiale soit homogène ou non, afin de garantir la précision des modèles. La TFPA correspond à une mesure de la température au niveau du capteur et constitue actuellement le seul paramètre utilisé pour les modèles de correction. Cette exigence de stabilité rend souvent la calibration longue et coûteuse, représentant ainsi un enjeu économique majeur.

Les caméras thermiques non-refroidies trouvent des applications dans de nombreux domaines industriels et scientifiques. Elles interviennent notamment dans la surveillance et la maintenance prédictive des installations industrielles, le contrôle qualité, la détection de fuites, ainsi que dans la sécurité, en permettant la détection d'intrusions ou d'anomalies en conditions de faible visibilité. Elles sont également utilisées en médecine, agriculture de précision, conduite assistée, lutte contre les incendies, et jouent un rôle clé dans les systèmes militaires embarqués sur drones pour la surveillance, la reconnaissance et les opérations tactiques (cf. figure 1.2).

Les contraintes industrielles majeures liées à ces caméras concernent :

— le coût de fabrication, influencé par la durée et la complexité de la calibration,

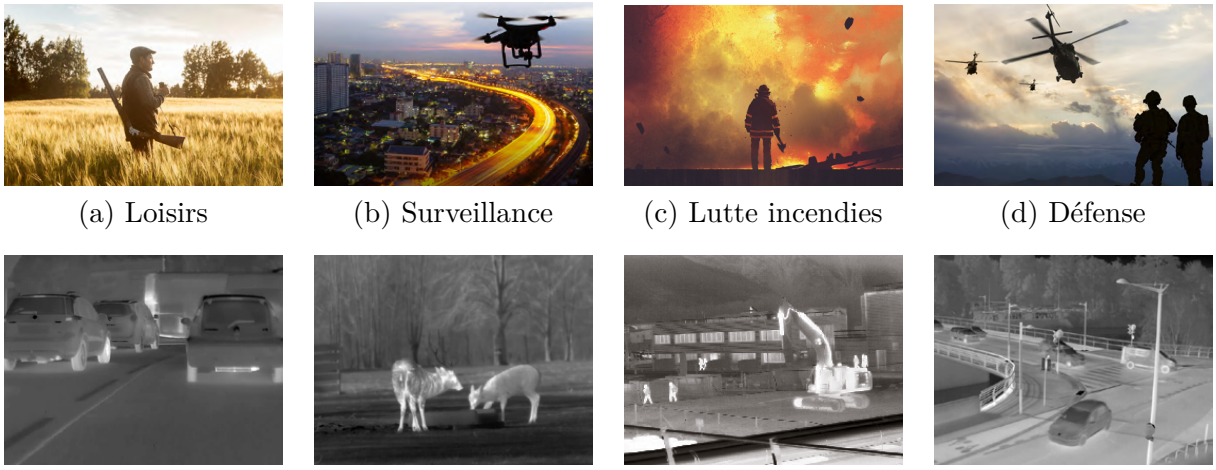
— la compacité et le poids, favorisés par la miniaturisation et l'absence d'obturateur mécanique,

— la consommation énergétique, limitée pour les applications embarquées,

— la qualité d'image, dépendante de la robustesse des modèles de correction face à des situations critiques, telles que des variations rapides de la température environnante, l'apparition d'un nouvel environnement thermique ou la dérive temporelle des mesures électroniques

— les ressources computationnelles, qui doivent rester compatibles avec les capacités limitées des *firmwares* embarqués.

Ces contraintes et besoins industriels justifient l'étude des modèles de correction existants. La section suivante présente l'état de l'art des modèles et méthodes employés pour la correction du FPN dans les images thermiques non-refroidies.



(a) Loisirs (b) Surveillance (c) Lutte incendies (d) Défense

(e) Automobile (f) Zoologie (g) *Monitoring* (g) Analyse de flux

FIGURE 1.2 – **Quelques domaines d'application de l'imagerie thermique non-refroidie.** Source brochure *Lynred*. Les quatre premières images (a, b, c, d) illustrent différents domaines d'application. Les quatre suivantes présentent des images thermiques corrigées dans le cadre d'applications spécifiques : (e) aide à la conduite autonome, (f) caméra thermique intégrée à des lunettes de visée pour observer des animaux, (g) maintenance prédictive de machines industrielles, (h) analyse des flux de circulation automobile.

2 État de l'art des modèles de correction existants

Dans cette section, nous présentons une synthèse des modèles et méthodes employés pour la correction du bruit fixe de motif (FPN) dans les images thermiques issues de capteurs non-refroidis. Nous décrivons d'abord la modélisation affine du capteur, couramment utilisée pour représenter la réponse des pixels, puis les méthodes de correction qui visent à en estimer les paramètres. Ces approches incluent les méthodes « *calibration-based* » et « *scene-based* », ainsi que les méthodes « mixtes » combinant ces deux stratégies. La section se conclut par une analyse des principales limites de ces approches.

2.1 Modélisation affine de la réponse du capteur

On considère un modèle affine paramétrique en θ, où θ regroupe l'ensemble des paramètres influençant la réponse du capteur, tels que la TFPA, les caractéristiques du circuit électronique de lecture, les flux thermiques émanant du boîtier, de l'optique ou du circuit électronique, ainsi que d'autres facteurs affectant la mesure, mais excluant les paramètres liés à la scène observée.

La réponse brute d'un pixel selon la scène s'exprime par le modèle affine [1, 2]

$$r_\theta(x, y) = \mathcal{R}_\theta(x, y)\, S(x, y) + \mathcal{D}_\theta(x, y), \tag{1.2}$$

où $\mathcal{R}_\theta$ représente l'image de responsivité, illustrée à la figure 1.3a, traduisant la sensibilité du pixel (x, y) au rayonnement thermique émis par la scène $S(x, y)$, et $\mathcal{D}_\theta$ correspond à l'image de décalage, présentée à la figure 1.3b, regroupant les contributions additives parasites souvent organisées en motifs colonnaires. L'inversion du modèle affine permet

d'estimer l'image corrigée S par

$$S(x,y) = \frac{r_\theta(x,y) - \mathcal{D}_\theta(x,y)}{\mathcal{R}_\theta(x,y)}. \tag{1.3}$$

Une reformulation couramment utilisée consiste à introduire une image de gain $\mathcal{G}_\theta$ et une image d'offset $\mathcal{O}_\theta$ définies par

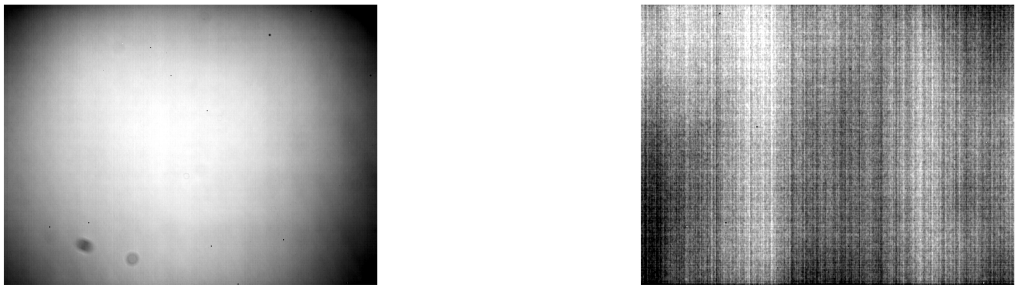
$$\mathcal{G}_\theta(x,y) = \frac{\mu(\mathcal{R}_\theta)}{\mathcal{R}_\theta(x,y)}, \tag{1.4a}$$

$$\mathcal{O}_\theta(x,y) = -\mathcal{G}_\theta(x,y)\,\mathcal{D}_\theta(x,y), \tag{1.4b}$$

où $\mu(\cdot)$ désigne l'opérateur moyenne spatiale. En posant $U = \mu(\mathcal{R}_\theta)\,S$, l'image corrigée normalisée, le modèle s'écrit alors

$$U(x,y) = \mathcal{G}_\theta(x,y)\,r_\theta(x,y) + \mathcal{O}_\theta(x,y). \tag{1.5}$$

La responsivité et le décalage présentent une variabilité importante selon les paramètres environnementaux θ. Bieszczad et al. [3] ont montré que lorsque les paramètres électroniques du capteur restent constants, la TFPA τ constitue le facteur prédominant influençant cette variabilité. Il est alors courant de reformuler le modèle de correction (1.3) en considérant que θ est réduit à la seule variable τ.



(a) Responsivité (b) Décalage

FIGURE 1.3 – **Images de responsivité et de décalage d'une caméra infrarouge non-refroidie.** La figure présente des données de la caméra *ATI-640b*. En (a) l'image de responsivité est spatialement lisse, reflétant principalement l'optique du système. En (b) l'image de décalage montre un fort effet colonnaire et contient des structures parasites assimilables au FPN observé sur la figure 1.1b.

2.2 Méthodes « *calibration-based* »

Les méthodes « *calibration-based* » reposent sur l'acquisition préalable de données calibrées dans des conditions contrôlées, généralement à l'aide de corps noirs à températures fixes. Elles ont pour objectif d'estimer avec précision les paramètres du modèle affine paramétrique en TFPA $\mathcal{R}_\tau$ et $\mathcal{D}_\tau$, ou de manière équivalente $\mathcal{G}_\tau$ et $\mathcal{O}_\tau$. En pratique, cette estimation repose très souvent sur une modélisation polynomiale des paramètres en fonction de la TFPA.

2.2.1 Calibration à deux points

La calibration à deux points consiste à déterminer les paramètres du modèle affine à partir de deux images brutes acquises sur des scènes uniformes [1, 4, 5]. Pour une TFPA donnée τ, on note $r_\tau^{(T_1)}$ et $r_\tau^{(T_2)}$ les images correspondant à deux corps noirs aux températures T_1 et T_2 distinctes. Le modèle affine s'écrit

$$\begin{cases} r_\tau^{(T_1)} = \mathcal{R}_\tau \, T_1 + \mathcal{D}_\tau, \\ r_\tau^{(T_2)} = \mathcal{R}_\tau \, T_2 + \mathcal{D}_\tau \end{cases} \tag{1.6}$$

et ce système admet une solution unique

$$\begin{cases} \mathcal{R}_\tau = \dfrac{r_\tau^{(T_2)} - r_\tau^{(T_1)}}{T_2 - T_1}, \\ \mathcal{D}_\tau = \dfrac{T_2 \, r_\tau^{(T_1)} - T_1 \, r_\tau^{(T_2)}}{T_2 - T_1} \end{cases} \tag{1.7}$$

On en déduit ensuite les paramètres gain et offset à partir de (1.4) (avec $\theta = \tau$).

2.2.2 Modélisation polynomiale en TFPA

Pour estimer les paramètres du modèle affine pour une TFPA quelconque τ, on acquiert N paires d'images correspondant à deux scènes uniformes de radiance distincte, typiquement un corps noir chaud et un corps noir froid, pour des TFPA $(\tau_1, \ldots, \tau_N)$. Pour chaque τ_i, les paramètres du modèle affine sont déterminés par la calibration à deux points, fournissant ainsi une série de $(\mathcal{R}_{\tau_i}, \mathcal{D}_{\tau_i})$ ou de $(\mathcal{G}_{\tau_i}, \mathcal{O}_{\tau_i})$.

Les paramètres associés à une TFPA τ quelconque sont obtenus par régression polynomiale, réalisée pixel par pixel à partir des séries d'images correspondantes [6]. Le degré des polynômes dépend de l'étendue de la plage thermique. Par exemple, Nugent et al. [7] utilisent deux polynômes d'ordre 1 pour couvrir des variations de 10°C. Tempelhahn et al. [8, 9] emploient un polynôme d'ordre 2 pour la responsivité et un polynôme d'ordre 3 pour le décalage sur une plage de 45°C. De son côté, l'entrepriss *Lynred* applique un polynôme d'ordre 2 pour la responsivité et un polynôme d'ordre 5 pour le décalage sur une plage de 115°C. Enfin, d'autres travaux fondés sur le modèle gain–offset montrent que le gain reste presque constant, tandis que l'offset est bien approché par un polynôme d'ordre 2 sur une plage de calibration de 70°C [10].

Cette approche corrige efficacement les non-uniformités spatiales lorsque la TFPA de l'image se situe dans la plage calibrée. En dehors de cette plage, dite extrapolation en TFPA, la correction peut se dégrader [1]. Malgré ces limites, cette méthode, parfois désignée sous le terme *« shutterless polynomial »*, reste une référence grâce à sa fiabilité. Elle permet de corriger les non-uniformités pour tout type de scène sans condition particulière, contrairement aux approches « *scene-based* » qui exploitent directement le contenu de l'image et que nous présentons dans la section suivante.

1. Des difficultés apparaissent également lorsque l'environnement de la caméra pendant l'utilisation diffère fortement de celui rencontré lors de la calibration. Par exemple, une calibration réalisée dans un environnement thermiquement homogène conduit à une correction médiocre si la caméra est utilisée dans un environnement très hétérogène, comme lorsqu'une source chaude se situe à proximité immédiate de la caméra. Autrement dit, une même valeur de TFPA peut traduire des environnements thermiques très différents.

2.3 Méthodes *scene-based*

Les méthodes *scene-based* ajustent les paramètres du modèle affine dans sa version non paramétrique en exploitant l'information contenue dans la scène elle-même. Elles reposent sur l'hypothèse d'une dérive lente des paramètres et d'une variabilité suffisante de la scène. On distingue deux grandes familles, celles fondées sur l'analyse statistique [11–14] et celles qui utilisent le recalage d'images [15,16] pour estimer directement les paramètres du modèle affine

2.3.1 Méthodes fondées sur l'analyse statistique

Les méthodes fondées sur l'analyse statistique exploitent l'information contenue dans la séquence d'images pour estimer les paramètres du modèle affine non paramétrique. L'idée centrale est que, sur une durée suffisamment longue, la scène fournit naturellement des variations et des motifs permettant de caractériser les dérives du capteur sans recourir à des images de calibration externes. Nous présentons ici trois méthodes de référence.

Les méthodes *constant statistics* (CS) supposent que, sur une séquence d'images suffisamment longue, chaque pixel du détecteur est exposé à un flux lumineux comparable à celui des autres. Le signal corrigé doit alors présenter le même comportement statistique, en particulier la même moyenne et le même écart-type temporel [11,13]. Cette hypothèse permet d'estimer les paramètres du modèle affine, supposés constants sur la séquence. Pour suivre leurs variations éventuelles, on peut recourir à des moyennes glissantes ou à des filtres de Kalman [12]. En pratique toutefois, la distribution radiométrique n'est pas identique d'un pixel à l'autre. La présence de sources ponctuelles ou d'objets en mouvement fausse l'estimation et engendre des artefacts de « *ghosting* » sous forme de traînées ou d'ombres. Cette faible robustesse face à des scènes peu variées limite l'efficacité de CS.

La méthode *adjacent differential statistics* (ADS) proposée par Geng et al. [14] repose sur une hypothèse plus souvent respectée que CS : un pixel et son voisin voient en moyenne la même scène. Les gradients spatiaux de l'image corrigée doivent ainsi avoir une moyenne temporelle nulle. Les auteurs simplifient le modèle affine en considérant un gain unitaire concentrant la correction sur l'offset. Cette approche permet d'estimer les gradients spatiaux de l'offset puis de reconstruire l'offset final par intégration. ADS réduit les artefacts de « *ghosting* » observés avec CS et fournit des résultats visuellement plus stables pour des scènes présentant un mouvement modéré, mais des artefacts de basse fréquence peuvent apparaître en raison de l'accumulation d'erreurs lors de l'intégration

Les méthodes *least mean square* (LMS) corrigent une séquence d'images en construisant d'abord une séquence d'images cibles [13]. Chaque image cible est obtenue en prenant l'image brute à laquelle on applique une correction affine initiale avec des paramètres qui ne sont pas forcément mis à jour, produisant ainsi une image encore dégradée par un motif de bruit considéré comme homogène spatialement. Cette image est ensuite lissée pour s'affranchir de ce bruit et obtenir l'image cible. Les paramètres du modèle affine sont estimés pour chaque pixel en minimisant la distance quadratique pixel par pixel entre l'image corrigée et l'image cible à l'aide d'une descente de gradient stochastique, typiquement l'algorithme de Robbins-Monro [17]. Cette approche repose sur l'hypothèse que la scène varie suffisamment et que les variations temporelles du capteur restent faibles. Même lorsque ces hypothèses sont respectées, des artefacts de « ghosting » peuvent subsister, ce qui rend

LMS moins performant qu'ADS pour certaines scènes.

2.3.2 Méthodes fondées sur le recalage d'images

Les méthodes qui reposent sur le recalage d'images s'inspirent des approches LMS en exploitant l'information extraite entre deux images successives pour définir une image cible [15, 16]. L'hypothèse commune est que chaque élément de la scène émet approximativement la même radiation entre deux images consécutives. Un exemple représentatif est donné par Zuo et al. [16], qui modélisent le mouvement de la scène par une simple translation, sans rotation ni changement d'échelle. A l'instant t, le vecteur de déplacement entre les images brutes au temps t et $t+1$ est estimé à l'aide de la transformée inverse de Fourier du spectre de puissance croisée (*cross-power spectrum*). Les paramètres de correction du modèle affine au temps t sont considérés comme connus ou estimés. L'image cible au temps $t+1$ est ensuite construite en appliquant cette correction à la version translatée de l'image brute précédente. Les paramètres du modèle affine au temps $t+1$ sont mis à jour pixel par pixel afin de minimiser la distance quadratique entre l'image corrigée au temps $t+1$ et l'image cible. Cette approche exploite l'information spatiale et temporelle pour réduire le FPN et les non-uniformités. Cette méthode reste très sensible aux erreurs de recalage et, malgré son intérêt conceptuel, elle n'améliore pas significativement la qualité finale des images corrigées.

2.4 Méthodes « mixtes »

Les méthodes « mixtes » visent à corriger les non-uniformités en combinant les informations issues de la calibration et celles extraites de la scène.

Un premier exemple de méthode « mixte » est utilisé dans le cadre industriel [19]. La correction se fait sur une seule image, sans hypothèse sur la variabilité temporelle de la scène. On considère un modèle affine avec un gain constant (identique selon l'environnement) et un décalage variable. Le décalage recherché est exprimé comme une combinaison affine de la composante colonnaire et de la composante « dispersion » (résidu de la composante colonnaire) d'un décalage déterminé à partir des données de calibration. Les coefficients de cette combinaison sont ajustés successivement selon des critères de qualité de l'image reconstruite. Cette approche permet d'adapter la correction à la scène tout en exploitant à la fois l'information de calibration et l'information extraite de la scène. L'utilisation d'une première référence peut suffire, mais l'ajout d'une seconde référence améliore la prise en compte des dérives thermiques.

Par rapport au modèle *shutterless polynomial*, la qualité peut être légèrement inférieure si la TFPA de l'image à corriger est éloignée des TFPA des images de référence. En revanche, les méthodes « mixtes » restent plus robustes dans des environnements thermiques non homogènes ou soumis à de fortes variations locales.

Un autre exemple de méthode « mixte » est proposé par Liu et al. [20]. Cette méthode ne fait également aucune hypothèse sur la variabilité temporelle de la scène et vise à corriger une image brute à partir d'une famille de scènes uniformes en imposant un critère de régularité sur l'image corrigée. La calibration repose sur un ensemble d'images de scènes uniformes acquises pour différentes valeurs de TFPA. Le modèle de correction est affine : l'image corrigée est obtenue en soustrayant un décalage estimé comme combinaison affine

des images de calibration, puis en multipliant par une image de gain supposée constante selon l'environnment.

Pour favoriser une image spatialement lisse, un critère de régularité est appliqué. Il consiste à minimiser la norme euclidienne de la différence entre l'image corrigée et une version localement moyennée sur une fenêtre de pixels voisins, ce qui revient à minimiser la norme du laplacien de l'image corrigée. Les coefficients de la combinaison affine sont déterminés à partir d'une zone uniforme préalablement « marquée », sur laquelle le laplacien de l'image corrigée est supposé minimal.

La résolution numérique utilise uniquement les quatre images de calibration dont la TFPA est la plus proche de celle de l'image à corriger. Cette résolution constitue une approximation qui, comme on le verra dans l'une de nos contributions, ne permet pas de traiter certains problèmes liés à la résolution numérique, en particulier le mauvais conditionnement du système dû à la forte corrélation spatiale entre les images de calibration.

Cette méthode exploite les données de calibration pour corriger l'image en imposant que la zone uniforme préalablement marquée soit reconstruite de manière la plus lisse possible, ce qui en fait un modèle « mixte » particulièrement intéressant. Cependant, elle présente plusieurs limites : la nécessité de disposer d'une zone uniforme préalablement marquée, le recours à un nombre limité d'images de calibration pour la résolution, et l'incapacité à restituer correctement un voile affine correspondant à une variation linéaire d'intensité selon une direction.

2.5 Limites des modèles de correction

Les méthodes *« calibration-based »* présentent une forte dépendance à un paramètre unique, la TFPA. Pour que l'ajustement polynomial soit fiable, la calibration doit être effectuée dans un environnement thermique strictement contrôlé afin de limiter l'influence d'autres composantes de l'environnement θ, comme les flux thermiques parasites liés à une distribution spatiale non homogène. Cette exigence entraîne des coûts importants et réduit la représentativité des conditions réelles. La correction n'est efficace que lorsque la TFPA se situe dans la plage de calibration. En dehors de cette plage ou lorsque certaines composantes de l'environnement diffèrent de celles rencontrées en calibration, la qualité de la correction se dégrade. Un effet indésirable typique est l'apparition d'artefacts verticaux appelés *« stripes »*, résultant de la dérive temporelle des mesures du circuit de lecture électronique. Ces bandes, généralement moins intenses que le FPN initial, dégradent la qualité radiométrique. Dans un traitement complet, un débruitage colonnaire est envisagé pour réduire ces artefacts tout en maintenant la qualité des images. Malgré ces limitations, les méthodes « calibration-based » restent fiables et le modèle *shutterless polynomial* constitue une référence en correction.

Les approches *« scene-based »* reposent sur des hypothèses fortes, telles que l'existence d'un mouvement suffisant de la scène ou la stabilité temporelle de l'environnement. En pratique, ces conditions ne sont pas toujours vérifiées, ce qui peut générer des artefacts comme le « ghosting ».

Enfin, les approches « mixtes », combinant données de calibration et informations extraites de la scène, constituent une solution prometteuse pour corriger les non-uniformités. Très souvent non paramétriques, elles imposent une régularité sur l'image reconstruite, ce qui leur confère une bonne robustesse face aux variations de l'environnement thermique,

contrairement à la méthode *shutterless polynomial* dont l'extrapolation en TFPA peut être inexacte. Les deux méthodes précédemment décrites présentent des limitations pratiques. La première repose sur un nombre limité d'images de calibration correspondant à au plus deux environnements thermiques [19]. La seconde nécessite une zone uniforme préalablement marquée, difficile à mettre en œuvre, et recourt à une résolution numérique approximative [20].

Après avoir identifié ces limites, il devient nécessaire de préciser la problématique de cette thèse ainsi que les enjeux et objectifs qui guideront le développement de nouvelles approches de correction.

3 Problématique et objectifs de la thèse

Les modèles existants ne permettent pas de concilier pleinement robustesse, qualité radiométrique et contraintes industrielles. Dans le contexte des caméras thermiques non-refroidies sans obturateur, les modèles paramétriques fondés sur la TFPA nécessitent une calibration longue et contrôlée pour garantir la fiabilité des paramètres, entraînant un coût et une durée significatifs.

La problématique centrale consiste à développer des modèles mathématiques de correction capables de réduire cette dépendance aux conditions de calibration strictes, tout en assurant une correction de qualité au moins équivalente à celle des modèles existants et robuste face aux variations thermiques et aux situations critiques.

3.1 Enjeux industriels

Les principaux enjeux industriels pour les caméras thermiques non-refroidies sont les suivants :

— réduire la dépendance à des procédures de calibration longues et coûteuses,

— garantir la robustesse des corrections face à des situations critiques, telles que des variations rapides de température ou des environnements thermiques éloignés des références de calibration,

— préserver la qualité radiométrique des images, incluant la correction de défauts structurés comme les « stripes »,

— limiter la complexité calculatoire afin de permettre l'exécution des algorithmes de correction sur le système embarqué sans surcharger le processeur ni la mémoire,

— minimiser le coût, la taille et la consommation afin de faciliter l'embarquabilité dans des systèmes compacts et autonomes.

Ces enjeux mettent l'accent sur la nécessité de développer des méthodes de correction efficaces, robustes aux variations d'environnement thermique, peu gourmandes en ressources et capables de préserver la qualité des images tout en respectant les contraintes industrielles de coût et d'embarquabilité.

3.2 Objectifs de la thèse

L'objectif est de concevoir des modèles mathématiques de correction capables :

— d'exploiter les données issues de la calibration et les informations contenues dans les images de scène,

— de réduire la dépendance aux conditions strictement contrôlées,

— de garantir une correction robuste aux variations thermiques et temporelles,

— de maintenir une qualité radiométrique élevée dans des conditions réalistes d'utilisation.

3.3 Démarche

Pour atteindre ces objectifs, cette thèse mobilise des outils issus de la statistique, du traitement d'images et de l'optimisation variationnelle afin de concevoir des algorithmes de correction fondés sur la qualité des images et l'exploitation des données de calibration. Aucune contrainte n'est imposée a priori sur la complexité des calculs en vue d'un déploiement embarqué, mais cette question sera abordée en conclusion pour évaluer la faisabilité d'un tel déploiement.

Dans un premier temps, les données issues de la calibration ont été exploitées afin de mieux comprendre le comportement du capteur. L'étape suivante a consisté à développer des modèles de correction « mixtes » combinant les informations de calibration et les observations issues des images de scène.

Les développements proposés dans ce cadre s'articulent autour de plusieurs contributions méthodologiques et expérimentales qui sont présentées dans la section suivante.

4 Résumé des travaux et contributions

Les travaux de cette thèse s'organisent autour de quatre contributions principales. Les deux premières fournissent un cadre d'analyse et de modélisation du capteur, tandis que les deux suivantes permettent de limiter la dépendance aux procédures de calibration et d'assurer une correction robuste et de qualité.

— **Modélisation polynomiale des images brutes** : ajustement précis du modèle affine en TFPA à partir de polynômes construits sur des images brutes de scènes uniformes issues des données de calibration.

— **Analyse statistique des données de calibration** : étude du comportement du capteur en fonction de la TFPA et de la scène, analyse de l'espace des corrections dans le cadre du modèle affine avec gain constant et biais variable.

— **Correction fondée sur une famille de biais** : proposition et étude d'un modèle variationnel « mixte » non paramétrique, exploitant simultanément les données de calibration et les informations issues de la scène. Un dépôt de brevet relatif à ce modèle est en cours.

— **Débruitage colonnaire** : post-traitement atténuant les artefacts structurés liés aux dérives temporelles du circuit de lecture tout en préservant la qualité radiométrique des images.

4.1 Modélisation polynomiale des images brutes

Dans cette partie, nous nous intéressons à la modélisation du capteur. Le cadre retenu repose sur un modèle affine paramétrique en fonction de la TFPA, considérée comme l'unique paramètre explicite parmi l'ensemble des composantes de l'environnement. Les autres facteurs, bien que présents, sont négligés car plus difficiles à mesurer ou à contrôler, les paramètres électroniques du circuit de lecture étant fixés. Ce choix correspond à l'hypothèse classiquement adoptée dans les approches « *calibration-based* ». Il suppose de disposer de données de calibration acquises dans un environnement contrôlé, de manière à limiter l'influence des composantes environnementales non prises en compte et à isoler l'effet de la TFPA.

Dans ce modèle, pour chaque pixel $(x, y) \in \Omega$, l'image corrigée $S(x, y)$ s'écrit

$$S(x, y) = \frac{r_\tau(x, y) - \mathcal{D}_\tau(x, y)}{\mathcal{R}_\tau(x, y)}, \tag{1.8}$$

où $r_\tau(x, y)$ désigne l'image brute délivrée par la caméra à la TFPA τ, et où $\mathcal{R}_\tau$ et $\mathcal{D}_\tau$ représentent respectivement les images de responsivité et de décalage.

Les données de calibration sont établies de façon à permettre l'estimation des composantes $\mathcal{R}_\tau$ et $\mathcal{D}_\tau$ pour toute valeur de TFPA τ.

4.1.1 Protocole de calibration et données acquises

La calibration consiste à acquérir des images brutes de scènes uniformes tout en faisant varier l'environnement thermique du capteur. Pour cela, la caméra est placée dans un milieu spatialement homogène, par exemple une étuve, où la température ambiante est réglée de manière contrôlée. Cette homogénéité réduit l'influence des autres paramètres environnementaux et justifie que l'état thermique du capteur soit caractérisé par la seule valeur de la TFPA.

Pour chaque configuration thermique, plusieurs images brutes de corps noirs uniformes sont acquises. Chaque image est notée r_τ^T, où T désigne la température du corps noir et τ la TFPA mesurée simultanément. Le jeu de calibration comprend typiquement des acquisitions sur un corps noir froid, un corps noir chaud, ainsi que sur un ou plusieurs corps noirs de température intermédiaire, ces derniers étant utilisés pour évaluer les performances de correction.

4.1.2 Méthode de la littérature et motivation de notre approche

Avant de présenter notre méthode d'estimation des composantes de responsivité et de décalage pour toute TFPA τ, il convient de rappeler le principe de la méthode *shutterless polynomial* [8], référence dans la littérature (voir section 2.2.2). Pour chaque environnement thermique, deux acquisitions de scènes uniformes, un corps noir chaud et un corps noir froid, sont utilisées. À partir de ces images brutes et de leur TFPA mesurée simultanément, la formule de calibration à deux points (1.7) permet de calculer, pour chaque pixel, la responsivité et le décalage correspondants.

En pratique, les TFPA mesurées pour le corps noir chaud et le corps noir froid peuvent légèrement différer. Pour appliquer la formule de calibration et résoudre le système, ces deux valeurs sont souvent remplacées par leur moyenne, ce qui fournit une estimation

unique des images de responsivité et de décalage pour cette TFPA moyenne. Ces images sont ensuite étendues à toute TFPA par interpolation polynomiale, par exemple selon la méthode *shutterless polynomial*, avec un degré de 2 pour la responsivité et de 3 pour le décalage [8].

Une limitation importante de cette approche réside dans l'approximation introduite par la TFPA moyenne. Selon la procédure de calibration, la différence entre les TFPA du corps noir chaud et du corps noir froid peut être faible ou plus marquée. Lorsque ces valeurs sont proches, l'approximation reste négligeable, mais si elles diffèrent plus sensiblement, elle peut engendrer une erreur dans l'estimation des composantes de responsivité et de décalage. La méthode que nous proposons s'affranchit de cette approximation en modélisant directement les images brutes des corps noirs chauds et froids par des polynômes distincts, puis en déduisant les images de responsivité et de décalage à partir de la formule de calibration à deux points.

4.1.3 Modélisation polynomiale des images brutes

Pour chaque pixel (x, y) et pour l'image brute d'un corps noir de température T (froide ou chaude), plusieurs observations issues des données de calibration sont disponibles, correspondant à différentes valeurs de TFPA. La réponse d'un pixel à température de scène fixée, observée en fonction des valeurs de TFPA mesurées lors de l'acquisition, présente une régularité qui motive l'adoption d'une modélisation polynomiale en TFPA à partir de ces points.

Le polynôme $P_{x,y}$ est déterminé par régression polynomiale. Il est néanmoins nécessaire de convenir au préalable d'un degré. Celui-ci est d'abord estimé pixel par pixel à l'aide d'un test d'*overfitting*, puis un degré global est retenu et appliqué à l'ensemble des pixels.

4.1.3.1 Choix du degré du polynôme à partir d'un test d'*overfitting*

Le choix du degré d'un polynôme dans le cadre d'une régression polynomiale est un problème classique. Les critères AIC [21] et BIC [22] pénalisent la complexité, mais l'AIC peut favoriser des modèles trop complexes et le BIC peut être inadapté si la fonction réelle n'appartient pas à la classe des polynômes [23, 24]. Les approches bayésiennes estiment conjointement coefficients, bruit et degré, mais nécessitent des simulations coûteuses [25].

Nous proposons une méthode alternative fondée sur un test d'*overfitting*. Pour un degré d, un polynôme est ajusté aux données réelles et sa RMSE (*root mean square error*) est calculée. À partir de ce polynôme ajusté, plusieurs jeux de données simulés sont générés en y ajoutant un bruit gaussien d'écart-type égal à cette RMSE. Sur chacun de ces jeux simulés, un polynôme de degré $d + 1$ est ensuite ajusté et sa RMSE est calculée. La distribution obtenue de ces RMSE sert de référence pour caractériser la situation où le degré $d + 1$ ne capture que du bruit.

La RMSE du polynôme de degré $d + 1$ ajusté sur les données réelles est comparée à cette distribution. Si elle est significativement plus faible, le degré $d + 1$ apporte une information réelle ; sinon, il y a risque d'*overfitting*. Le degré final d est choisi comme le plus grand pour lequel ce risque est détecté, maximisant l'information utile sans sur-ajuster le modèle.

Le procédé industriel vise à produire des capteurs dont plus de 99,9 % des pixels présentent une réponse homogène aux flux thermiques. Il est ainsi possible de définir un

degré global D pour la modélisation polynomiale des pixels en fonction de la TFPA, par exemple en prenant le vote majoritaire issu de la distribution des degrés individuels.

4.1.3.2 Masque de pixels défectueux

Les pixels défectueux, cassés ou présentant un comportement atypique, sont exclus de l'analyse. La détection se fait en ajustant un polynôme de degré D à chaque pixel et en comparant la RMSE aux autres pixels. Les pixels dont la RMSE est très éloignée de la majorité sont exclus, constituant ainsi un masque $\mathbf{M} \subset \Omega$ de pixels non-défectueux.

Après exclusion des pixels défectueux, on analyse les erreurs de *fit* sur l'ensemble des pixels restants.

4.1.4 Observation des erreurs de *fit* et correction delta

4.1.4.1 Analyse des erreurs de *fit*

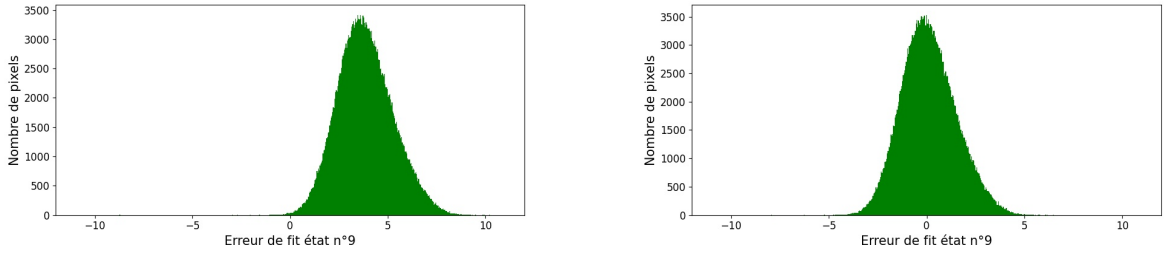
Pour chaque pixel $(x, y) \in \mathbf{M}$ et pour un état thermique de calibration j associé à la TFPA τ_j, l'erreur de *fit* est définie par

$$\varepsilon_{x,y}(j) = r_j^{(T)}(x, y) - P_{x,y}(\tau_j), \tag{1.9}$$

où $r_j^{(T)}$ désigne l'image brute du corps noir de température T et $P_{x,y}$ le polynôme de degré au plus D obtenu par régression. L'ensemble des erreurs de *fit* pour l'état j est alors donné par

$$\mathscr{E}(j) = \{\varepsilon_{x,y}(j),\ (x, y) \in \mathbf{M}\}. \tag{1.10}$$

Si les erreurs n'étaient que du bruit, on s'attendrait à ce que la distribution $\mathscr{E}(j)$ soit centrée autour de zéro. Cependant, comme illustré sur la figure 1.4a, on observe une tendance commune au sein des pixels. La moyenne de $\mathscr{E}(j)$ peut être significativement non nulle, révélant la présence d'un biais. Nous émettons l'hypothèse que ce biais résulte uniquement d'une imprécision de mesure de la TFPA.

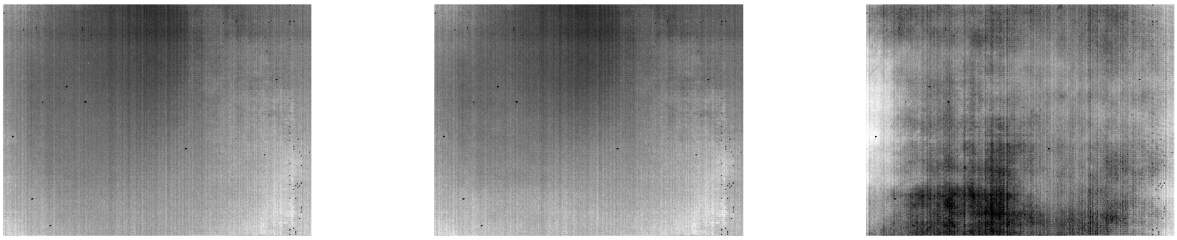


(a) Sans correction delta (b) Avec correction delta

FIGURE 1.4 – **Histogramme des erreurs de *fit* pour un état avant et après correction delta.** Données de calibration issues de la caméra *ATTO 640a*. On considère les erreurs de *fit* pour un certain état utilisé lors de la calibration au sein des pixels non-défectueux. (a) Sans correction delta, la moyenne est significativement non nulle, ce qui indique la présence d'un biais alors qu'une erreur de *fit* devrait correspondre à un bruit centré. (b) Après correction delta, la distribution des erreurs de *fit* est centrée autour de zéro, confirmant l'efficacité de la correction pour éliminer le biais.

Pour corriger ce biais, la correction delta consiste à estimer simultanément tous les polynômes de degré au plus D pour chaque pixel et les TFPA corrigées, delta faisant référence à l'imprécision de la mesure de TFPA. L'estimation se fait en minimisant la somme des carrés des erreurs de *fit*, d'abord sommée sur tous les pixels de l'image, puis sommée sur tous les états thermiques de calibration. Si les TFPA corrigées correspondent effectivement aux TFPA réelles, on peut espérer obtenir des polynômes mieux ajustés aux données.

Après application de cette correction delta, la moyenne des erreurs devient proche de zéro et la RMSE globale diminue d'un facteur trois, comme illustré sur la figure 1.4b.



(a) *Shutterless polynomial* (b) Polynômes images brutes (b) Correction delta

FIGURE 1.5 – **Correction de scènes uniformes en extrapolation de TFPA.** La figure montre la correction d'une scène uniforme selon trois méthodes, sa TFPA étant éloignée de l'intervalle mesuré lors de la calibration. En (a) la méthode *Shutterless polynomial* et en (b) la méthode polynomiale donnent des corrections très similaires, produisant une image assez uniforme avec quelques effets de colonne et de basse fréquence (écarts-types à peu égaux à 5.1). En (c) la correction delta donne une image nettement moins uniforme (écart-type égal à 12.9), avec des artefacts plus prononcés tels que des effets de colonne et de basse fréquence plus visibles, reflétant la précision moindre des paramètres du modèle affine de responsivité et de décalage obtenus avec cette méthode par rapport aux deux autres corrections.

4.1.5 Évaluation des performances sur un sous-ensemble de test

Pour comparer les performances, trois méthodes de correction ont été considérées : la référence *shutterless polynomial*, la modélisation polynomiale des images brutes, et la modélisation polynomiale des images brutes avec correction delta. Pour ces deux dernières méthodes, la responsivité et le décalage ont été déterminés à partir des polynômes ajustés sur les corps noirs froid et chaud. L'évaluation a été effectuée sur des images de corps noir dites de « test », dont la température de scène diffère de celle des corps noirs froid et chaud. Pour chaque image, une TFPA corrigée a été recherchée de manière à minimiser la variance spatiale de l'image corrigée, garantissant ainsi une comparaison équitable entre les méthodes et restant cohérent avec l'hypothèse d'imprécision dans la mesure de la TFPA.

Les résultats mettent en évidence trois constats principaux. Premièrement, la correction delta donne des performances inférieures à celles de la modélisation polynomiale des images brutes. En conséquence, les polynômes ajustés ne sont pas plus précis, ce qui invalide l'hypothèse initiale selon laquelle les erreurs seraient dues uniquement à des imprécisions dans la mesure de la TFPA. Le non-centrage observé des erreurs de *fit* pour

un état donné (figure 1.4a) suggère plutôt la présence de flux parasites supplémentaires, variables d'un état de calibration à l'autre, malgré les mesures mises en place pour limiter cette variabilité. Comme ces flux sont difficiles à mesurer, un modèle paramétrique en TFPA ne peut pas les représenter, ce qui incite à explorer des approches non paramétriques afin d'améliorer la correction.

Deuxièmement, la modélisation polynomiale des images brutes n'est jamais moins performante que la méthode de référence *shutterless polynomial*, offrant parfois de très légères améliorations tout en restant globalement équivalente en termes de qualité de correction.

Troisièmement, l'application des trois méthodes de correction à des images dont la TFPA se situe en dehors de la plage de calibration conduit à des performances nettement dégradées (voir figure 1.5). Cela illustre une limite inhérente à la modélisation polynomiale, car on cherche à approcher des fonctions inconnues par des polynômes ajustés sur un intervalle restreint et l'extrapolation au-delà de cet intervalle peut entraîner des écarts significatifs. Cette situation met en évidence le manque de robustesse des modèles face à des conditions thermiques non rencontrées lors de la calibration.

Pour approfondir la compréhension de la réponse du capteur, nous nous intéressons à présent à l'analyse statistique des données de calibration. Cette analyse se fait d'abord sans a priori de modèle affine paramétrique en TFPA afin d'observer la réponse des pixels selon la température de scène et la TFPA.

4.2 Analyse statistique des données de calibration

Les ensembles de données de calibration permettent de déterminer les composantes d'un modèle affine paramétré par la TFPA. Il est naturel de s'interroger sur l'exploitation complète de l'information qu'ils contiennent, d'où une analyse statistique des données de calibration. Cette section se divise en deux parties. La première procède à une analyse en composantes principales pour étudier la réponse du capteur en fonction des deux facteurs d'intérêt, la température de scène T et la TFPA τ, sans faire référence à la modélisation affine paramétrée par la TFPA (voir (1.2) avec $\theta = \tau$). La seconde propose une analyse de l'espace des corrections dans le cadre d'un modèle affine simplifié.

4.2.1 Étude du comportement des pixels à partir d'une ACP

Lors de la calibration, chaque pixel est exposé à plusieurs corps noirs de températures T variées et à différents états thermiques de l'étuve correspondant à des TFPA τ. La réponse de chaque pixel à l'ensemble de ces situations peut être représentée par un vecteur Y_p de taille N, avec N égal au nombre de combinaisons de corps noirs et d'états thermiques. L'ensemble des vecteurs $(Y_1, \dots, Y_M)$, avec M le nombre de pixels non-défectueux, constitue le nuage des réponses des pixels et reflète leur comportement face à toutes les situations de calibration. Pour analyser la structure de ce nuage, nous appliquons une analyse en composantes principales (ACP) à la matrice $Y \in \mathbb{R}^{N \times M}$, où chaque colonne correspond à Y_p, la réponse d'un pixel à l'ensemble des situations de calibration. L'ACP permet de décomposer ce nuage en axes principaux, chacun expliquant une fraction de la variabilité totale, mesurée par l'inertie associée. Les premiers axes principaux capturent

ainsi la majeure partie de l'information contenue dans les données, révélant les tendances dominantes de la réponse des pixels.

La figure 1.6 montre que plus de 99.8% de la variabilité est expliquée par les trois premiers axes principaux. Les surfaces associées au centre du nuage et aux axes principaux sont suffisamment régulières pour être approchées par des polynômes de deux variables (T, τ).

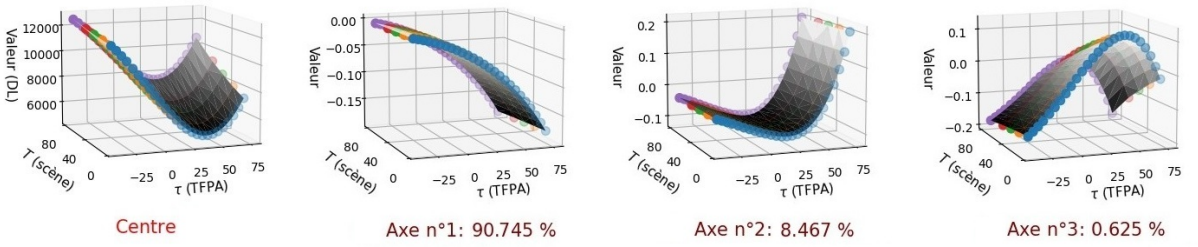


FIGURE 1.6 – **Comportement général du capteur en fonction de la température de scène et de la TFPA.** La figure montre le centre du nuage et les trois premiers axes principaux issus de l'ACP appliquée à 130 situations. Ces éléments peuvent être vus comme des fonctions de la température de scène T et de la TFPA τ, cumulant plus de 99,8% de l'inertie explicative. Les surfaces sont particulièrement lisses, évoquant des lois physiques, ce qui suggère qu'elles peuvent être modélisées par des polynômes en T et τ.

Cette observation conduit à introduire un « modèle direct ». La réponse d'un pixel p est exprimée comme une combinaison linéaire de polynômes à deux variables

$$r_p(T, \tau) = F_0(T, \tau) + \sum_{k=1}^{n} W_{p,k} F_k(T, \tau), \tag{1.11}$$

où F_0 et F_k $(1 \leq k \leq n)$ sont des polynômes en (T, τ) représentant respectivement le centre du nuage et les axes principaux, et les coefficients $W_{p,k}$ proviennent de l'ACP. Le nombre n correspond au nombre d'axes principaux retenus et est choisi en fonction de l'inertie cumulée afin de capturer l'essentiel de la variance. Ce modèle direct décrit ainsi la réponse des pixels en fonction de (T, τ). Si les polynômes considérés sont d'ordre 1 en température de scène, on retombe alors sur le modèle *shutterless polynomial* [8]. Il devient alors possible d'inverser le modèle, c'est-à-dire d'exprimer T en fonction de la réponse du capteur et de la TFPA. Plus l'ordre en T est élevé, plus l'inversion devient complexe.

Après avoir exploré la réponse des pixels sans a priori de modèle, nous revenons maintenant au cadre plus restreint du modèle affine simplifié, dans lequel le gain reste constant pour chaque pixel tandis que le biais peut varier. Cette analyse de l'espace des biais prépare la proposition d'un modèle de correction, présentée dans la section suivante.

4.2.2 Analyse de l'espace des biais du modèle affine

Dans un premier temps, il convient de définir le modèle affine à gain constant et biais.

4.2.2.1 Modèle affine avec gain constant et biais variable

On part du modèle affine gain/offset (1.5) paramétré par $\theta = \tau$, et l'on considère le cas particulier d'un gain constant en TFPA $\mathcal{G}_\tau = \mathcal{G}_0$, hypothèse fréquemment adoptée

dans la littérature [10, 19, 20, 26, 27]. Le biais est défini à partir de l'offset $\mathcal{O}_\tau$ par

$$B_\tau = -\mathcal{O}_\tau + \mu(\mathcal{O}_\tau), \tag{1.12}$$

où $\mu(\cdot)$ désigne la moyenne spatiale sur les pixels du masque. La correction d'une image brute r_τ s'écrit alors

$$U = \mathcal{G}_0\, r_\tau - B_\tau, \tag{1.13}$$

où U représente l'image corrigée.

4.2.2.2 Construction de la famille de biais

Un biais B_τ peut être déterminé à partir de l'image brute $r_\tau^{(T)}$ d'un corps noir de température T, acquise à la TFPA τ par la formule

$$B_\tau = \mathcal{G}_0\, r_\tau^{(T)} - \mu(\mathcal{G}_0\, r_\tau^{(T)}). \tag{1.14}$$

En appliquant cette formule à un ensemble de calibration de cardinal N, on obtient une famille de biais

$$\mathscr{B} = (B_1, \ldots, B_N), \quad B_i \in \mathbb{R}^M.$$

La figure 1.7 montre que ces biais sont fortement corrélés. L'objectif est de caractériser l'espace engendré par ces biais.

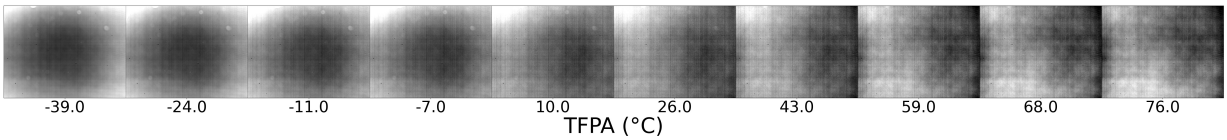


FIGURE 1.7 – **Famille des biais.** Les biais ont été construits à partir des données de calibration. Deux biais issus de TFPA proches présentent des ressemblances marquées, ce qui suggère une forte corrélation au sein de la famille. Les TFPA basses produisent des biais globalement lisses, tandis que les TFPA élevées révèlent des motifs davantage haute fréquence, structurés en lignes et en colonnes.

4.2.2.3 Analyse de l'espace des biais

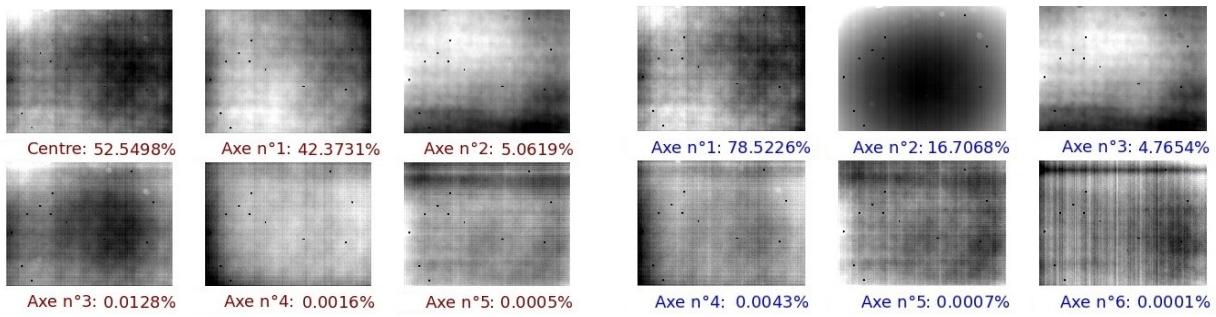
On regroupe les biais dans la matrice

$$Z = (B_1, \ldots, B_N) \in \mathbb{R}^{M \times N}.$$

Nous avons considéré deux approches pour caractériser l'espace engendré par la famille des biais :

— **ACP centrée : sous-espace affine.** Le centre du nuage est considéré comme le biais moyen et les directions principales autour de ce centre définissent un sous-espace affine.

— **ACP non centrée : sous-espace vectoriel.** Les biais sont traités comme des vecteurs et la décomposition fournit directement un espace vectoriel engendré par les biais.



(a) Approche ACP (b) Approche ACP non centrée

FIGURE 1.8 – **Axes principaux obtenus par ACP centrée et ACP non centrée.** (a) Centre et cinq premiers axes principaux obtenus par ACP centrée, avec leur part de variance expliquée. (b) Les six axes principaux obtenus par ACP non centrée, avec leur part de variance expliquée. À nombre de composantes retenues égal, l'ACP non centrée restitue mieux la variabilité de la famille de biais que l'ACP centrée, cette dernière comprenant en outre le centre du nuage comme composante supplémentaire.

Dans les deux cas, on peut associer à chaque composante, qu'il s'agisse du centre pour l'approche ACP centrée ou des directions principales, une part de variance expliquée, ce qui permet de quantifier l'information contenue en fonction du nombre de composantes considérées. La figure 1.8 montre que l'ACP non centrée rend mieux compte de la variabilité de la famille de biais que l'ACP centrée. On considère alors qu'une famille de biais peut être approchée par un espace vectoriel. Cela nécessite que la calibration couvre une plage de TFPA suffisamment large et comporte un nombre suffisant d'acquisitions à des TFPA distinctes.

Les deux premières contributions ont porté sur la modélisation du capteur à partir des données de calibration. La première propose une légère amélioration du modèle de référence *shutterless polynomial* et met en évidence les limites d'une modélisation polynomiale paramétrique en TFPA, notamment sa faible robustesse lorsque l'environnement d'acquisition diffère de celui utilisé pour la calibration. La seconde contribution analyse la structure de l'espace des biais correctifs dans le cadre d'un modèle affine, montrant que ces derniers peuvent être approchés par un espace vectoriel. Ces travaux fournissent un cadre conceptuel pertinent mais n'aboutissent pas à un modèle de correction opérationnel. La prochaine contribution s'inscrit dans ce prolongement en proposant un modèle de correction innovant, non paramétrique, qui combine les données de calibration et une information a priori sur la scène, afin de produire une image corrigée présentant une régularité spatiale cohérente.

4.3 Correction d'une image à partir d'une famille de biais

4.3.1 Modèle variationnel générique

On suppose que les pixels défectueux ont été préalablement corrigés et que les images appartiennent à $\mathbb{R}^{\Omega}$. Le modèle de correction retenu est le modèle affine à gain constant

et biais (1.13), que l'on exprime sous forme non paramétrique

$$U = v - B, \tag{1.15}$$

où $v = \mathcal{G}_0\, r$ désigne l'image pseudo-brute associée à l'image observée r, $\mathcal{G}_0$ l'image du gain unique, et B le biais à estimer, soumis à la contrainte de centrage $\mu(B) = 0$.

Le biais B est recherché dans l'espace vectoriel engendré par une famille de biais $\mathscr{B} = (B_1, \ldots, B_N)$, construite par exemple à partir de scènes uniformes selon (1.14). L'estimation de B se fait en minimisant une métrique de qualité image F, adaptée aux objectifs de régularité souhaités, de sorte que

$$\widehat{B} \in \underset{B \in \mathrm{Vect}(\mathscr{B})}{\operatorname{argmin}} \; F(v - B). \tag{1.16}$$

Trois déclinaisons particulières de ce modèle générique seront considérées. La première, ℓ^2, correspond à la métrique de l'écart-type spatial. La seconde, $\mathscr{H}^{\mathbf{1}}$, utilise la norme euclidienne du gradient spatial. La troisième, TV (*total variation*), repose sur la norme ℓ^1 du gradient spatial. Chacune de ces déclinaisons sera présentée du point de vue du modèle et de sa résolution numérique, cette dernière nécessitant un soin particulier en raison de la forte corrélation existant au sein de la famille des biais.

4.3.2 Modèle ℓ^2

Le modèle ℓ^2 consiste à déterminer un biais B dans l'espace vectoriel engendré par la famille $\mathscr{B}$, de sorte que l'écart-type spatial de l'image corrigée soit minimal. Puisque les biais sont centrés, cela revient au problème

$$\widehat{B} \in \underset{B \in \mathrm{Vect}(\mathscr{B})}{\operatorname{argmin}} \; \|v - B\|_2 \,. \tag{1.17}$$

Pour expliciter la solution, on peut reformuler le problème en cherchant un vecteur de coefficients $\widehat{\alpha}$ dans la base $(B_1, \ldots, B_N)$, (1.17) est équivalent à

$$\begin{cases} \widehat{B} = \mathbf{B}\,\widehat{\alpha}\,, & \text{(1.18a)} \\ \widehat{\alpha} \in \underset{\alpha \in \mathbb{R}^N}{\operatorname{argmin}} \|v - \mathbf{B}\,\alpha\|_2^2\,, & \text{(1.18b)} \end{cases}$$

où $\mathbf{B}$ est la matrice dont les colonnes sont les biais $B_1, \ldots, B_N$. La solution analytique de (1.18b) est

$$\widehat{\alpha} = (\mathbf{B}^*\mathbf{B})^{-1}\mathbf{B}^* v\,. \tag{1.19}$$

En pratique, l'inversion [2] de $\mathbf{B}^*\mathbf{B}$ peut se révéler délicate en raison de la forte corrélation des biais. Pour contourner cet écueil, on construit une base orthonormée $\mathbf{O} = \mathbf{B}\,\mathbf{P}$ à l'aide d'une matrice de passage $\mathbf{P}$. On cherche alors un vecteur $\widehat{\beta}$ solution de

$$\widehat{\beta} \in \underset{\beta \in \mathbb{R}^N}{\operatorname{argmin}} \|v - \mathbf{O}\,\beta\|_2^2\,, \tag{1.20}$$

et on obtient

$$\widehat{\beta} = \mathbf{O}^* v.$$

2. Plus généralement, il s'agit de résoudre $(\mathbf{B}^*\mathbf{B})\widehat{\alpha} = \mathbf{B}^* v$.

La solution de (1.18b) s'écrit alors

$$\widehat{\alpha} = \mathbf{P}\,\widehat{\beta}. \tag{1.21}$$

Cette formulation dans la base orthonormée permet d'éviter les problèmes numériques liés au mauvais conditionnement de $\mathbf{B}^*\mathbf{B}$.

4.3.3 Modèle $\mathscr{H}^{\mathbf{1}}$

Le modèle $\mathscr{H}^{\mathbf{1}}$ consiste à déterminer un biais $B \in \mathrm{Vect}(\mathscr{B})$ minimisant la norme euclidienne du gradient de l'image corrigée

$$\widehat{B} \in \operatorname*{argmin}_{B \in \mathrm{Vect}(\mathscr{B})} \|\nabla(v - B)\|_{2,2}\,. \tag{1.22}$$

La recherche des coefficients dans la base originale $\mathscr{B}$ implique à nouveau l'inversion d'une matrice souvent très mal conditionnée en raison de la forte corrélation des gradients des biais. On contourne cet écueil une nouvelle fois en orthonormalisant la famille des gradients via une matrice de passage $\mathbf{Q}$. Dans cette base orthonormée, on détermine la combinaison optimale des coefficients, puis la combinaison correspondante dans la base originale s'obtient en appliquant la matrice de passage $\mathbf{Q}$.

4.3.4 Modèle TV

Le modèle TV consiste à déterminer un biais $B \in \mathrm{Vect}(\mathscr{B})$ minimisant la norme $\ell^{1,2}$ du gradient spatial de l'image corrigée :

$$\widehat{B} \in \operatorname*{argmin}_{B \in \mathrm{Vect}(\mathscr{B})} \|\nabla(v - B)\|_{1,2}\,. \tag{1.23}$$

Une approche consiste à reformuler le problème en termes de coefficients dans la base originale des biais et à appliquer le schéma primal-dual de Chambolle–Pock [28]. La convergence de cette approche est très lente, voire instable, l'énergie à minimiser oscillant au cours des itérations. Chambolle et Pock [29] ont proposé une variante du schéma, dans laquelle les pas scalaires de descente sont remplacés par des matrices symétriques définies positives, adaptées aux directions primales et duales. Cependant, même cette variante ne permet pas d'atteindre une précision double en moins d'un million d'itérations.

Comme pour le problème $\mathscr{H}^{\mathbf{1}}$, on contourne les difficultés numériques liées à la forte corrélation de la famille des gradients des biais en orthonormalisant cette famille. On recherche alors le vecteur de coefficients dans cette base orthonormée, et la combinaison optimale dans la base originale s'obtient via la matrice de passage. Cette procédure assure une convergence très rapide.

On peut également résoudre le problème de manière intrinsèque en cherchant directement l'image du biais $\widehat{B}$ dans le schéma itératif primal-dual de Chambolle–Pock [28]. À chaque itération, la mise à jour de la variable primal, c'est-à-dire l'image du biais, est projetée orthogonalement dans l'espace $\mathrm{Vect}(\mathscr{B})$. Cette approche permet d'atteindre une précision double en un nombre raisonnable d'itérations.

4.3.5 Étude comparative

Nous comparons nos modèles ℓ^2, $\mathscr{H}^1$ et TV au modèle de référence *shutterless polynomial* ainsi qu'à la méthode de Liu et al. [20].

La méthode de Liu et al. repose sur un principe très proche de nos modèles. Le biais y est exprimé comme une combinaison linéaire des biais de calibration et de l'image du gain, et les coefficients sont déterminés en minimisant la norme euclidienne de la différence entre l'image corrigée et sa version lissée. Autrement dit, pour une certaine fenêtre de lissage, la métrique utilisée correspond à la norme euclidienne du Laplacien de l'image désirée. Cette approche reste difficilement applicable en pratique car elle repose sur une zone uniforme préalablement marquée. En outre, elle ne restitue pas correctement un voile affine et propose une résolution approximative utilisant un nombre limité d'images de calibration ne tenant pas compte de la possible très forte corrélation d'une famille de biais. Nos modèles reprennent la même idée générale mais déterminent les coefficients sans recourir à une zone préalablement marquée, utilisent une métrique du gradient spatial ($\mathscr{H}^1$ ou TV) qui restitue correctement les voiles affines et la résolution numérique tient compte de la corrélation de la famille des biais sans être limitée à un petit nombre d'images de calibration.

La comparaison avec le modèle *shutterless polynomial* permet d'évaluer l'efficacité réelle de la correction sur les images, ce modèle constituant la référence opérationnelle utilisée par l'entreprise *Lynred*. Pour les modèles $\mathscr{H}^1$ et TV, la correction est comparable à celle du modèle *shutterless polynomial* sur des scènes réelles, mais la dépasse en extrapolation, c'est-à-dire lorsque l'environnement thermique de la caméra diffère de ceux dans lesquels les biais de la famille ont été établis (cf. figure 1.9). Le modèle ℓ^2 reste adapté aux scènes uniformes mais produit de nombreux artefacts sur des scènes réelles. Habituellement, la régularisation $\mathscr{H}^1$ rend l'image plus lisse avec des contours légèrement atténués tandis que la régularisation TV préserve mieux les contours et peut générer des zones constantes par morceaux. Dans notre approche, les corrections obtenues avec $\mathscr{H}^1$ et TV sont pratiquement équivalentes car la recherche d'une image dans un espace restreint limite le nombre de degrés de liberté et empêche l'apparition de ces différences attendues.



(a) Modèle ℓ^2 (b) Modèle $\mathscr{H}^1$ (c) Modèle TV (d) *Shutterless polynomial*

FIGURE 1.9 – **Correction d'une image à partir d'une famille de biais et comparaison avec la méthode de référence *Shutterless polynomial*** La TFPA de l'image à corriger se situe en dehors de la plage des TFPA de calibration. Le modèle ℓ^2 (a) génère des artefacts de lignes et de colonnes et réduit les contrastes sur certaines zones sombres, notamment l'arrière-plan en haut à droite. Les modèles $\mathscr{H}^1$ (b) et TV (c) restituent globalement les détails fins, la texture et les contrastes, tout en introduisant un léger voile basse fréquence qui éclaircit le centre-gauche de l'image. La méthode *Shutterless polynomial* (d) produit des artefacts de colonnes et des variations basse fréquence marquées, dégradant fortement la lisibilité de la scène.

Le modèle que nous avons proposé est « mixte » et non paramétrique. Il améliore sensiblement la correction par rapport aux méthodes existantes, notamment en termes de robustesse face aux variations d'environnement thermique, et fait l'objet d'un dépôt de brevet en cours. Malgré ces performances, des artefacts colonnaires peuvent subsister, typiques des méthodes fondées sur des données de calibration et susceptibles de compromettre la qualité de l'image corrigée. Il devient alors nécessaire de compléter le traitement par un débruitage colonnaire en post-traitement.

4.4 Débruitage colonnaire

Le bruit colonnaire se manifeste par des bandes verticales affectant tous les pixels d'une même colonne. Son amplitude, bien que généralement inférieure à celle du *fixed pattern noise* (FPN) de l'image brute, peut atteindre plusieurs degrés Celsius et dégrader la qualité radiométrique de la correction. Il résulte essentiellement d'une dérive temporelle des caractéristiques électroniques du circuit de lecture de la caméra entre la calibration et l'acquisition.

4.4.1 État de l'art des modèles de débruitage colonnaire

Avant de présenter nos contributions, il est utile de rappeler les approches existantes pour corriger le bruit colonnaire. Trois grandes familles sont généralement distinguées, méthodes statistiques, filtrages spatial ou fréquentiel et les approches variationnelles.

4.4.1.1 Approches statistiques

Les méthodes statistiques visent à harmoniser les colonnes par « *moment matching* » ou par ajustement des histogrammes colonnaires. Les méthodes de « *moment matching* » corrigent le bruit colonnaire (ou *stripes*) en ajustant la moyenne et l'écart-type de chaque colonne, soit globalement, soit en effectuant un lissage entre colonnes voisines [30,31]. Ces méthodes peuvent toutefois générer des artefacts dans les zones à fort contraste.

Les méthodes par harmonisation des histogrammes colonnaires ajustent la répartition des intensités de chaque colonne pour qu'elle soit proche de celle des colonnes voisines. Pour chaque colonne x, on calcule la fonction de répartition des intensités, puis on obtient une fonction de référence en moyennant les fonctions de répartition des colonnes voisines dans une fenêtre de plusieurs colonnes. La transformation de la distribution initiale vers cette référence se fait par un transport monotone des intensités, assurant un ajustement global de la colonne. L'approche « MIRE » [32,33] sera utilisée comme méthode de comparaison pour évaluer les performances des méthodes que nous proposons.

4.4.1.2 Méthodes de filtrage spatial ou fréquentiel

Le filtrage pour réduire le bruit colonnaire peut se faire dans le domaine spatial ou fréquentiel. Dans le domaine spatial, des filtres 1-D guidés par l'image lissent les variations horizontales des *stripes* tout en préservant les détails verticaux [34]. Dans le domaine fréquentiel, les variations verticales indésirables apparaissent comme des composantes haute fréquence que l'on peut atténuer par transformée en ondelettes discrète (*discrete wavelet transform*, DWT) ou filtres fréquentiels [35]. Cette approche réduit efficacement

le bruit colonnaire, mais certaines structures verticales réelles peuvent également être atténuées, entraînant une perte partielle de détails.

4.4.1.3 Méthodes variationnelles

Les méthodes variationnelles considèrent le débruitage du bruit colonnaire comme un problème d'optimisation régularisé. L'image observée est modélisée comme la somme de l'image réelle et du bruit de *stripes*. L'objectif est de réduire les oscillations verticales tout en préservant les détails de la scène [36,37]. Ces approches appliquent des régularisations spécifiques au bruit colonnaire afin de favoriser des *stripes* quasi-constantes le long des colonnes tout en autorisant certaines variations locales.

Les méthodes de la littérature supposent souvent que les *stripes* ont une amplitude inférieure aux variations naturelles de la scène, ce qui limite leur efficacité dans les zones à faible contraste. Les approches que nous allons présenter ne reposent pas sur cette hypothèse.

4.4.2 Modèle de perturbation

Nous ferons l'hypothèse que la perturbation colonnaire résulte d'un offset commun appliqué à tous les pixels d'une colonne. Pour un pixel $(x, y) \in \Omega$, le modèle de perturbation s'écrit

$$U_0(x, y) = V(x, y) + C(x), \tag{1.24}$$

où U_0 désigne l'image observée, V l'image réelle et C le vecteur des offsets colonnaires. La correction s'effectue simplement par

$$V(x, y) = U_0(x, y) - C(x). \tag{1.25}$$

Le débruitage d'une image affectée par un bruit colonnaire est un problème mal posé car plusieurs images originales différentes peuvent produire la même image observée. Lorsque des structures verticales réelles sont présentes, il peut exister une ambiguïté visuelle pour notre œil, comme illustré figure 1.10.

4.4.3 Modèle TVx

Le modèle TVx est variationnel et repose sur l'hypothèse que les pixels voisins d'une image naturelle ont des intensités proches, sauf aux contours. Selon le modèle de perturbation (1.24), le bruit colonnaire n'affecte que le gradient horizontal. L'objectif est d'estimer un vecteur d'offsets $\widehat{C}$ en minimisant la norme ℓ_1 du gradient horizontal de l'image corrigée :

$$\widehat{C} \in \underset{C \in \mathbb{R}^{N_c},\, C(1)=0}{\operatorname{argmin}} \left\| \nabla_x \left(U_0 - C \right) \right\|_1, \tag{1.26}$$

où N_c est le nombre de colonnes. Le vecteur C est assimilé à une image dont les valeurs sont constantes dans chaque colonne. La première colonne est fixée à zéro pour que le problème soit bien posé.

La résolution entraîne la propriété suivante : la médiane du gradient horizontal de l'image corrigée est nulle pour chaque colonne x. La figure 1.11c montre que le modèle TVx supprime efficacement les *stripes*. Cependant, des discontinuités horizontales apparaissent dans certaines colonnes. Dans une colonne contenant une ou plusieurs structures verticales, le gradient horizontal de l'image réelle n'a pas nécessairement une médiane nulle.

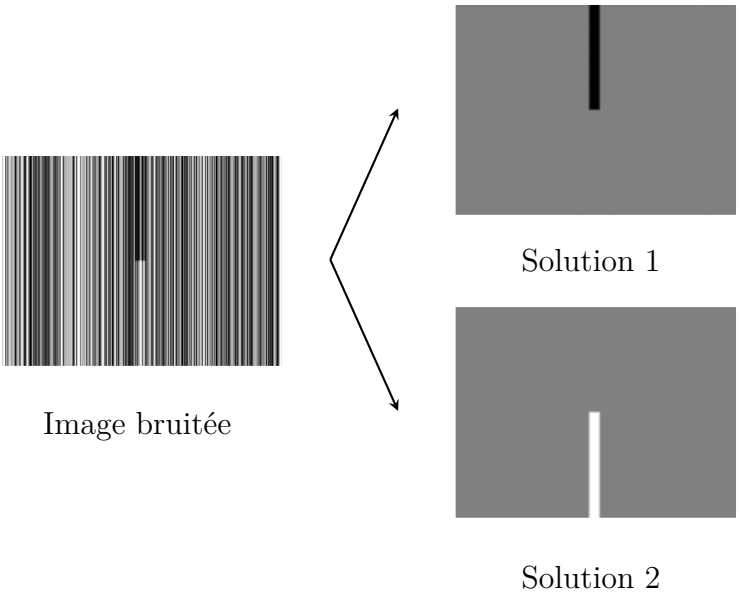


FIGURE 1.10 – **Caractère mal posé du débruitage colonnaire.** L'image de gauche est bruitée. À droite, deux corrections différentes sont proposées. À la perception humaine, les deux apparaissent plausibles, et d'autres solutions pourraient également exister, ce qui illustre le caractère mal posé du problème.

4.4.4 De l'analyse du gradient horizontal à la formulation d'un modèle variationnel

Pour chaque colonne x, on analyse le gradient horizontal de l'image bruitée $\nabla_x U_0(x, \cdot)$ et on identifie des segments où le signal oscille faiblement. L'idée est d'espérer trouver parmi ces segments une zone uniforme dans l'image réelle. Dans ce cas, comme le gradient horizontal de l'image corrigée devrait y être proche de zéro, la valeur autour de laquelle le signal oscille fournit une estimation du gradient de la perturbation colonnaire $\nabla_x \widehat{C}(x)$. L'intégration le long des colonnes permet de reconstruire $\widehat{C}$ en fixant $\widehat{C}(1) = 0$

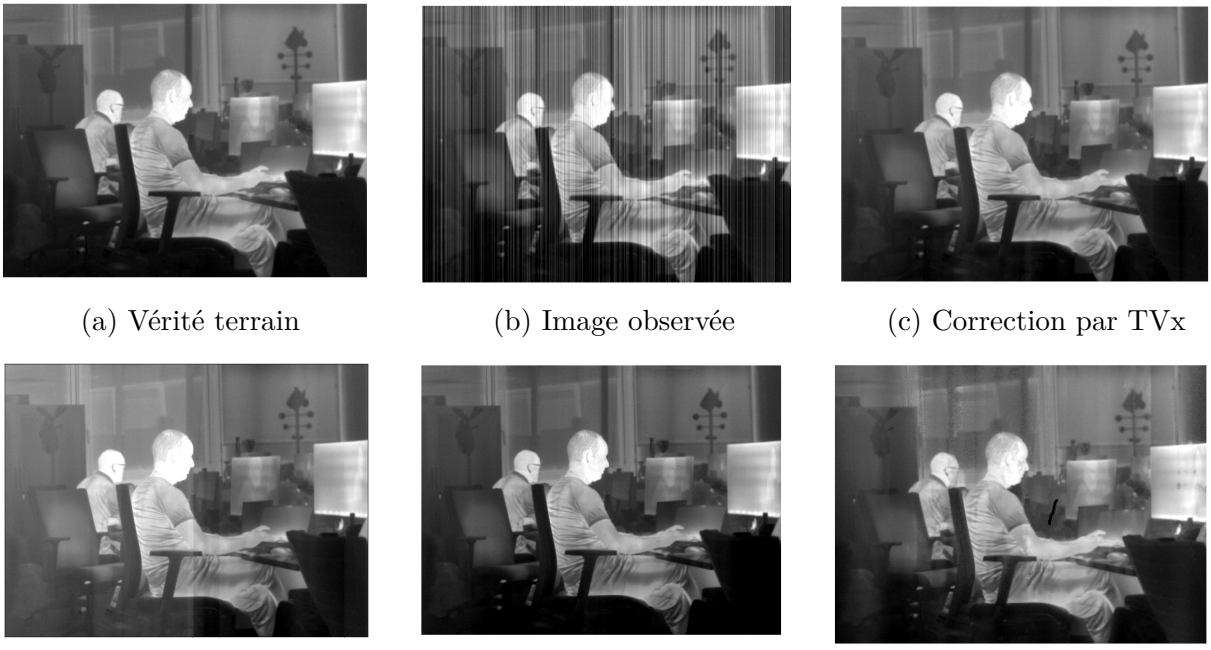
4.4.4.1 Détection des segments homogènes au sein d'une colonne du gradient horizontal

Pour une colonne x, un segment homogène $[\![a, b]\!]$ est un intervalle où les variations du signal $\nabla_x U_0(x, \cdot)$ restent faibles. Sa détection repose sur un test statistique du χ^2 qui considère la somme des variances renormalisées par une valeur de référence représentant le bruit de mesure et une faible texture. Cette somme suit une loi du χ^2 à $b - a$ degrés de liberté, ce qui permet de calculer la p-valeur $p_{a,b}$ mesurant l'homogénéité du segment. Le score

$$\gamma_{a,b} = (b - a) \log \left((1 - p_{a,b})^{-1} \right) \tag{1.27}$$

favorise les segments longs et peu variables.

Pour hiérarchiser les segments d'une colonne, on commence par retenir celui ayant le score le plus élevé puis on sélectionne successivement les segments restants qui ne se chevauchent pas et qui présentent les meilleurs scores. La moyenne d'oscillation de chaque



(a) Vérité terrain (b) Image observée (c) Correction par TVx

(d) Meilleur score (e) Problème variationnel (f) MIRE

FIGURE 1.11 – **Correction d'une image par débruitage colonnaire selon différentes méthodes.** (a) La vérité terrain. (b) L'image bruitée. (c–f) Comparaison des méthodes : le modèle TVx supprime les *stripes* verticales tout en conservant une bonne régularité globale ; la sélection par score γ réduit partiellement le bruit mais génère de fortes discontinuités horizontales ; le problème variationnel offre une correction globalement satisfaisante, avec un gradient horizontal lisse et de petits artefacts résiduels ; la méthode MIRE [32] restitue correctement la scène dans l'ensemble, mais certains contrastes sont mal restitués.

segment fournit alors une estimation possible de $\nabla_x \widehat{C}(x)$.

Une approche simple consiste à considérer pour chaque colonne le segment ayant le score maximal. La figure 1.11d montre que cette stratégie peut créer des discontinuités horizontales marquées lorsque le segment choisi correspond à une structure verticale réelle. Pour y remédier, on formule un problème variationnel visant à sélectionner des segments avec un score élevé tout en assurant un gradient horizontal de l'image corrigée le plus lisse possible.

4.4.4.2 Problème variationnel

Pour chaque colonne, on dispose d'un ou plusieurs candidats pour estimer le gradient de la correction colonnaire, chacun associé à un score issu de la détection de segments homogènes. Le problème consiste à choisir un candidat par colonne de manière à privilégier ceux ayant un score élevé tout en assurant un gradient horizontal de l'image corrigée le plus lisse possible, ce qui se traduit par une pénalisation de la norme euclidienne de la dérivée seconde horizontale. On cherche ainsi une sélection optimale qui équilibre deux objectifs : maximiser la confiance dans les segments choisis et limiter les variations brusques entre colonnes pour réduire les discontinuités horizontales. Cette formulation définit un problème variationnel dont la solution peut être trouvée par programmation dynamique.

La figure 1.11e montre que la méthode supprime efficacement les *stripes* et produit une image corrigée sans discontinuités horizontales marquées. La correction est visuellement satisfaisante, avec une homogénéité globale améliorée par rapport à la sélection systématique du candidat au meilleur score ou à la régularisation TVx. On note néanmoins de légers artefacts de basse fréquence dans la direction horizontale, résultant de l'accumulation des petites erreurs locales lors de l'intégration du gradient estimé $\nabla_x \widehat{C}$, qui se traduisent par un voile ou une dérive progressive sur l'image corrigée.

Après avoir présenté nos contributions principales, le plan général de ce manuscrit de thèse est présenté ci-dessous.

5 Organisation de la thèse

La thèse se compose de quatre chapitres principaux. Le présent chapitre constitue l'introduction, présentant le contexte, les enjeux industriels et scientifiques ainsi que les objectifs et une synthèse des contributions, et il est suivi de trois chapitres détaillant le cœur de l'étude.

Le chapitre 2 présente les principes physiques de l'imagerie thermique, la formation de l'image brute, la formulation du modèle affine qui relie la réponse du capteur à la scène ainsi que l'état de l'art des modèles de correction.

Le chapitre 3 présente la modélisation du capteur et l'analyse statistique des données de calibration. Il inclut la modélisation polynomiale des images brutes, une alternative au modèle *shutterless polynomial* de la littérature, l'estimation du degré via un test d'*overfitting*, l'analyse des erreurs de *fit* et la détection des pixels défectueux. Il traite également de l'étude du comportement des pixels selon la température de scène et la TFPA, ainsi que de l'exploration de l'espace de correction dans le cadre du modèle affine à gain constant et biais variable.

Le chapitre 4 présente les modèles de correction développés. Il comprend un modèle variationnel non paramétrique « mixte » qui exploite simultanément les données de calibration et les informations issues de la scène. Une étude comparative de correction est effectuée avec le modèle *shutterless polynomial*. Le chapitre présente également le débruitage colonnaire en post-traitement, pour lequel on propose un modèle variationnel élaboré à partir de l'analyse du gradient horizontal.

La thèse se conclut par un bilan général et des perspectives, mettant en regard les résultats obtenus avec les objectifs et les enjeux industriels identifiés.

Chapitre 2

Principes physiques de l'imagerie thermique et modèles de correction existants

Tous les objets de la vie quotidienne émettent du rayonnement thermique, y compris les glaçons. L'intensité de ce rayonnement dépend de deux grandeurs physiques principales : la température de l'objet et son émissivité (capacité à rayonner thermiquement). Ainsi deux objets à température égale mais d'émissivités différentes n'émettent pas la même quantité de rayonnement. Par ailleurs, plus un objet est chaud, plus il émet de rayonnement thermique. Le contraste d'une scène thermique est ainsi déterminé par la répartition spatiale de la température et de l'émissivité des surfaces observées.

Une caméra thermique capte ce rayonnement infrarouge et en restitue une image monochrome. Bien qu'elle ne permette généralement pas de voir à travers les objets opaques, elle rend visible la chaleur émise, stockée ou transportée par les matériaux, et permet de visualiser des flux thermiques.

La technologie des capteurs thermiques est en plein essor, avec des applications variées : véhicules autonomes, lutte contre l'incendie, inspection de bâtiments, médecine, contrôle de procédés industriels, veille sanitaire, mesure de pollution, défense, ou encore étude de populations animales.

Dans ce chapitre, nous présentons d'abord les principes physiques du rayonnement infrarouge émis par les corps. Nous décrirons ensuite le fonctionnement des caméras thermiques non-refroidies et la manière dont elles convertissent ce rayonnement en un signal numérique que nous appellerons image brute. Enfin, puisque l'image brute est fortement dégradée par un bruit de structure spatiale non uniforme, nous nous intéresserons aux principaux modèles de correction proposés dans la littérature, développés pour compenser ces non-uniformités.

1 Principes physiques et électroniques de l'imagerie thermique

Une caméra infrarouge est sensible à un rayonnement thermique. Avant de nous pencher sur la technologie des caméras thermiques non-refroidies, nous allons nous intéresser aux principes physiques qui expliquent le rayonnement émis par un objet selon sa température et son émissivité. Pour expliquer ce phénomène, nous commencerons par énoncer

les lois physiques relatives à un objet idéal (d'émissivité totale), désigné sous le terme de corps noir. Nous nous intéresserons ensuite au rayonnement émis par un objet réel appelé corps gris, dont l'émissivité est partielle.

1.1 Rayonnement du corps noir

Avant de nous concentrer sur le rayonnement thermique, rappelons brièvement qu'il existe trois modes principaux de transfert thermique : la conduction, la convection et le rayonnement. La conduction correspond à la transmission d'énergie (généralement par contact direct) à travers un matériau sans déplacement de matière. La convection est induite par un transport d'énergie lié au mouvement d'un fluide. Le rayonnement se produit par émission d'ondes électromagnétiques. Il s'agit donc du seul mode de transfert qui ne requiert aucun support matériel. C'est ce mécanisme que nous allons étudier plus en détail puisqu'il est à la base de l'image thermique.

1.1.1 Définition d'un corps noir

En physique, un corps noir [1] est un objet idéal capable d'absorber intégralement toute l'énergie électromagnétique incidente, sans en réfléchir ni en transmettre aucune (on dit que son émissivité est égale à 1). Selon le principe d'équilibre thermodynamique, cette absorption s'accompagne nécessairement d'une émission de rayonnement électromagnétique, appelé rayonnement du corps noir. Ce rayonnement dépend uniquement de la température de l'objet et constitue une référence théorique fondamentale pour l'étude des transferts radiatifs.

Considérons une cavité fermée, dont les parois internes sont recouvertes de graphite, et percée d'une petite ouverture. Lorsque l'ensemble est placé dans un environnement calorifugé et maintenu à température constante, le rayonnement émis par l'ouverture constitue une excellente approximation de celui d'un corps noir. L'ouverture se comporte alors comme un émetteur quasi-idéal, présentant une répartition angulaire lambertienne (également qualifiée d'orthotrope) [2].

Pourquoi des parois en graphite ? Le graphite est choisi pour tapisser les parois internes de la cavité en raison de ses excellentes propriétés optiques : il présente une forte absorbance sur une large gamme spectrale, une bonne stabilité thermique, et une grande résistance à l'oxydation. Ces caractéristiques en font un matériau particulièrement adapté pour simuler le comportement d'un corps noir réaliste.

Le corps noir est donc un objet théorique dont le spectre d'émission électromagnétique est exclusivement fonction de la température. Nous allons à présent introduire les lois physiques qui permettent de le modéliser.

1. L'expression « corps noir » a été introduite par Gustav Kirchhoff en 1860.
2. L'intensité radiante suit la loi du cosinus de Lambert : $I = I_0 \cos(\phi)$, où ϕ désigne l'angle entre la direction du flux et la normale à la surface émettrice.

1.1.2 Lois physiques

L'étude du rayonnement du corps noir a suscité un vif intérêt dès la fin du XIXe siècle. La première loi formulée dans ce cadre est celle de Rayleigh-Jeans, qui exprime $L_\lambda(T)$, la luminance spectrale (en W/m^3/sr) du rayonnement émis par un corps noir, selon

$$L_\lambda(T) = \frac{2ckT}{\lambda^4}, \tag{2.1}$$

où T désigne la température absolue du corps noir (en K), λ la longueur d'onde du rayonnement (cf. figure 2.1), $c = 2.998 \times 10^8\,\mathrm{m}\cdot\mathrm{s}^{-1}$ la vitesse de la lumière dans le vide et $k = 1.381 \times 10^{-23}\,\mathrm{J}\cdot\mathrm{K}^{-1}$ la constante de Boltzmann.

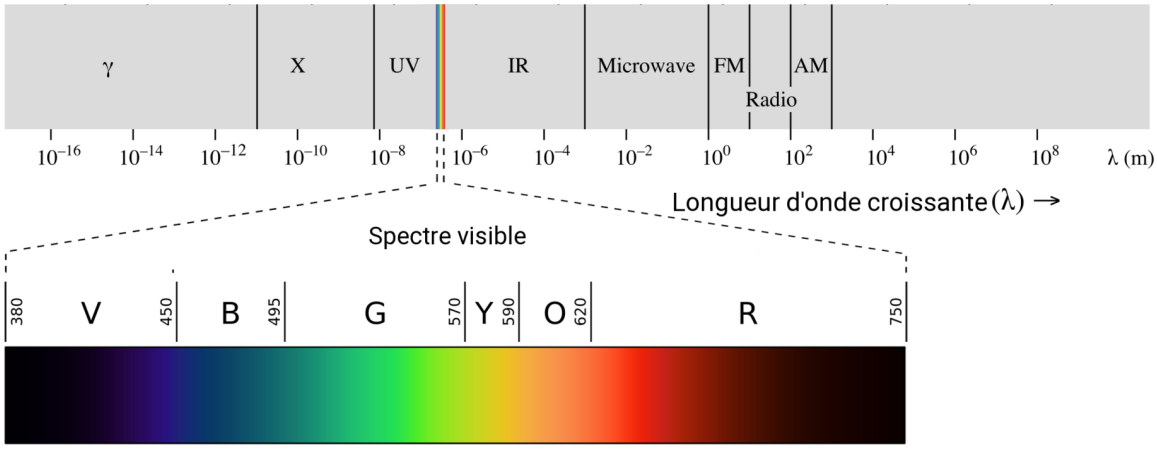


FIGURE 2.1 – **Spectre électromagnétique.** Source Wikipédia. Des plus petites longueurs d'onde aux plus grandes, on distingue : les rayons gamma, les rayons X, les ultraviolets (UV), le spectre visible par l'œil humain, les infrarouges (IR), les micro-ondes et les ondes radio. Notre domaine d'étude est l'infrarouge, dont les longueurs d'onde s'étendent approximativement de $0.75 \cdot 10^{-6}$ à $1000 \cdot 10^{-6}\,\mathrm{m}$.

Cette loi fournit une approximation satisfaisante aux grandes longueurs d'onde, notamment en domaine radio. En revanche, elle diverge pour les courtes longueurs d'onde (par exemple dans l'ultraviolet), où elle prédit une luminance tendant vers l'infini. Ce désaccord avec les mesures expérimentales est désigné sous le nom de « catastrophe ultraviolette » (cf. figure 2.2).

Ce paradoxe fut résolu par l'introduction des concepts fondateurs de la physique quantique. En 1900, Max Planck proposa une nouvelle expression du rayonnement, fondée sur l'hypothèse de quantification de l'énergie. Cette formulation, connue sous le nom de loi de Planck, décrit avec précision le spectre d'émission du corps noir sur l'ensemble des longueurs d'onde et marque une rupture majeure avec les approches classiques.

La loi de Planck exprime la luminance spectrale $L_\lambda(T)$ d'un corps noir selon

$$L_\lambda(T) = \frac{2hc^2}{\lambda^5\left(\exp\left(\frac{hc}{\lambda kT}\right) - 1\right)}, \quad \lambda \in]0, +\infty[,\ T \in]0, +\infty[, \tag{2.2}$$

où $h = 6.626 \times 10^{-34}\,\mathrm{J}\cdot\mathrm{s}$ est la constante de Planck, $c = 2.998 \times 10^8\,\mathrm{m}\cdot\mathrm{s}^{-1}$ la vitesse de la lumière dans le vide, et $k = 1.381 \times 10^{-23}\,\mathrm{J}\cdot\mathrm{K}^{-1}$ la constante de Boltzmann.

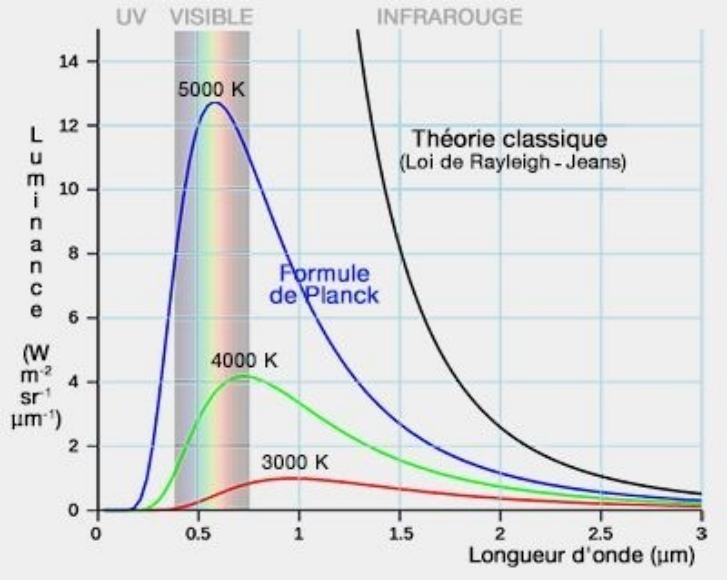


FIGURE 2.2 – **De la théorie classique à la théorie quantique, la catastrophe ultraviolette n'a heureusement pas lieu.** Source Wikipédia. Luminance spectrale en fonction de la longueur d'onde d'émission pour différentes températures de corps noir. Pour des températures élevées comme celle de la surface du Soleil (5000 K), la loi de Rayleigh-Jeans (2.1) prédit une augmentation indéfinie de la luminance spectrale lorsque la longueur d'onde diminue, ce qui impliquerait une émission extrêmement intense dans l'ultraviolet, les rayons X et les rayons gamma. Ce comportement, qualifié de « catastrophe ultraviolette » par le physicien Paul Ehrenfest, révèle l'échec de la physique classique pour les courtes longueurs d'onde. Max Planck introduisit alors une nouvelle formule, fondée sur la quantification de l'énergie, qui marqua la naissance de la physique quantique. La formule de Planck (2.2) remplace celle de Rayleigh-Jeans.

Cette loi constitue un tournant majeur dans la physique : elle repose sur l'hypothèse révolutionnaire que l'énergie électromagnétique n'est pas émise de manière continue, mais par paquets discrets appelés quanta. Elle fournit une description précise de la distribution spectrale du rayonnement d'un corps noir à toute température, et correspond exactement aux mesures expérimentales, y compris à courte longueur d'onde.

Remarque 1 (Équivalence pour les grandes longueurs d'onde)**.** *Pour les grandes longueurs d'onde (ou basses fréquences), on peut montrer que la loi de Planck* (2.2) *tend vers la loi de Rayleigh-Jeans* (2.1)*. Cette convergence justifie l'usage de la loi classique dans le domaine radiofréquence.*

Pour une température T donnée, la luminance spectrale $L_\lambda(T)$ (considérée comme une fonction de λ) présente un maximum en un point $\lambda_{\max}$. L'expression de ce maximum se déduit de la loi de Planck (2.2) et est connue sous le nom de loi du déplacement de Wien. Elle s'écrit

$$\lambda_{\max} = \frac{2.89 \cdot 10^{-3}}{T}. \tag{2.3}$$

La formule de Planck permet, après intégration sur l'ensemble du spectre des longueurs d'onde, de déduire la loi de Stefan (également nommée loi de Stefan-Boltzmann), selon laquelle la puissance surfacique W_S émise par un corps noir de température T est

proportionnelle à T^4. Elle s'écrit

$$W_S = \sigma_S T^4, \tag{2.4}$$

où W_S désigne la puissance totale rayonnée par unité de surface dans le demi-espace libre (en $\mathrm{W/m^2}$), et $\sigma_S = 5.67 \cdot 10^{-8}$ $\mathrm{W/m^2/K^4}$ est la constante de Stefan.

1.2 Corps gris, émissivité

Le corps noir est un objet idéal dont le comportement radiatif ne dépend que de sa température. En pratique, aucun matériau réel ne se comporte exactement comme un corps noir. On parle alors de corps gris pour désigner un objet qui n'absorbe qu'une fraction de l'énergie incidente. Selon le principe de Kirchhoff du rayonnement thermique, à l'équilibre, un objet émet dans les mêmes proportions qu'il absorbe : un corps qui n'absorbe qu'une fraction de l'énergie incidente n'en émet également qu'une fraction. Ainsi, deux objets portés à la même température peuvent présenter des luminances spectrales différentes, en raison de leur émissivité respective, c'est-à-dire leur capacité à émettre un rayonnement par rapport à celle d'un corps noir.

L'émissivité $\varepsilon \in]0, 1]$ d'un corps gris est définie comme le rapport entre la puissance $\widehat{P}$ effectivement émise par une surface S à la température T et la puissance théorique émise par un corps noir dans les mêmes conditions. Elle s'écrit :

$$\varepsilon = \frac{\widehat{P}}{\sigma_S S T^4}. \tag{2.5}$$

La loi de Stefan (2.4) se généralise alors à un corps gris d'émissivité ε :

$$W_S^{(\varepsilon)} = \varepsilon\, \sigma_S T^4, \tag{2.6}$$

où $W_S^{(\varepsilon)}$ désigne la puissance surfacique effectivement rayonnée.

De manière analogue, la loi de Planck s'adapte aux corps gris : leur luminance spectrale devient

$$L_\lambda^{(\varepsilon)}(T) = \varepsilon\, \frac{2hc^2}{\lambda^5 \left(\exp\left(\frac{hc}{\lambda k T}\right) - 1\right)}. \tag{2.7}$$

La loi du déplacement de Wien (2.3), quant à elle, reste inchangée puisqu'elle est indépendante de l'émissivité.

Remarque 2 (Émissivité de quelques objets)**.** *La plupart des substances organiques ont une émissivité ε supérieure à 0.95 : le corps humain atteint 0.98, l'eau 0.96. Les matériaux non métalliques comme le verre, la pierre (0.93), le bois naturel (0.90) ou le carton (0.85) présentent également des émissivités élevées. À l'inverse, les matériaux métalliques réfléchissants ont des émissivités beaucoup plus faibles : l'argent (0.03), l'aluminium non oxydé (moins de 0.1), l'acier inoxydable (0.2). Lorsqu'un objet présente une émissivité très faible (proche de 0), on parle de corps blanc. En thermographie (ce qui dépasse le cadre du présent travail), les caméras infrarouges peuvent intégrer des algorithmes de détection d'objets. Une fois un objet détecté, l'utilisateur peut renseigner la nature du matériau, à laquelle on associe une émissivité. Ces informations permettent alors à un modèle de correction d'estimer avec précision la température de l'objet.*

1.3 Rayonnement infrarouge, irradiance reçue par un pixel

Le rayonnement infrarouge (IR) est un rayonnement électromagnétique dont la longueur d'onde est comprise entre 0.75 et 1000 µm (cf. figure 2.1). Il n'est donc pas perceptible par l'homme. L'existence de ce rayonnement a été mise en évidence en 1800 par l'astronome anglais William Herschel. En étudiant le spectre solaire après dispersion à l'aide d'un prisme, il observa, en déplaçant un thermomètre au-delà du rouge visible, une élévation de température inattendue. Il venait ainsi de détecter pour la première fois le rayonnement infrarouge émis par le Soleil.

Le rayonnement infrarouge est généralement associé à une émission de chaleur. En effet, lorsque ce rayonnement atteint un matériau (par exemple la peau humaine), il induit des vibrations au niveau des liaisons moléculaires, générant ainsi une énergie thermique que nous percevons sous forme de chaleur.

D'après la loi de Planck (2.7) (cf. figure 2.3), on comprend aisément que l'intensité du rayonnement infrarouge augmente avec la température T de l'objet émetteur.

Le spectre du rayonnement infrarouge peut être divisé en trois grandes bandes, dont les limites varient légèrement selon le domaine d'application [3]. En thermographie infra-

3. En astronomie par exemple, le proche infrarouge (*Near InfraRed*, NIR) se situe entre 0.75 et 2.5 µm, l'infrarouge moyen (*Mid InfraRed*, MIR) entre 2.5 et 40 µm et l'infrarouge lointain (*Far InfraRed*, FIR) entre 40 et 800 µm.

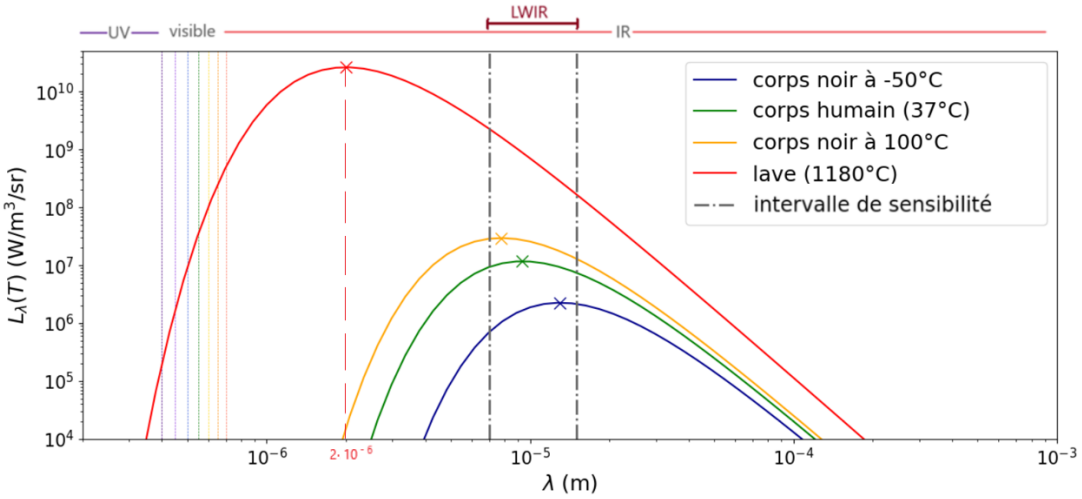


FIGURE 2.3 – **Courbe de luminance spectrale émise par des objets à différentes températures.** Les axes sont en échelle logarithmique. Les courbes sont issues de la loi de Planck (2.7). Pour chaque température, le maximum d'émission se situe à une longueur d'onde donnée, conformément à la loi du déplacement de Wien (2.3) ; il est indiqué par une croix sur la courbe correspondante. Lorsqu'un objet chauffe, sa luminance spectrale augmente, notamment dans la bande LWIR (*Long-Wavelength Infrared*), qui s'étend de $7 \cdot 10^{-6}$ à $15 \cdot 10^{-6}$,m et correspond à la sensibilité spectrale des capteurs des caméras thermiques non-refroidies. Le corps noir à $-50\,°C$ (bleu), le corps humain (vert) et le corps noir à $100\,°C$ (orange) émettent principalement dans le LWIR. La lave (rouge) présente un pic autour de $2 \cdot 10^{-6}\,m$; elle émet donc à la fois dans le LWIR, détectable par les caméras thermiques, et dans le visible, où l'œil humain perçoit une dominante rouge. La bande spectrale LWIR est ainsi bien adaptée pour observer les objets courants émettant dans l'infrarouge thermique.

rouge, on travaille généralement dans une bande spectrale comprise entre 0.9 et 16 µm, et plus particulièrement dans les fenêtres spectrales [2 ; 5 µm] (*Middle Wavelength InfraRed*, MWIR) et [7 ; 15 µm] (*Long Wavelength InfraRed*, LWIR).

Une caméra thermique est composée d'un capteur sensible au rayonnement infrarouge ; son rôle est de capter le rayonnement émis par les objets d'une scène. Ce capteur est constitué d'une grille (ou matrice) de pixels situé dans le plan focal de l'optique, et est désigné par l'expression *focal-plane array* (FPA).

On introduit la notion d'irradiance $I_{T,\varepsilon}$ pour désigner la puissance surfacique du rayonnement infrarouge reçue en un point d'un pixel (en $\mathrm{W} \cdot \mathrm{m}^{-2}$), émise par une source de température T et d'émissivité ε. On suppose cette irradiance uniforme sur toute la surface du pixel. Elle s'écrit

$$I_{T,\varepsilon} = \Omega \int_0^\infty L_\lambda^{(\varepsilon)}(T)\, D(\lambda)\, \mathrm{d}\lambda\,, \tag{2.8}$$

où, pour toute longueur d'onde $\lambda > 0$, $L_\lambda^{(\varepsilon)}(T)$ est la luminance spectrale émise par la source (voir équation (2.7)), $D(\lambda)$ la fonction d'absorption spectrale du matériau détecteur (cf. figure 2.4) et Ω l'angle solide sous lequel l'optique forme l'image de la scène sur un pixel du FPA.

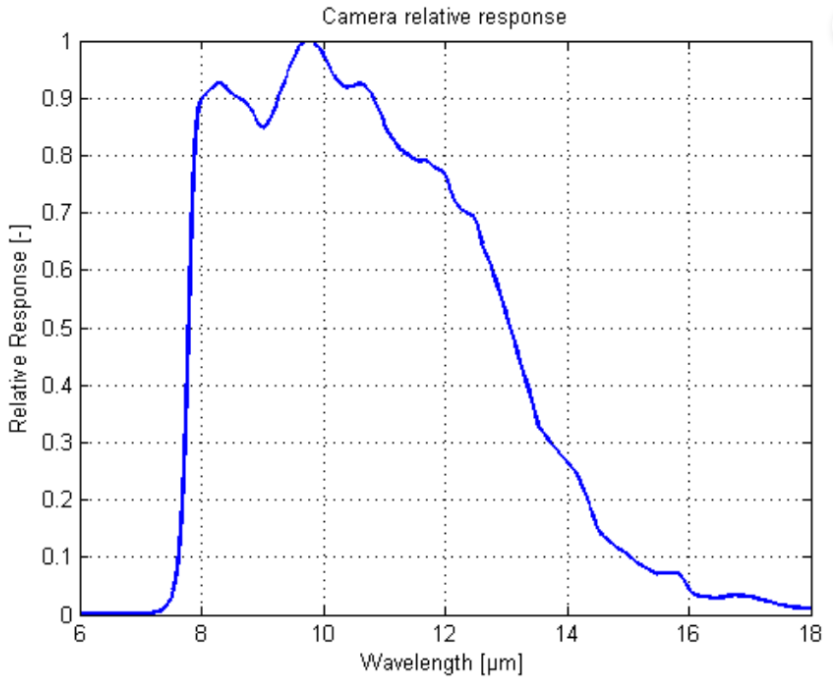


FIGURE 2.4 – **Absorption spectrale d'un pixel de caméra thermique.** Données issues de Flir (courbes typiques de réponse spectrale). La sensibilité du microbolomètre à un flux infrarouge dépend de la longueur d'onde. Dans cet exemple, le rayonnement IR compris dans la bande $[8, 16]$ µm est majoritairement absorbé. L'irradiance reçue par le pixel correspond à l'intégrale, sur l'ensemble des longueurs d'onde, des luminances spectrales incidentes pondérées par cette courbe d'absorption spectrale (équation (2.8)).

Remarque 3 (Relation entre irradiance et température de scène.)**.** *En imagerie thermique, on confond souvent les notions d'irradiance et de température de scène. La température est une grandeur plus intuitive : on parle ainsi de sensibilité thermique ou encore*

de température d'un corps noir. Les modèles de correction évoqués par la suite consistent à retrouver l'irradiance de la scène à partir d'une image observée. Par abus de langage, nous parlerons donc fréquemment de température de scène, bien qu'il ne s'agisse pas, à proprement parler, de la température physique réelle des objets observés.

Il est néanmoins possible de retrouver cette température physique à partir de l'irradiance mesurée, comme nous l'expliquons ci-dessous.

Un objet (corps gris) d'émissivité ε porté à une température T émet, selon la loi de Planck généralisée (2.7), une luminance spectrale $L_\lambda^{(\varepsilon)}(T)$. Cette luminance contribue à l'irradiance $I_{T,\varepsilon}$ reçue par un pixel via l'équation (2.8). En connaissant l'émissivité ε de l'objet, il est alors possible d'établir une correspondance entre l'irradiance mesurée et la température de l'objet (cf. figure 2.5).

Cette relation, complexe en raison de l'intégration sur le spectre de fonctions planckiennes, peut être approchée à l'aide des équations de Sakuma-Hattori [38–40]. Ces équations établissent un lien fonctionnel entre la température T d'un objet d'émissivité ε et l'irradiance $I_{T,\varepsilon}$ reçue par un pixel.

Dans la section suivante, nous nous intéressons à la technologie des caméras thermiques non-refroidies. Le capteur est constitué d'une matrice de pixels, chacun étant associé à un microbolomètre. Ce composant détecte le rayonnement infrarouge en convertissant l'irradiance reçue en une variation de résistance, laquelle est ensuite traduite en une tension électrique. Cette tension est finalement convertie en un signal numérique à l'aide d'un circuit électronique de lecture (*readout integrated circuit*, ROIC).

1.4 Technologie des caméras thermiques non-refroidies et génération de l'image brute

On distingue deux grandes catégories de caméras thermiques : les caméras à refroidissement intégré et les caméras non-refroidies. Deux paramètres qualitatifs essentiels caractérisent les capteurs thermiques : la sensibilité thermique et le temps de réponse thermique. La sensibilité thermique, ou *noise equivalent temperature difference* (NETD), correspond à la plus petite différence de température (de scène) que la caméra est capable de détecter. Le temps de réponse, quant à lui, mesure le délai nécessaire au capteur pour réagir à un changement de température (plus précisément, de luminance) provenant de la scène observée.

Dans ce manuscrit, notre attention se porte exclusivement sur les caméras thermiques non-refroidies et les modèles de correction qui leur sont associés. Nous présentons toutefois brièvement, à titre de comparaison, la technologie des caméras refroidies.

Les caméras thermiques refroidies reposent sur une détection photonique : elles mesurent directement le flux de photons d'énergie dans une bande spectrale spécifique, généralement le domaine infrarouge moyen (MWIR). Leur capteur, qualifié de *photodétecteur*, convertit les photons incidents en électrons, lesquels sont ensuite stockés dans un condensateur d'intégration. Ces dispositifs offrent une sensibilité thermique élevée, typiquement de l'ordre de 20 mK, et un temps de réponse très court, de l'ordre de 50 µs, ce qui permet de capturer des phénomènes thermiques rapides avec une grande précision. Les images produites par ces caméras présentent un excellent niveau de détail, rendant visibles les

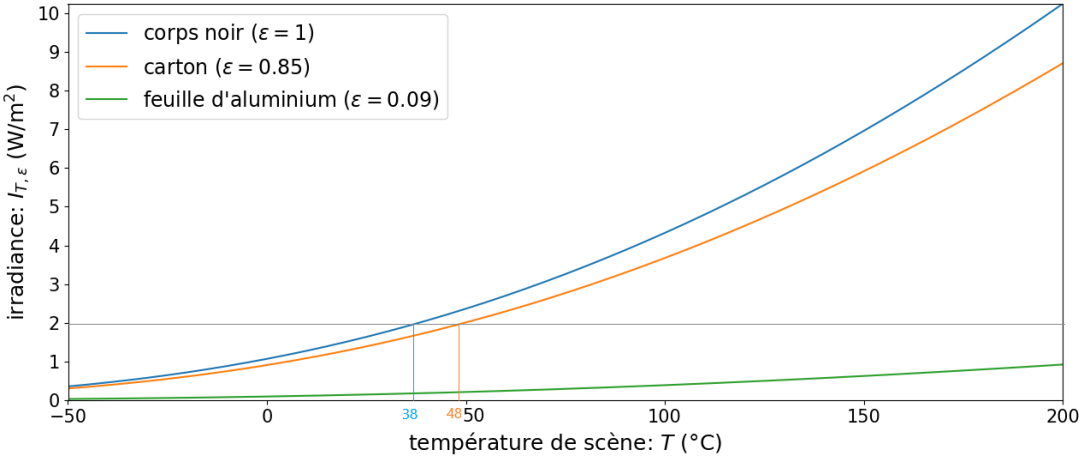


FIGURE 2.5 – **Lien entre température de scène et irradiance reçue.** La scène est composée d'un objet d'émissivité ε et de température T. Cet objet émet, selon la loi de Planck, une luminance spectrale qui, après passage dans l'optique de la caméra, impacte le pixel. L'irradiance $I_{T,\varepsilon}$ correspond à la puissance surfacique reçue par le pixel, issue du rayonnement infrarouge émis par l'objet. Il existe donc, connaissant l'émissivité ε, une relation entre température de scène T et irradiance $I_{T,\varepsilon}$. Cette relation est illustrée par les trois courbes, chacune correspondant à un objet d'émissivité différente. À irradiance identique, la température de l'objet dépend de son émissivité : par exemple, une irradiance de 2 W/m^2 (ligne horizontale grise) correspond à une température de 38 °C pour un corps noir (en bleu), 48 °C pour un objet en carton (en orange) et 339 °C pour une feuille d'aluminium (en vert). Les modèles de correction présentés ultérieurement visent à relier l'irradiance à la réponse du capteur. Or, les données de calibration disponibles proviennent de corps noirs dont seule la température est fournie, et non leur irradiance. Dans ce manuscrit, les résultats seront donc illustrés à l'aide d'images dites « en température de scène », même si ces températures ne sont pas toujours exactes. Notre objectif étant la qualité de reconstruction des images, une estimation précise des températures n'est pas centrale ici. Pour les applications en thermographie, il sera toutefois possible de retrouver des températures plus précises (connaissant l'émissivité) à partir des équations de Sakuma-Hattori (cf. remarque 3).

moindres variations thermiques. Par ailleurs, les images brutes sont généralement directement exploitables, nécessitant peu, voire aucune correction. En contrepartie, ces systèmes sont coûteux et encombrants [41].

1.4.1 Capteur d'une caméra thermique non-refroidie, microbolomètres, circuit électronique de lecture.

La technologie des caméras thermiques non-refroidies repose sur la thermo-détection plutôt que sur la photo-détection : on ne mesure pas le nombre de photons émis par un flux lumineux, mais l'élévation de température d'un matériau qui absorbe le rayonnement infrarouge reçu. Un pixel est composé d'un microbolomètre (figure 2.6), un composant électronique dont la résistance électrique varie avec la température.

Pour comprendre le signal délivré par un pixel d'une caméra non-refroidie, on peut représenter l'ensemble des flux thermiques qui l'impactent comme illustré à la figure 2.7.

On fait l'hypothèse que l'ensemble des flux thermiques est stable temporellement, et que la résistance du microbolomètre vaut $R(T_0)$, où T_0 est la température stabilisée du

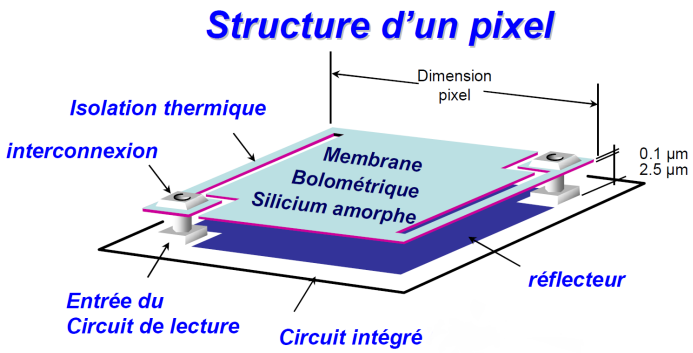


FIGURE 2.6 – **Schéma d'un microbolomètre associé à un pixel.** Source *Lynred*. La taille moyenne d'un microbolomètre est de l'ordre de 10 µm×10 µm. Les photons issus d'une source de chaleur viennent impacter la membrane absorbante du microbolomètre, entraînant une élévation de sa température et, par conséquent, une variation de sa résistance électrique. Certaines radiations pouvant traverser la couche absorbante, un réflecteur est positionné sous la membrane afin d'augmenter l'efficacité d'absorption du dispositif, permettant ainsi de maximiser le signal détecté. Afin de garantir une isolation thermique optimale entre la membrane et le substrat (contenant l'électronique de lecture), tout en assurant une bonne conductivité électrique, l'interconnexion est réalisée à l'aide de deux contacts en or.

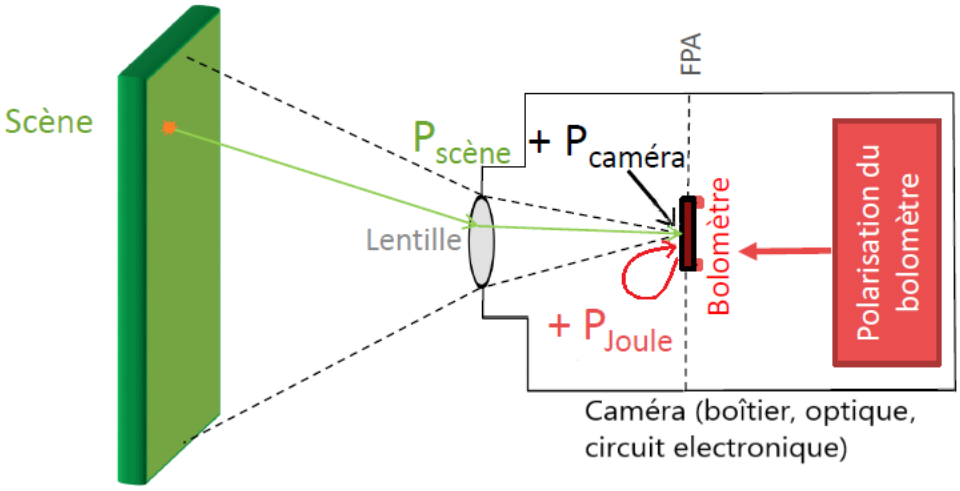


FIGURE 2.7 – **Bilan des flux thermiques sur un microbolomètre.** Un pixel est constitué d'un microbolomètre dont la résistance varie en fonction des flux thermiques qu'il reçoit. Ces flux proviennent de la scène (en vert), de la caméra elle-même (en noir), et de l'effet Joule induit par le courant électrique traversant le bolomètre (en rouge). Ce dernier phénomène domine largement, notamment par rapport aux flux provenant de la scène. C'est la combinaison de ces différents apports qui modifie la température du microbolomètre, et donc sa résistance, à l'origine du signal délivré par la caméra.

microbolomètre. Imaginons maintenant que la température (et donc la luminance) de la scène, représentée par le point orange sur la figure 2.7, change instantanément, par exemple lorsqu'un corps noir froid est soudainement remplacé par un corps noir chaud. Ce changement de température de scène implique une variation de la puissance reçue par

le microbolomètre, égale au produit de l'irradiance par la surface du pixel. Cela induit une très légère élévation δT de la température du bolomètre.

Cette élévation δT n'est pas immédiate (cf. figure 2.8) : le temps de réponse thermique d'un microbolomètre est d'environ 40 ms. Une fois ce délai écoulé, la température du microbolomètre passe de T_0 à $T_0 + \delta T$, et sa résistance évolue de $R(T_0)$ à $R(T_0 + \delta T)$. Au premier ordre, on a la relation suivante

$$R(T_0 + \delta T) = R(T_0)(1 + \alpha\, \delta T), \tag{2.9}$$

où α désigne le coefficient thermique de résistance. Il dépend du matériau semi-conducteur qui compose la membrane du microbolomètre. Ce coefficient est toujours négatif, autrement dit, plus le microbolomètre chauffe, plus sa résistance diminue [42]. Pour qu'une faible variation de température entraîne une variation significative de résistance, α doit être élevé en valeur absolue. Les matériaux utilisés sont généralement du silicium amorphe ou de l'oxyde de vanadium, avec des coefficients thermiques de résistance de l'ordre de $\alpha = -2\%/\mathrm{K}$. Ces matériaux doivent également présenter une faible résistance nominale afin de manière à engendrer le moins d'effet Joule possible lorsqu'un courant lui sera appliqué.

La variation de résistance formulée par (2.9) induit ensuite une variation de la tension aux bornes du microbolomètre. Cette tension est mesurée à l'aide d'un amplificateur opérationnel, qui la compare à celle d'un pixel aveugle (figure 2.9). Elle est ensuite amplifiée, échantillonnée, puis convertie en signal numérique. L'ensemble des valeurs ainsi obtenues pour tous les pixels forme ce que nous appelons une image brute.

Une variation de la température de scène induit donc, par ce processus, une variation de la réponse du capteur. La sensibilité thermique (NETD) d'un capteur non-refroidi est évaluée à environ 40 mK.

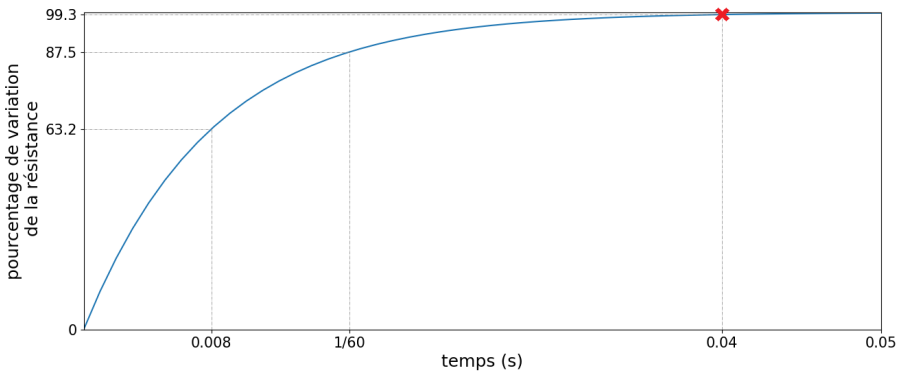


FIGURE 2.8 – **Temps de réponse thermique d'un microbolomètre.** Lorsqu'un pixel observe une variation soudaine d'irradiance, la résistance de son microbolomètre évolue progressivement selon une loi exponentielle. La constante de temps associée est de 0.008 s. L'état stable est atteint au bout de cinq constantes de temps : après 0.04 s (croix rouge), plus de 99 % de la variation est accomplie. Or, à 60 images par seconde (1/60 s), seule une partie de cette variation est prise en compte (environ 87.5 %). Ce temps de réponse limite donc la fréquence du flux vidéo. En cas de transition rapide, la mesure des irradiances devient imprécise. Un arrêt sur image montre alors des objets flous ou des artefacts de « *ghosting* » liés au mouvement de la scène.

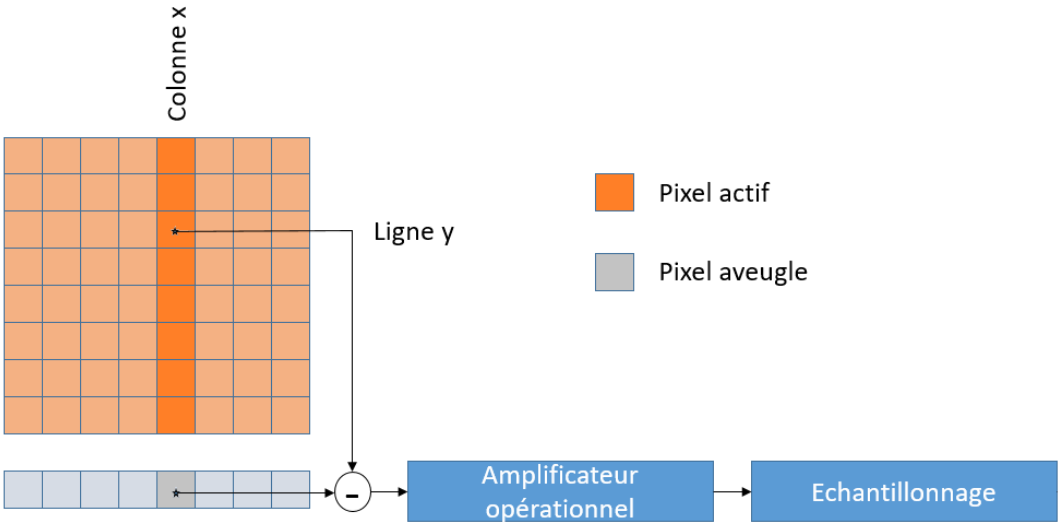


FIGURE 2.9 – **Schéma du circuit de lecture d'une matrice de microbolomètres**. Pour chaque colonne, la tension mesurée sur un pixel actif (sensible à la scène) est soustraite à celle d'un pixel aveugle (insensible à la scène) situé dans la même colonne. Le courant ainsi obtenu est ensuite amplifié, échantillonné, puis converti en signal numérique. Cette soustraction permet de compenser partiellement l'effet Joule, causé par le courant de polarisation qui parcourt le microbolomètre, et qui constitue la principale contribution au flux thermique (cf. figure 2.7). Toutefois, elle introduit (ou accentue) un artefact appelé effet colonnaire, systématiquement observable dans les images brutes issues de capteurs non-refroidis (cf. figure 2.10b).

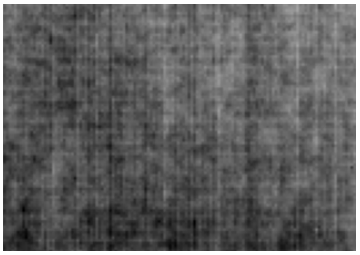
Les caméras thermiques non-refroidies délivrent des images brutes très fortement bruitées (cf. figure 2.10b). Cela s'explique par deux raisons principales. En premier lieu, chaque pixel possède sa propre géométrie : la surface d'un pixel peut varier, de même que son coefficient de résistance thermique. En second lieu, les flux thermiques autres que ceux provenant de la scène ont une influence prépondérante sur la réponse du capteur. C'est un peu comme si une caméra classique, captant la lumière dans le domaine du visible, était dotée d'un boîtier translucide et d'éléments électroniques internes éblouissant son capteur.

Il est donc nécessaire de corriger le signal capté afin d'obtenir une image exploitable. Une difficulté supplémentaire réside dans le fait que le signal de sortie a tendance à dériver dans le temps, sous l'effet des variations thermiques du capteur lui-même, des fluctuations de température de l'environnement de la caméra ainsi que des instabilités des paramètres physiques et électroniques (parfois mal connus ou absents du modèle de correction). Il devient alors indispensable d'effectuer régulièrement une mise à jour des paramètres de correction.

Parmi les caméras non-refroidies, on distingue généralement celles qui sont équipées d'un obturateur mécanique (*shutter*) de celles qui ne le sont pas (dites « *shutterless* »). L'obturateur mécanique, en se plaçant périodiquement devant le capteur, permet une recalibration, c'est à dire une mise à jour des paramètres de correction à appliquer à chaque pixel. Deux inconvénients sont toutefois associés à ce dispositif : d'une part, il complique la miniaturisation de la caméra, et d'autre part, il engendre une discontinuité du flux vidéo. Dans ce manuscrit, nos études et expérimentations sont menées à partir de caméras thermiques non-refroidies et sans *shutter* (*shutterless*).



(a) Caméra refroidie (observée)



(b) Non-refroidie (observée)

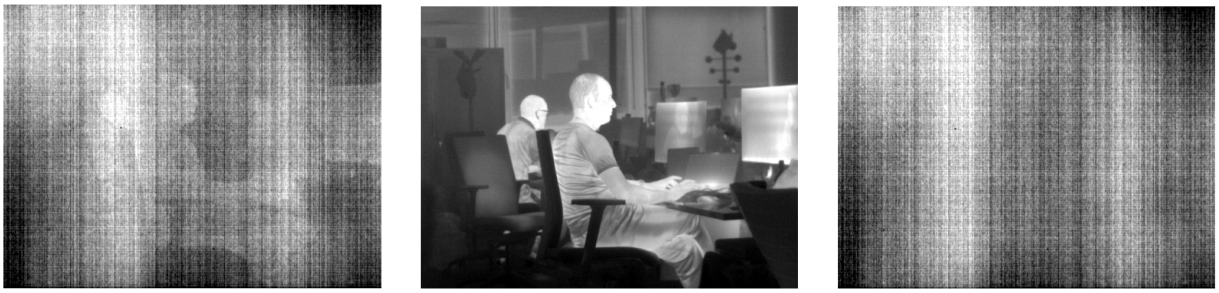


(c) Non refroidie (corrigée)

Type de caméra	Refroidie	Non-refroidie
Technologie d'un pixel	Photodiode	Microbolomètre
Domaine spectral	MWIR	LWIR
Sensibilité thermique (NETD)	20 mK	40 mK
Temps de réponse thermique	50 µs	40 ms
Image brute	Exploitable	Non exploitable
Prix approximatif	100000€	5000€
Taille de la caméra	De l'ordre de 30 cm	De l'ordre de 4 cm

(d) Tableau comparatif entre les technologies refroidies et non-refroidies

FIGURE 2.10 – **Comparaison entre les caméras thermiques refroidies et non-refroidies.** Les scènes considérées ici ont relativement peu de mouvement. Toutes les images ont la même résolution (130 par 90). (a) L'image issue d'une caméra refroidie est parfaitement lisible et ne nécessite que peu de correction. L'excellente sensibilité thermique du capteur (20 mK) permet de percevoir avec précision les détails thermiques de tous les objets présents dans la scène (e.g. les plis sur les vêtements). (b) L'image observée (ou brute) issue d'une caméra thermique non-refroidie (il s'agit ici de la caméra *Smart IR320*) ne permet pas de distinguer la scène. On observe de fortes non-uniformités spatiales (effets de ligne, de colonne et de basse fréquence). Ce bruit fixe (*fixed pattern noise*, FPN) provient des flux thermiques parasites (autres que la scène) et du circuit de lecture électronique. Un modèle de correction est donc nécessaire. (c) L'image corrigée de la sous-figure (b), obtenue à l'aide d'un modèle de correction issu de la littérature (le modèle *shutterless polynomial*), permet de distinguer les objets. Toutefois, le niveau de détail reste inférieur à celui fourni par une caméra refroidie : certains éléments fins, comme les plis des vêtements, apparaissent de manière moins nette. Cette perte de précision visuelle s'explique par la nature du capteur, un thermodétecteur, dont la sensibilité thermique (40 mK) est inférieure à celle des photodétecteurs utilisés dans les caméras refroidies. (d) Les caméras thermiques refroidies ont une meilleure sensibilité thermique, un bien meilleur temps de réponse thermique (permettant de faire d'excellents arrêts sur images lors de phases de transition rapides) et délivrent une image brute ne nécessitant que peu de corrections. En revanche, elles sont beaucoup plus onéreuses et encombrantes. Elles ne feront cependant l'objet d'aucune étude ou expérimentation dans ce manuscrit.



(a) Image brute (b) Image corrigée de la scène (c) FPN

FIGURE 2.11 – **Image brute, image corrigée, motif du bruit FPN.** Les images proviennent de la caméra *ATI-640b*. (a) L'image brute est difficilement lisible. Elle est fortement dégradée par un bruit à structure spatiale non homogène, caractérisé par des motifs de colonne et des composantes de basse fréquence. (b) L'image corrigée a été obtenue à l'aide d'un modèle de correction qui permet d'estimer la scène observée. (c) Le FPN (« fixed-pattern noise ») correspond à une image du bruit spatialement non uniforme qui se superpose à la scène dans l'image brute. Il peut être vu comme la différence entre les images (a) et (b). Les effets de colonne y sont particulièrement marqués, ce qui s'explique notamment par l'utilisation d'un même pixel aveugle pour toutes les lignes d'une colonne donnée (cf. figure 2.9). Le motif du FPN diffère d'une caméra à l'autre.

2 Modélisation du capteur et état de l'art des modèles de correction

L'image brute délivrée par une caméra thermique non-refroidie est difficilement exploitable. Elle est fortement altérée par un bruit de non-uniformité spatiale, aussi appelé *fixed-pattern noise* (FPN). Ce bruit génère un motif qui se superpose à l'image de la scène (cf. figure 2.11a). Il résulte notamment des variations introduites lors de la fabrication des microbolomètres constituant le capteur, de la présence de flux thermiques parasites (autres que ceux provenant de la scène), ainsi que de l'électronique du circuit de lecture. La structure spatiale du FPN varie peu dans le temps, mais son intensité dépend fortement de l'environnement thermique de la caméra, en particulier de la température du plan focal (TFPA) [43]. Le FPN diffère d'un capteur à un autre.

En raison de l'hétérogénéité spatiale du FPN, les algorithmes classiques de débruitage sont peu efficaces. Une modélisation affine de la réponse de chaque pixel est couramment utilisée pour corriger ces non-uniformités (*non uniformity correction*, NUC) [1, 2]. Ce modèle établit, pixel par pixel, une relation entre la valeur observée et l'irradiance issue de la scène. Il repose sur des principes physiques et électroniques [8].

Les techniques de correction proposées dans la littérature visent à estimer les deux composantes de ce modèle affine. On distingue principalement deux grandes familles de méthodes : les méthodes *calibration-based*, qui déterminent les paramètres du modèle à partir de données de calibration, et les méthodes *scene-based*, qui exploitent directement l'information contenue dans un flux temporel d'images.

Avant de proposer une première formulation du modèle affine, introduisons θ comme l'ensemble des paramètres externes à la scène mais influençant la réponse du capteur.

2.1 Modélisation affine du capteur

2.1.1 Paramètres environnementaux qui influent sur la réponse du capteur

On définit l'environnement θ comme un vecteur regroupant l'ensemble des paramètres (autres que ceux liés à la scène) ayant une influence sur la réponse du capteur. Ce vecteur peut être scindé en deux sous-ensembles : l'un associé à l'environnement thermique du capteur, l'autre aux paramètres électroniques du circuit de lecture.

Parmi les composantes de θ, certaines sont contrôlables (comme les paramètres du circuit électronique de lecture), d'autres sont mesurables (comme la température du plan focal, TFPA), et certaines restent difficilement mesurables (comme les flux thermiques non homogènes issus du boîtier de la caméra ou la dérive temporelle du circuit de lecture), voire non identifiables.

Au sein des composantes du sous-ensemble thermique, la TFPA constitue le paramètre le plus influent sur la réponse du capteur. C'est pourquoi la plupart des caméras thermiques sont munies d'une sonde située au niveau du plan focal, permettant d'en fournir une mesure notée τ.

2.1.2 Modèle affine multi-paramétrique

2.1.2.1 Modèle affine direct avec responsivité et décalage

Soit $r_\theta \in \mathbb{R}^\Omega$ l'image brute (*raw*) délivrée par la caméra dans un environnement θ, et

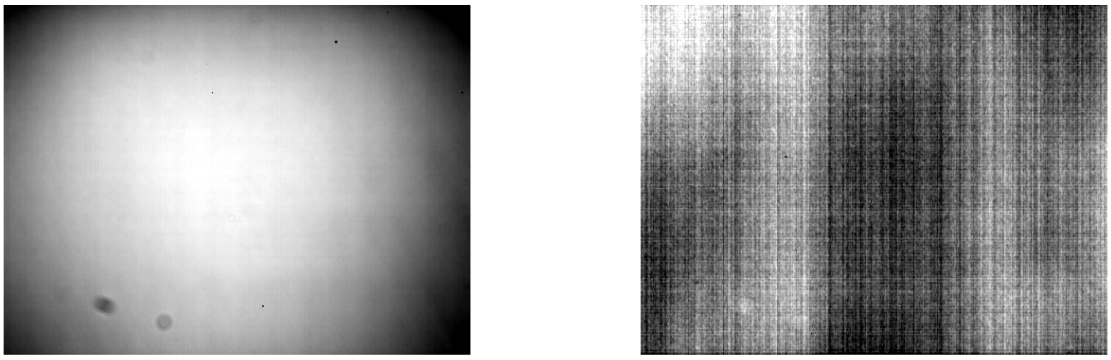
soit $(x, y) \in \Omega$ un pixel donné. Le modèle affine direct s'écrit

$$r_\theta(x, y) = \mathcal{R}_\theta(x, y)\, S(x, y) + \mathcal{D}_\theta(x, y)\,, \tag{2.10}$$

où $\mathcal{R}_\theta \in \mathbb{R}^\Omega$ et $\mathcal{D}_\theta \in \mathbb{R}^\Omega$ désignent respectivement l'image de responsivité et l'image de décalage en niveau de gris, et $S \in \mathbb{R}^\Omega$ représente l'image des irradiances de la scène.

La responsivité d'un pixel quantifie la sensibilité de sa réponse en tension par rapport à l'irradiance de scène. Elle dépend principalement de l'optique du système. La responsivité est toujours strictement positive, sauf dans le cas d'un pixel défectueux [4]. L'image $\mathcal{R}_\theta$ est généralement lisse spatialement (cf. figure 2.12a).

Le décalage correspond à une composante additive liée à tous les flux thermiques parasites présents lors de l'acquisition, c'est-à-dire autres que ceux provenant de la scène. En raison de l'électronique de lecture, le décalage en un pixel (x, y) dépend fortement de sa position colonnaire x. L'image $\mathcal{D}_\theta$ présente ainsi un fort effet de colonne et une structure spatiale irrégulière (cf. figure 2.12b). Il existe une forte analogie entre le décalage et le FPN : le décalage peut en effet être interprété comme l'image brute obtenue en observant une scène uniforme d'irradiance nulle.



(a) Responsivité (b) Décalage

FIGURE 2.12 – **Images de responsivité et de décalage d'une caméra infrarouge non-refroidie.** Les données ont été obtenues à partir de la caméra *ATI-640b*. (a) L'image de responsivité présente des variations spatiales lentes et apparaît globalement lisse. On distingue une forme circulaire inscrite, correspondant à la zone utile de l'optique, à l'extérieur de laquelle les valeurs sont plus faibles. Deux taches sombres apparaissent également en bas de l'image, imputables à des défauts optiques. La responsivité d'un pixel dépend essentiellement de l'optique du système. (b) L'image de décalage présente un fort effet de colonne, ce qui la rend très irrégulière spatialement. Elle montre une forte ressemblance avec l'image du FPN (cf. figure 2.11c) observée avec la même caméra.

Nous écrirons aussi le modèle affine direct (2.10) de manière plus compacte

$$r_\theta = \mathcal{R}_\theta\, S + \mathcal{D}_\theta\,. \tag{2.11}$$

Lorsque les images de responsivité $\mathcal{R}_\theta$ et de décalage $\mathcal{D}_\theta$ sont connues (ce qui est l'objet des méthodes de correction présentées par la suite), il est possible de reconstruire l'image

4. Nous reviendrons sur la notion de pixel défectueux au chapitre suivant (cf. 1.4.1).

de la scène S à partir de l'image brute observée r_θ. En effet, lorsque $\mathcal{R}_\theta$ est strictement positive, on peut inverser (2.11) et obtenir

$$S = \frac{r_\theta - \mathcal{D}_\theta}{\mathcal{R}_\theta}\,. \tag{2.12}$$

Ce modèle inverse est fréquemment reformulé dans la littérature sous la forme d'un modèle affine de correction, faisant intervenir un gain et un offset, que nous présentons à présent.

2.1.2.2 Reformulation du modèle affine multi-paramétrique avec gain et offset

On suppose connues les images de responsivité $\mathcal{R}_\theta$ et de décalage $\mathcal{D}_\theta$. À partir de celles-ci et de l'image brute S issue du modèle inverse (2.12), on définit l'image corrigée U (exprimée en *digital level* (DL)), ainsi que les images de gain $\mathcal{G}_\theta$ et d'offset $\mathcal{O}_\theta$, selon

$$\begin{cases} \mathcal{G}_\theta = \dfrac{\mu(\mathcal{R}_\theta)}{\mathcal{R}_\theta}, & \text{(2.13a)} \\ \mathcal{O}_\theta = -\mathcal{G}_\theta\,\mathcal{D}_\theta, & \text{(2.13b)} \\ U = \mu(\mathcal{R}_\theta)\,S, & \text{(2.13c)} \end{cases}$$

où $\mu(\cdot)$ désigne la moyenne spatiale d'une image.

Par substitution, la formule (2.12) devient alors :

$$U = \mathcal{G}_\theta\, r_\theta + \mathcal{O}_\theta\,. \tag{2.14}$$

En effet, partant de (2.12) et en supposant que $\forall (x,y) \in \Omega, \mathcal{R}_\theta(x,y) > 0$, on a

$$\begin{aligned} & S = \frac{r_\theta - \mathcal{D}_\theta}{\mathcal{R}_\theta} \\ \iff & \mu(\mathcal{R}_\theta)\ S = \frac{\mu(\mathcal{R}_\theta)}{\mathcal{R}_\theta}\,(r_\theta - \mathcal{D}_\theta) \\ \iff & U = \mathcal{G}_\theta\, r_\theta - \mathcal{G}_\theta\,\mathcal{D}_\theta \quad \text{d'après (2.13c)} \\ \iff & U = \mathcal{G}_\theta\, r_\theta + \mathcal{O}_\theta \quad \text{d'après (2.13b)}\,. \end{aligned}$$

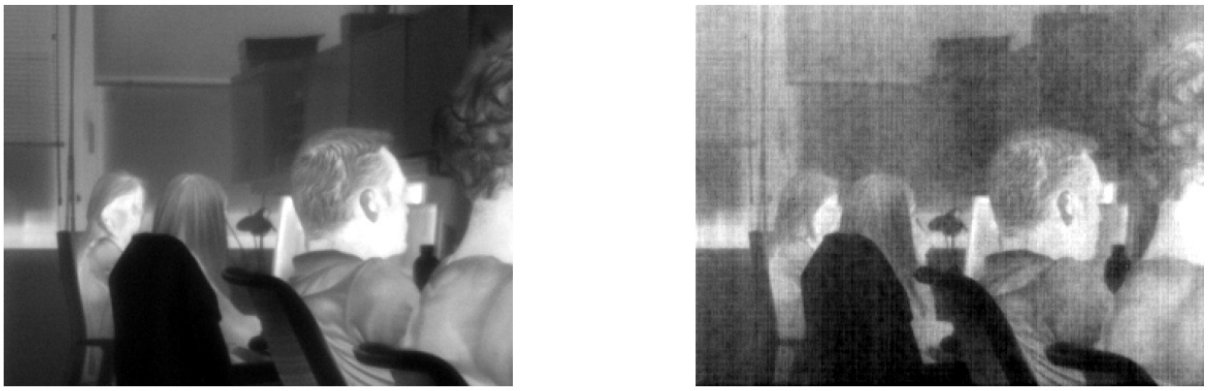
2.1.2.3 Dérive de la réponse du capteur selon l'environnement (*thermal drift*)

La correction des non-uniformités spatiales (NUC) d'une image brute r_θ délivrée par la caméra dans un environnement θ nécessite de disposer d'estimations fiables des composantes $\mathcal{R}_\theta$ et $\mathcal{D}_\theta$ (ou de $\mathcal{G}_\theta$ et $\mathcal{O}_\theta$ selon la formulation adoptée du modèle de correction). Or, ces composantes ont une propension certaine à varier au cours du temps, et l'utilisation d'une estimation obsolète peut entraîner une correction fortement dégradée.

La figure 2.13a montre la NUC d'une image brute r_θ acquise dans un environnement θ, corrigée à l'aide des paramètres $\mathcal{R}_\theta$ et $\mathcal{D}_\theta$. L'image corrigée est de bonne qualité. En

revanche, la figure 2.13b illustre la NUC d'une image brute $r_{\widetilde{\theta}}$ acquise dans un environnement différent $\widetilde{\theta} \neq \theta$, mais corrigée à l'aide des paramètres précédents $\mathcal{R}_\theta$ et $\mathcal{D}_\theta$. La correction obtenue est alors de mauvaise qualité, et c'est une conséquence directe de la variation temporelle du couple responsivité/décalage. On parle alors de *thermal drift* : la variation des conditions environnementales (principalement thermiques) entraîne une dérive de la réponse du capteur.

En pratique, il est très difficile de modéliser l'ensemble des composantes de l'environnement θ, car certaines d'entre elles sont difficilement identifiables. Bieszczad et al. [3, 44] se sont intéressés à l'impact des flux thermiques parasites (température ambiante, auto-échauffement) qui influencent la réponse du capteur. Ils ont montré que la TFPA permet de caractériser efficacement ces flux lorsque la caméra est placée dans une enceinte climatique à température et humidité stabilisées. On peut alors réduire θ à un seul paramètre, la TFPA.



(a) NUC sans *thermal drift* (b) NUC après *thermal drift*

FIGURE 2.13 – **Illustration du *thermal drift*, variabilité de la responsivité et du décalage.** La caméra *Smart IR320* est mise sous tension et acquiert une séquence d'environ cinq minutes. En début de séquence, elle acquiert deux scènes uniformes dans un environnement donné θ, ce qui permet d'estimer les images de responsivité $\mathcal{R}_\theta$ et de décalage $\mathcal{D}_\theta$ selon la procédure de calibration à deux points [1], qui sera présentée plus en détail dans la section 2.2.1.1. Ces deux images, affichées sur une échelle de niveaux de gris commune, correspondent à la correction NUC d'une image brute par le modèle affine (2.12) à partir de $\mathcal{R}_\theta$ et $\mathcal{D}_\theta$. (a) L'image brute est acquise une vingtaine de secondes après la scène uniforme : l'environnement thermique θ a peu évolué, et la NUC produit une image de bonne qualité. (b) L'image brute est acquise quatre minutes plus tard : l'environnement thermique $\widetilde{\theta}$ n'est plus le même, i.e. $\widetilde{\theta} \neq \theta$. La TFPA, paramètre environnemental dominant, a par exemple augmenté de près de 4 °C. La NUC devient alors moins efficace : des motifs colonnaires apparaissent sur l'image corrigée, signe d'une dérive de la réponse du capteur, phénomène connu sous le nom de *thermal drift*. Ce changement de conditions, principalement dû ici à l'échauffement interne de la caméra après sa mise sous tension, rend les images de responsivité et de décalage obtenues dans l'environnement initial inadéquates pour corriger les images en fin de séquence.

2.1.3 Modélisation affine à un seul paramètre (TFPA)

Pour un pixel $(x, y) \in \Omega$ donné, on considère le modèle affine direct (2.10) en restreignant l'environnement θ à la seule mesure de la TFPA en ce pixel, notée $\tau_{x,y}$. Par souci de simplification, on suppose que la TFPA est uniforme sur l'ensemble du capteur, hypothèse raisonnable dans un environnement thermique homogène autour de la matrice, bien qu'approximative en pratique. On note alors $\tau \in \mathbb{R}$ la mesure fournie par la sonde située au niveau du FPA, et l'on considère $\theta = \tau$. Le modèle se réduit ainsi à un seul paramètre, les paramètres électroniques du circuit de lecture étant supposés constants.

Soit $r_\tau \in \mathbb{R}^\Omega$ l'image brute délivrée par la caméra à une TFPA τ. Le modèle affine direct à un seul paramètre s'écrit

$$r_\tau = \mathcal{R}_\tau \, S + \mathcal{D}_\tau \,, \tag{2.15}$$

où $\mathcal{R}_\tau$ et $\mathcal{D}_\tau$ désignent respectivement les images de responsivité et de décalage associées à la TFPA τ, et S l'image de la scène (en irradiance ou en température, selon l'usage).

Le modèle inverse, permettant de retrouver S à partir de l'image brute r_τ (à condition de connaître $\mathcal{R}_\tau$ et $\mathcal{D}_\tau$), s'écrit

$$S = \frac{r_\tau - \mathcal{D}_\tau}{\mathcal{R}_\tau} \,. \tag{2.16}$$

Comme pour le modèle multi-paramétrique, ce modèle inverse peut être reformulé sous forme affine en TFPA, avec un gain $\mathcal{G}_\tau$ et un offset $\mathcal{O}_\tau$ (définis par (2.13) avec $\theta = \tau$)

$$U = \mathcal{G}_\tau \, r_\tau + \mathcal{O}_\tau \,, \tag{2.17}$$

où U est l'image corrigée dans une échelle radiométrique en DL.

2.1.4 Modèle affine avec gain unique

Certains modèles de correction supposent que l'image du gain reste constante. Par exemple, Cao et al. [10] ont observé que le gain d'un pixel varie peu avec la TFPA. On suppose donc qu'à paramètres électroniques fixés, le gain est constant, indépendamment de l'environnement de la caméra. On note $\mathcal{G}_0$ l'image de ce gain unique. Le modèle affine multi-paramétrique (2.14) devient alors

$$U = \mathcal{G}_0 \, r_\theta + \mathcal{O}_\theta \,. \tag{2.18}$$

Lorsqu'on considère $\theta = \tau$, on obtient le modèle affine à un seul paramètre avec gain unique

$$U = \mathcal{G}_0 \, r_\tau + \mathcal{O}_\tau \,. \tag{2.19}$$

2.2 État de l'art des modèles de correction

Les modèles de correction en imagerie thermique ont évolué en réponse aux avancées technologiques et aux besoins croissants des caméras dans des domaines variés. Au départ, les détecteurs photoniques refroidis étaient privilégiés, mais leur coût et encombrement ont conduit à l'émergence de détecteurs bolométriques non-refroidis. Ces derniers ont toutefois introduit des dérives thermiques, nécessitant l'utilisation d'un obturateur (en anglais

shutter) pour réactualiser périodiquement les paramètres de correction. L'introduction de traitements embarqués a permis de réduire ou d'éliminer cet usage, ouvrant ainsi la voie aux approches *shutterless*. Avant d'explorer les modèles de correction qui prédominent aujourd'hui, il est pertinent de retracer brièvement cette évolution historique.

2.2.1 Évolution des corrections thermiques, des capteurs refroidis aux méthodes *less shutter* et *shutterless*

Dans les années 1980, les détecteurs refroidis permettaient une correction affine (avec un gain et un offset) stable dans le temps [45]. Cependant, l'arrivée des détecteurs bolométriques non-refroidis dans les années 1990 a introduit des dérives thermiques, nécessitant une correction périodique par *shutter*, qui permet d'acquérir une image de référence pour ajuster régulièrement les paramètres de correction.

2.2.1.1 Méthode de la calibration à deux points

Avant d'expliquer brièvement le principe du modèle de correction *shutter*, nous présentons celui de la calibration à deux points. Il s'agit d'une méthode élémentaire et essentielle, consistant à déterminer les paramètres du modèle affine (responsivité/décalage ou gain/offset) à partir d'une paire d'images brutes acquises sur des scènes uniformes [1,4,5].

Le principe est le suivant : on applique le modèle affine (2.11) à deux images brutes $r_\theta^{(T_1)}$ et $r_\theta^{(T_2)}$, acquises sur des scènes uniformes constituées de corps noirs à des températures distinctes T_1 et T_2, dans un même environnement θ, on obtient

$$\begin{cases} r_\theta^{(T_1)} = \mathcal{R}_\theta\, T_1 + \mathcal{D}_\theta, \\ r_\theta^{(T_2)} = \mathcal{R}_\theta\, T_2 + \mathcal{D}_\theta. \end{cases} \tag{2.20}$$

Puisque $T_1 \neq T_2$, ce système admet une unique solution

$$\begin{cases} \mathcal{R}_\theta = \dfrac{r_\theta^{(T_2)} - r_\theta^{(T_1)}}{T_2 - T_1}, \\ \mathcal{D}_\theta = \dfrac{T_2\, r_\theta^{(T_1)} - T_1\, r_\theta^{(T_2)}}{T_2 - T_1}. \end{cases} \tag{2.21}$$

En utilisant la définition des composantes gain/offset donnée en (2.13), on obtient

$$\begin{cases} \mathcal{G}_\theta = \dfrac{\mu(\mathcal{R}_\theta)}{\mathcal{R}_\theta}, \\ \mathcal{O}_\theta = -\mathcal{G}_\theta\, \mathcal{D}_\theta. \end{cases} \tag{2.22}$$

Ces deux dernières formules correspondent à la calibration à deux points. Elles sont exprimées ici dans le cadre général du modèle multi-paramétrique, mais leur mise en œuvre pratique se fait le plus souvent en supposant $\theta = \tau$, c'est-à-dire en se plaçant dans le cadre à un seul paramètre, la TFPA.

2.2.1.2 Modèles de correction pour une caméra équipée d'un *shutter*

Les méthodes de correction *shutter* s'appuient sur le modèle affine avec gain et offset dépendant d'un seul paramètre (la TFPA).

Une première méthode simple considère que le gain est constant (2.19). Le modèle retenu s'écrit donc comme suit : soit r_τ l'image brute acquise par la caméra à la TFPA τ, l'image corrigée U s'exprime alors selon

$$U = \mathcal{G}_0\, r_\tau + \mathcal{O}_\tau\,. \tag{2.23}$$

L'image du gain unique $\mathcal{G}_0$ est obtenue lors d'une calibration à deux points, à partir de laquelle on déduit une image de gain via la formule (2.22). L'offset $\mathcal{O}_\tau$ est quant à lui déterminé à partir d'une image brute $r_\tau^{(T_s)}$ acquise avec le *shutter* fermé à la TFPA τ, selon

$$\mathcal{O}_\tau = -\mathcal{G}_0\, r_\tau^{(T_s)} + T_s, \tag{2.24}$$

où T_s désigne la température du *shutter*. Toutefois, cette température n'étant pas toujours mesurée, l'offset et, par conséquent, l'image corrigée sont souvent déterminés avec une erreur d'une image constante près. Pendant l'utilisation de la caméra, l'offset est mis à jour lorsque la TFPA évolue de quelques dixièmes de degré.

D'autres méthodes de correction *shutter* prennent en compte une variabilité du gain. Krupinsky et al. ont par exemple développé une approche intégrant cette variabilité à l'aide d'un refroidisseur thermoélectrique placé au niveau du FPA [44]. Pendant la phase où le *shutter* est en position fermée, la caméra acquiert plusieurs images à différentes valeurs de TFPA. Un algorithme permet ensuite de déterminer les paramètres de gain et d'offset mis à jour à partir de ces acquisitions multiples.

Le principal inconvénient de cette technologie réside dans le fait que le *shutter*, lorsqu'il est positionné devant l'optique, interrompt le flux vidéo pendant plusieurs secondes. Pour atténuer cet effet, Olbrycht et al. ont proposé l'intégration d'un *shutter* semi-opaque [46], réduisant ainsi l'impact de cette interruption sur l'observation de la scène.

2.2.1.3 Vers des stratégies *less shutter* et *shutterless*

L'usage du *shutter* permet d'obtenir des corrections satisfaisantes, mais entraîne des interruptions du flux vidéo. Pour contourner cette limitation, l'intégration de systèmes de calcul embarqués a permis de réduire l'usage du shutter (approche *less shutter*) ou de l'éliminer complètement (approche *shutterless*). Ces solutions reposent sur des méthodes *calibration-based* ou *scene-based*, ainsi que sur des stratégies hybrides combinant les deux. Aujourd'hui, ces solutions sont pleinement mises en œuvre dans les techniques de correction en imagerie thermique non-refroidie.

2.2.2 Méthodes principalement « *calibration-based* »

Les méthodes *calibration-based* considèrent la TFPA comme le principal paramètre environnemental influençant la réponse du capteur. Elles reposent sur l'acquisition de N paires d'images brutes représentant deux scènes uniformes, typiquement un corps noir chaud et un corps noir froid, dans des environnements correspondant à N valeurs mesurées de TFPA, notées $(\tau_1, \ldots, \tau_N)$. À partir de ces N paires, on détermine les paramètres du modèle affine responsivité/décalage, selon la formule de calibration à deux points (2.21), ou les paramètres gain/offset, selon (2.22), pour chacune des valeurs de TFPA $(\tau_1, \ldots, \tau_N)$.

2.2.2.1 Modélisation polynomiale en TFPA

Dans un premier temps, examinons les méthodes proposées pour déterminer les composantes de responsivité et de décalage en fonction de la TFPA. Le modèle le plus simple repose sur une relation linéaire. Par exemple, Nugent et al. modélisent, pour chaque pixel, $\mathcal{R}_\tau$ et $\mathcal{D}_\tau$ par des polynômes de degré 1 en fonction de la TFPA [7], ce qui donne

$$\forall (x,y) \in \Omega \quad \begin{cases} \mathcal{R}_\tau(x,y) = a(x,y)\,\tau + b(x,y), \\ \mathcal{D}_\tau(x,y) = c(x,y)\,\tau + d(x,y), \end{cases} \tag{2.25}$$

Les paramètres a, b, c et d sont estimés pixel par pixel par une méthode des moindres carrés, à partir des N paires responsivité/décalage issues de calibrations à deux points. Cette modélisation est valable sur une plage restreinte de températures, où la dépendance des composantes de responsivité et de décalage à la TFPA peut être approximée par une relation linéaire. Nugent et al. l'appliquent, par exemple, à des variations de TFPA de l'ordre d'une dizaine de degrés.

Cependant, cette hypothèse de linéarité ne tient plus lorsque l'on considère une gamme plus large de TFPA. Tempelhahn et al. ont ainsi estimé les images de responsivité et de décalage pour des TFPA comprises entre 5 et 50°C [8,9]. Ils observent que la responsivité suit un comportement polynomial bien modélisé par un polynôme d'ordre 2, tandis que le décalage est mieux représenté par un polynôme d'ordre 3. Budzier et al. retiennent également ces degrés, bien qu'ils ne s'appuient sur aucune donnée expérimentale [47].



(a) $\tau_0 = -20$°C (b) $\tau_0 = 38$°C (c) $\tau_0 = 66$°C (d) $\tau_0 = 80$°C

FIGURE 2.14 – **Correction d'une image brute par le modèle *shutterless polynomial*.** Données issues de la caméra *ATI-640c*. La responsivité et le décalage sont modélisés par des polynômes de degré 2 et 3, ajustés à partir de plusieurs calibrations à deux points (2.21), réalisées pour 8 valeurs de TFPA quasi uniformément réparties entre 5 et 50°C, à l'aide de paires d'images de corps noirs. L'intervalle de calibration est noté $\mathbb{I} = [5, 50\text{°C}]$. La correction d'une image brute r_{τ_0} est ensuite effectuée selon (2.16). L'affichage des images corrigées est centré sur la moyenne et couvre cinq fois l'écart-type associé. (a) τ_0 est située 25°C en deçà de $\mathbb{I}$: correction très dégradée, avec forts effets de colonne et voile basse fréquence. (b) $\tau_0 \in \mathbb{I}$: correction optimale, avec bonne restitution des détails et des textures. (c) τ_0 est située 16°C au-delà de $\mathbb{I}$: correction moyenne, avec apparition d'un voile basse fréquence et d'artefacts de colonne. (d) τ_0 est située 30°C au-delà de $\mathbb{I}$: correction fortement altérée, avec artefacts amplifiés similaires au cas (a). Plus la TFPA de l'image à corriger s'éloigne de l'intervalle de calibration $\mathbb{I}$, plus la qualité de correction se dégrade, en raison des limites de l'extrapolation polynomiale hors de la plage de référence.

L'ajustement polynomial permet de corriger efficacement les non-uniformités spatiales d'une image brute lorsque sa TFPA se situe dans l'intervalle des TFPA de calibration

que l'on appelle « plage de calibration »[5]. En revanche, lorsque la TFPA de l'image à corriger est située en dehors de cet intervalle, la correction peut être significativement détériorée (cf. figure 2.14). Malgré cette limite, cette méthode, parfois appelée *« shutterless polynomial »*, fait aujourd'hui référence.

2.2.2.2 Modèle polynomial avec ré-estimation de la mesure de TFPA

D'autres études se sont intéressées au modèle affine avec gain et offset (2.17). Cao et Tisse [10] ont, par exemple, estimé les images de gain et d'offset à partir d'images brutes uniformes acquises pour des TFPA comprises entre -5 et 65°C. Ils observent que le gain est relativement indépendant de la TFPA, tandis que l'offset suit un comportement bien modélisé par un polynôme d'ordre 2. Leur modèle s'écrit, pour tout pixel $(x, y) \in \Omega$

$$\forall (x,y) \in \Omega \quad \begin{cases} \mathcal{G}_\tau(x,y) = \mathcal{G}_0(x,y) \\ \mathcal{O}_\tau(x,y) = a(x,y)\,\tau^2 + b(x,y)\,\tau + c(x,y). \end{cases} \tag{2.26}$$

Les images des coefficients a, b, et cs sont obtenues par une méthode des moindres carrés. Cao et Tisse proposent une amélioration de leur modèle polynomial fondée sur un critère de qualité d'image, donnant lieu à une méthode mixte, à la fois *« calibration-based »* et *« scene-based »*. On suppose alors connues les images des coefficients a, b, c, et l'on considère r_{τ_0}, une image brute quelconque acquise à une TFPA τ_0. Les auteurs soulignent que cette mesure peut être imprécise, en particulier dans des environnements thermiquement instables. Ils proposent alors d'estimer une TFPA corrigée $\widehat{\tau_0}$, solution du problème suivant

$$\widehat{\tau_0} \in \underset{\tau \in \mathbb{R}}{\operatorname{argmin}} \, ||\nabla(\mathcal{G}_0\, r_{\tau_0} + a\tau^2 + b\,\tau + c)||_{2,2}, \tag{2.27}$$

où $\nabla(\cdot)$ désigne le gradient spatial estimé à l'aide d'un schéma aux différences finies. L'image corrigée U s'obtient alors par

$$U = \mathcal{G}_0\, r_{\tau_0} + \mathcal{O}_{\widehat{\tau_0}}. \tag{2.28}$$

Cette ré-estimation de la TFPA ne remet pas en cause la valeur de la TFPA elle-même, mais tient compte des variations des autres variables cachées contenues dans θ et non mesurées. Comme la puissance totale correspond à la somme des contributions de ces variables, il est possible de déterminer une TFPA équivalente reproduisant la même puissance totale, ce qui permet de compenser partiellement les effets locaux liés aux flux thermiques parasites.

Les méthodes *« calibration-based »* sont reconnues pour leur simplicité et leur précision dans la correction du bruit de non-uniformité. Elles consistent à approximer les paramètres du modèle affine par des fonctions polynomiales de la TFPA. Le degré de ces polynômes dépend de l'intervalle de TFPA couvert, sans reposer a priori sur une loi physique. Rien ne garantit que ce modèle soit exact, mais les polynômes, en raison de leurs bonnes propriétés de régularité, sont couramment utilisés pour approximer une fonction à partir d'observations.

5. Les TFPA de calibration sont en général nombreuses et réparties uniformément, ce qui permet de considérer toute TFPA de l'intervalle comme étant assez proche d'une TFPA de calibration.

2.2.3 Méthodes « scene-based »

Dans les approches « scene-based », l'objectif est d'exploiter les informations issues de la scène pour ajuster en temps réel les paramètres du modèle affine de correction. On distingue principalement deux grandes familles de méthodes. La première repose sur l'analyse statistique des images acquises sur un intervalle de temps donné. La seconde se base sur le recalage d'images successives pour estimer les paramètres du modèle affine.

Ces méthodes s'appuient sur deux hypothèses principales :

— Une dérive lente des paramètres au cours de la séquence, ce qui suggère une faible variation de l'environnement thermique.

— Une variabilité spatio-temporelle importante de la scène, impliquant du mouvement dans la scène.

La manière dont ces hypothèses sont appliquées varie selon la méthode. Certaines approches exigent un mouvement plus marqué dans la scène pour être efficaces, tandis que d'autres peuvent tolérer des scènes avec moins de mouvement tout en maintenant de bonnes performances.

Dans cette partie, nous notons $(r_1, \dots, r_N)$ un flux temporel d'images brutes, où N désigne le nombre total d'images dans le flux. La moyenne temporelle d'un tel flux est une fonction $E : (\mathbb{R}^\Omega)^N \to \mathbb{R}^\Omega$ définie par

$$E(r_1, \dots, r_N) = \frac{r_1 + \cdots + r_N}{N}, \tag{2.29}$$

et l'écart-type temporel est une fonction $\sigma : (\mathbb{R}^\Omega)^N \to \mathbb{R}^\Omega$ donnée par

$$\sigma(r_1, \dots, r_N) = \sqrt{E(r_1^2, \dots, r_N^2) - E(r_1, \dots, r_N)^2}. \tag{2.30}$$

2.2.3.1 Méthodes fondées sur les statistiques

On distingue principalement trois classes de méthodes statistiques : *constant statistics* (CS), *adjacent differential statistics* (ADS) et *least mean square* (LMS).

Constant statistics

L'hypothèse fondamentale sur laquelle reposent les méthodes CS est que, sur une séquence d'images suffisamment longue, chaque pixel du détecteur est statistiquement exposé à un flux lumineux comparable à celui des autres pixels. Autrement dit, on suppose que tous les pixels « voient » en moyenne la même scène, avec des variations similaires d'intensité dans le temps. Cette hypothèse implique que, dans l'image corrigée, les séries temporelles associées aux pixels doivent présenter une moyenne nulle et une variance unitaire [11, 13].

Formellement, soit $(r_1, ..., r_N)$ un flux temporel d'images brutes, on cherche $(S_1, ..., S_N)$ le flux d'images corrigées tel que

$$\begin{cases} r_t = \mathcal{R}\, S_t + \mathcal{D}, \quad t \in [\![1, N]\!] \\ \text{avec } E(S_1, ..., S_N) = \mathbf{0} \quad \text{et} \quad \sigma(S_1, ..., S_N) = \mathbf{1}, \end{cases} \tag{2.31}$$

La responsivité $\mathcal{R}$ et le décalage $\mathcal{D}$ sont ici des images constantes (temporellement). On déduit facilement de (2.31) que

$$\begin{cases} \mathcal{D} = E(r_1, ..r_N) \\ \mathcal{R} = \sigma(r_1, ..r_N). \end{cases} \tag{2.32}$$

Le flux d'images corrigées $(S_1, ..., S_N)$ s'obtient alors par la formule

$$S_t = \frac{r_t - E(r_1, ..r_N)}{\sigma(r_1, ..r_N)}, \quad t \in [\![1, N]\!]. \tag{2.33}$$

Pour tenir compte de la variabilité de la responsivité et du décalage au cours d'une séquence (voir *thermal drift* illustré figure 2.13), plusieurs méthodes ont été développées. Une approche fréquemment utilisée repose sur une moyenne glissante temporelle à fenêtre exponentielle [13, 48], où les moyennes et écarts-types sont remplacés par leurs versions fenêtrées. Un écueil majeur de cette méthode, commun à toutes les approches « *scene-based* », est l'apparition d'artefacts de « *ghosting* », visibles sous forme de traînées ou d'ombres dues à la rémanence de l'information de scènes antérieures. Ces artefacts se manifestent particulièrement lorsque la scène présente peu de mouvement, ce qui entraîne une faible variabilité temporelle des signaux des pixels. Cette faiblesse fausse les estimations des paramètres de correction, perturbant ainsi le signal corrigé. Pour limiter ces artefacts, Hardie *et al.* proposent de ne mettre à jour les paramètres statistiques que lorsque la variation du signal brut dépasse un certain seuil [13]. Toutefois, ce seuil est souvent trop élevé, ce qui rend l'actualisation des paramètres peu fréquente, notamment dans des scènes où le mouvement est insuffisant. Ce manque d'actualisation régulière est un problème courant, rendant la correction inefficace, comme le montrent les artefacts de « *ghosting* » visibles en figure 2.15a.

Une autre approche consiste à diviser la séquence en segments où la méthode CS peut être appliquée, puis à affiner les estimations des paramètres de responsivité et de décalage à l'aide d'un filtre de Kalman [12].

Les méthodes CS reposent sur l'hypothèse que chaque pixel observe globalement la même scène, ce qui est souvent irréaliste. Par exemple, une source chaude peut n'affecter qu'un petit groupe de pixels, générant ainsi des artefacts de « *ghosting* » visibles (figure 2.15a).

Adjacent differential statistics

La méthode ADS est proposée par Geng et al. [14]. ADS repose sur une hypothèse plus faible que CS : un pixel et son voisin voient en moyenne la même chose. Ainsi, les signaux temporels entre deux pixels voisins doivent avoir la même moyenne. En conséquence, les gradients spatiaux de l'image corrigée doivent présenter une moyenne temporelle égale à 0.

Les auteurs simplifient le modèle affine en supposant que les non-uniformités sont uniquement dues à l'offset, avec un gain constant et unitaire

$$U_t = r_t + \mathcal{O}, \quad t \in [\![1, N]\!], \tag{2.34}$$

où $(r_1, ..., r_N)$ représente les flux d'images brutes, $(U_1, ..., U_N)$ le flux corrigé, N le nombre d'images, et $\mathcal{O}$ l'offset.

Si on note les opérateurs de gradient horizontal et vertical $\nabla_x(\cdot)$ et $\nabla_y(\cdot)$ (comme ils sont définis dans (4.32)), on obtient

$$\forall t \in [\![1, N]\!] \quad \begin{cases} \nabla_x U_t = \nabla_x r_t + \nabla_x \mathcal{O}, \\ \nabla_y U_t = \nabla_y r_t + \nabla_y \mathcal{O}. \end{cases} \tag{2.35}$$

L'hypothèse d'ADS permet d'écrire

$$\begin{cases} E(\nabla_x U_1, ..., \nabla_x U_N) = \mathbf{0}, \\ E(\nabla_y U_1, ..., \nabla_y U_N) = \mathbf{0}. \end{cases} \tag{2.36}$$

En combinant (2.35) et (2.36), on obtient

$$\begin{cases} \nabla_x \mathcal{O} = -E(\nabla_x r_1, ..., \nabla_x r_N), \\ \nabla_y \mathcal{O} = -E(\nabla_y r_1, ..., \nabla_y r_N). \end{cases} \tag{2.37}$$

Pour estimer le gradient de l'offset, on calcule une moyenne temporelle du signal $(\nabla r_1(x,y), \ldots, \nabla r_N(x,y))$ en tout pixel $(x,y) \in \Omega$. Ce signal peut contenir des perturbations ponctuelles non représentatives de l'offset, notamment lorsque le pixel (x,y) se situe, à un moment donné, sur le contour d'un objet. Dans ce cas, le gradient reflète une discontinuité liée à la scène plutôt qu'une variation réelle de l'offset, ce qui peut entraîner une valeur anormalement élevée. Pour filtrer ces valeurs aberrantes, Geng et al. proposent un estimateur robuste fondé sur la médiane des écarts absolus. Cependant, après retrait de ces valeurs extrêmes, le champ de gradient moyen associé à $\mathcal{O}$ peut perdre sa structure conservative. Un champ vectoriel est dit *conservatif* s'il existe une fonction scalaire dont il constitue le gradient (on rappelle qu'une image est une fonction scalaire $\Omega \to \mathbb{R}$). Lorsque cette propriété n'est plus satisfaite, les composantes horizontale et verticale deviennent incompatibles entre elles, et le système (2.37) peut alors ne pas admettre de solution exacte.

Par ailleurs, afin de tenir compte de la variabilité temporelle de l'offset, les équations (2.37) sont évaluées à l'aide d'une moyenne glissante à fenêtre exponentielle.

L'offset estimé $\widehat{\mathcal{O}}$ est reconstruit par intégration spatiale du champ de gradient $\nabla\mathcal{O}$, calculé selon (2.37). La valeur initiale est fixée en $(1,1)$ par $\widehat{\mathcal{O}}(1,1) = 0$. Ensuite, l'intégration se propage le long de la première colonne et de la première ligne

$$\widehat{\mathcal{O}}(1, y+1) = \widehat{\mathcal{O}}(1,y) + \nabla_y \mathcal{O}(1,y), \quad 1 \le y \le N_l - 1,$$

$$\widehat{\mathcal{O}}(x+1, 1) = \widehat{\mathcal{O}}(x,1) + \nabla_x \mathcal{O}(x,1), \quad 1 \le x \le N_c - 1.$$

Une fois ces bords définis, le reste de l'image est reconstruit ligne par ligne avec la récurrence suivante

$$\widehat{\mathcal{O}}(x+1, y+1) = \frac{1}{2}\left[\widehat{\mathcal{O}}(x, y+1) + \nabla_x \mathcal{O}(x, y+1) + \widehat{\mathcal{O}}(x+1, y) + \nabla_y \mathcal{O}(x+1, y)\right],$$

pour $1 \le x \le N_c - 1$ et $1 \le y \le N_l - 1$.

L'idée est donc d'exploiter la demi-somme des deux champs de gradients pour atténuer les erreurs spécifiques à chacun. La solution $\widehat{\mathcal{O}}$ ainsi obtenue ne vérifie pas forcément le système (2.37), mais constitue un compromis visant à reconstruire l'offset à partir d'un champ moyenné, moins sensible aux erreurs directionnelles. Toutefois, l'intégration de ce champ peut engendrer des décalages locaux, qui s'accumulent et se traduisent par des artefacts de basse fréquence, perceptibles sous forme de motifs de « *ghosting* » dans l'image reconstruite. Malgré ces défauts, la méthode ADS fournit des résultats visuellement plus satisfaisants que ceux de la méthode CS (cf. figures 2.15b–c).

Least mean square

En ce qui concerne les méthodes LMS, l'idée est de corriger une image bruitée de façon à ce que la correction estimée par le modèle affine soit proche, au sens des moindres carrés, d'une image cible dépourvue de non-uniformité spatiale. Cela nécessite une hypothèse de mouvement de scène important et une hypothèse de faible variabilité temporelle du gain et de l'offset.

Soit $(r_1, ..., r_N)$ une séquence d'images brutes. Il faut, pour chaque image brute r_t ($t \in [\![1, N]\!]$), définir une image cible associée T_t. Scribner et al. considèrent l'image brute à laquelle ils appliquent un filtre de moyenne spatio-temporelle entre pixels voisins [49], tandis que Hardie et al. utilisent un filtre spatial gaussien [13]. Ces choix sont valides lorsque le bruit est spatialement i.i.d., ce qui n'est pas le cas en imagerie thermique non-refroidie, où le bruit (FPN) présente des non-uniformités fortement corrélées (cf. figure 2.11c).

Une première correction « grossière », par exemple avec une image de gain et une image d'offset, permet de construire une séquence d'images cibles $(T_1, ..., T_N)$ [13].

On suppose dans un premier temps que le gain $\mathcal{G}$ et l'offset $\mathcal{O}$ ne varient pas au cours de la séquence. L'image corrigée U_t s'écrit alors

$$U_t = \mathcal{G}\, r_t + \mathcal{O}, \quad t \in [\![1, N]\!]. \tag{2.38}$$

On note ε_t l'erreur entre l'image corrigée et l'image cible. Pour chaque pixel $(x, y) \in \Omega$, on obtient

$$\forall (x,y) \in \Omega \begin{cases} \mathcal{G}(x,y)\, r_1(x,y) + \mathcal{O}(x,y) - T_1(x,y) = \varepsilon_1(x,y) \\ \vdots \\ \mathcal{G}(x,y)\, r_N(x,y) + \mathcal{O}(x,y) - T_N(x,y) = \varepsilon_N(x,y). \end{cases} \tag{2.39}$$

Les observations sont les T_t, expliquées par une régression affine des r_t. On suppose que les erreurs $\varepsilon_t(x, y)$ sont centrées i.i.d., ce qui est raisonnable si la scène varie suffisamment.

Il est alors possible d'estimer $\widehat{\mathcal{G}}_t(x, y)$ et $\widehat{\mathcal{O}}_t(x, y)$ en minimisant la somme des carrés des erreurs. L'approche généralement adoptée dans la littérature s'inspire de l'algorithme stochastique de Robbins-Monro [17] pour résoudre le problème (2.39). Le schéma itératif correspondant s'écrit alors

$$\forall t \geq 1 \quad \begin{cases} \widehat{\mathcal{G}}_{t+1}(x,y) = \widehat{\mathcal{G}}_t(x,y) - \alpha_t(x,y)\, r_t(x,y) \left(\widehat{\mathcal{G}}_t(x,y)\, r_t(x,y) + \widehat{\mathcal{O}}_t(x,y) - T_t(x,y) \right) \\ \widehat{\mathcal{O}}_{t+1}(x,y) = \widehat{\mathcal{O}}_t(x,y) - \alpha_t(x,y) \left(\widehat{\mathcal{G}}_t(x,y)\, r_t(x,y) + \widehat{\mathcal{O}}_t(x,y) - T_t(x,y) \right), \end{cases} \tag{2.40}$$

où $\alpha_t(x, y)$ est le pas de descente. On initialise les images de gain et d'offset à 1 et 0. Si le pas est bien réglé (e.g. $\alpha_t = c\, t^{-\gamma}$ avec $0.5 < \gamma \leq 1$), la convergence est assurée [50].

Ce schéma est utile pour suivre des variations lentes du gain et de l'offset (*thermal drift*), à condition qu'elles restent faibles. Torres et al. [51] ont proposé une version spatialement adaptative où $\alpha_t(x, y)$ est inversement proportionnel à la variance locale de l'image cible, plus faible dans les zones à forte variance.

Pour être utilisables sur des capteurs non-refroidis, les méthodes LMS nécessitent donc une étape initiale de correction. Malgré cela, elles peuvent produire des artefacts de « *ghosting* », comme visible figure 2.15c.

2.2.3.2 Méthodes fondées sur le recalage d'images

Les méthodes qui reposent sur le recalage d'images s'inspirent des approches LMS en exploitant l'information extraite du recalage entre deux images successives pour définir l'image cible.

L'hypothèse de base est que la radiation émise par chaque élément de la scène reste approximativement constante entre deux images consécutives d'un même flux vidéo [15,16, 18]. Un exemple représentatif est proposé par Zuo et al. [16], qui modélisent le mouvement de la scène par une simple translation (sans rotation ni changement d'échelle). Une zone de recouvrement $\widetilde{\Omega_t} \subset \Omega$ est définie, et les auteurs posent

$$\forall t \in [\![2, N]\!],\ \forall (x,y) \in \widetilde{\Omega_t},\ \exists (d_x, d_y) \in \mathbb{R}^2, \quad r_t(x,y) = r_{t-1}(x - d_x, y - d_y). \tag{2.41}$$

Le vecteur de translation (d_x, d_y) est estimé à partir de la transformée de Fourier spatiale inverse du spectre de puissance croisée (*cross-power spectrum*) entre les images r_t et r_{t-1}, cette transformée atteignant un maximum en (d_x, d_y). Pour simplifier, Zuo et al. supposent que $(d_x, d_y) \in \mathbb{Z}^2$, mais une précision *sub-pixellique* est possible [52].

Une fois ce déplacement connu, la technique LMS est appliquée (cf. section 2.2.3.1), en définissant l'image cible T_t par

$$\begin{cases} \forall (x,y) \in \widetilde{\Omega_t}, & T_t(x,y) = \widehat{\mathcal{G}}_t(x - d_x, y - d_y)\, r_{t-1}(x - d_x, y - d_y) + \widehat{\mathcal{O}}_t(x - d_x, y - d_y) \\ \forall (x,y) \notin \widetilde{\Omega_t}, & T_t(x,y) = 0 \quad \text{(ou toute autre valeur arbitraire)} \end{cases} \tag{2.42}$$

L'image cible est ainsi obtenue par translation de l'estimation courante de la correction appliquée à r_{t-1}. Le schéma d'optimisation reprend celui des méthodes LMS (2.40), réécrit ici

$$t \geq 1 \quad \begin{cases} \widehat{\mathcal{G}}_{t+1}(x,y) = \widehat{\mathcal{G}}_t(x,y) - \alpha_t(x,y)\, r_t(x,y) \left(\widehat{\mathcal{G}}_t(x,y)\, r_t(x,y) + \widehat{\mathcal{O}}_t(x,y) - T_t(x,y) \right) \\ \widehat{\mathcal{O}}_{t+1}(x,y) = \widehat{\mathcal{O}}_t(x,y) - \alpha_t(x,y) \left(\widehat{\mathcal{G}}_t(x,y)\, r_t(x,y) + \widehat{\mathcal{O}}_t(x,y) - T_t(x,y) \right) \end{cases} \tag{2.43}$$

avec $\alpha_t(x,y) = 0$ si $(x,y) \notin \widetilde{\Omega_t}$, et initialisation à $\widehat{\mathcal{G}}_2(x,y) = 1$, $\widehat{\mathcal{O}}_2(x,y) = 0$ pour tout $(x,y) \in \Omega$.

Malgré leur intérêt, ces approches présentent des limitations importantes. En présence d'un rapport signal sur bruit trop faible, l'estimation du déplacement peut devenir instable ou erronée, menant à des résultats aberrants.

2.2.4 Méthodes « mixtes »

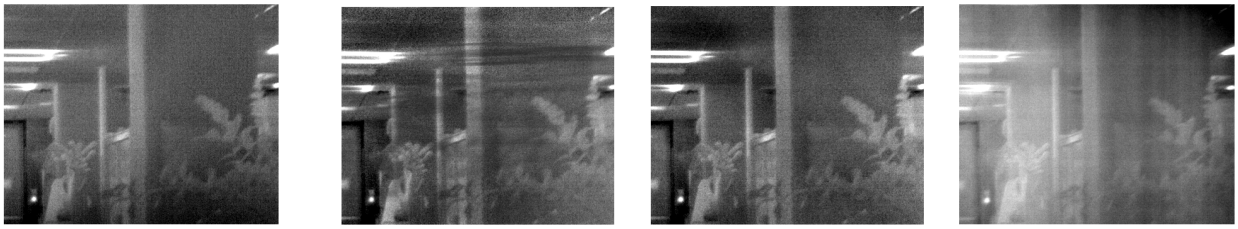
2.2.4.1 Modèle de correction avec une ou deux références

Une famille de méthodes « mixtes » est couramment utilisée en industrie. Elle consiste à déterminer la correction à partir d'images de calibration, selon un critère de qualité d'image [19].

Le modèle de correction retenu est un modèle affine à gain constant $\mathcal{G}_0$, qui s'écrit de manière non paramétrique sous la forme

$$U = \mathcal{G}_0(r - \mathcal{D}), \tag{2.44}$$

où r désigne l'image brute courante délivrée par la caméra, U l'image corrigée, et $\mathcal{D}$ l'image de décalage.



(a) *Shutterless polynomial* (b) CS (c) ADS (d) LMS

FIGURE 2.15 – **Comparatif de quelques modèles de correction de l'état de l'art. Illustration du phénomène de « *ghosting* ».** Les images proviennent d'une séquence avec du mouvement de scène acquise par la caméra *ATI-640b*. L'affichage des images corrigées est centré sur la moyenne et couvre cinq fois l'écart-type associé. (a) L'image corrigée à l'aide du modèle *shutterless polynomial*, une méthode « *calibration-based* », présente une restitution particulièrement fidèle des détails fins et des textures, sans artefact visible lié à la scène. Cela en fait une méthode de référence en matière de correction, malgré l'inconvénient majeur d'exiger une phase de calibration longue et coûteuse. (b) La méthode CS, dans sa version à moyenne glissante exponentielle avec seuil proposée par Hardie et al. [13], produit une image fortement dégradée par des effets de « *ghosting* », visibles sous forme de traînées horizontales marquées. Ces artefacts résultent de la mémoire de scène contenue dans les images précédentes, qui influence de manière incorrecte la correction des pixels dans les nouvelles images. (c) La méthode ADS offre une correction a priori satisfaisante, mais présente également des artefacts de très basse fréquence, localisés notamment en haut à gauche et en bas à droite de l'image. Ces défauts sont également attribués à des « *ghosting artifacts* » d'origine scénique. (d) Enfin, la méthode LMS nécessite une pré-correction pour l'imagerie thermique non-refroidie. Ici, un *thermal drift* résiduel (cf. figure 2.13b) impose un filtrage intense des hautes fréquences, surtout latérales, pour éviter qu'il n'apparaisse dans l'image cible. Ce traitement reste toutefois insuffisant : des artefacts de basse et moyenne fréquence, principalement latéraux, se manifestent dans l'image corrigée, compromettant assez nettement la qualité de la correction. En dépit du mouvement de scène, des effets de « *ghosting* » persistent avec les méthodes CS, ADS et LMS, confirmant que le modèle *shutterless polynomial* demeure la référence en matière de correction.

2.2.4.1.1 Modèle de correction avec une seule référence

Une première version, relativement simple, repose sur l'acquisition d'une paire d'images de corps noir chaud et froid, obtenue à une même TFPA τ_0, ce qui signifie qu'il n'existe qu'une seule référence liée aux conditions thermiques ambiantes. À partir des formules de calibration à deux points (2.21) et (2.22), on déduit une image de gain $\mathcal{G}_0$ ainsi qu'une image de décalage $\mathcal{D}_0$.

Ce décalage $\mathcal{D}_0$ est considéré comme la réalisation du motif du FPN à la TFPA τ_0. Le FPN présentant de très forts effets de colonne (cf. figure 2.11c), l'idée est de décomposer $\mathcal{D}_0$ en la somme d'une composante colonnaire C_0, obtenue par moyennage de chaque colonne de $\mathcal{D}_0$, et d'un résidu appelé « dispersion », noté D_0, image d'un bruit moins structuré que le FPN, dépourvue d'effet colonnaire mais présentant des structures basse fréquence avec de légers effets de ligne

$$\mathcal{D}_0 = C_0 + D_0.$$

Pour déterminer l'image corrigée U dans (2.44), l'idée est d'estimer l'image de décalage

$\mathcal{D}$ sous la forme

$$\mathcal{D} = \widehat{\alpha}\, C_0 + \widehat{\beta}\, D_0 + \widehat{\gamma},$$

où les coefficients réels $\widehat{\alpha}$, $\widehat{\beta}$ et $\widehat{\gamma}$ sont déterminés en résolvant successivement

$$\begin{cases} \widehat{\beta} \in \underset{\beta\in\mathbb{R}}{\operatorname{argmin}}\, F\Big(r - \beta D_0\Big), & \text{(2.45a)} \\ \widehat{\alpha} \in \underset{\alpha\in\mathbb{R}}{\operatorname{argmin}}\, G\Big(r - \widehat{\beta} D_0 - \alpha C_0\Big), & \text{(2.45b)} \\ \widehat{\gamma} \in \underset{\gamma\in\mathbb{R}}{\operatorname{argmin}}\, H\Big(\mathcal{G}_0(r - \widehat{\beta} D_0 - \widehat{\alpha} C_0 - \gamma)\Big), & \text{(2.45c)} \end{cases}$$

où les fonctions $F(\cdot)$, $G(\cdot)$ et $H(\cdot)$ désignent des métriques de qualité d'image qui peuvent être choisies comme suit :

— $F(\cdot)$ mesure une norme du gradient spatial, par exemple la variation totale si l'on utilise la norme ℓ^1. Le résidu $r - \widehat{\beta} D_0$ issu de (2.45a) correspond alors à une image brute partiellement corrigée, débarrassée de la composante « dispersion » du FPN.

— $G(\cdot)$ correspond à la norme euclidienne d'un filtre passe-haut horizontal. Le résidu $r - \widehat{\beta} D_0 - \widehat{\alpha} C_0$ issu de (2.45b) représente l'image brute débarrassée des composantes colonnaires et de « dispersion » du FPN, dans laquelle on espère retrouver uniquement les informations de scène.

— $H(\cdot)$, définie comme la variance de l'image, est utilisée dans (2.45c) pour corriger les variations résiduelles de réponse des pixels à la scène, dues à leur sensibilité optoélectronique individuelle. L'estimation du paramètre $\widehat{\gamma}$ permet de compenser ces écarts de sensibilité sous forme d'un offset scalaire, appliqué avant multiplication par le gain $\mathcal{G}_0$. Il s'agit donc d'une correction de gain, qui permet de restituer les structures de scène de manière cohérente dans l'image corrigée.

2.2.4.1.2 Modèle de correction avec deux références

Ce modèle est une extension du précédent : on dispose en plus d'une image brute de scène uniforme $r_1^{(T)}$, acquise à une TFPA $\tau_1 \neq \tau_0$, où T désigne une température de scène choisie de manière quelconque. On détermine alors une image de décalage $\mathcal{D}_1$ selon le modèle de correction à une seule référence, ce qui conduit à

$$\mathcal{D}_1 = \widehat{\alpha_1}\, C_0 + \widehat{\beta_1}\, D_0 + \widehat{\gamma_1},$$

où les coefficients $\widehat{\alpha_1}$, $\widehat{\beta_1}$, $\widehat{\gamma_1}$ sont obtenus en résolvant le système (2.45) en utilisant l'image brute $r_1^{(T)}$.

On définit alors un résidu R', calculé avant correction du gain, par

$$R' = r_1^{(T)} - \widehat{\alpha_1}\, C_0 - \widehat{\beta_1}\, D_0.$$

Ce résidu R' représente la composante de « thermal drift » (cf. figure 2.13b) de la correction du modèle affine entre les TFPA des deux références τ_0 et τ_1. À l'instar de la décomposition du décalage de la première référence, ce résidu est décomposé en une composante colonnaire C' et une composante « dispersion » D', telles que

$$R' = C' + D'.$$

Soit r l'image brute délivrée par la caméra, l'image corrigée U s'obtient en résolvant successivement

$$\begin{cases} \widehat{\beta} \in \underset{\beta\in\mathbb{R}}{\operatorname{argmin}} F\Big(r - \beta D_0\Big), & \text{(2.46a)} \\ \widehat{\alpha} \in \underset{\alpha\in\mathbb{R}}{\operatorname{argmin}} G\Big(r - \widehat{\beta} D_0 - \alpha C_0\Big), & \text{(2.46b)} \\ \widehat{\beta}' \in \underset{\beta\in\mathbb{R}}{\operatorname{argmin}} F\Big(r - \widehat{\beta} D_0 - \widehat{\alpha} C_0 - \beta D'\Big), & \text{(2.46c)} \\ \widehat{\alpha}' \in \underset{\alpha\in\mathbb{R}}{\operatorname{argmin}} G\Big(\mathcal{G}_0(r - \widehat{\beta} D_0 - \widehat{\alpha} C_0 - \widehat{\beta}' D' - \alpha C')\Big), & \text{(2.46d)} \\ \widehat{\gamma} \in \underset{\gamma\in\mathbb{R}}{\operatorname{argmin}} H\Big(\mathcal{G}_0(r - \widehat{\beta} D_0 - \widehat{\alpha} C_0 - \widehat{\beta}' D' - \widehat{\alpha}' C' - \gamma)\Big), & \text{(2.46e)} \\ U = \mathcal{G}_0(r - \widehat{\beta} D_0 - \widehat{\alpha} C_0 - \widehat{\beta}' D' - \widehat{\alpha}' C' - \widehat{\gamma}). & \text{(2.46f)} \end{cases}$$

Les critères de qualité d'image s'appliquent normalement à une image d'irradiance réelle. Or ici, l'optimisation des coefficients α et β s'effectue sur une expression qui n'est pas strictement une image physique, car elle ne tient pas compte du gain $\mathcal{G}_0$, c'est-à-dire des sensibilités individuelles des pixels. Toutefois, comme ce gain est spatialement lisse et proche de l'unité (voir figure 4.1a), cette approximation reste acceptable dans le cadre de l'application visée.

Cette famille de méthodes « mixtes », de nature non paramétrique, s'avère plus robuste que la méthode *shutterless polynomial* face aux changements d'environnement, notamment lorsque l'environnement de calibration est très différent de celui de l'image courante à corriger, comme c'est le cas en extrapolation de TFPA (voir figure 2.14a, b, c et d). En effet, la modélisation polynomiale peut conduire à des valeurs de correction ne permettant pas une compensation satisfaisante. Cependant, cette méthode peut produire une qualité visuelle légèrement inférieure lorsque, par exemple, la TFPA de l'image courante s'écarte significativement de celles associées aux images de référence.

2.2.4.2 Correction d'une image à partir d'une famille de scenes uniformes

Une méthode « mixte » a été développée par Liu et al. [20]. Elle vise à corriger une image brute à partir d'une famille de scènes uniformes en introduisant un critère appliqué à l'image corrigée. La calibration repose sur un ensemble de N images brutes de scènes uniformes acquises pour différentes valeurs de TFPA $\tau_1, \ldots, \tau_N$, notées $r^{(1)}, \ldots, r^{(N)}$. Le modèle de correction adopté est affine et s'écrit

$$V = \mathcal{G}_0\,(r - \mathcal{D}), \tag{2.47}$$

où r désigne l'image brute, V l'image corrigée, $\mathcal{G}_0$ l'image de gain supposée constante et $\mathcal{D}$ l'image de décalage à estimer. Les auteurs font l'hypothèse que cette image de décalage peut être représentée comme une combinaison affine des images de référence $r^{(i)}$

$$\mathcal{D} = c_0 + \sum_{i=1}^{N} c_i\, r^{(i)}, \tag{2.48}$$

où $c_0, c_1, \ldots, c_N$ sont des coefficients réels à déterminer.

À l'image corrigée V on associe une image $\widetilde{V}$ définie par

$$\forall (x,y) \in \Omega, \quad \widetilde{V}(x,y) = \frac{1}{\mathrm{card}(\mathscr{W}(x,y))} \sum_{(i,j)\in\mathscr{W}(x,y)} V(i,j), \tag{2.49}$$

où $\mathscr{W}(x,y)$ désigne une fenêtre de pixels voisins autour de (x,y). Ainsi, $\widetilde{V}$ correspond à une version lissée spatialement de l'image corrigée V, obtenue en effectuant une moyenne locale sur chaque voisinage.

Les auteurs proposent d'estimer les coefficients $(\hat{c}_0, \ldots, \hat{c}_N)$ en résolvant

$$(\hat{c}_0, \ldots, \hat{c}_N) = \underset{c_0,\ldots,c_N}{\mathrm{argmin}} \, \|V - \widetilde{V}\|^2, \tag{2.50}$$

où V est l'image corrigée et $\widetilde{V}$ sa version lissée par moyenne locale. La minimisation pénalise les écarts entre V et $\widetilde{V}$, ce qui revient à imposer que V soit spatialement lisse. C'est précisément pour exploiter ce critère que les auteurs introduisent une zone uniforme Λ préalablement « marquée », dans laquelle l'image corrigée est supposée être uniforme donc lisse. Dans la zone uniforme Λ, cette approche permet de déterminer la combinaison affine appropriée d'images de scène uniformes afin de supprimer le bruit fixe (FPN) présent dans l'image brute et d'obtenir une image lisse conforme à l'objectif de produire une zone uniforme.

Dans la résolution proposée par les auteurs, le problème est traité numériquement en considérant le paramètre TFPA comme une métadonnée. Pour corriger une image acquise avec une TFPA τ, ils n'utilisent que les quatre images de scène uniforme issues de la calibration dont la TFPA est la plus proche de τ.

Cette méthode présente l'avantage de corriger une image à partir d'une famille de données de calibration en considérant que le décalage recherché appartient à l'espace affine des images brutes de scène uniforme acquises lors de la calibration. Le critère de régularité appliqué à l'image corrigée, visant à obtenir une image la plus lisse possible sur la zone uniforme Λ préalablement marquée, est pertinent. Il s'agit d'une méthode mixte, non paramétrique, même si, pour la résolution, les auteurs proposent de considérer la TFPA comme métadonnée.

Cependant, cette approche présente deux limites importantes. La première concerne la zone uniforme Λ, supposée préalablement « marquée », sans que les auteurs ne précisent comment cette information est obtenue. Cela constitue une contrainte pratique, car il faut connaître à l'avance une région supposée uniforme dans l'image à corriger. Même si l'on imaginait une méthode automatique pour détecter une telle zone, rien ne garantit qu'elle existe dans toutes les images ni qu'elle soit de taille suffisante pour contraindre correctement l'estimation des coefficients. Une zone trop petite rend le système sous-déterminé et peut conduire à une correction correcte dans Λ mais arbitraire dans le reste de l'image. La seconde limite est que, dans leur résolution numérique, les auteurs ne considèrent que quatre images de calibration dont la TFPA est la plus proche de celle de l'image à corriger. Ce choix leur permet de s'affranchir des problèmes de résolution liés à une famille plus large d'images. Ces difficultés, liées à l'utilisation d'un nombre plus important d'images de calibration, seront analysées et traitées dans une future contribution.

Dans une contribution ultérieure, nous proposerons une approche inspirée du modèle des auteurs. Contrairement à leur méthode, nous n'aurons pas besoin d'une zone Λ préalablement « marquée ». Nous imposerons un critère de qualité de l'image reconstruite

différent et proposerons une méthode de résolution adaptée, plus complète que celle des auteurs. En particulier, nous pourrons exploiter l'ensemble des images de calibration plutôt que de nous limiter à quatre.

2.2.5 Méthodes de correction par apprentissage avec réseaux de neurones

Une image corrigée peut être vue comme une image brute débarrassée du bruit fixe de type FPN. La plupart des méthodes basées sur les réseaux de neurones pour corriger ce bruit reposent sur un apprentissage supervisé. Quelques travaux explorent néanmoins des approches dites « non-supervisées », comme celle proposée par Torres [53]. Il s'agit d'un réseau relativement simple, inspiré de la structure de la rétine, et fondé sur la méthode LMS présentée en sous-section 2.2.3.1. Dans ce cadre, chaque neurone représente un pixel du capteur, et applique une transformation affine à l'intensité mesurée pour produire la valeur corrigée. Les paramètres de gain et d'offset sont ajustés de manière adaptative à l'aide de la méthode LMS.

Les approches supervisées de correction du FPN reposent généralement sur des réseaux neuronaux convolutifs profonds (CNN, pour *convolutional neural network*), entraînés à partir de paires d'images brutes et d'images corrigées. Cela suppose de connaître le FPN réel du capteur, ce qui est rarement le cas. Pour contourner cette difficulté, la plupart des travaux s'appuient sur des données synthétiques obtenues via des simulateurs de capteurs infrarouges [54]. Ceux-ci modélisent les principales composantes du FPN (colonnes, dérives thermiques, variations de gain, etc.), soit de manière générique, soit à partir d'estimations obtenues sur des séquences réelles corrigées.

Cao et al. [55] proposent par exemple un réseau profond pour atténuer les structures colonnaires, entraîné sur des images artificiellement bruitées. De même, Guan et al. [56] introduisent une approche *scene-based* utilisant un CNN, mais également fondée sur un FPN simulé. Ces méthodes dépendent donc fortement de la qualité de la simulation, ce qui limite leur transférabilité à d'autres capteurs ou environnements.

En résumé, les approches supervisées fondées sur les réseaux de neurones présentent de bonnes performances, mais reposent sur des paires d'images brute et corrigée, le plus souvent obtenues par simulation. Comme le FPN varie en fonction des capteurs et de l'environnement de la caméra (température, humidité, orientation, etc.), cette dépendance aux données simulées limite la généricité des modèles : un réseau entraîné sur un capteur donné est rarement transférable tel quel à un autre. Pour pallier cette limitation, des méthodes hybrides ou semi-supervisées ont émergé. Celles-ci combinent un apprentissage supervisé initial à une phase d'adaptation auto-supervisée sur les données réelles du capteur cible, réduisant ainsi la dépendance aux données simulées. Une contribution notable est celle de Li *et al.* [57], qui propose un modèle capable d'estimer la carte de FPN et de s'ajuster progressivement aux conditions de l'environnement de la caméra.

3 Conclusion

Une scène est composée d'objets qui émettent de la radiation thermique en fonction de leur température et de leur émissivité. Cette radiation est captée par la caméra thermique à travers son système optique, qui la focalise sur un détecteur. Le détecteur convertit cette radiation en un signal électrique, produisant ainsi une image brute. Cependant, cette

image brute est dégradée par un bruit non uniforme spatialement (FPN). Un modèle de correction affine est couramment utilisé, et l'objectif des méthodes de correction est d'estimer les deux composantes de ce modèle afin de réduire les effets du FPN.

Dans la littérature, on distingue deux grandes classes de méthodes de correction. Les méthodes *« calibration-based »* reposent sur une modélisation polynomiale en TFPA des deux composantes du modèle affine, dérivées à partir d'images brutes de scènes uniformes, avec plusieurs TFPA considérées lors de la calibration. Ces méthodes sont efficaces, mais en extrapolation de TFPA, lorsque le TFPA de l'image courante est en dehors des plages de calibration, la correction peut être fortement dégradée. Un modèle de méthode « mixte » s'appuyant sur la modélisation polynomiale réestime la TFPA de l'image courante de manière à obtenir une image de meilleure qualité et peut s'avérer concluant lorsque la caméra sera soumise à des conditions très changeantes de l'environnement thermique.

Les méthodes *« scene-based »* estiment les deux composantes à partir des informations de scène. Elles se fondent sur des hypothèses relativement fortes (scène en mouvement, variabilité temporelle de l'environnement thermique relativement faible), qui, en pratique, ne sont pas toujours vérifiées, et des effets de *« ghosting »* apparaissent souvent. En outre, les méthodes *« scene-based »* nécessitent un réglage parfois assez fastidieux de divers paramètres, qui ne seront pas toujours adaptés à tous les types de scène.

Les méthodes « mixtes » combinent les informations de calibration avec celles de la scène observée pour estimer le décalage courant. Une première approche s'appuie sur un maximum de deux images de calibration. Elle est relativement robuste aux changements d'environnement thermique, mais le nombre limité de références de calibration peut limiter la qualité de la correction lorsque l'environnement de l'image à corriger s'écarte de celui associé aux deux références. Une autre méthode repose sur la détermination d'une combinaison affine d'images de calibration uniformes afin de déterminer le décalage. Pour cela, il est nécessaire de disposer d'une zone uniforme préalablement identifiée, sur laquelle les coefficients de la combinaison sont optimisés pour rendre l'image corrigée aussi lisse que possible. La principale limite de cette approche réside dans la difficulté à identifier de telles zones uniformes dans toutes les images, ce qui restreint la généralisation de la méthode.

En termes de qualité de correction, la méthode *shutterless polynomial* fait référence. Cette approche *« calibration-based »* permet de corriger une image brute sans formuler d'hypothèse sur la scène observée. Elle nécessite toutefois une série d'images de calibration acquises pour une large gamme de températures du plan focal (TFPA). La correction d'une image correspondant à une TFPA extrapolée en dehors de la plage de calibration peut en effet être fortement dégradée par des artefacts de modélisation.

Dans les deux chapitres suivants, nous présentons les principales contributions de cette thèse. Le premier est consacré à la modélisation du capteur et à l'analyse des données de calibration, tandis que le second porte sur la proposition de modèles de correction appliqués aux images issues d'une caméra thermique non-refroidie.

Chapitre 3

Modélisation mathématique du capteur et observation statistique des données de calibration

Nous nous plaçons ici dans le cadre des modèles de correction *« calibration-based »*. La correction d'une image brute délivrée par une caméra thermique dans un environnement θ se fait à partir du modèle affine multi-paramétrique (2.12). Afin d'établir une estimation des composantes responsivité $\mathcal{R}_\theta$ et décalage $\mathcal{D}_\theta$ de ce modèle[1], une calibration préalable de chaque pixel est nécessaire.

L'environnement θ est composé de plusieurs paramètres :

— La température de chaque pixel est le paramètre le plus influent. Une sonde de température proche du capteur en donne une mesure approximative, commune à tous les pixels. Cette mesure est appelée TFPA.

— Les paramètres électroniques du circuit de lecture sont également très influents et contrôlables. Au cours de la calibration, ces paramètres sont fixés.

— Les autres paramètres sont plus problématiques, car non mesurables voire non identifiables (par exemple, les flux thermiques parasites provenant du boîtier, les flux thermiques parasites liés à la non-homogénéité de l'environnement thermique, ou encore la dérive temporelle des mesures du circuit de lecture).

Une calibration consiste à mesurer le signal numérique brut de chaque pixel en fonction de la variation contrôlée et mesurée de certaines des composantes de l'environnement θ. En pratique, une calibration usuelle consiste à ne faire varier que la TFPA. On place par exemple la caméra devant une scène uniforme (par exemple, un corps noir) dans une étuve où la température ambiante est stabilisée avant chaque mesure.

Dans ce chapitre, nous faisons l'hypothèse communément admise dans la littérature [3, 44] que la réponse du capteur en un pixel donné ne dépend que de la température de scène[2] et de la TFPA (c'est-à-dire θ est réduit à la seule TFPA τ). La figure 3.1 illustre la réponse d'un pixel en fonction de ces deux facteurs.

1. Si le modèle de correction affine considéré est (2.14), les composantes sont le gain $\mathcal{G}_\theta$ et l'offset $\mathcal{O}_\theta$.

2. Il s'agit plus exactement de l'irradiance, irradiance et température de scène étant reliés par la formule 2.8.

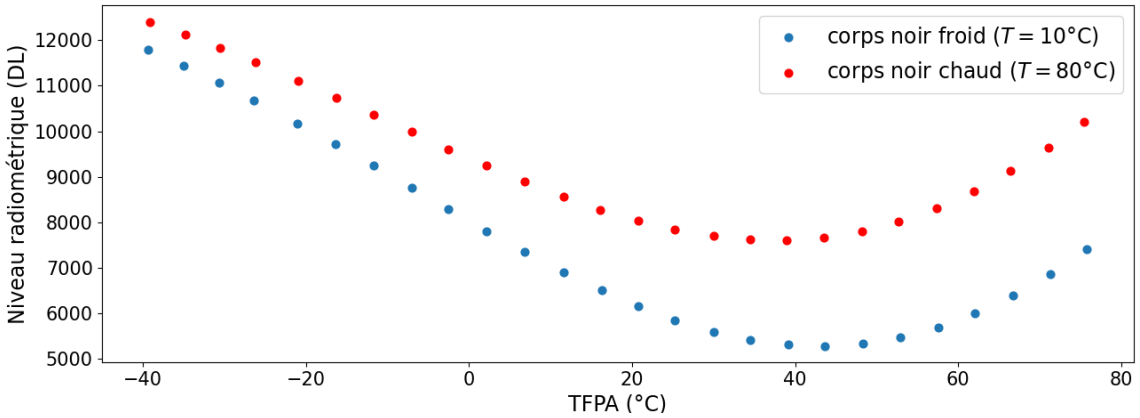


FIGURE 3.1 – **Variabilités induites par des changements de TFPA et de scène en un pixel représentatif.** Les points bleus représentent le niveau radiométrique délivré par un pixel du capteur selon 26 mesures de TFPA pour un corps noir froid de température $T = 10$°C, les points rouges correspondent à un corps noir chaud ($T = 80$°C). Les observations constatées ci-dessous sont les mêmes pour une centaine de pixels tirés au hasard. À température de scène fixée T, la réponse du pixel varie beaucoup selon la TFPA. À TFPA fixée, la réponse du corps noir chaud est supérieure (donc différente) à celle du corps noir froid. Pour une température de scène donnée ($T \in \{10°\text{C}, 80°\text{C}\}$), si l'on souhaite établir le niveau radiométrique en ce pixel pour toute TFPA, la modélisation polynomiale semble ici adaptée. En effet, les points d'observation (rouges ou bleus) semblent placés sur une courbe assez lisse que l'on peut approcher par un polynôme de degré peu élevé.

Comme nous l'avions vu dans le chapitre précédent, les méthodes *« calibration-based »* présentes dans la littérature, et en particulier la méthode *shutterless polynomial* (section 2.2.2.1), consistent à déterminer en premier lieu les composantes responsivité $\mathcal{R}_\tau$ et décalage $\mathcal{D}_\tau$ à partir d'une paire de scènes uniformes et de mesures de TFPA, en utilisant la formule de calibration à deux points (2.21). En pratique, ces composantes sont déterminées par une approximation, puisque les TFPA des deux scènes uniformes utilisées ne sont pas identiques.

Afin d'éviter cette approximation, nous proposons dans un premier temps une nouvelle manière de calibrer le modèle affine paramétrique en TFPA. Les composantes responsivité $\mathcal{R}_\tau$ et décalage $\mathcal{D}_\tau$ seront établies après avoir modélisé les images brutes de deux scènes uniformes par des polynômes, en utilisant la formule de calibration à deux points. Dans une deuxième partie, nous procéderons à une analyse statistique des données de calibration, ce qui nous permettra de mieux comprendre le comportement du capteur en fonction de la scène et de la TFPA. Ensuite, nous réaliserons une analyse de l'espace des corrections dans le cadre du modèle affine avec gain constant et biais.

1 Modèles de correction fondé sur une modélisation polynomiale des images brutes

Dans cette partie, nous considérons que la correction d'une image brute r_τ acquise à une TFPA τ se fait par le modèle affine en TFPA (2.16), qui pour rappel s'écrit

$$S = \frac{r_\tau - \mathcal{D}_\tau}{\mathcal{R}_\tau}\,, \tag{3.1}$$

où S désigne l'image corrigée, $\mathcal{R}_\tau$ et $\mathcal{D}_\tau$ désignent respectivement les images de responsivité et de décalage en niveau de gris. Comme nous l'avions vu dans le chapitre précédent (cf. 2.2.2), le modèle *shutterless polynomial* communément considéré dans la littérature consiste à calculer ces deux composantes $\mathcal{R}_\tau$ et $\mathcal{D}_\tau$ pour quelques TFPA (à partir de l'acquisition d'images brutes de scènes uniformes par la formule de calibration à deux points (2.21)). Les deux composantes sont ensuite modélisées par des polynômes en TFPA. La première étape de calcul des composantes pouvant se faire via une approximation, nous proposons ici une méthode sensiblement différente qui consiste à modéliser par des polynômes en TFPA chacune des images brutes de scène uniforme, puis de déterminer les composantes responsivité et décalage.

1.1 Établissement d'un jeu de données de calibration

Afin de calibrer correctement les composantes responsivité et décalage du modèle affine, il convient de construire un jeu de données fiable. En ce sens, des images brutes sont acquises en faisant varier la température de scène T et la TFPA τ. En pratique, on fait varier la scène en considérant N_T scènes uniformes provenant de corps noirs portés à des températures distinctes, et on fait varier la TFPA en plaçant la caméra dans N_e environnements thermiques distincts (qu'on appellera états). Chaque enregistrement d'une image brute est accompagné d'une mesure τ de la TFPA.

1.1.1 Calibration en « étuve stabilisée »

Les jeux de données sur lesquels se sont faites nos études et expérimentations ont été établis lors d'une procédure de calibration dite en « étuve stabilisée ». Lors d'une telle procédure, la caméra est placée dans une étuve à chaleur tournante au cours de N_e états distincts. Pour chacun de ces états, une consigne de température est donnée au système de régulation thermique, et on attend que température et humidité soient stabilisées temporellement. On peut ainsi dire que l'environnement thermique de chacun de ces états est spatialement homogène. La caméra fait alors l'acquisition de N_T images brutes de corps noirs distincts, accompagnées chacune d'une mesure de la TFPA. Une acquisition en « étuve stabilisée » essaie de réduire au mieux la variabilité des paramètres autres que la scène et la TFPA. En effet, sans stabilisation ou lors de la transition entre deux états thermiques distincts, les dynamiques variées des flux parasites de la caméra complexifient la modélisation. Dans le cas d'une procédure d'acquisition en « étuve stabilisée », ces flux parasites ont une influence sur la réponse du capteur relativement invariable (on verra cependant par la suite que ce n'est pas parfaitement le cas, mais ce type de procédure est celui qui permet de réduire au mieux les variabilités indésirables).

1.1.2 Jeu de données de calibration

Ceci étant dit, chaque élément d'intérêt de nos jeux de données se limite donc à un triplet : le premier élément est l'image brute d'une scène uniforme d'un corps noir, le deuxième est la mesure de la TFPA associée, le troisième est la température du corps noir associé. On dispose ainsi de la réponse du capteur selon les deux facteurs que sont la température de scène et la TFPA. Pour le corps noir $i \in [\![1, N_T]\!]$ au cours d'un état $j \in [\![1, N_e]\!]$, l'image brute est notée $r_j^{(i)}$, la mesure de TFPA $\tau_{i,j}$, et la température du corps noir T_i. Un ensemble de données de calibration $\mathscr{C}_k$ est donc défini par

$$\mathscr{C}_k = \{(r_j^{(i)}, \tau_{i,j}, T_i),\ i \in [\![1, N_T]\!],\ j \in [\![1, N_e]\!]\},$$

où k désigne l'indice associé au k^{ème} jeu de données.

Pour un corps noir i_0, nous considérerons la suite des mesures de TFPA $\left(\tau_{i_0,j}\right)_{j\in[\![1,N_e]\!]}$ comme strictement croissante. On considèrera la suite des températures de scène relatives aux corps noirs $\left(T_i\right)_{i\in[\![1,N_T]\!]}$ comme strictement croissante également.

Remarque 4 (Sélection des corps noirs pour déterminer les variables du modèle affine)**.** *Le k^{ème} jeu de données de calibration $\mathscr{C}_k$ est composé de N_T corps noirs distincts. Puisque deux corps noirs suffisent pour déterminer les composantes $\mathcal{R}_\tau$ et $\mathcal{D}_\tau$ par la formule (2.21), nous choisissons le plus froid et le plus chaud, soit $f = 1$ et $c = N_T$, où T_f et T_c correspondent respectivement aux températures des corps noirs froid et chaud. Le sous-ensemble de calibration $\mathscr{C}'_k$ est donc défini par*

$$\mathscr{C}'_k = \mathscr{C}_k^f \cup \mathscr{C}_k^c, \tag{3.2}$$

où

$$\begin{cases} \mathscr{C}_k^f = \{(r_j^{(f)}, \tau_{f,j}, T_f),\ j \in [\![1, N_e]\!]\} & \textit{pour le corps noir froid,} \\ \mathscr{C}_k^c = \{(r_j^{(c)}, \tau_{c,j}, T_c),\ j \in [\![1, N_e]\!]\} & \textit{pour le corps noir chaud.} \end{cases} \tag{3.3}$$

Les données des autres corps noirs seront utilisées comme données de test pour évaluer la performance d'un modèle de correction, et sont définies par

$$\mathscr{C}_k^t = \{(r_j^{(t)}, \tau_{t,j}, T_t),\ j \in [\![1, N_e]\!]\}, \quad t \in [\![2, N_T - 1]\!]. \tag{3.4}$$

Dans ce chapitre, nous ferons nos expérimentations sur les deux ensembles de calibration suivants.

$\mathscr{C}_1 :=$

Modéle de caméra	*ATTO-640a*
Résolution du capteur (Ω)	480×640
Nombre d'états (N_e)	46
Intervalle des TFPA mesurées	[0,46°C]
Nombre de corps noirs (N_T)	3
Températures des corps noirs	T_1=10°C
	T_2=30°C
	T_3=50°C
Date	20 avril 2021
Procédure	Étuve stabilisée
Quantification (N_{bits})	16 bits
Sous-ensemble corps noir froid ($\mathscr{C}_1^f$)	$f = 1$
Sous-ensemble corps noir chaud ($\mathscr{C}_1^c$)	$c = 3$
Sous-ensemble corps noir de test ($\mathscr{C}_1^t$)	$t = 2$

$\mathscr{C}_2 :=$

Modéle de caméra	*ATI-640a*
Résolution du capteur (Ω)	480×640
Nombre d'états (N_e)	26
Intervalle des TFPA mesurées	[-39,76°C]
Nombre de corps noirs (N_T)	5
Températures des corps noirs	T_1=10°C
	T_2=25°C
	T_3=40°C
	T_4=55°C
	T_5=80°C
Date	21 janvier 2022
Procédure	Étuve stabilisée
Quantification (N_{bits})	14 bits
Sous-ensemble corps noir froid ($\mathscr{C}_2^f$)	$f = 1$
Sous-ensemble corps noir chaud ($\mathscr{C}_2^c$)	$c = 5$
Sous-ensemble corps noir de test ($\mathscr{C}_2^t$)	$t = 3$

Avant de développer notre proposition de modèle, nous allons revenir sur la méthode *shutterless polynomial* communément utilisée dans la littérature. Il s'agit d'étudier son application pratique à partir d'un jeu de données de calibration $\mathscr{C}_k$.

1.2 Rappel de la méthode *shutterless polynomial* utilisée à titre comparatif

On considère la méthode *shutterless polynomial*, qui fait office de référence dans l'état de l'art et que nous avions présentée dans le chapitre précédent (cf. section 2.2.2). Les

composantes du modèle affine $\mathcal{R}_\tau$ et $\mathcal{D}_\tau$ sont dans un premier temps établies à partir des images brutes de deux scènes uniformes. En pratique, lorsque l'on considère un sous-ensemble de calibration $\mathscr{C}'_k$, pour le $j^{\text{ème}}$ état ($j \in [\![1, N_e]\!]$), nous disposons de deux triplets : $(r_j^{(f)}, \tau_{f,j}, T_f)$ pour le corps noir froid et $(r_j^{(c)}, \tau_{c,j}, T_c)$ pour le corps noir chaud.

Le modèle affine paramétrique en TFPA (2.15) permet d'écrire

$$\begin{cases} r_j^{(f)} = \mathcal{R}_{\tau_{f,j}} T_f + \mathcal{D}_{\tau_{f,j}} \\ r_j^{(c)} = \mathcal{R}_{\tau_{c,j}} T_c + \mathcal{D}_{\tau_{c,j}}. \end{cases} \tag{3.5}$$

Dans les faits, les deux mesures de TFPA $\tau_{f,i}$ et $\tau_{c,j}$ diffèrent de quelques dixièmes de degré, i.e. $\tau_{f,i} \neq \tau_{c,j}$. Le système (3.5) admet alors une infinité de solutions. Une technique fréquemment utilisée consiste à remplacer ces deux TFPA par leur moyenne τ_j et à considérer les approximations $\tau_{f,j} \approx \tau_j$ et $\tau_{c,j} \approx \tau_j$. On obtient alors

$$\begin{cases} r_j^{(f)} \approx \mathcal{R}_{\tau_j} T_f + \mathcal{D}_{\tau_j} & \text{(3.6a)} \\ r_j^{(c)} \approx \mathcal{R}_{\tau_j} T_c + \mathcal{D}_{\tau_j}, & \text{(3.6b)} \end{cases}$$

où

$$\tau_j = \frac{\tau_{f,j} + \tau_{c,j}}{2}. \tag{3.7}$$

Puisque $T_f \neq T_c$, le système (3.6) admet une solution unique $(\mathcal{R}_{\tau_j}, \mathcal{D}_{\tau_j})$. En découle la formule de la calibration à deux points qui permet d'établir les composantes $\mathcal{R}_{\tau_j}$ et $\mathcal{D}_{\tau_j}$, elle s'écrit

$$\begin{cases} \mathcal{R}_{\tau_j} \approx \dfrac{r_j^{(c)} - r_j^{(f)}}{T_c - T_f} & \text{(3.8a)} \\ \mathcal{D}_{\tau_j} \approx \dfrac{T_c r_j^{(f)} - T_f r_j^{(c)}}{T_c - T_f}. & \text{(3.8b)} \end{cases}$$

Une fois déterminés $\mathcal{R}_{\tau_j}$ et $\mathcal{D}_{\tau_j}$ pour $j \in [\![1, N_e]\!]$, une modélisation polynomiale permet de donner une estimation de ces composantes pour une TFPA quelconque.

1.3 Modélisation polynomiale des images brutes de scène uniforme

Plutôt que de déterminer après une approximation les composantes $\mathcal{R}_{\tau_j}$ et $\mathcal{D}_{\tau_j}$ à partir des images brutes du $j^{\text{ème}}$ état, puis de modéliser ces composantes par des polynômes (méthode *shutterless polynomial*), nous proposons d'adopter une modélisation polynomiale des images brutes relatives aux deux corps noirs. Nous déterminerons ensuite $\mathcal{R}_\tau$ et $\mathcal{D}_\tau$ à partir des polynômes obtenus (cf. figure 3.2).

Dans la littérature, il est rapporté que la méthode *shutterless polynomial* utilise des degrés de polynômes différents selon l'étendue des TFPA mesurées lors de la calibration, généralement de degré 2 pour la responsivité et de degré 3 pour le décalage, comme dans le cas d'une étendue d'environ 45°C rapportée dans [8, 9]. La méthode que nous proposons consiste d'abord à estimer le « bon » degré, c'est-à-dire celui qui s'adapte aux données spécifiques. Lors de la comparaison avec la méthode *shutterless polynomial*, nous choisirons pour cette dernière le même degré que celui estimé pour notre modèle.

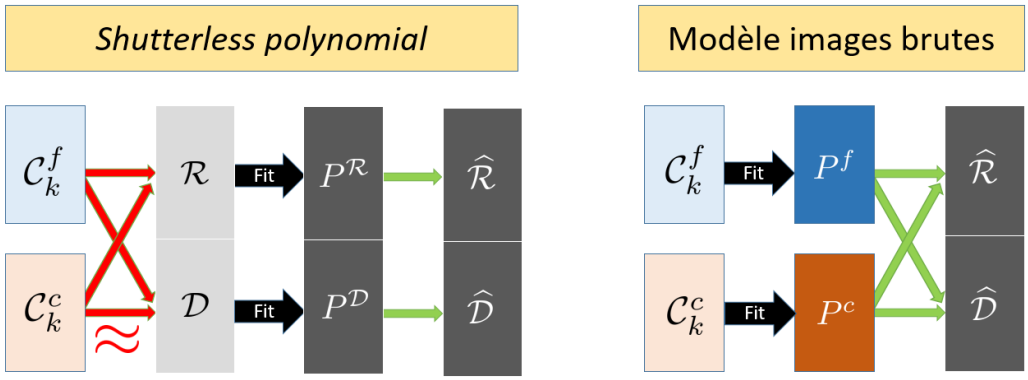


FIGURE 3.2 – **Schéma fonctionnel comparatif pour établir responsivité et décalage à partir de données de calibration.** $\mathscr{C}_k^f$ et $\mathscr{C}_k^c$ désignent les sous-ensembles du $k^{\text{ème}}$ jeu de calibration relatifs aux corps noir froid et chaud. Les flèches noires correspondent à une étape de modélisation polynomiale (« fit »). Ces deux schémas illustrent deux manières d'établir la responsivité $\widehat{\mathcal{R}}$ et le décalage $\widehat{\mathcal{D}}$ pour toute TFPA à partir des deux sous-ensembles $\mathscr{C}_k^f$ et $\mathscr{C}_k^c$. Sur le schéma de gauche, représentant la méthode *shutterless polynomial* (état de l'art), la responsivité $\mathcal{R}$ et le décalage $\mathcal{D}$ (en gris clair) sont d'abord déterminés pour un certain nombre de TFPA à partir de la formule (3.8). $\widehat{\mathcal{R}}$ (respectivement $\widehat{\mathcal{D}}$) sont ensuite obtenus à partir des polynômes $P^{\mathcal{R}}$ (respectivement $P^{\mathcal{D}}$) issus d'une régression linéaire. Sur le schéma de droite, on établit d'abord le polynôme P^f (respectivement P^c) par régression polynomiale à partir des images brutes de l'ensemble $\mathscr{C}_k^f$ (respectivement $\mathscr{C}_k^c$). $\widehat{\mathcal{R}}$ et $\widehat{\mathcal{D}}$ s'obtiennent ensuite par la formule de calibration à deux points (2.21) à partir de ces deux polynômes.

Au vu de la réponse d'un pixel à température de scène fixée selon les TFPA mesurées au cours de l'acquisition (cf. figure 3.1), et étant donné que cette réponse est représentative de la très grande majorité des pixels (comme l'ont montré nos observations sur plusieurs centaines de pixels tirés au hasard), une modélisation polynomiale en fonction des TFPA à partir des points de mesure semble tout à fait appropriée. En d'autres termes, la modélisation polynomiale en TFPA des images brutes de scène uniforme pour chaque pixel est une approche justifiée.

1.3.1 Présentation du problème

Considérons un sous-ensemble de calibration $\mathscr{C}_k' = \mathscr{C}_k^f \cup \mathscr{C}_k^c$ tel que défini dans (3.2). Il s'agit de modéliser, en chaque pixel, les images brutes obtenues à partir du corps noir froid, puis du corps noir chaud, à l'aide de polynômes dépendant de la TFPA, notée τ. Nous détaillons ci-dessous la construction du polynôme à partir des données du corps noir froid, $\mathscr{C}_k^f = \{(r_j^{(f)}, \tau_{f,j}, T_1) \,|\, j \in [\![1, N_e]\!]\}$; la démarche est strictement analogue pour le corps noir chaud.

Pour un pixel donné $(x, y) \in \Omega$, on cherche à déterminer $P_{x,y}$, polynôme réel de degré au plus d (i.e. $P_{x,y} \in \mathbb{R}_d[X]$), solution du problème suivant :

$$P_{x,y} \in \underset{P \in \mathbb{R}_d[X]}{\operatorname{argmin}} \sum_{j=1}^{N_e} \left(r_j^{(f)}(x,y) - P(\tau_{f,j}) \right)^2. \tag{3.9}$$

À partir de la solution $P_{x,y}$ du problème (3.9), on définit la racine carrée de l'erreur quadratique moyenne (*root mean square error*, notée *RMSE*) relative au jeu de données

$\mathscr{C}_k^f$ par

$$\sigma_{\mathscr{C}_k^f}(P_{x,y}) = \sqrt{\frac{\sum_{j=1}^{N_e}\left(r_j^{(f)}(x,y) - P_{x,y}(\tau_{f,j})\right)^2}{N_e}}. \tag{3.10}$$

Remarque 5 (Existence et unicité d'une régression polynomiale)**.** *Le problème* (3.9) *correspond à une régression polynomiale. Comme les valeurs de TFPA* $\{\tau_{f,j} \mid j \in [\![1, N_e]\!]\}$ *sont toutes distinctes, il admet une solution unique dès lors que le degré maximal* d *est strictement inférieur au nombre de données, soit* $d \leq N_e - 1$*.*

La résolution du problème (3.9) nécessite donc, en amont, de fixer un degré maximal d pour chaque pixel. L'objectif est ensuite de proposer un choix de degré global pertinent à l'échelle de l'ensemble des pixels.

1.3.2 Détermination du degré polynomial

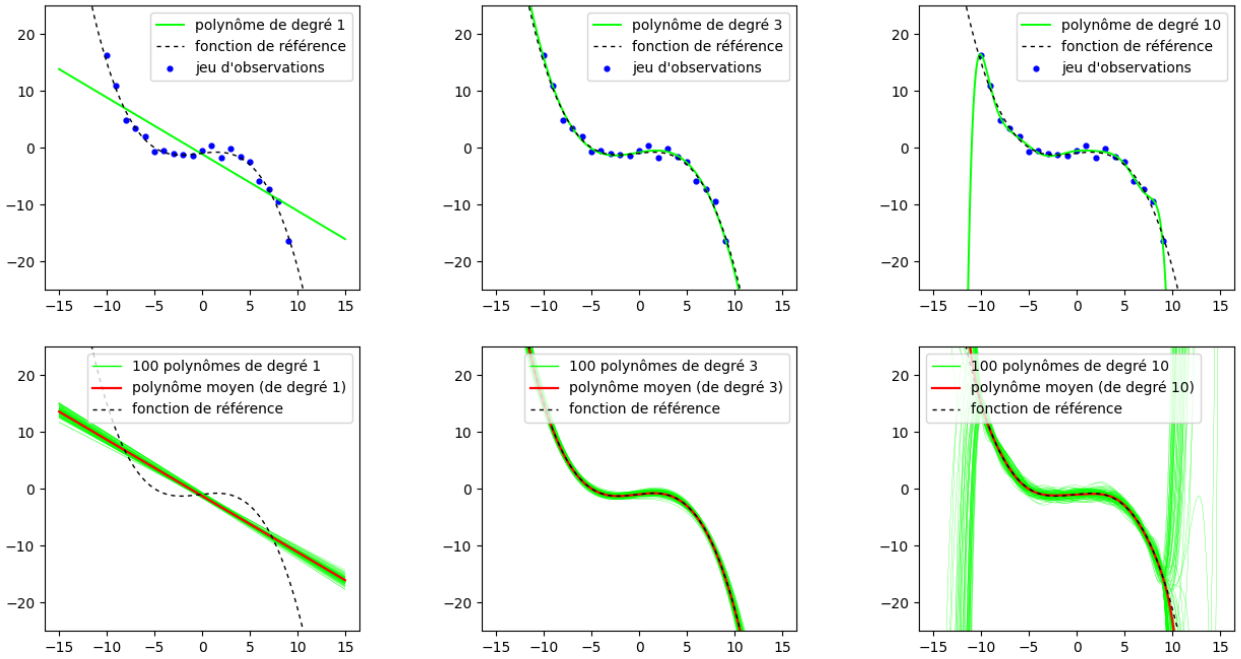
Le choix du degré idoine n'est pas une chose simple, un degré trop petit pourrait donner une modélisation polynomiale qui ne s'ajuste pas assez aux données (*underfitting*, figure 3.3a), un degré trop élevé permet certes un meilleur ajustement par rapport aux données mais peut donner un modèle incorrect (*overfitting*, figure 3.3c). Il existe un compromis à trouver entre un bon ajustement par rapport aux données et un degré pas trop élevé.

1.3.2.1 Revue des pratiques sur le choix du degré polynômial

Ce thème a été traité dans la littérature. Dans le domaine des statistiques, on peut citer les critères AIC [21] et BIC [22]. Ils reviennent à ajouter un terme pénalisant la dimension de l'espace de paramétrisation (dans notre cas : le degré) dans un problème de minimisation. Les méthodes AIC peuvent être problématiques lorsque le nombre d'échantillons est trop grand, un modèle de trop grande dimension peut être préféré et la probabilité de se retrouver en situation d'*overfitting* devient non négligeable [23, 58]. Les méthodes BIC, malgré leur consistance (aptitude à s'approcher du vrai modèle lorsque le nombre d'échantillons devient très grand), peuvent s'avérer problématiques lorsqu'il s'agit d'approcher une fonction ne faisant pas partie de la classe des polynômes [24]. Dans une approche de statistique bayésienne, il est possible d'estimer conjointement les lois de chacun des coefficients, du bruit et du degré. L'établissement de ces lois jointes, souvent non déterminées d'un point de vue analytique, nécessite des méthodes relativement lourdes pour les simuler (Markov Chain Monte Carlo) [25]. Dans un autre registre, Pakdemirli [59] a proposé un critère pour déterminer le degré optimal fondé sur le coefficient du terme de plus haut degré (après une renormalisation des données). Cependant, ce critère dépend d'un seuil difficile à ajuster. Nous proposons une méthode alternative pour déterminer le degré optimal, fondée sur un test d'*overfitting*. Nous comparerons les résultats obtenus avec cette méthode et ceux issus du critère BIC.

1.3.2.2 Détermination du degré propre à un pixel à partir d'un test d'*overfitting*

Pour chaque pixel, on cherche le plus grand degré du polynôme $P_{x,y}$ solution de (3.9) qui garantisse un très faible risque d'*overfitting*.



(a) Degré 1 : *underfitting* (b) Degré 3 : meilleur *fit* (c) Degré 10 : *overfitting*

FIGURE 3.3 – **Illustration de l'*underfitting* et de l'*overfitting* selon le degré choisi par une régression polynomiale.** La fonction de référence est un polynôme de degré 3 (courbe discontinue noire). Sur la première ligne on dispose de $n = 20$ observations (points bleus) : les abscisses sont régulièrement réparties entre -10 et 9 et les ordonnées sont des échantillons bruités de la fonction de référence. La courbe verte illustre le résultat de la régression polynomiale pour différents degrés. Sur la deuxième ligne 100 jeux d'observations ont été obtenus de manière aléatoire selon le même procédé que décrit précédemment. Pour chaque degré, 100 polynômes sont alors obtenus (courbes vertes). La moyenne de ces polynômes est illustrée par une courbe rouge. On définit le biais comme étant l'écart moyen entre le polynôme obtenu et la fonction de référence, il peut se voir comme l'écart entre la courbe rouge et la courbe discontinue noire. La variance représente la variabilité des polynômes obtenus, elle est illustrée par la largeur du faisceau des 100 courbes vertes sur les graphiques qui composent la deuxième ligne. (a) Degré 1 : la courbe du polynôme obtenu est loin de celle de la fonction de référence, i.e. le biais est grand. Les polynômes obtenus sont assez proches, i.e la variance est assez peu élevée. De plus la position des observations par rapport au polynôme obtenu n'a pas un comportement d'indépendance attendu (sur le graphique du haut, les 2 premiers points sont au dessus, les 9 suivants en dessous etc). Le degré choisi est trop petit puisque le biais est excessivement élevé, cette situation porte le nom d'*underfitting*. (b) Degré 3 : le biais est très petit (la courbe rouge et la courbe discontinue noire apparaissent confondues) et la variance relativement peu élevée. Le *fit* est tout à fait convenable dans ce cas. (c) Degré 10 : le biais est relativement petit mais la variance est élevée puisque le polynôme obtenu peut prendre des valeurs très différentes en extrapolation, c'est à dire en dehors de l'intervalle contenant les abscisses des points d'observation. Ce phénomène que l'on appelle *overfitting* se produit lorsque le degré considéré est trop élevé. Il y a un compromis biais-variance à trouver, il se manifeste ici par le choix d'un degré ni trop petit ni trop grand.

De manière générique, on considère un jeu d'observations [3] $\mathbb{O} = \{(X_i, Y_i) \in \mathbb{R}^2;\ i \in [\![1, n]\!]\}$ et on cherche le polynôme réel P_d de degré au plus d solution de

$$P_d \in \underset{Q \in \mathbb{R}_d[X]}{\operatorname{argmin}} \sum_{i=1}^{n} \Big(Y_i - Q(X_i)\Big)^2. \tag{3.11}$$

On définit alors pour ce polynôme P_d solution de (3.11) la RMSE

$$\sigma_{\mathbb{O}}(P_d) = \sqrt{\frac{\sum_{i=1}^{n} \Big(Y_i - P_d(X_i)\Big)^2}{n}}. \tag{3.12}$$

Si on note $\widehat{\beta}$ le vecteur colonne de taille $d+1$ constitué par les coefficients de P_d ordonnés par degré croissant, la formule de la régression linéaire s'écrit

$$\widehat{\beta} = ({}^tMM)^{-1}\ {}^tMY, \tag{3.13}$$

où $M = \left(X_i^j\right)_{\substack{1 \le i \le n \\ 0 \le j \le d}}$ est une matrice de Vandermonde et $Y = (Y_i)_{1 \le i \le n}$ un vecteur colonne.

Remarque 6 (Formule régression linéaire, existence et unicité du problème (3.11))**.** *La formule (3.13) n'est valable que si la matrice de Vandermonde M est de rang plein, i.e. $d+1 \le card(\{X_i, i \in [\![1, n]\!]\})$; hypothèse déjà rencontrée dans la remarque 5 et que nous considérerons comme toujours vérifiée.*

Dans ce cas, il y a existence et unicité de la solution pour le problème (3.11). Pour éviter des erreurs numériques suite à l'inversion de ${}^tMM$ et sans rien changer fondamentalement au problème de minimisation des moindres carrés, il est en pratique préférable de renormaliser l'échelle des abscisses (ici des TFPA) en remplaçant par exemple X_i par $\tilde{X}_i = \dfrac{X_i}{\max(|X_j|,\ j \in [\![1, n]\!])}$ dans la formule (3.13) (cf. figure 3.4).

La procédure du test d'*overfitting* est détaillée ci-dessous. On commence par tester le degré $d = 0$, puis nous incrémenterons par pas de 1 jusqu'à tester le degré $d = n - 2$.

Pour tester le degré d, on détermine le polynôme P_d de degré au plus d solution de (3.11) grâce à la formule (3.13). On lui associe la RMSE $\sigma_{\mathbb{O}}(P_d)$ par (3.12). On construit alors un ensemble $\mathbb{S}_d = \left(\mathbb{S}_d^{(1)}, \ldots, \mathbb{S}_d^{(n_s)}\right)$ composé de n_s jeux de données simulées.
Le $k^{\text{ème}}$ jeu de données

$$\mathbb{S}_d^{(k)} = \{(X_i, Z_i^{(k)}) \in \mathbb{R}^2;\ i \in [\![1, n]\!]\}, \quad k \in [\![1, n_s]\!]$$

est tel que

$$Z_i^{(k)} = P_d(X_i) + \varepsilon_i, \tag{3.14}$$

où les ε_i sont des réalisations i.i.d. de la loi $\mathcal{N}\Big(0,\ \sigma_{\mathbb{O}}^2(P_d)\Big)$, avec $i \in [\![1, n]\!]$.

Chacun de ces jeux de données simulées $\mathbb{S}_d^{(k)}$ ne contient que l'information présente dans $\mathbb{O}$ relative aux degrés inférieurs ou égaux à d.

On détermine ensuite le polynôme $Q_{d+1}^{(k)}$ de degré au plus $d+1$ qui vérifie

$$Q_{d+1}^{(k)} = \underset{Q \in \mathbb{R}_{d+1}[X]}{\operatorname{argmin}} \sum_{i=1}^{n} \Big(Z_i^{(k)} - Q(X_i)\Big)^2, \tag{3.15}$$

3. On prend $n \ge 2$.

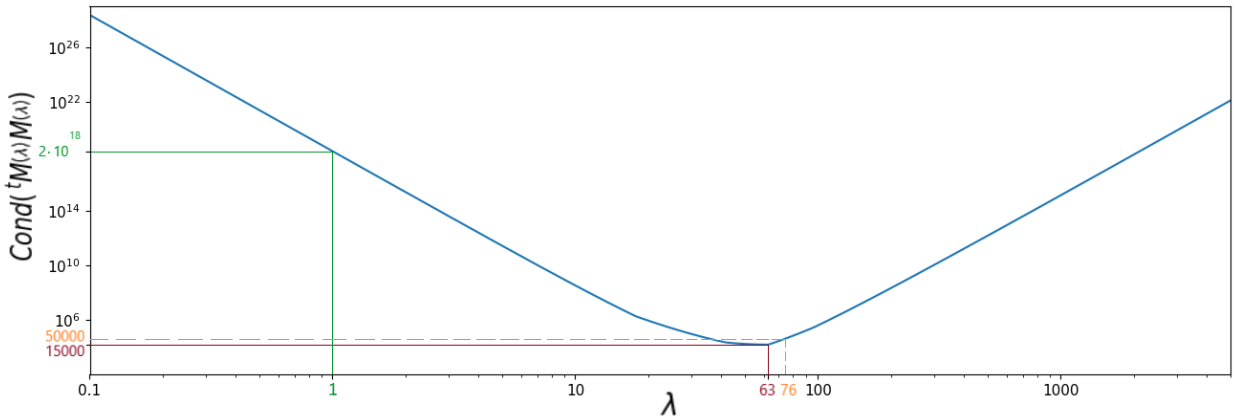


FIGURE 3.4 – **Conditionnement de la matrice à inverser dans la formule de régression linéaire** (3.13)**.** La famille des $(X_i)_{1\leq i\leq n}$ provient des mesures de TFPA du jeu de données $\mathscr{C}_2$, nous avons $n = 26$ et $\max_{1\leq i\leq n}(|X_i|) = 76$°C. Pour $\lambda > 0$, on définit la matrice de Vandermonde $M(\lambda) = \left((\frac{X_i}{\lambda})^j\right)_{\substack{1\leq i\leq n \\ 0\leq j\leq d}}$ en prenant $d = 5$ et nous nous intéressons au conditionnement relativement à la norme euclidienne de ${}^tM(\lambda)\,M(\lambda)$ en fonction de λ (courbe bleue). Le conditionnement est minimal pour $\lambda = 63$ (15000) conduisant à une perte de précision modérée lors de l'évaluation de (3.13). En effet, les erreurs numériques peuvent engendrer une perte maximale d'environ $\log_{10}(15000) \approx 4$ chiffres significatifs, ce qui est raisonnable pour un calcul en précision double pour lesquels les nombres sont représentés avec environ 16 chiffres significatifs. Lorsque $\lambda = 1$, le conditionnement est très élevé (2.10^{18}) par rapport à la précision double (erreurs pouvant atteindre 18 chiffres significatifs). Ce conditionnement excessif peut engendrer des erreurs numériques très importantes lors du calcul direct de (3.13). Lorsque $\lambda = \max_{1\leq i\leq n}(|X_i|)$, le conditionnement est relativement peu élevé (50000), équivalent à une perte maximale d'environ 5 chiffres significatifs. Lorsque l'on dispose d'une famille $(X_i)_{1\leq i\leq n}$ et que l'on doit appliquer la formule de régression linéaire (3.13), le conditionnement de la matrice à inverser est excessif (cas où $\lambda = 1$), ce qui engendre de potentielles importantes erreurs numériques. Renormaliser l'échelle des $(X_i)_{1\leq i\leq n}$ en divisant chacun de ses éléments par $\max_{1\leq i\leq n}(|X_i|)$ permet de réduire très fortement le conditionnement de la matrice à inverser (cas ou $\lambda = 76$). Dans ce cas, les erreurs numériques lors d'un calcul en précision double sont relativement négligeables.

auquel on associe la RMSE

$$\sigma_{\mathbb{S}^{(k)}}(Q_{d+1}^{(k)}) = \sqrt{\frac{\sum_{i=1}^{n}\left(Z_i^{(k)} - Q_{d+1}^{(k)}(X_i)\right)^2}{n}}. \tag{3.16}$$

On s'intéresse ensuite à la distribution empirique $\mathbb{D}(d)$ de ces n_s RMSE sur données simulées, qui se définit par

$$\mathbb{D}(d) = \{\sigma_{\mathbb{S}^{(k)}}(Q_{d+1}^{(k)}),\ k \in [\![1, n_s]\!]\}.$$

On calcule la moyenne μ_d et l'écart type ρ_d de $\mathbb{D}(d)$.

On détermine aussi la RMSE $\sigma_{\mathbb{O}}(P_{d+1})$ du polynôme P_{d+1} de degré $d+1$ obtenu à partir des vraies observations $\mathbb{O}$ par (3.12). Cette dernière est-elle une anomalie dans la distribution des RMSE sur données simulées $\mathbb{D}(d)$? Si tel est le cas, on conclut qu'il existe une information pertinente dans le jeu d'observation $\mathbb{O}$ relative au degré $d+1$ non présente dans chacun des éléments de l'ensemble des jeux de données simulées $\mathbb{S}_d$. Par conséquent, le risque d'*overfitting* au degré $d+1$ est faible et on continue la procédure pour tester le degré $d+1$. Dans le cas contraire, il y a un risque d'*overfitting*. Le pseudo-code 1 résume le test d'*overfitting* relatif au degré d (si le résultat T_d est « Vrai », l'*overfitting* est possible au degré $d+1$).

Une première idée naïve consiste à arrêter le procédé et choisir le degré d dès lors qu'il existe un risque d'*overfitting* à considérer le degré suivant. En prenant par exemple [4] $K = 4$ (et n_s suffisamment grand), la première approche naïve se formule par

$$\begin{cases} \text{si } \left|\dfrac{\sigma_{\mathbb{O}}(P_{d+1}) - \mu_d}{\rho_d}\right| \leq K, \text{ le degré choisi est } d, & \text{(3.17a)} \\ \text{sinon, on passe au degré } d+1. & \text{(3.17b)} \end{cases}$$

La figure 3.5 (respectivement 3.6) illustre les résultats de cette procédure naïve à partir des données de calibration $\mathscr{C}_1$ (respectivement $\mathscr{C}_2$). Il en résulte que, sur le jeu $\mathscr{C}_2$, le degré optimal obtenu en appliquant le critère d'arrêt (3.17a) conduit à sélectionner le degré $d = 0$, qui correspond pourtant à une situation manifeste d'*underfitting*.

En fait, le test nous permet de dire que si la condition

$$\left|\frac{\sigma_{\mathbb{O}}(P_{d+1}) - \mu_d}{\rho_d}\right| > K \tag{3.18}$$

est vérifiée pour le degré d, la probabilité qu'il y ait *overfitting* au degré $d+1$ est infime. Mais dans le cas contraire, l'*overfitting* n'est pas certain. Le meilleur candidat garantissant un faible risque d'*overfitting* est donc le plus grand degré pour lequel la condition (3.18) n'est pas respectée. Il convient alors de changer la procédure naïve (3.17) en

$$\begin{cases} \text{si } \forall d' \in [\![d, n-2]\!],\ \left|\dfrac{\sigma_{\mathbb{O}}(P_{d'+1}) - \mu_{d'}}{\rho_{d'}}\right| \leq K \text{ , le degré choisi est } d, \\ \text{sinon, on passe au degré } d+1. \end{cases} \tag{3.19}$$

4. Pour une loi normale centrée réduite, la probabilité d'être en dehors de l'intervalle $[-4, 4]$ est d'environ 1/16000. Lorsque le cardinal de l'ensemble des pixels vaut par exemple 307200 (résolution de 480 par 640), on s'autorise donc en moyenne une estimation erronée du degré pour une vingtaine (307200/16000) d'entre eux.

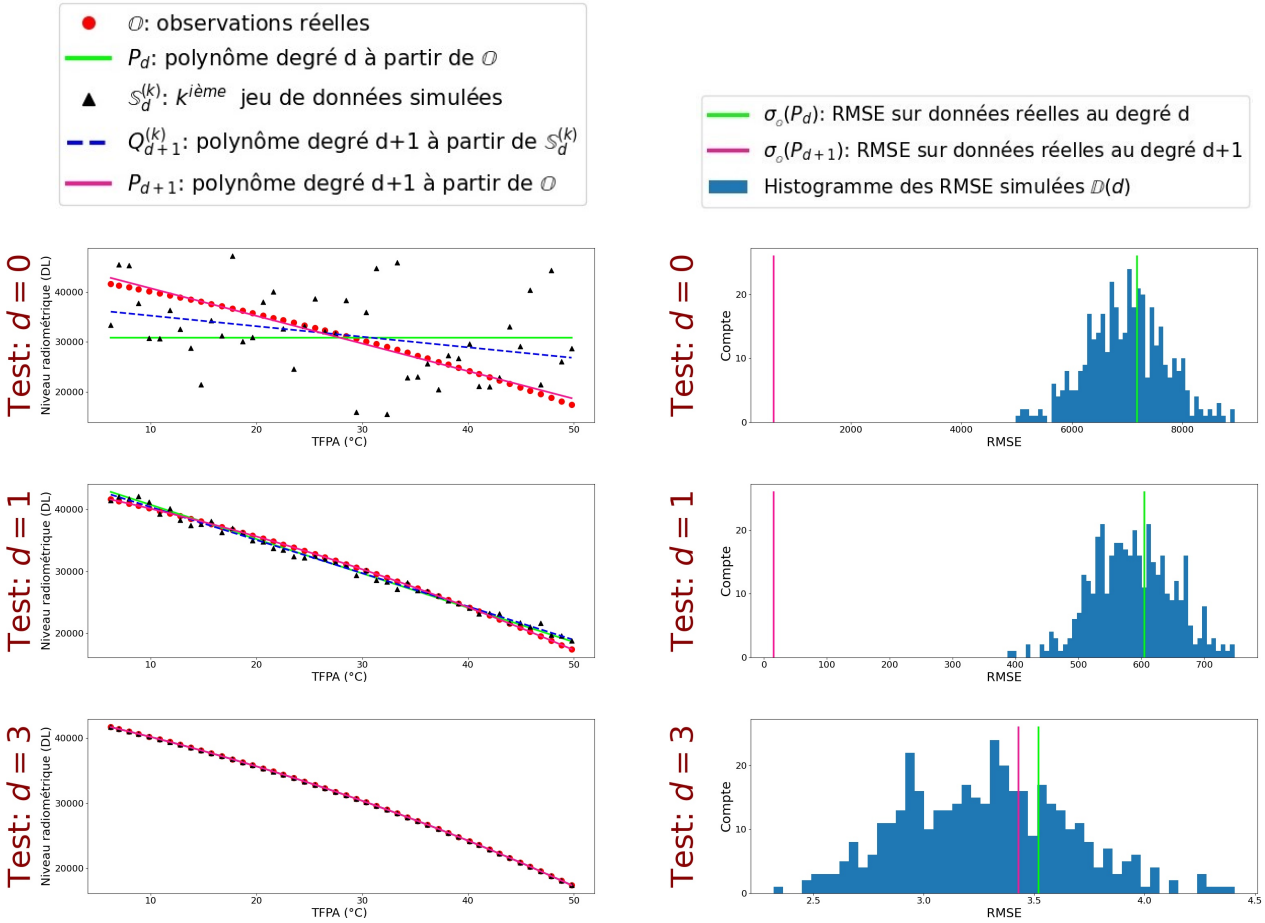


FIGURE 3.5 – **Explication de la procédure fondée sur un test d'*overfitting* selon une première approche naïve.** Les $n = 46$ observations réelles $\mathbb{O}$ (points rouges sur les figures de la colonne de gauche) proviennent du jeu $\mathscr{C}_1^f$ et sont relatives à un seul pixel. Le nombre de simulations n_s vaut 400. Chaque ligne illustre la procédure de test pour un certain degré d. Sur le graphe de gauche qui représente la réponse du pixel selon la TFPA (à température de scène fixée), la courbe continue verte représente le résultat de la régression polynomiale au degré au plus d à partir des observations réelles $\mathbb{O}$ (polynôme P_d), la courbe continue rose foncé désigne le résultat de la régression polynomiale au degré au plus $d+1$ (P_{d+1}). On simule alors un $k^{\text{ème}}$ nouveau jeu de données $\mathbb{S}_d^{(k)}$ (triangles noirs) à partir de P_d et la RMSE associée $\sigma_{\mathbb{O}}(P_d)$ par (3.14). On peut voir ce nouveau jeu de données simulées comme ne contenant que l'information présente dans les observations $\mathbb{O}$ relative aux degrés inférieurs ou égaux à d. La régression polynomiale au degré au plus $d+1$ à partir de ce jeu de données simulées $\mathbb{S}_d^{(k)}$ donne le polynôme $Q_{d+1}^{(k)}$ (courbe discontinue bleue), on lui associe une RMSE. On répète cette opération pour $k \in [\![1, n_s]\!]$ afin d'obtenir la distribution $\mathbb{D}(d)$ constituée de ces n_s RMSE. Le graphe de droite montre l'histogramme relatif à $\mathbb{D}(d)$ (barres bleues). Si la ligne verticale rose foncé désignant la RMSE $\sigma_{\mathbb{O}}(P_{d+1})$ se trouve en dehors de l'histogramme $\mathbb{D}(d)$ (comme sur les 2 premières lignes pour les degrés 0 et 1), nous dirons que le risque d'*overfitting* au degré $d+1$ est très peu probable. Il existe dans le jeu réel $\mathbb{O}$ une information pertinente qui justifie que l'on doive considérer le degré $d+1$, nous devons donc tester le degré $d+1$ (pour savoir si $d+2$ présente un risque etc). Dans le cas contraire (comme sur la dernière ligne pour le degré 3), dans une première approche naïve, nous arrêterons de tester les degrés suivants et considérerons le degré d comme étant le bon. Ici, le degré $d = 3$ est choisi pour ce pixel. Il nous faudra cependant rectifier la procédure lorsque nous ferons ce test sur un autre jeu de données (cf. figure 3.6).

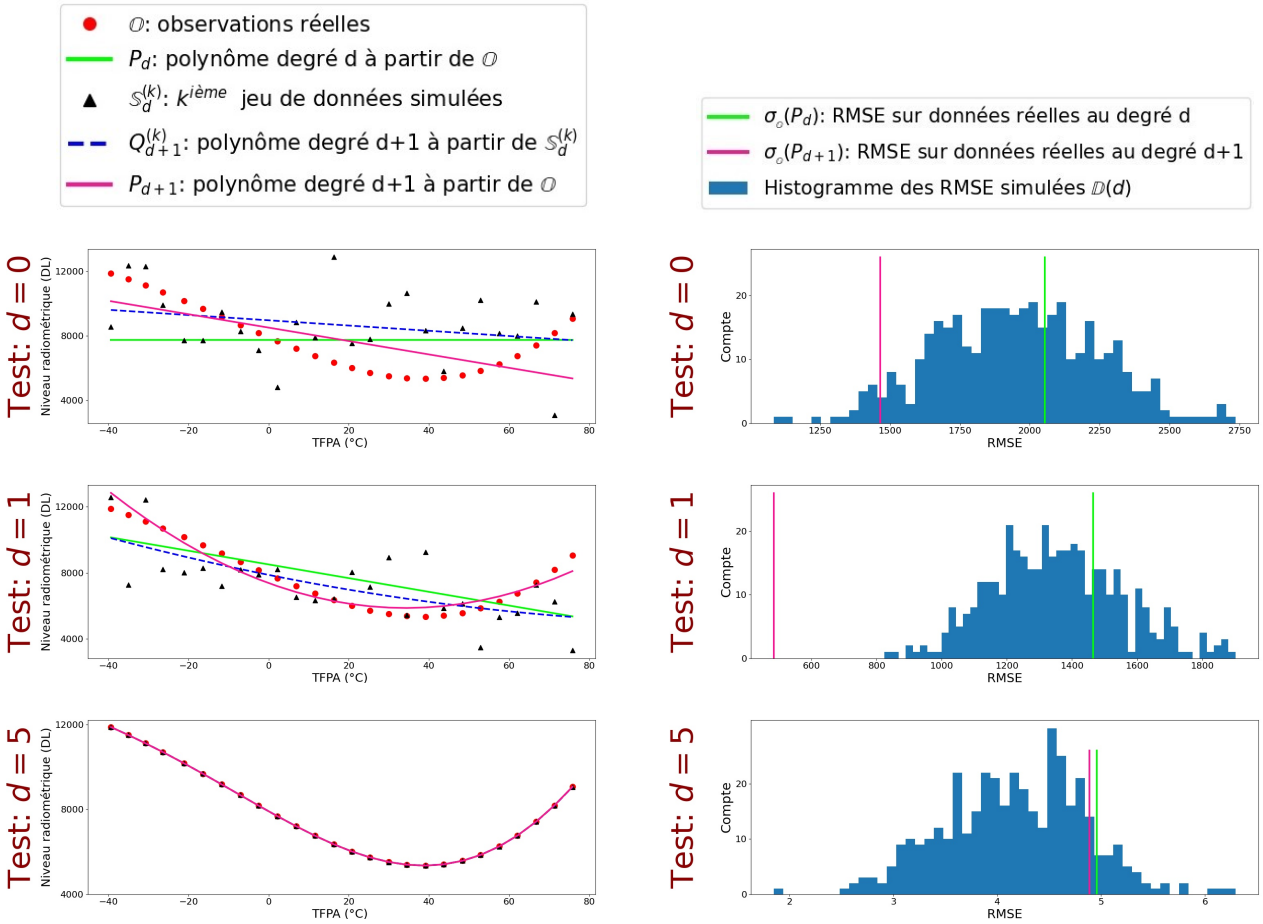


FIGURE 3.6 – **Explication de la procédure fondée sur un test d'*overfitting*, rectification de l'approche naïve.**. Les $n = 26$ observations (points rouges) proviennent des données de calibration $\mathscr{C}_2^f$ pour un certain pixel. En reprenant la procédure naïve que nous avons détaillée figure 3.5, sur la première ligne pour le degré $d = 0$, le fait d'envisager le degré suivant présente un risque d'*overfitting* et il faudrait arrêter la procédure et considérer comme degré définitif $d = 0$. Or ce degré est un cas d'*underfitting* manifeste. En effet la position relative entre les observations (points rouges) et le polynôme obtenu P_0 (courbe verte) n'a pas un comportement d'indépendance attendu (les 8 premières observations et les 2 dernières sont au dessus, les 16 autres en dessous). En forçant le passage à la ligne suivante pour tester le degré $d = 1$, le test nous dit cependant que le risque d'*overfitting* est très peu probable pour le degré 2. On devrait donc considérer un degré au moins égal à 2. Lorsque l'on teste un degré d, sur le graphe de droite, lorsque la ligne verticale rose foncée est située hors de l'histogramme (barres bleues), le test d'*overfitting* nous permet de dire que le degré $d + 1$ ne présente qu'un risque infime mais dans le cas contraire, l'*overfitting* au degré $d + 1$ n'est pas certain. Il est alors préférable de tester tous les degrés et considérer comme degré définitif le plus grand ne présentant pas de risque d'*overfitting* (au sens de notre test). Ici le degré $d = 5$ est choisi (dernière ligne) car dans les faits, le risque d'*overfitting* restera avéré pour les degrés $6, 7, ..., 24$. La procédure rectifiée est telle que définie par (3.19). Elle garantit un faible risque d'*overfitting* tout en évitant l'*underfitting* (en privilégiant le plus grand des degrés possibles parmi ceux non sujets à l'*overfitting*).

Algorithme 1 : Test d'*overfitting* relatif au degré d

Données : $X \in \mathbb{R}^n$, $Y \in \mathbb{R}^n$, $d \in \mathbb{N}$, $n_s \in \mathbb{N}$, $K \in \mathbb{R}_+$.
Résultat : $T_d \in \{\text{Faux,Vrai}\}$

Fonction *Polynôme(X,Y,d)*
  Retourne la solution de (3.11) associée à $X = (X_i)_{1 \leq i \leq n}$, $Y = (X_i)_{1 \leq i \leq n}$ et d
Fonction $RMSE(X, Y, P)$
  Retourne $\sqrt{\dfrac{\sum_{i=1}^{n} \left(Y_i - P(X_i)\right)^2}{n}}$
$P \leftarrow \text{Polynôme}(X, Y, d)$
$\sigma \leftarrow RMSE(X, Y, P)$
$\tilde{P} \leftarrow \text{Polynôme}(X, Y, d+1)$
$\tilde{\sigma} \leftarrow RMSE(X, Y, \tilde{P})$
pour $k = 1$ **à** n_s **faire**
  pour $i := 1$ **à** n **faire**
    $Z_i \sim \mathcal{N}(0, \sigma^2) + P(X_i)$
  $Q \leftarrow \text{Polynôme}(X, Z, d+1)$
  $\mathbb{D}_k \leftarrow RMSE(X, Z, Q)$
$\mu \leftarrow \sum_{k=1}^{n_s} \mathbb{D}_k / n_s$
$\rho \leftarrow \sqrt{\sum_{k=1}^{n_s} (\mathbb{D}_k - \mu)^2 / n_s}$
si $\left|\dfrac{\mu - \tilde{\sigma}}{\rho}\right| \leq K$ **alors**
  $T_d = \text{Vrai}$
sinon
  $T_d = \text{Faux}$

Ainsi modifié, le procédé décisionnel (3.19) nous permet de trouver le plus haut degré d garantissant une très faible probabilité d'*overfitting*. La figure 3.7 montre quelques résultats obtenus par la procédure d'*overfitting* sur des données simulées.

1.3.2.3 Carte des degrés

Une fois choisi, pour chaque pixel, le meilleur degré par le test d'*overfitting*, nous pouvons construire une image $\mathbf{C} \in \mathbb{N}^{\Omega}$ où chaque pixel aura comme valeur le degré choisi. Nous l'appellerons carte des degrés (cf. figure 3.8a).

1.3.2.4 Détermination d'un degré global

Le procédé industriel de fabrication vise à produire des capteurs dont la très grande majorité des pixels (plus de 99.9%) a une réponse selon les différents flux thermiques incidents la plus homogène possible. En cela on peut considérer qu'il existe un degré global relatif à la modélisation polynomiale de la réponse du capteur à température de scène fixée selon la TFPA. Ce degré global, noté D, peut par exemple être déterminé par

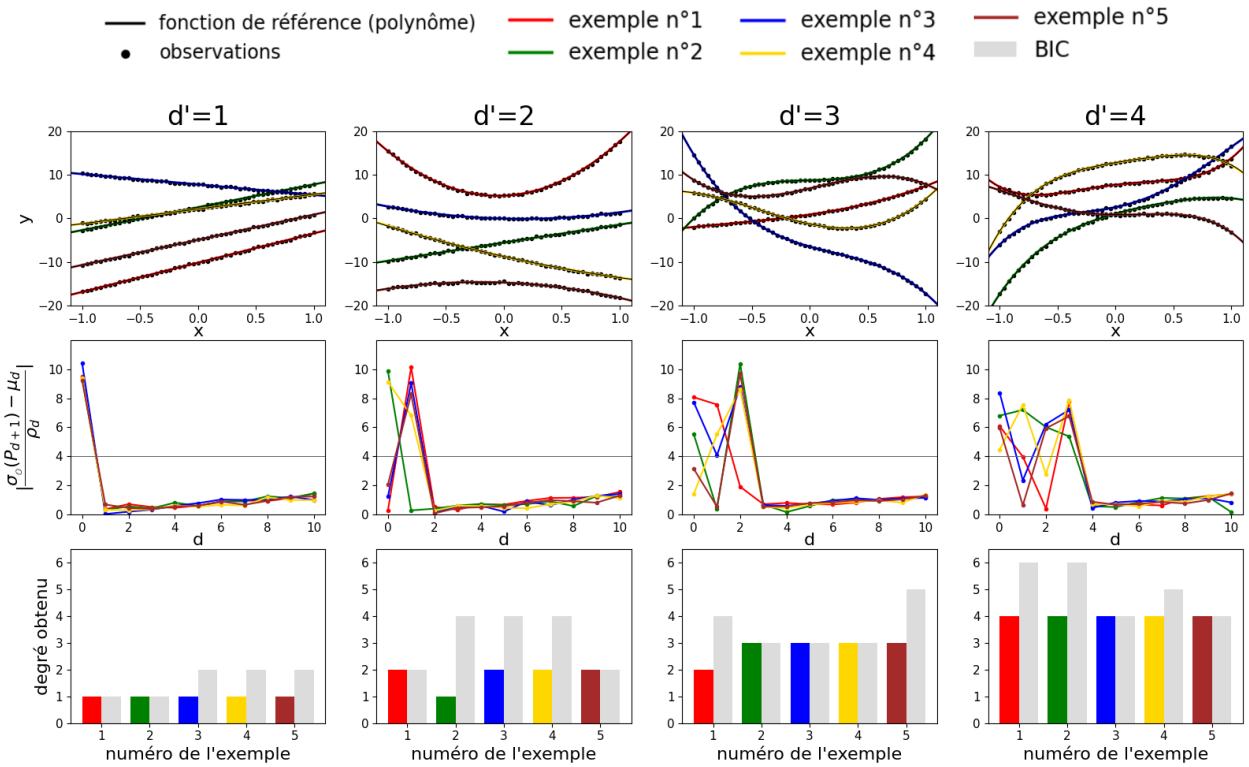


FIGURE 3.7 – **Illustration de la procédure de détermination du degré par un test d'*overfitting* à partir de données simulées, comparaison avec BIC.** Sur chaque colonne on illustre un degré $d' \in \{1, 2, 3, 4\}$ avec cinq exemples (associés chacun à une couleur différente). Chacun de ces exemples est constitué de 51 observations bruitées d'un polynôme de degré d' dont les coefficients ont été tirés au hasard (par une loi uniforme continue sur [-20,20]). Toutes les observations ont des abscisses régulièrement réparties dans l'intervalle [-1,1]. La première ligne montre les observations (points noirs) ainsi que les courbes des polynômes obtenus par régression linéaire (le degré choisi étant celui obtenu par la procédure d'*overfitting*). La deuxième ligne montre l'évolution du score $\left|\frac{\sigma_{\mathbb{O}}(P_{d+1})-\mu_d}{\rho_d}\right|$ en fonction de d (le degré testé lors du test d'*overfitting*) pour chacun des exemples (associés à leur couleur respective). En appliquant le procédé (3.19), le degré obtenu est le plus petit degré d à partir duquel le score est inférieur à 4. Ce degré obtenu est symbolisé par une barre verticale (dont la couleur correspond à celle de l'exemple) sur le diagramme de la troisième ligne. Une comparaison est faite avec le degré obtenu par la méthode BIC, ce dernier est symbolisé par une barre verticale grise. Sur ces vingt exemples, le degré obtenu par la procédure d'*overfitting* correspond au degré du polynôme de référence sauf dans deux cas. Le deuxième exemple (couleur verte) du degré $d' = 2$ donne un degré égal à un, ce qui peut s'expliquer par le fait que les observations semblent placées sur une droite (courbe verte sur le deuxième graphique du haut). Le premier essai (couleur rouge) du degré $d' = 3$ donne un degré 2. Sur ces exemples le degré obtenu par la méthode BIC est différent dans une moitié des cas du degré du polynôme de référence. Sur l'ensemble de ces exemples, la procédure d*overfitting* est plus fiable que la méthode BIC. En outre le degré estimé n'est jamais strictement supérieur au degré du polynôme de référence, ce qui n'est pas le cas pour BIC.

une votation (ce qui revient à déterminer le mode de la distribution des degrés au sein de tous les pixels). Sur la figure 3.8a (relative à l'ensemble de calibration $\mathscr{C}_1^f$), on constate qu'une écrasante majorité des pixels a obtenu le degré 3, considérer comme degré global $D = 3$ semble alors raisonnable.

Pour l'ensemble de calibration $\mathscr{C}_2^f$, une écrasante majorité des pixels obtient le degré 5, il s'en suit que $D = 5$. Comme nous l'avions vu dans la section relative à l'état de l'art (2.2.2), le degré choisi dépend de l'intervalle dans lequel sont situées les TFPA de la calibration.

1.4 Détection des pixels défectueux

Afin d'avoir une meilleure compréhension du capteur, il convient d'en exclure les pixels anormaux (défectueux). Nous proposons ici une méthode assez expérimentale de détection des pixels défectueux, l'objectif n'étant pas d'obtenir le meilleur détecteur possible mais de tous les détecter. On pourra donc se permettre de classer un pixel comme anormal sans qu'il le soit de facto (faux négatif) mais il faudra éviter, tant que faire se peut, qu'un pixel anormal ne soit pas détecté (faux positif).

1.4.1 Définition d'un pixel défectueux

Un pixel défectueux est soit un pixel au comportement atypique soit un pixel cassé. Un pixel a un comportement atypique lorsque le niveau radiométrique qu'il délivre à température de scène constante selon la TFPA est irrégulier (dérivée non continue ou très grande en valeur absolue). Un pixel est dit cassé lorsque le niveau radiométrique qu'il délivre (en DL) à température de scène constante selon la TFPA est soit constant, soit atteint la saturation basse (valeur 0) ou haute ($2^{N_{bits}} - 1$). Un pixel au comportement atypique peut également être cassé.

1.4.2 Construction du masque de pixels non défectueux

Considérons le sous-ensemble de calibration relatif au corps noir froid $\mathscr{C}^f$. Une fois choisi le degré global D par vote, on détermine pour chaque pixel $(x, y) \in \Omega$ la RMSE $\sigma_{\mathscr{C}^f}(P_{x,y})$ définie par (3.10) associée au polynôme $P_{x,y}$ obtenu par régression polynomiale de degré au plus D. On considère alors la distribution empirique de ces RMSE au sein de tous les pixels, ensemble que nous notons $\mathbb{H}$ et qui se définit par

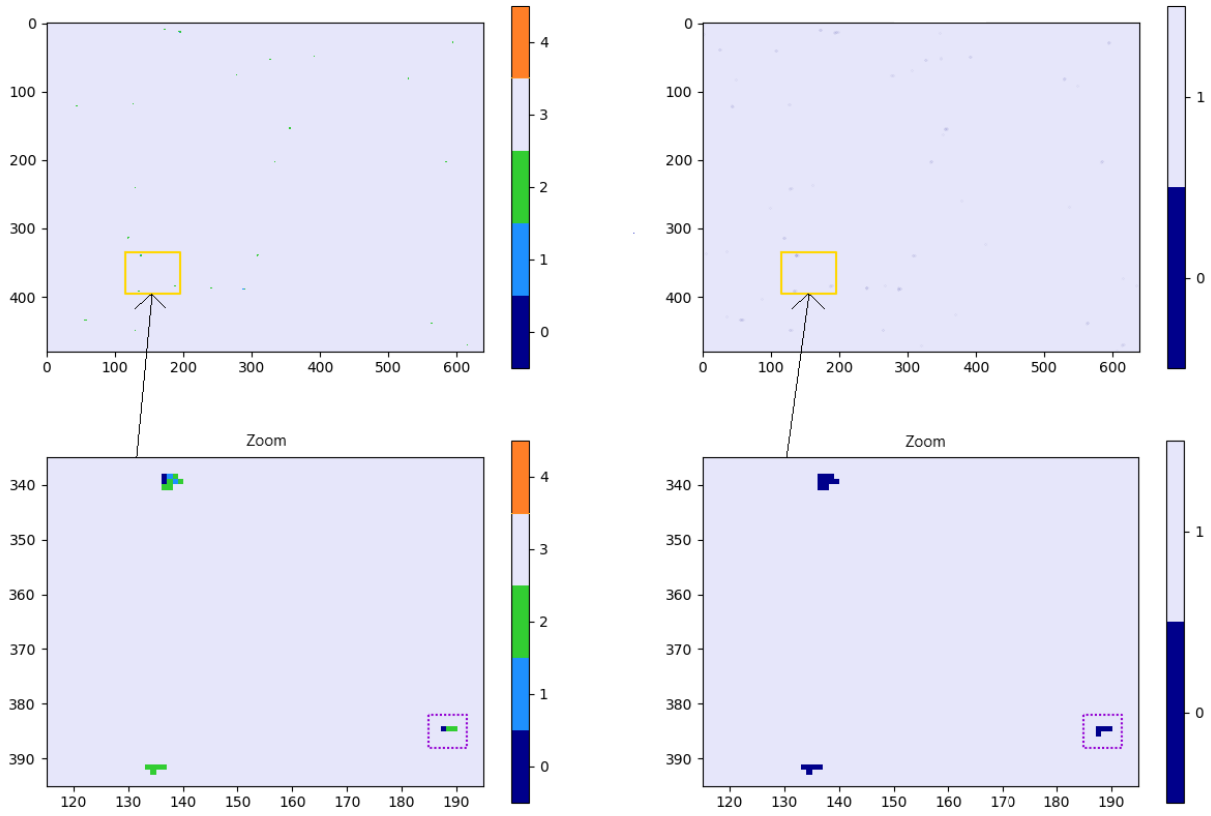
$$\mathbb{H} = \{\sigma_{\mathscr{C}^f}(P_{x,y}),\ (x, y) \in \Omega\}.$$

Pour détecter les pixels défectueux, nous proposons de considérer l'histogramme des RMSE $\mathbb{H}$ et d'exclure les quelques pixels dont la RMSE se distingue de la majorité.

La figure 3.9 montre l'histogramme des RMSE obtenu à partir du jeu $\mathscr{C}_1^f$. On constate qu'une très grande majorité des pixels a obtenu une RMSE qui se situe dans un certain intervalle $\mathbb{I}$, correspondant à la zone colorée en indigo clair, il s'agit de les considérer comme normaux, tous les autres pixels sont alors estampillés défectueux.

On définit un sous-ensemble $\mathbf{M} \subset \Omega$, qu'on appelle masque, tel que $\forall (x, y) \in \Omega$

$$(x, y) \in \mathbf{M} \iff \sigma_{\mathscr{C}^f}(P_{x,y}) \in \mathbb{I}. \tag{3.20}$$



(a) Carte des degrés

(b) Image du masque

FIGURE 3.8 – **Comparatif entre l'image de la carte des degrés et l'image du masque, choix du degré global.** Les données proviennent de l'ensemble $\mathscr{C}_1^f$. (a) En haut l'image de carte des degrés **C**, en bas un zoom sur la partie encadrée en jaune. Une très grande majorité des pixels a obtenu par la procédure fondée sur un test d'*overfitting* le degré $d = 3$. Le degré global obtenu après un vote est donc $D = 3$. (b) Le masque **M** (ensemble des pixels non défectueux) est obtenu à partir de l'histogramme de la figure 3.9. C'est un sous-ensemble de Ω auquel on peut associer une image de booléens (à droite), un pixel appartient au masque si et seulement si le niveau radiométrique de cette image vaut 1 (0 sinon). En haut l'image du masque, en bas un zoom sur la partie encadrée en jaune. Les pixels défectueux apparaissent souvent sous la forme de *clusters* (on en distingue 3 en bleu foncé sur l'image du bas). Parmi ces derniers, une majorité avait obtenu un degré $d \leq D-1$ (comme on peut l'observer sur l'image en bas à gauche). Dans la partie encadrée en pointillé pourpre (ligne du bas), on perçoit un pixel défectueux qui a obtenu un degré $d = D$. Il faudra analyser son comportement de manière plus précise (cf fig. 3.10) pour voir s'il est vraiment défectueux. Une très grande majorité des pixels (99.97%) ayant permis de déterminer le degré global fait partie du masque, ce qui corrobore a posteriori le choix du degré global.

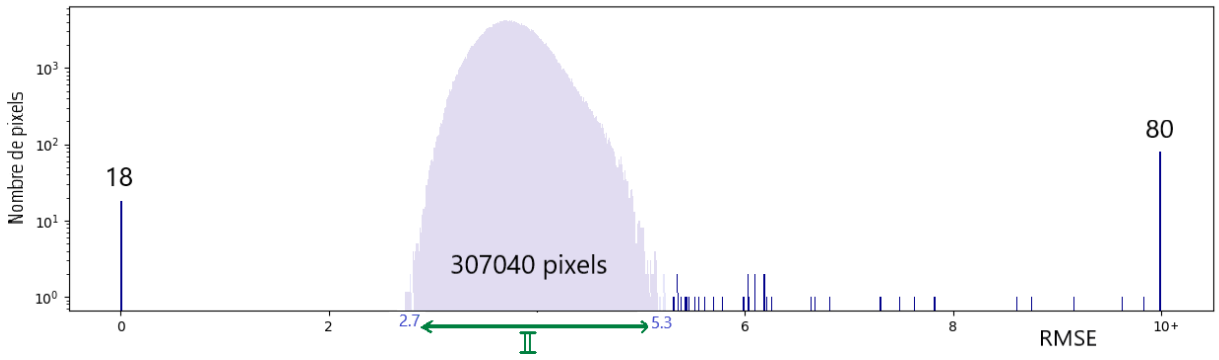


FIGURE 3.9 – **Création du masque de pixels non défectueux à partir de l'histogramme des RMSE.** Histogramme représentant la répartition des RMSE au sein de tous les pixels suite à une modélisation polynomiale de degré au plus D. L'échelle en ordonnée est logarithmique. Les données proviennent du sous-ensemble de calibration $\mathscr{C}_1^f$, le degré global $D = 3$ a été obtenu par vote. On dénombre en tout 307200 pixels. Une très grande majorité des pixels a obtenu une RMSE dans l'intervalle $\mathbb{I} = [2.7, 5.3]$ (représenté en vert), nous allons les considérer comme normaux. Les 160 autres sont considérés comme défectueux, on peut en effet exclure les pixels qui ont un comportement (observé) qui s'éloigne trop de leur modèle polynomial et ceux qui ont un comportement qui s'approche trop parfaitement de leur polynôme alors qu'il existe un certain bruit d'observation (en l'occurrence ces pixels ont ici des polynômes qui s'ajustent exactement aux observations, i.e. de RMSE nulle). Parmi ces pixels estampillés défectueux, 18 pixels ont obtenu une RMSE nulle (ce sont en fait des pixels cassés), et 80 ont obtenu une RMSE relativement excessive (au delà de 10). Un masque $\mathbf{M}$ est ainsi constitué contenant les 307040 pixels non défectueux (correspondant à la zone colorée en indigo clair).

Par conséquent un pixel défectueux $(x, y) \in \Omega$ est tel que $(x, y) \notin \mathbf{M}$. La figure 3.9 illustre le processus de création du masque à partir de l'histogramme des RMSE $\mathbb{H}$, la figure 3.8b représente le masque résultant. Nous ne pouvons garantir qu'il n'y ait aucuns faux positifs mais l'étude du comportement de tous les pixels (figure 3.10) nous laisse présager que tous les pixels défectueux ont bien été détectés.

Remarque 7 (Comment déterminer l'intervalle $\mathbb{I}$)**.** *La détermination de l'intervalle $\mathbb{I}$ est délicate à automatiser, et compte tenu du faible nombre de jeux de données à traiter nous avons préféré le faire par inspection visuelle. Nous avons essayé de modéliser la distribution des RMSE (fig. 3.9) par des lois (e.g. normale, log-normale, Gamma, inverse Gamma, Cauchy) mais selon le sous-ensemble de calibration considéré les résultats pouvaient être assez différents. L'établissement d'un intervalle de confiance à partir de moyenne/écart type ou médiane/écart médian absolu ne nous a pas permis également d'établir une automatisation fiable quant à l'établissement de l'intervalle (ce qui est logique puisque ces méthodes sont fondées sur l'hypothèse que la distribution est gaussienne).*

1.4.3 Étude du détecteur par analyse du comportement de l'ensemble des pixels

Dans cette section, nous tentons de savoir si l'information contenue dans la carte des degrés $\mathbf{C}$ permet de détecter les pixels défectueux et/ou de trouver des pixels défectueux

présents dans le masque $\mathbf{M}$ (faux positifs). Pour un degré donné d, on s'intéresse à la classe

$$C_d = \{(x, y) \in \Omega, \text{ t.q. } \mathbf{C}(x, y) = d\}.$$

Les figures 3.8 et 3.10 illustrent les résultats obtenus à partir du jeu de calibration $\mathscr{C}_1^f$. On a pu détecter au sein de tous ces ensembles C_d (pourvu qu'ils soient non vides) des pixels cassés ou au comportement irrégulier. Il n'est donc pas possible de détecter avec certitude tous les pixels défectueux à partir de la seule carte des degrés. En outre l'analyse du comportement des pixels de chaque classe (figure 3.10) n'a pas permis pour le jeu de données $\mathscr{C}_1^f$ de trouver des faux positifs (i.e. éléments défectueux présents dans masque). La création d'un masque $\mathbf{M}$ à partir de l'histogramme $\mathbb{H}$ est une méthode de détection des pixels défectueux a priori fiable. Dans la suite de ce chapitre, nous ne considérerons que les pixels du masque $\mathbf{M}$.

1.5 Modèle de correction par prise en compte des erreurs de mesure de la TFPA

Soient un pixel $(x, y) \in \mathbf{M}$ et un état $j \in [\![1, N_e]\!]$, on définit l'erreur de *fit* $\varepsilon_{x,y}(j) \in \mathbb{R}$ par

$$\varepsilon_{x,y}(j) = r_j^{(f)}(x, y) - P_{x,y}(\tau_{f,j}), \tag{3.21}$$

où $r_j^{(f)}$ désigne l'image brute du corps noir froid acquise lors du j$^{\text{ème état}}$, $\tau_{f,j}$ la mesure de TFPA associée, $P_{x,y}$ désigne le polynôme de degré au plus D solution de (3.9), le degré global D étant connu. L'erreur de *fit* s'apparente au bruit d'observation, pour un état donné j, on s'attend à ce que la distribution des erreurs de *fit* au sein des pixels du masque

$$\mathscr{E}(j) = \{\varepsilon_{x,y}(j), (x, y) \in \mathbf{M}\} \tag{3.22}$$

soit indépendante de l'état et de moyenne nulle.

La figure 3.11 montre cependant le contraire, $\mathscr{E}(j)$ dépend de l'état j et peut être de moyenne significativement non nulle (sont illustrés des résultats sur le jeu $\mathscr{C}_1^f$, de mêmes constatations ont été faites sur le jeu $\mathscr{C}_2^f$). Nous faisons l'hypothèse que cette dépendance résulte d'une erreur de la mesure de la TFPA $\tau_{f,j}$ établie à l'état j.

Dans la littérature, il a déjà été proposé une ré-estimation de la TFPA [10] en considérant un critère de qualité relatif à l'image corrigée (donc « scene-based »), notre idée est de chercher la TFPA corrigée τ_j avec des considérations « calibration-based ». Pour le $j^{\text{ème}}$ état nous écrivons

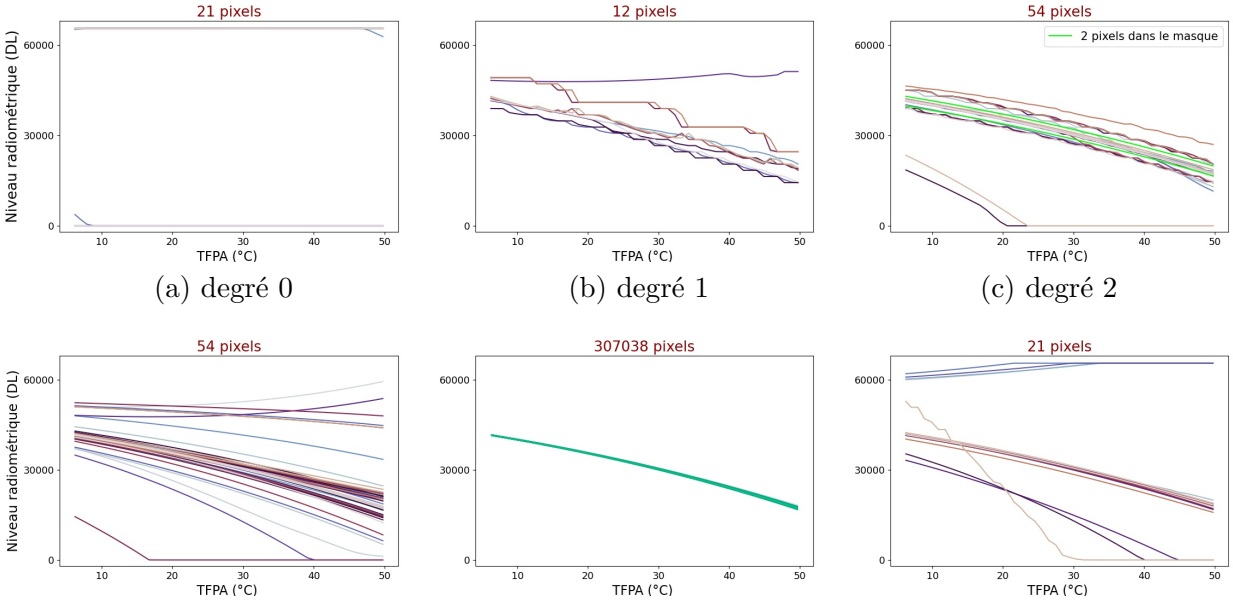
$$\tau_j = \tau_{f,j} + \Delta_j, \tag{3.23}$$

$\tau_{f,j}$ est la mesure de la TFPA et Δ_j désigne l'erreur de mesure associée à $\tau_{f,j}$.

L'idée est de s'intéresser à une nouvelle erreur de *fit* $\varepsilon'_{x,y}(j)$ définie par

$$\varepsilon'_{x,y}(j) = r_j^{(f)}(x, y) - P_{x,y}(\tau_j), \quad (x, y) \in \mathbf{M},\ j \in [\![1, N_e]\!], \tag{3.24}$$

et de déterminer l'ensemble des TFPA corrigées et l'ensemble des polynômes associés à chacun des pixels du masque en résolvant un problème de type minimisation des moindres carrés. Le qualificatif « delta » (en référence à l'imprécision définie dans (3.23)) sera utilisé



(a) degré 0 (b) degré 1 (c) degré 2

(d) degré 3 (hors du masque) (e) degré 3 (dans le masque) (f) degré supérieur à 4

FIGURE 3.10 – **Comportement des pixels selon le degré obtenu par le test d'*overfitting*.** Les données proviennent de l'ensemble $\mathscr{C}_1^f$. Les graphiques représentent la réponse de tous les pixels en fonction de la TFPA selon le degré obtenu à l'issue du test d'*overfitting* (valeur du pixel sur la carte des degrés **C**). On dénombre en tout 307200 pixels. Lors de la création du masque **M** (cf figure 3.9), 160 ont été estampillés défectueux, leurs courbes sont associées à des couleurs non vertes, les autres pixels font partie du masque et sont associés aux courbes à dominante verte. (a) Les 21 pixels ayant obtenu le degré 0 sont cassés (saturés), ils ont été détectés comme anormaux (défectueux). (b) Les 12 pixels ayant obtenu le degré 1 ont un comportement irrégulier, ils ont été détectés comme défectueux. (c) La moitié environ des 54 pixels ayant obtenu le degré 2 ont un comportement irrégulier, ils ont été détectés comme défectueux, parmi eux deux sont cassés. L'autre moitié (une vingtaine) a un comportement a priori régulier, deux pixels font partie du masque (courbes vertes), les autres (courbes non vertes) correspondent probablement à des pixels ayant été considérés à tord comme anormaux (faux négatifs). (d) La plupart des pixels défectueux (hors du masque) ayant obtenu le degré 3 ont un comportement régulier (une cinquantaine). Les 2 courbes les plus en bas montrent cependant 2 pixels cassés, celle juste au dessus montre un pixel qui arrive presque à saturation basse, il nous semble potentiellement défectueux. (e) Tous les pixels faisant partie du masque et ayant obtenu le degré 3 ont un comportement régulier et particulièrement similaire. Ils représentent 99.94% du nombre total de pixels. (f) Les 21 pixels ayant obtenu un degré supérieur ou égal à 4 ont été détectés comme défectueux, quelques uns sont cassés, une dizaine a un comportement régulier. Les degrés considérés ici sont 4, 5 et 6, on a pu observer des pixels cassés pour chacun des ces degrés. Parmi les 160 pixels estampillés défectueux (hors du masque **M**), une moitié environ a un comportement a priori régulier, on peut les considérer comme des faux négatifs. Parmi les pixels du masque **M**, nous n'avons a fortiori observé aucun pixel au comportement irrégulier ou cassé. Le nombre de faux positifs est donc nul, ce qui est un gage de fiabilité. De manière générale, on a pu observer, à tous les degrés, des pixels soit au comportement irrégulier soit cassés. Il serait alors impossible de détecter tous les pixels défectueux à partir de la seule carte des degrés **C**. La création du masque **M** à partir de l'histogramme des RMSE est un moyen sûr pour détecter tous les pixels défectueux.

pour désigner cette correction de la mesure de la TFPA. On souhaite que la distribution des erreurs de *fit* corrigées

$$\mathscr{E}'(j) = \{\varepsilon'_{x,y}(j),\ (x,y) \in \mathbf{M}\} \tag{3.25}$$

ne dépende plus de l'état j.

1.5.1 Modèle de la correction delta

Le degré global D étant fixé, on propose de trouver conjointement $\mathscr{T} = (\tau_j)_{j\in[\![1,N_e]\!]}$ et $\mathbf{P} = (P_{x,y})_{(x,y)\in\mathbf{M}}$ en minimisant la somme des carrés de toutes les erreurs de *fit* corrigées définies par (3.24).

Le problème d'optimisation relatif à la correction delta d'ordre D s'écrit

$$(\mathbf{P}, \mathscr{T}) \in \underset{(\mathbf{Q},\mathscr{T})\in\Gamma}{\operatorname{argmin}} \sum_{(x,y)\in\mathbf{M}} \sum_{j=1}^{N_e} \Big(r_j^{(f)}(x,y) - Q_{x,y}(\tau_j)\Big)^2, \tag{3.26}$$

avec

$$\Gamma = \Big(\mathbb{R}_D[X]\Big)^{\mathbf{M}} \times \mathbb{V} \tag{3.27}$$

$$\mathbb{V} = \{\mathscr{T} \in \mathbb{R}^{N_e},\ \text{t.q. } \tau_1 = \tau_{1,1}\}. \tag{3.28}$$

La contrainte $\tau_1 = \tau_{f,1}$ dans (3.28) est présente pour éviter que le problème (3.26) ne soit mal posé, elle équivaut à dire que la TFPA mesurée lors du premier état est correcte (elle sert donc de référence). La correction delta consiste à modéliser les images brutes à l'aide de polynômes dépendant de la TFPA, en introduisant des valeurs de TFPA d'observation corrigées. Si l'on admet que la dépendance en fonction de l'état j de l'erreur de *fit* $\mathscr{E}(j)$, définie par (3.22), résulte d'une imprécision sur la mesure de la TFPA, les polynômes issus de la correction delta (3.26) devraient fournir une estimation plus précise que ceux obtenus par la modélisation polynomiale « directe » des images brutes (3.9).

1.5.2 Résolution algorithmique de la correction delta

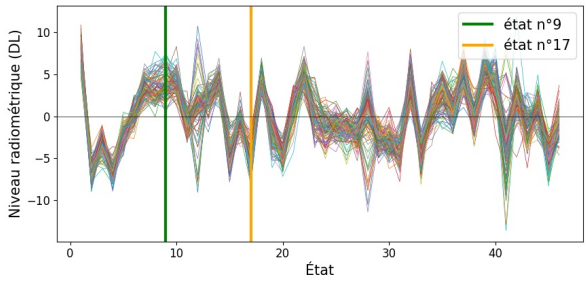
La résolution de (3.26) se fait par un schéma alterné. On initialise [5] $\mathscr{T}^{(0)}$ puis on itère sur k.

A l'étape $k \geq 1$, on suppose déterminé $\mathscr{T}^{(k-1)}$ et on résout

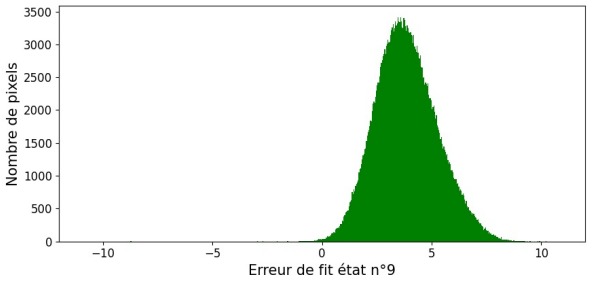
$$\begin{cases} P_{x,y}^{(k)} \in \underset{Q\in\mathbb{R}_D[X]}{\operatorname{argmin}} \sum_{j=1}^{N_e} \Big(r_j^{(f)}(x,y) - Q(\tau_j^{(k-1)})\Big)^2, \quad \forall (x,y) \in \mathbf{M} & \text{(3.29a)} \\ \mathscr{T}^{(k)} \in \underset{\mathscr{T}\in\mathbb{V}}{\operatorname{argmin}} \sum_{(x,y)\in\mathbf{M}} \sum_{j=1}^{N_e} \Big(r_j^{(f)}(x,y) - P_{x,y}^{(k)}(\tau_j)\Big)^2, & \text{(3.29b)} \end{cases}$$

La résolution de (3.29a) se fait avec la formule (3.13) puisqu'il s'agit d'une régression polynomiale.

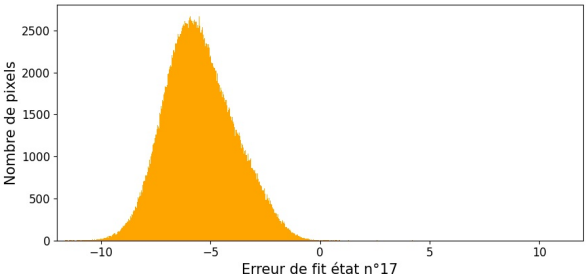
5. $\mathscr{T}^{(0)}$ doit avoir des composantes distinctes.



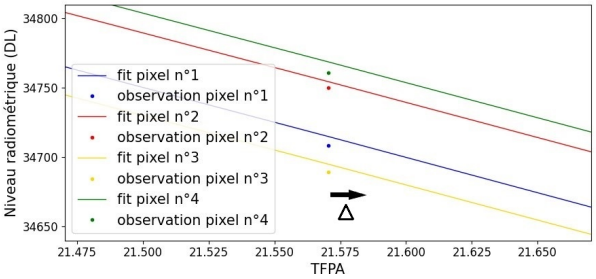
(a) Erreur de *fit* pour 100 pixels selon l'état



(b) Histogramme des erreurs de *fit* (état 9)



(c) Histogramme des erreurs de *fit* (état 17)



(d) Explication de la correction delta (état 17)

FIGURE 3.11 – **Analyse de l'erreur de *fit* selon l'état avant correction delta.** Les données proviennent de l'ensemble $\mathscr{C}_1^f$, Les images du corps noir froid ont été acquises pour $N_e = 46$ états. (a) Erreur de *fit* pour 100 pixels non défectueux et choisis au hasard selon l'état. On remarque une tendance commune au sein de tous ces pixels ; l'erreur de *fit* d'un pixel semble dépendre de l'état. Ainsi pour l'état 9 (ligne verticale verte), l'erreur de *fit* pour ces 100 pixels est toujours positive. Pour l'état 17 (verticale orange), l'erreur de *fit* est toujours négative. On se serait attendu à ce que les erreurs de *fit* obéissent à une loi de probabilité indépendante de l'état et de moyenne nulle. (b) Histogramme des erreurs de *fit* de tous les pixels non défectueux pour l'état 9. La très grande majorité des pixels a une erreur de *fit* positive. (c) Histogramme des erreurs de *fit* pour l'état 17. Tous les pixels ont une erreur de *fit* négative. (d) Valeurs radiométriques observées et polynôme de degré 3 obtenu par régression polynomiale en fonction de la TFPA pour 4 pixels (zoom autour de la valeur de la TFPA correspondant à l'état 17). La tendance commune constatée précédemment est confirmée. Pour cet état, les observations ont tendance à se situer en déca de la courbe du polynôme (i.e. erreurs de *fit* négatives). Nous faisons l'hypothèse que ces anomalies (erreur de *fit* dépendant de l'état) sont dues à une imprécision de la mesure de la TFPA réalisée à chaque état. Si, dans la figure (d), on translate les 4 observations par un même vecteur horizontal Δ correspondant à l'imprécision de la mesure TFPA, cette tendance commune disparaitrait. En faisant cette translation en TFPA sur tous les pixels, l'histogramme (c) deviendrait centré. En cherchant pour chacun des états les translations adéquates en TFPA, nous espérons retrouver les vraies TFPA et ne plus observer cette anormale dépendance de l'erreur de *fit* selon l'état. Nous avons baptisé « correction delta » cette correction de la mesure de TFPA.

La résolution de (3.29b) est équivalente, pour des raisons de séparabilité des variables, au système

$$\begin{cases} \tau_2^{(k)} & \in \underset{z\in\mathbb{R}}{\operatorname{argmin}} \sum_{(x,y)\in\mathbf{M}} \left(r_2^{(f)}(x,y) - P_{x,y}^{(k)}(z)\right)^2, \\ \vdots & \\ \tau_{N_e}^{(k)} & \in \underset{z\in\mathbb{R}}{\operatorname{argmin}} \sum_{(x,y)\in\mathbf{M}} \left(r_{N_e}^{(f)}(x,y) - P_{x,y}^{(k)}(z)\right)^2, \end{cases} \tag{3.30}$$

où l'on a posé $\mathscr{T}^{(k)} = (\tau_j^{(k)})_{j\in[\![1,N_e]\!]}$.

La résolution de (3.30) se fait ligne par ligne. Pour chaque ligne, il faut trouver le minimum d'un polynôme. On s'intéresse donc à ses points critiques en utilisant l'algorithme de Newton.

Formellement on définit le fonction $F_j : \mathbb{R} \to \mathbb{R}$ par

$$\forall X \in \mathbb{R}, \quad F_j(X) = \sum_{(x,y)\in\mathbf{M}} \left(r_j^{(f)}(x,y) - P_{x,y}^{(k)}(X)\right)^2, \tag{3.31}$$

F_j est une fonction polynôme de degré au plus $2D$, elle est donc dérivable et sa dérivée F_j' est un polynôme de degré au plus $2D-1$. Pour $j = 2, \ldots, N_e$, la résolution de la $j^{\text{ème}}$ ligne de (3.30) s'écrit

$$\tau_j^{(k)} \in \underset{X\in\mathbb{R}}{\operatorname{argmin}} F_j(X), \tag{3.32}$$

ce qui équivaut à résoudre

$$\tau_j^{(k)} \in \underset{X\in\mathbb{R}}{\operatorname{argmin}} \{F_j(X),\ F_j'(X) = 0\}. \tag{3.33}$$

Il nous faut donc résoudre

$$F_j'(X) = 0, \tag{3.34}$$

F_j' étant un polynôme de degré impair, il admet au moins une racine. Après calculs, il est possible d'expliciter les coefficients du polynôme F_j'. F_j' étant de degré au plus $2D-1$, on peut l'écrire

$$F_j'(X) = a_{0,j} + a_{1,j}X + \ldots + a_{2D-1,j}X^{2D-1} \quad X \in \mathbb{R}, \tag{3.35}$$

où les coefficients $a_{i,j}$ $(0 \le i \le 2D-1)$ sont tels que

$$\begin{cases} a_{i,j} = (i+1)(\sum_{l=0}^{i+1} C_l\, C_{i+1-l}) & \text{pour } 0 \le i \le D-1, \\ a_{i,j} = (i+1)(\sum_{l=0}^{2D-i-1} C_{D-l}\, C_{i+1-D+l}) & \text{pour } D \le i \le 2D-1. \end{cases} \tag{3.36}$$

Les coefficients C_l présents dans (3.36) sont définis par

$$\begin{cases} C_0 = \sum_{(x,y)\in\mathbf{M}} \left((P_{x,y}^{(k)})_0 - r_j^{(f)}(x,y)\right), \\ C_l = \sum_{(x,y)\in\mathbf{M}} (P_{x,y}^{(k)})_l, \quad 1 \le l \le D, \end{cases} \tag{3.37}$$

où $(P_{x,y}^{(k)})_l$ désigne le coefficient de degré l du polynôme $P_{x,y}^{(k)}$.

Une racine du polynôme F_j' est ensuite déterminée grâce à l'algorithme de Newton dont le schéma numérique [6] s'écrit

$$\begin{cases} X^{(0)} = \tau_j^{(k-1)}, \\ X^{(p+1)} = X^{(p)} - \dfrac{F_j'(X^{(p)})}{F_j''(X^{(p)})}, \quad p \geq 0. \end{cases} \tag{3.38}$$

On note $\overline{X}$ la valeur de convergence de ce schéma, $\overline{X}$ est un point critique de F_j mais rien ne nous garantit qu'il s'agisse d'un point en lequel F_j atteint son minimum. Pour contourner cet écueil nous proposons de ne considérer $\overline{X}$ comme solution de (3.32) que si $F_j(\overline{X}) \leq F_j(\tau_j^{(k-1)})$.

Pour l'étape k, la résolution de la j^{ème} ligne de (3.30) se fait donc de la manière suivante

$$\begin{cases} \tau_j^{(k)} = \overline{X} & \text{si} \quad F_j(\overline{X}) \leq F_j(\tau_j^{(k-1)}), \quad \text{(3.39a)} \\ \tau_j^{(k)} = \tau_j^{(k-1)} & \text{sinon.} \quad \text{(3.39b)} \end{cases}$$

En conclusion, on obtient $\mathscr{T}^{(k)} = (\tau_j^{(k)})_{j \in [\![1, N_e]\!]}$ comme solution de (3.29b).

Remarque 8 (Minimum du polynôme F_j)**.** *Sur les différents jeux de données dont nous disposons, le cas* (3.39b) *ne s'est jamais présenté, ce qui signifie qu'en pratique* $\overline{X}$ *correspond à un minimum local de* F_j*.*

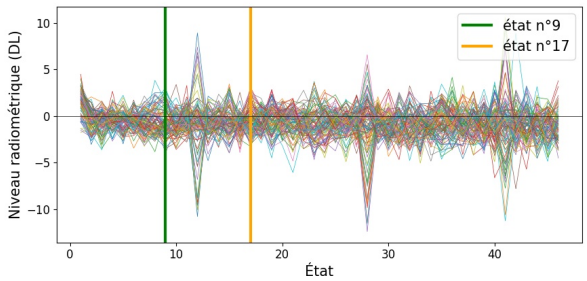
1.5.3 Résultats expérimentaux de la correction delta, insuffisance du modèle paramétrique.

Les résultats obtenus et leur interprétation sont différents selon le jeu de données considéré.

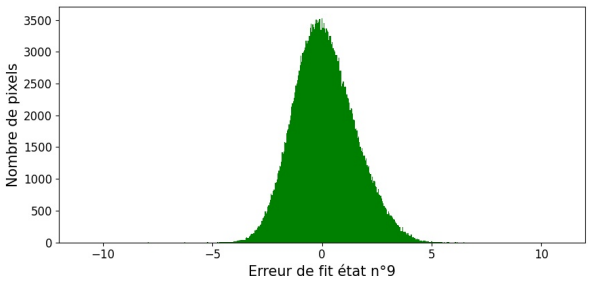
- Pour l'ensemble de calibration $\mathscr{C}_1^f$, la figure 3.12 montre que des anomalies constatées dans la figure 3.11 ont disparu, ainsi la distribution des erreurs de *fit* corrigées $\mathscr{E}'(j)$ définie par (3.25) est centrée. Par ailleurs l'amplitude des corrections delta est à comparer à la précision de la sonde de température (0.1°C), la figure 3.13a montre que les corrections d'imprécision de mesure TFPA sont plausibles. Le problème relatif à la correction delta (3.26), en considérant $Ne - 1 = 45$ paramètres supplémentaires au problème (3.9) qui en compte $(D+1)\,\text{card}(\mathbf{M}) \approx 10^6$, permet de réduire l'erreur quadratique de *fit* globale d'un facteur d'environ 3.
- En ce qui concerne le jeu de données $\mathscr{C}_2^f$, si l'erreur quadratique de *fit* globale est réduite d'un facteur d'environ 3, les imprécisions de mesures obtenues sont bien au-delà du seuil d'erreur fourni par le fabriquant et dépendent anormalement de l'état (cf. figure 3.13b).

Les résultats obtenus sur le jeu $\mathscr{C}_2^f$ posent question. Sans doute les anomalies illustrées par la figure 3.11 [7] ne sont pas uniquement le fruit d'une imprécision de la mesure de la TFPA, des flux parasites (inévitables malgré une procédure en « étuve stabilisée » censée

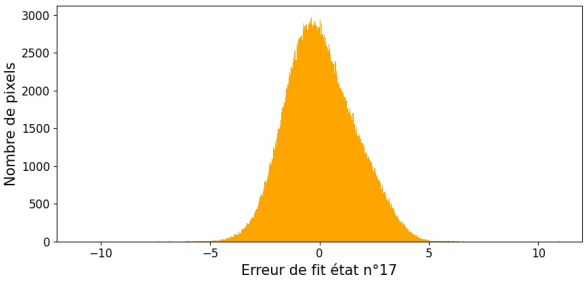
6. On suppose que $F_j''(X^{(p)}) \neq 0$, ce qui en pratique sera vérifié.
7. Sur le jeu $\mathscr{C}_2^f$, les mêmes anomalies ont été observées.



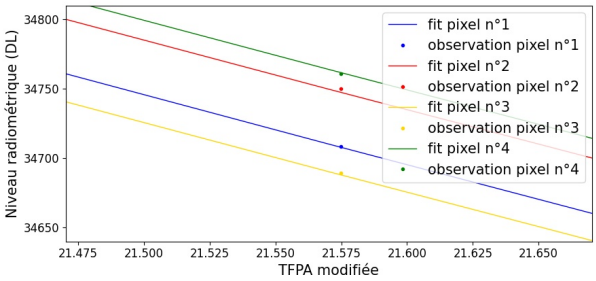
(a) Erreur de *fit* pour 100 pixels selon l'état



(b) Histogramme des erreurs de *fit* (état 9)

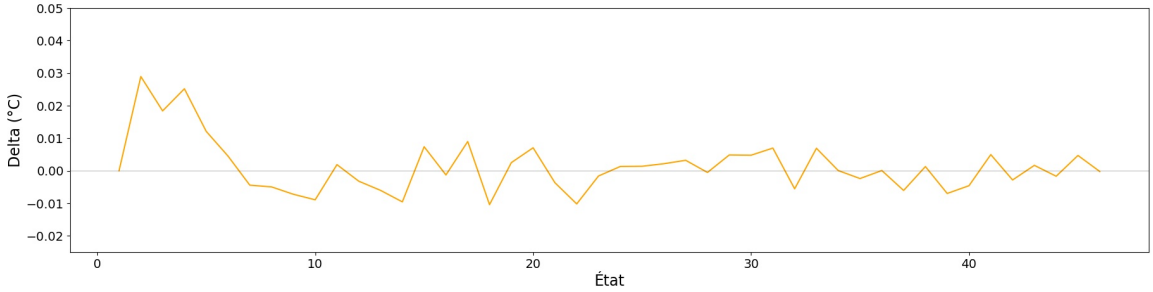


(c) Histogramme des erreurs de *fit* (état 17)

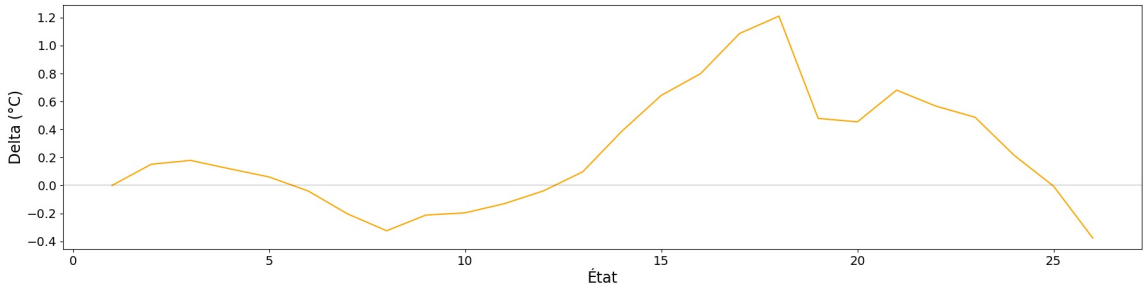


(d) Explication de la correction delta (état 17)

FIGURE 3.12 – **Erreur de *fit* après correction delta.** Les données proviennent de l'ensemble $\mathscr{C}_1^f$, Les images du corps noir froid ont été acquises pour $N_e = 46$ états. Cette figure est à comparer avec la figure 3.11. (a) Erreur de *fit* de 100 pixels non défectueux selon l'état. Pour un état donné, l'erreur de *fit* peut être aussi bien positive que négative. On ne peut pas néanmoins dire que les erreurs de *fit* sont indépendante d'un état à un autre puisque on distingue quelques états (e.g. 12, 28 ou 41) où leur amplitude est plus grande. (b) Histogramme des erreurs de *fit* composées de tous les pixels non défectueux pour l'état 9. (c) Histogramme des erreurs de *fit* pour l'état 17. Les histogrammes sont désormais centrés. (d) Valeurs radiométriques observées et polynôme de degré 3 obtenu par correction delta en fonction de la TFPA pour 4 pixels (zoom autour de la valeur de TFPA correspondant à l'état 17). Les observations ne sont plus systématiquement situées en deçà des courbes associées aux polynômes déterminés par la correction delta (qui sont très légèrement différents des polynômes déterminés par la modélisation polynomiale des images brutes). En comparant avec la figure 3.11, il est manifeste que la correction delta aura permis de recentrer les erreurs de *fit* relatives à chaque état.



(a) Ensemble de calibration $\mathscr{C}_1$



(b) Ensemble de calibration $\mathscr{C}_2$

FIGURE 3.13 – **Valeurs de la correction delta obtenues pour deux ensembles de calibration. Remise en question de la correction delta.** Valeurs de l'imprécision de la mesure TFPA déterminées par la correction delta selon les différents états pour 2 sous-ensembles de données de calibration. (a) Pour l'ensemble $\mathscr{C}_1^f$, l'imprécision de la mesure TFPA (delta) se situe entre -0.02°C et +0.03°C. Selon le fabriquant, l'imprécision maximale de la mesure de la sonde TFPA est de 0.1°C en valeur absolue. (b) Pour l'ensemble $\mathscr{C}_2^f$, l'imprécision de mesure se situe entre -0.6°C et +1°C. On est nettement au-delà de l'imprécision maximale communiquée par la fabriquant. En outre on observe que les valeurs de la correction delta sont positives entre les états 13 et 25, la correction delta semble donc dépendre de l'état. S'il s'agissait d'une imprécision de mesure, cela ne serait pas le cas. Une telle régularité pourrait être le fruit d'un degré considéré pour le choix des polynômes pas adapté (*overfitting*) mais le test avec tous les autres degrés a donné le même genre de constatations (valeurs trop grandes et régularité selon l'état). Pour le jeu $\mathscr{C}_2^f$, l'interprétation de la correction delta comme correction d'une mesure de TFPA est peu plausible.

réduire leur influence) différents d'un état à un autre ont pu influencer l'erreur de *fit* et contribuer à cette dépendance selon l'état.

La seule considération de la TFPA comme paramètre peut s'avérer problématique car insuffisante. Il faudrait pouvoir considérer d'autres paramètres mais cela semble impossible tant l'ensemble des paramètres θ est grand et mal défini. Nous proposerons dans le chapitre suivant un modèle de correction non paramétrique qui sera capable de prendre en compte cet ensemble potentiellement infini de paramètres θ.

Dans la suite nous allons observer et analyser les résultats obtenus en terme de correction d'une image brute de scène uniforme acquise à une certaine TFPA τ_0. Si on garde l'hypothèse selon laquelle la TFPA mesurée τ_0 est imprécise, nous déterminerons la TFPA corrigée $\widehat{\tau_0}$ de sorte que l'image corrigée ait une certaine régularité. Nous pourrons ainsi comparer les méthodes que nous venons de vous présenter (modélisation polynomiale des images brutes avec ou sans correction delta) avec la méthode *shutterless polynomial* qui fait office de référence dans la littérature.

1.6 Correction d'une scène uniforme, comparaison des différentes méthodes

Considérons un k^{ème} ensemble de calibration $\mathscr{C}_k$, comme nous l'avions vu précédemment (cf. remarque 4), $\mathscr{C}_k$ est composé des données relatives à un corps noir froid $\mathscr{C}_k^f$, un corps noir chaud $\mathscr{C}_k^c$ mais également celles d'un autre corps noir $\mathscr{C}_k^t$ appelé sous-ensemble de test. Le sous-ensemble $\mathscr{C}_k' = \mathscr{C}_k^f \cup \mathscr{C}_k^c$ permet de déterminer les composantes du modèle affine $\mathcal{R}_\tau$ et $\mathcal{D}_\tau$, le sous-ensemble $\mathscr{C}_k^t$ permet d'établir des mesures de la performance de correction.

Nous allons comparer trois méthodes de correction d'une scène uniforme :

— la méthode *shutterless polynomial* qui fait office de référence dans l'état de l'art [8],

— la modélisation polynomiale des images brutes sans correction delta (équation (3.9)),

— la modélisation polynomiale des images brutes avec correction delta (équation (3.26)).

Pour ces trois méthodes le même degré global D sera choisi (comme décrit dans le paragraphe 1.3.2.4) et le même masque de pixels non défectueux $\mathbf{M}$ (comme décrit dans 1.4.2) sera considéré..

1.6.1 Détermination de la responsivité et décalage selon la TFPA

On considère ainsi trois méthodes pour établir les composantes du modèle affine responsivité $\mathcal{R}_\tau$ et décalage $\mathcal{D}_\tau$. Pour la première (*shutterless polynomial*), responsivité et décalage sont établis pour N_e valeurs de TFPA $\tau_1, ..., \tau_{N_e}$ par (3.8), et une régression polynomiale de degré D permet de déterminer $\mathcal{R}_\tau$ et $\mathcal{D}_\tau$ pour tout τ. En ce qui concernent les deux autres, on détermine pour chacun des pixels $(x, y) \in \mathbf{M}$, les polynômes $P_{x,y}^f$ (pour le corps noir froid) et $P_{x,y}^c$ (pour le corps noir chaud) en résolvant le problème (3.9) pour la modélisation polynomiale des images brutes ou bien en résolvant (3.26) pour la modélisation polynomiale des images brutes avec correction delta (cf. figure 3.2).

On applique alors la formule de la calibration à deux points (2.21) pour déterminer

responsivité $\mathcal{R}_\tau$ et offset $\mathcal{D}_\tau$ pour toute TFPA τ, ce qui donne

$$\begin{cases} \mathcal{R}_\tau(x,y) = \dfrac{P^c_{x,y}(\tau) - P^f_{x,y}(\tau)}{T_c - T_f}, & \text{(3.40a)} \\ \mathcal{D}_\tau(x,y) = \dfrac{T_c\, P^f_{x,y}(\tau) - T_f\, P^c_{x,y}(\tau)}{T_c - T_f}. & \text{(3.40b)} \end{cases}$$

1.6.2 Correction de l'image brute d'une scène uniforme

Considérons un sous-ensemble de test $\mathscr{C}^t_k = \{(r^{(t)}_j, \tau_{t,j}, T_t),\ 1 \leq j \leq N_e\}$ tel que défini par (3.4). La formule (3.1) permet de corriger tout élément de $\mathscr{C}^t_k$. Pour $j \in [\![1, N_e]\!]$, l'image corrigée $S^{(t)}_j$ de l'image brute de la scène uniforme $r^{(t)}_j$ est

$$S^{(t)}_j(x,y) = \frac{r^{(t)}_j(x,y) - \mathcal{D}_{\tau_{t,j}}(x,y)}{\mathcal{R}_{\tau_{t,j}}(x,y)}, \quad (x,y) \in \mathbf{M}. \tag{3.41}$$

Sur les figures 3.14 et 3.15, nous avons constaté des valeurs de responsivité sensiblement différentes selon la méthode envisagée. Ces valeurs différentes de responsivité ont une répercussion sur l'échelle radiométrique de l'image corrigée $S^{(t)}_j$ dans (3.41), ce qui fausse la comparaison entre les différentes méthodes. Pour compenser ces différences, il nous semble plus judicieux de multiplier l'image corrigée $S^{(t)}_j$ par la moyenne (au sein des pixels du masque $\mathbf{M}$) de la responsivité[8]. Le modèle de correction s'écrit donc

$$U^{(t)}_j(x,y) = \mu(\mathcal{R}_{\tau_{t,j}}) \frac{r^{(t)}_j(x,y) - \mathcal{D}_{\tau_{t,j}}(x,y)}{\mathcal{R}_{\tau_{t,j}}(x,y)}, \quad (x,y) \in \mathbf{M}, \tag{3.42}$$

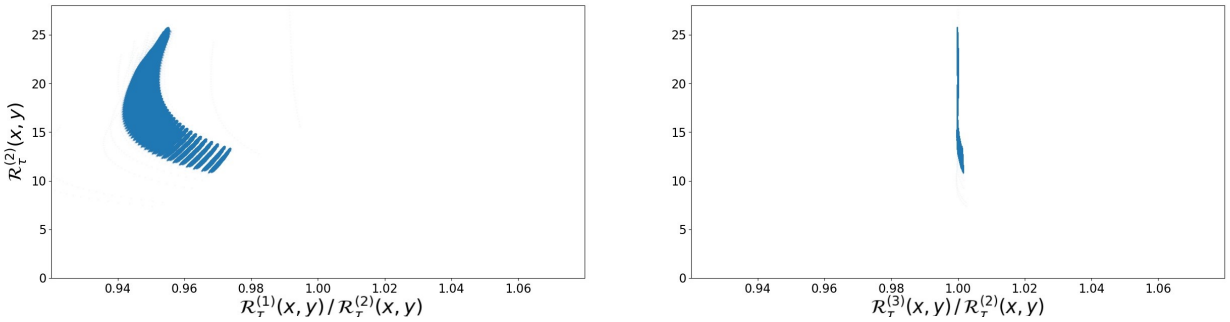
où $U^{(t)}_j$ désigne l'image corrigée en niveau de gris (unité DL). (3.42) consiste donc à corriger l'image brute $r^{(t)}_j$ acquise à la TFPA $\tau_{t,j}$ avec les composantes du modèle affine responsivité et décalage estimées pour la TFPA $\tau_{t,j}$. Or nous ne sommes pas certains que la mesure de $\tau_{t,j}$ soit précise. Il est préférable d'essayer de déterminer une TFPA corrigée et ce d'autant plus que l'on considère le modèle polynomial avec correction delta qui permet d'estimer responsivité et décalage après avoir corrigé des mesures de TFPA. Pour comparer la correction d'une image par un modèle, il est utile de définir une certaine métrique de qualité d'image.

Métrique associée à la qualité d'une correction de scène uniforme

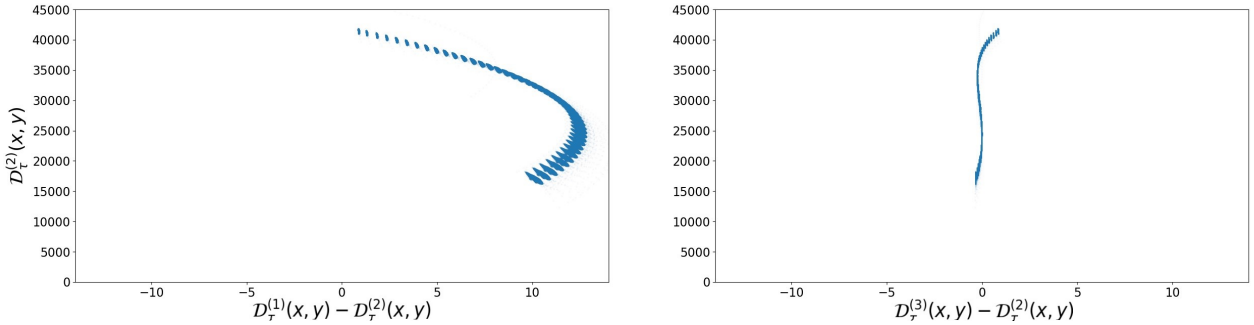
Soit $F : \mathbb{R}^{\mathbf{M}} \to \mathbb{R}$ une métrique de qualité d'image. S'agissant de la correction d'une scène uniforme, nous prendrons ici F comme étant l'écart-type spatial (au seins des pixels du masque), i.e.

$$\forall U \in \mathbb{R}^{\mathbf{M}}, \quad F(U) = \sqrt{\frac{\sum\limits_{(x,y)\in\mathbf{M}} U(x,y)^2}{\mathrm{card}(\mathbf{M})} - \left(\frac{\sum\limits_{(x,y)\in\mathbf{M}} U(x,y)}{\mathrm{card}(\mathbf{M})}\right)^2}, \tag{3.43}$$

8. Ce qui revient à considérer modèle de correction affine avec gain et offset (2.17)

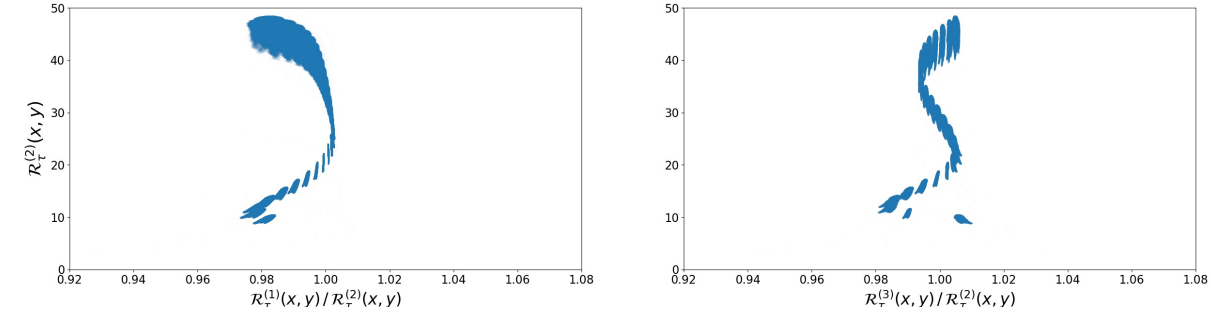


(a) Comparaison des responsivités (quotient)

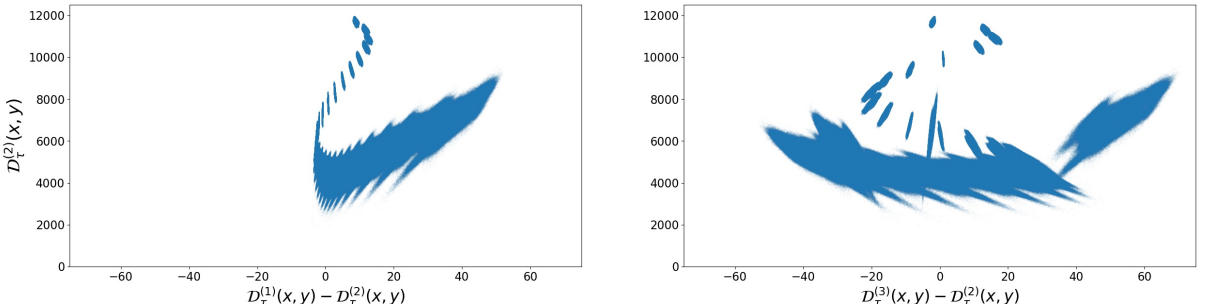


(b) Comparaison des décalages (différence)

FIGURE 3.14 – **Comparaisons des responsivités et décalages obtenus par trois méthodes (jeu de données $\mathscr{C}_1$).** Soient un pixel (x, y) du masque $\mathbf{M}$ et une TFPA $\tau \in \{\tau_1, .., \tau_{N_e}\}$, $\mathcal{R}_\tau^{(1)}(x,y)$ et $\mathcal{D}_\tau^{(1)}(x,y)$ désignent respectivement responsivité et décalage du pixel (x, y) à la TFPA τ déterminés par la méthode *shutterless polynomial*, $\mathcal{R}_\tau^{(2)}(x,y)$ et $\mathcal{D}_\tau^{(2)}(x,y)$ déterminés par la modélisation polynomiale des images brutes, $\mathcal{R}_\tau^{(3)}(x,y)$ et $\mathcal{D}_\tau^{(3)}(x,y)$ déterminés par la modélisation polynomiale des images brutes avec correction delta. Sur les 4 graphiques, le nuage est composé de N_e card($\mathbf{M}$) points bleus (avec $Ne = 46$ et card($\mathbf{M}$)=307040). (a) Sur le graphique de gauche (respectivement de droite) chaque point bleu à pour ordonnée $\mathcal{R}_\tau^{(2)}(x,y)$ et pour abscisse le quotient $\frac{\mathcal{R}_\tau^{(1)}(x,y)}{\mathcal{R}_\tau^{(2)}(x,y)}$ (respectivement $\frac{\mathcal{R}_\tau^{(3)}(x,y)}{\mathcal{R}_\tau^{(2)}(x,y)}$). Sur le graphique de gauche le nuage est composé de points dont les abscisses sont comprises entre 0.94 et 0.98. La responsivité établie par la méthode *shutterless polynomial* est donc entre 2 et 6% inférieure à la responsivité établie par la modélisation des images brutes. Sur le graphique de droite, les abscisses des points sont très proches de 1. La responsivité établie par la modélisation des images brutes est donc très proche de la responsivité établie par la correction delta. (b) Sur le graphique de de gauche (respectivement de droite) chaque point a pour ordonnée $\mathcal{D}_\tau^{(2)}(x,y)$ et pour abscisse la différence des décalages $\mathcal{D}_\tau^{(1)}(x,y) - \mathcal{D}_\tau^{(2)}(x,y)$ (respectivement $\mathcal{D}_\tau^{(3)}(x,y) - \mathcal{D}_\tau^{(2)}(x,y)$). Sur le graphique de gauche les points ont des abscisses non centrées autour zéro. Le décalage obtenu par la méthode *shutterless polynomial* diffère de celui obtenu par la modélisation polynomiale des images brutes (entre 0 et 13 DL donc parfois au delà du bruit de mesure qui vaut environ 1.5 DL). Sur le graphique de droite, les points ont des abscisses proches de zéro (inférieures en valeur absolue au bruit de mesure). Le décalage établi par la modélisation des images brutes est donc très proche de celui établi par la modélisation polynomiale des images brutes avec correction delta. En conclusion, pour le jeu de données $\mathscr{C}_1$, responsivité et décalage ne sont que très peu différents selon si ils sont obtenus après modélisation polynomiale des images brutes avec correction delta ou sans. Il existe cependant des différences non négligeables lorsque l'on compare avec le modèle *shutterless polynomial*.



(a) Comparaison des responsivités (quotient)



(b) Comparaison des décalages (différence)

FIGURE 3.15 – **Comparaisons des responsivités et décalages obtenus par trois méthodes (jeu de données $\mathscr{C}_2$).** Sur les quatre graphiques le nuage de point est établi selon le même procédé que décrit sur la figure 3.14 avec $N_e = 26$ et card($\mathbf{M}$) = 307104. En reprenant la même analyse, nous pouvons conclure que, pour le jeu de données $\mathscr{C}_2$, responsivité et décalage sont différents selon si ils ont été déterminés par l'une des trois méthodes (*shutterless polynomial*, modélisation polynomiale des images brutes, correction delta).

plus $F(U)$ sera petit, meilleure sera la qualité de l'image corrigée U.

Pour déterminer la TFPA corrigée, nous nous inspirons de l'article de Cao et Tisse [10] dont nous avions parlé dans le chapitre précédent (cf. (2.27)), le problème consiste à calculer

$$\widehat{\tau_{t,j}} \in \underset{\tau\in\mathbb{R}}{\operatorname{argmin}}\ F\Big(\mu(\mathcal{R}_\tau)\frac{r_j^{(t)}-\mathcal{D}_\tau}{\mathcal{R}_\tau}\Big). \tag{3.44}$$

L'image corrigée $\widehat{U_j^{(t)}}$ s'obtient alors par

$$\widehat{U_j^{(t)}} = \mu(\mathcal{R}_{\widehat{\tau_{t,j}}})\frac{r_j^{(t)}-\mathcal{D}_{\widehat{\tau_{t,j}}}(x,y)}{\mathcal{R}_{\widehat{\tau_{t,j}}}}. \tag{3.45}$$

$\sigma_j^{(t)}$, une mesure de performance de la correction de l'image brute du corps noir t acquise au j^{ème} état, est donnée par

$$\sigma_j^{(t)} = F(\widehat{U_j^{(t)}}). \tag{3.46}$$

La figure 3.16 illustre des résultats obtenus à partir du jeu $\mathscr{C}_1$. La modélisation polynomiale des images brutes donne présente des performances de correction très légèrement meilleures que la méthode *shutterless polynomial*. La modélisation polynomiale des images brutes avec correction delta donne une correction identique à la modélisation polynomiale des images brutes (sans delta), cela s'explique par le fait que la responsivité et l'offset estimés par les deux méthodes sont quasiment identiques (cf. 3.14). Des résultats obtenus sur le jeu $\mathscr{C}_2$ seront présentés dans la section suivante.

1.6.3 Étude des performances sur un jeu de calibration restreint, extrapolation en TFPA

Pour tester la robustesse du modèle de correction affine (3.1) aux changements de TFPA, nous allons corriger des images brutes de scènes uniformes (du sous-ensemble de test $\mathscr{C}_k^t$) acquises à différentes TFPA, en particulier à des TFPA éloignées de celles ayant permis de calibrer les variables responsivité et offset. Nous pourrons alors comparer les trois méthodes. Nous proposons de travailler sur un sous-ensemble de travail restreint aux états de $a \geq 1$ à $b \leq N_e$, nous le noterons $\tilde{\mathscr{C}}_k'$ et il est défini par

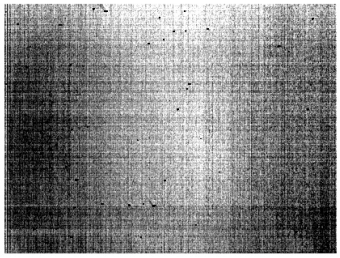
$$\begin{cases} \tilde{\mathscr{C}}_k^f = \{(r_j^{(f)}, \tau_{f,j}, T_f),\ a \leq j \leq b\} \text{ pour le corps noir froid,} \\ \tilde{\mathscr{C}}_k^c = \{(r_j^{(c)}, \tau_{c,j}, T_c),\ a \leq j \leq b\} \text{ pour le corps noir chaud,} \\ \tilde{\mathscr{C}}_k' = \tilde{\mathscr{C}}_k^f \cup \tilde{\mathscr{C}}_k^c \text{ sous-ensemble de données restreint.} \end{cases} \tag{3.47}$$

Nous appellerons intervalle de calibration restreint $\tilde{\mathbb{J}}$, l'intervalle dans lequel sont situées toutes les mesures de TFPA du sous ensemble restreint $\tilde{\mathscr{C}}'$. Il se définit donc par

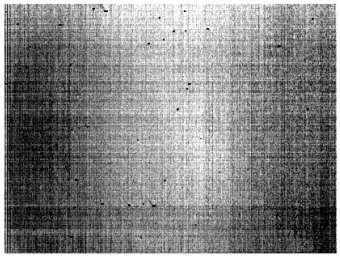
$$\tilde{\mathbb{J}} = [\min(\tau_{f,a}, \tau_{c,a}), \max(\tau_{f,b}, \tau_{c,b})].$$

Lorsque la TFPA de l'image de test sera située en dehors de $\tilde{\mathbb{J}}$, nous parlerons d'extrapolation en TFPA.

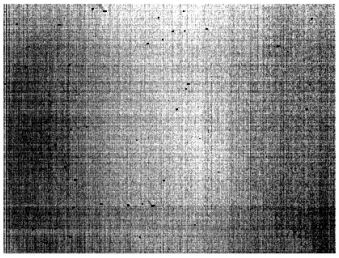
L'établissement des variables $\mathcal{R}_\tau$ et $\mathcal{D}_\tau$ se fait alors par exactement le même cheminement que décrit dans le paragraphe 1.6.1 en considérant $\tilde{\mathscr{C}}_k'$ à la place de $\mathscr{C}_k'$, à savoir la



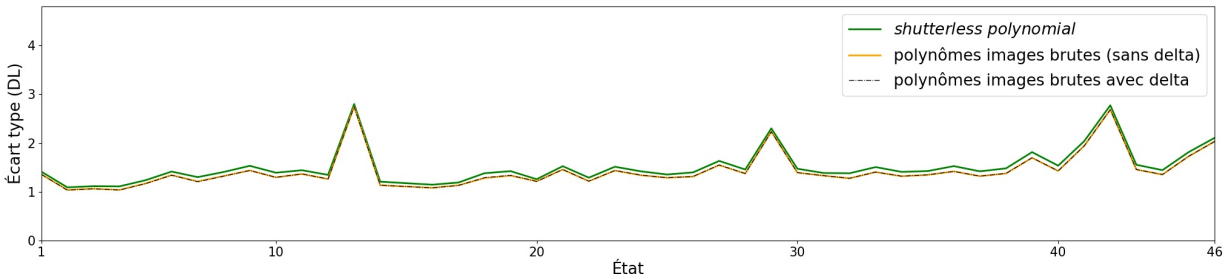
(a) *Shutterless polynomial*



(b) Polynômes images brutes



(c) Avec correction delta



(d) Mesures de qualité de correction des différentes méthodes

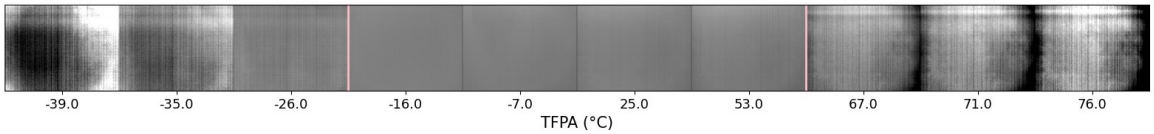
FIGURE 3.16 – **Correction d'une scène uniforme selon trois méthodes, mesures de qualité de correction.** Les données proviennent de l'ensemble $\mathscr{C}_1$. Les trois images correspondent à la correction d'une image brute de corps noir de test établie pour la TFPA de 30°C (25ème état) par la formule (3.45) (après correction de la TFPA associée à l'image de test), la dynamique des niveaux de gris affichés ici est de trois fois le bruit temporel de mesure (soit 4.5 DL). Les points noirs sur les images correspondent aux pixels défectueux. Les composantes responsivité et décalage ont été établies par : (a) La méthode *shutterless polynomial*. (b) La méthode relative à la modélisation des images brutes.(c) La méthode modélisation des images brutes avec correction delta. La différence entre ces trois images est imperceptible, les trois modèles de correction se valent. (d) Pour définir une mesure de qualité de correction, nous avons opté pour l'écart-type spatial de l'image corrigée (tel que défini par (3.46)). Plus ce dernier est petit, meilleure est la correction. La courbe continue verte correspond à la correction du modèle *shutterless polynomial*, la courbe continue jaune à la méthode de modélisation polynomiale des images brutes, la courbe discontinue noire à la modélisation polynomiale des images brutes avec correction delta. Ces deux dernières sont confondues. Par contre la courbe correspondant à la correction *shutterless polynomial* est légèrement au dessus des 2 autres. Les écarts-types de l'image corrigée sont légèrement supérieurs pour la méthode *shutterless polynomial* (entre 0 et 5% environ). En conclusion la correction d'une scène uniforme est très légèrement meilleure pour les méthodes fondées sur la modélisation des images brutes (avec ou sans delta) lorsque l'on compare avec la méthode *shutterless polynomial*, mais ces différences ne sont que très peu perceptibles. La correction delta donne des résultats identiques à la méthode fondée uniquement sur la modélisation polynomiale des images brutes.

détermination du degré global, la modélisation polynomiale des images brutes des deux corps noirs en fonction de τ puis établissement de la responsivité et du décalage par (3.40). On peut alors corriger toutes les images brutes d'un corps noir t issues d'un sous-ensemble de test $\mathscr{C}_k^t$ (tel que défini par (3.4)) avec la formule (3.45). Étant donnée la trop petite longueur de l'intervalle de calibration non restreint pour le jeu de données $\mathscr{C}_1$ ($[0, 46°\text{C}]$), nos expérimentations ont été réalisées à partir du jeu $\mathscr{C}_2$ (où les TFPA évoluent dans $[-39, 76°\text{C}]$).

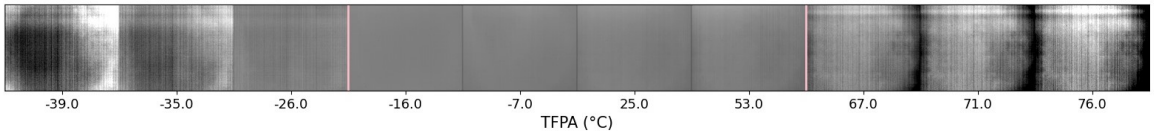
La figure 3.17 montre quelques résultats obtenus, nous en tirons deux enseignements. En premier lieu lorsqu'il s'agit de comparer la correction delta à la modélisation polynomiale des images brutes, la correction delta n'offre aucune amélioration en terme de qualité d'image, bien au contraire la qualité de la correction peut parfois être plus dégradée que celle apportée par la modélisation polynomiale des images brutes (sans delta). Nous avons étudié deux jeux de données. Pour le jeu $\mathscr{C}_1$ la correction delta permet d'obtenir les mêmes composantes responsivité et décalage (cf. figure 3.14) et en conséquence donner les mêmes corrections d'images (cf figure 3.16). Pour le jeu $\mathscr{C}_2$, les composantes sont différentes (cf. figure 3.15) et les corrections de moins bonne qualité (cf figure 3.17). Nous ne pouvons pas justifier pourquoi il existe de telles différences selon le jeu de données. Pour la plupart des jeux de données, on cherche à ce que les mesures de TFPA couvrent un spectre assez large ($[-40, 80°\text{C}]$). En cela le jeu $\mathscr{C}_2$ est représentatif. Si elle permet de réduire d'un facteur 3 l'erreur quadratique de *fit*, la correction delta ne peut pas être interprétée comme une correction de la mesure de TFPA. Les anomalies constatées sur les erreurs de fit (figure 3.11) sont sans doute également la conséquence de flux parasites qui sont difficiles à prendre en compte dans un modèle n'ayant la TFPA que pour seul paramètre.

En deuxième lieu, la modélisation polynomiale des images brutes donne des résultats en terme de correction au moins aussi bons que la méthode *shutterless polynomial*.

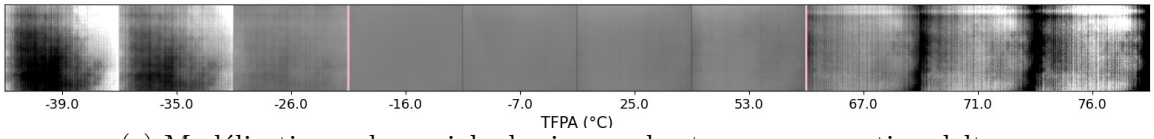
Les composantes du modèle affine, à savoir responsivité ét décalage, sont établies grâce à une partie de l'information présente dans un ensemble de calibration $\mathscr{C}_k$. On peut se demander si toute l'information a bien été exploitée, c'est l'objet de la partie suivante où nous allons procéder à des analyses statistiques à partir du jeu de données $\mathscr{C}_2$.



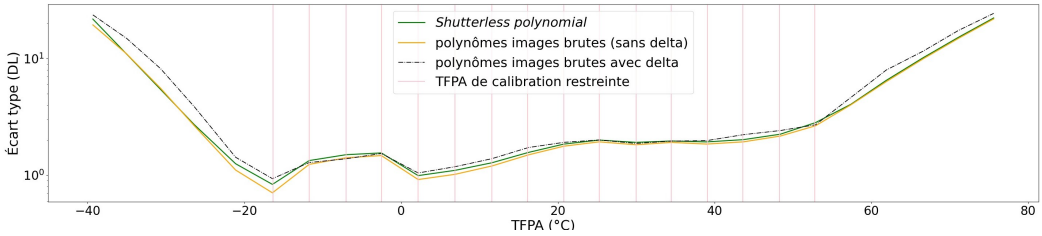
(a) *Shutterless polynomial*



(b) Modélisation polynomiale des images brutes (sans delta)



(c) Modélisation polynomiale des images brutes avec correction delta



(d) Performances de correction des trois méthodes.

FIGURE 3.17 – **Correction de scènes uniformes à partir d'un jeu de calibration restreint selon trois méthodes.** Les données proviennent de l'ensemble $\mathscr{C}_2$. L'intervalle de calibration restreint en TFPA est $\tilde{\mathbb{J}} = [-16, 53°\text{C}]$, sur les trois pellicules d'images il est caractérisé par deux verticales roses. L'ensemble de calibration restreint est composé de 16 états, le degré global est $D = 4$. L'affichage en niveaux de gris se fait sur étendue commune de 60 DL (soit 40 fois le bruit temporel de mesure). Dix images brutes provenant d'un sous-ensemble de test ont été corrigées par trois méthodes : (a) Méthode *shutterless polynomial*, (b) Modélisation polynomiale des images brutes, (c) Correction delta, leur TFPA se lit en abscisse. Chaque pellicule d'images montre la correction obtenue après estimation de la TFPA par (3.44) et (3.45). Pour ces trois méthodes, lorsque la TFPA de l'image de test se situe dans l'intervalle $\tilde{\mathbb{J}}$, la correction donne des images faiblement contrastées d'apparence relativement uniforme. Plus la TFPA de l'image de test est située loin de $\tilde{\mathbb{J}}$, plus l'image corrigée est dégradée (d'apparence peu uniforme), c'est davantage le cas pour la correction delta (c). On observe des images corrigées assez similaires entre *shutterless polynomial* (a) et polynômes des images brutes (b). (d) Écarts-types obtenus par (3.46) de 26 images de test corrigées en fonction de la TFPA selon les trois méthodes (échelle logarithmique). Les 16 TFPA de l'intervalle de calibration restreint $\tilde{\mathbb{J}}$ sont représentées par des lignes verticales roses. La méthode modélisation des images brutes (courbe orange) donne des performances presque identiques à celles de la méthode *shutterless polynomial* (courbe verte). Ces performances semblent toutefois sensiblement meilleures pour quelques TFPA froides (-39°C et -16°C). La modélisation polynomiale des images brutes avec correction delta donne les plus mauvaises performances puisque sa courbe (en discontinue noire) est située au dessus des deux autres. En extrapolation (TFPA en dehors de $\tilde{\mathbb{J}}$), les performances des trois méthodes sont nettement moins bonnes qu'en interpolation et la correction delta se distingue par les plus mauvaises performances. En conclusion, la modélisation polynomiale des images brutes (sans delta) offre une qualité de correction au moins aussi bonne que la méthode *shutterless polynomial*, la modélisation polynomiale des images brutes avec correction delta offre par contre des corrections d'une qualité moindre.

2 Analyse statistique des données de calibration

Les données de calibration que nous avons analysées sont issues d'une procédure d'acquisition en « étuve stabilisée ». Cette procédure est celle qui permet de réduire le plus l'influence des paramètres autres que la TFPA. Dans un premier temps, nous dresserons quelques caractéristiques du comportement général du capteur selon les deux facteurs d'intérêt que sont la TFPA et la température (irradiance) de scène. Dans un deuxième temps nous allons observer l'espace des corrections dans le cadre du modèle affine avec gain constant et biais que nous définirons.

2.1 Étude du comportement général d'un pixel à partir d'une ACP

2.1.1 Réponse du capteur face à un ensemble de situations

Nous n'allons ne considérer que les pixels non défectueux (appartenant au masque $\mathbf{M}$), un pixel sera ici noté $p = (x, y) \in \mathbf{M}$ et nous ferons l'assimilation $\mathbf{M} \approx [\![1, M]\!]$ où $M = \text{card}(\mathbf{M})$. Au cours d'une calibration, l'ensemble des pixels est confronté à différentes températures de scènes T_i et à différentes TFPA $\tau_{i,j}$ (i désigne le corps noir, j l'état). Soit un ensemble de calibration $\mathscr{C}_k$, nous pouvons en extraire

$$\mathbf{N} = \Big((T_i, \tau_{i,j}) \Big)_{i \in [\![1, N_T]\!],\, j \in [\![1, N_e]\!]},$$

nous l'appelons ensemble des situations. Il est commun à tous les pixels du capteur. On définit l'entier naturel $N = N_T \times N_e = \text{card}(\mathbf{N})$ et on fait les assimilations $\mathbb{R}^{N_e} \times \mathbb{R}^{N_T} \approx \mathbb{R}^N$ et $\mathbb{R}^{\mathbf{M}} \approx \mathbb{R}^M$. Soit un pixel $p \in \mathbf{M}$, nous disposons de la réponse de ce pixel face à l'ensemble des situations $\mathbf{N}$. Ce point $Y_p \in \mathbb{R}^N$ se définit par

$$Y_p = \Big(r_j^{(i)}(p) \Big)_{i \in [\![1, N_T]\!],\, j \in [\![1, N_e]\!]}.$$

La figure 3.18 permet d'observer le comportement général des pixels du capteur face à un ensemble de quatre situations. La TFPA semble être un facteur qui prédomine sur le facteur température de scène. Pour étudier le comportement général des pixels face à un ensemble de situations par rapport à un comportement moyen, nous allons considérer le nuage de points $(Y_1, ..., Y_M)$ et dresser quelques caractéristiques de l'espace dans lequel il évolue. Les points sont-ils disposés de manière totalement aléatoire autour de leur centre ou bien sont-ils plus ou moins concentrés spatialement selon une droite, un plan, un sous-espace affine ? L'analyse par composantes principales (ACP) permet de définir les axes principaux de ces sous-espaces affines tout en quantifiant leur degré d'explication.

2.1.2 Analyse par composantes principales

On suppose $N \ll M$ et on considère la matrice Y de taille $N \times M$ définie par

$$\begin{pmatrix} \vdots & & \vdots \\ Y_1 & \dots & Y_M \\ \vdots & & \vdots \end{pmatrix}. \tag{3.48}$$

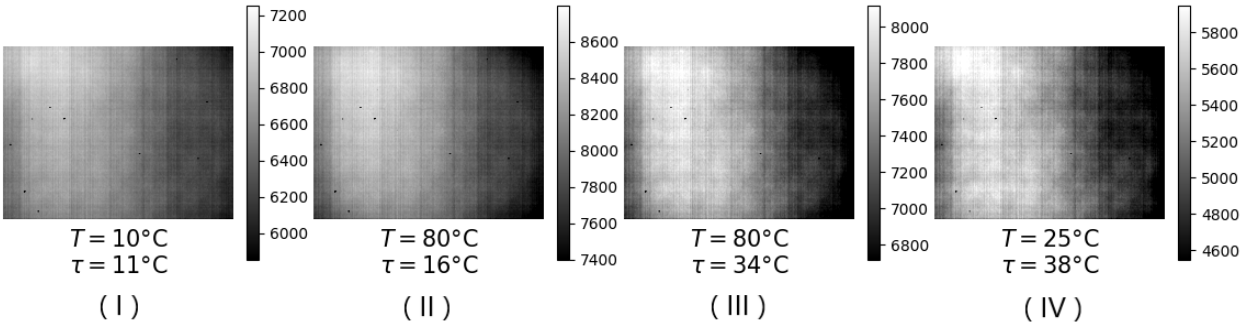


FIGURE 3.18 – **Comportement des pixels dans quatre situations.** Les données proviennent de l'ensemble $\mathscr{C}_2$. Chaque image représente l'image d'un corps noir de température de scène T acquise à une TFPA τ. Les points noirs correspondent aux pixels défectueux. L'affichage en niveau de gris se fait sur une étendue commune (1400 DL), il est centré autour de la moyenne spatiale de chaque image. Une situation est caractérisée par le couple (T, τ) et on dispose ici de quatre situations différentes. Chaque image illustre donc la réponse du capteur face à une des quatre situations. On distingue sur les quatre images des effets de ligne, de colonne et de basse fréquence communs qui se superposent spatialement et ces motifs sont d'autant plus semblables que les TFPA τ sont proches. En effet les deux images de gauche (I) et (II) se ressemblent assez fortement, leur TFPA τ est relativement proche et leur température de scène éloignée. Les images (III) et (IV) se ressemblent également (TFPA proches et température de scène éloignées également). Les deux images du centre (II) et (III) sont relativement différentes, leur TFPA est assez différent, leur température de scène identique. Ainsi la TFPA est un facteur plus influent que la scène sur la réponse du capteur.

Soit $\overline{Y}$ le centre du nuage de points, il est défini par

$$\overline{Y} = \frac{\sum_{p=1}^{M} Y_p}{M}. \tag{3.49}$$

L'ACP consiste à faire la décomposition en valeurs singulières (*singular value decomposition*, SVD) de $Y - \overline{Y}$, ce qui donne[9]

$$\exists U \in \mathbf{O}(N),\ \exists \Sigma \in \mathbf{D}(N, M), \exists V \in \mathbf{O}(M), \quad Y = \overline{Y} + U\,\Sigma\,V^t, \tag{3.50}$$

$\mathbf{O}(N)$ désigne le groupe des matrices orthogonales d'ordre N, $\mathbf{D}(N, M)$ désigne les matrices de taille $N \times M$ dont tous les éléments sont nuls sauf d'éventuels éléments sur sa diagonale, ces derniers étant positifs et ordonnés de manière décroissante.
Les vecteurs colonnes (aussi dénommés atomes) $(U_1, ..., U_N)$ de U représentent les axes principaux, ils forment une base orthonormée de $\mathbb{R}^N$. Au $k^{\text{ème}}$ atome U_k est associée l'inertie I_k définie par

$$I_k = \frac{\sigma_k^2}{\sum_{l=1}^{N} \sigma_l^2}, \quad k \in [\![1, N]\!], \tag{3.51}$$

où σ_k désigne le k^{ème} élément sur la diagonale de la matrice Σ. σ_k^2 désigne la variance explicative (ou erreur quadratique) associée à l'axe U_k. Exprimée en pourcentage l'inertie

9. Il faudra voir ici $\overline{Y}$ comme un élément de taille $N \times M$.

I_k représente donc la proportion de la variabilité expliquée par l'axe U_k. Les inerties $(I_k)_{k\in[\![1,N]\!]}$ forment une suite décroissante à valeur dans $[0, 100\%]$ et leur somme $\sum_{k=1}^{N} I_k$ vaut 100%. Pour k donné, on peut voir l'axe U_k comme l'un des facteurs qui explique (à I_k pour-cent) le comportement général du capteur face à l'ensemble des situations $\mathbf{N}$.

On définit alors un sous-espace affine $\mathbb{A}_n$ de dimension $n \leq N$ par

$$\mathbb{A}_n = \overline{Y} + \text{vect}(U_1, ..., U_n),$$

auquel on associe l'inertie explicative cumulée

$$\tilde{I_n} = \sum_{k=1}^{n} I_k.$$

On peut alors considérer que l'espace affine $\mathbb{A}_n$ explique le comportement général du capteur selon la valeur de l'inertie cumulée $\tilde{I_n}$ (plus elle est élevée, plus l'information relative au comportement général du capteur est prise en considération).

2.1.3 Observations, interprétations et perspectives d'étude du modèle direct

La figure 3.19 présente les résultats obtenus à partir du jeu de calibration $\mathscr{C}_2$, qui offre un échantillonnage étendu des conditions expérimentales. L'analyse met en évidence la prédominance de la TFPA dans la variation des réponses du capteur, ce qui rend le signal utile lié à la scène relativement faible par rapport à l'effet de la TFPA.

Par ailleurs, les surfaces associées au centre et aux composantes principales du nuage de points présentent une régularité remarquable, ce qui suggère que ces composantes peuvent être représentées par des polynômes de deux variables (T, τ), les degrés de ces polynômes restant à déterminer. La réponse d'un pixel p face à une situation (T, τ) peut alors être exprimée comme une combinaison linéaire de ces polynômes, conduisant au « modèle direct »

$$r_p(T, \tau) = F_0(T, \tau) + \sum_{k=1}^{n} W_{p,k} F_k(T, \tau), \tag{3.52}$$

où F_0 représente le polynôme associé au centre du nuage et les F_k $(1 \leq k \leq n)$ les polynômes correspondant aux axes principaux. Les coefficients $W_{p,k}$, spécifiques à chaque pixel, sont directement issus de l'ACP et quantifient l'importance de chaque axe dans la réponse du pixel. Le nombre n est choisi en fonction de l'inertie cumulée afin de capturer l'essentiel de la variance explicative.

Ce modèle direct décrit ainsi la réponse des pixels en fonction de (T, τ). Si les polynômes considérés sont d'ordre 1 en température de scène, on retombe alors sur le modèle *shutterless polynomial* [8]. Il devient alors possible d'inverser le modèle, c'est-à-dire d'exprimer T en fonction de la réponse du capteur et de la TFPA. Plus l'ordre en T est élevé, plus l'inversion devient complexe.

Nous nous sommes intéressés au comportement général des pixels selon la scène et la TFPA sans hypothèse de modèle affine sous-jacent. Nous allons désormais examiner l'espace des corrections dans le cadre du modèle affine avec gain constant et biais que nous définirons ci-dessous.

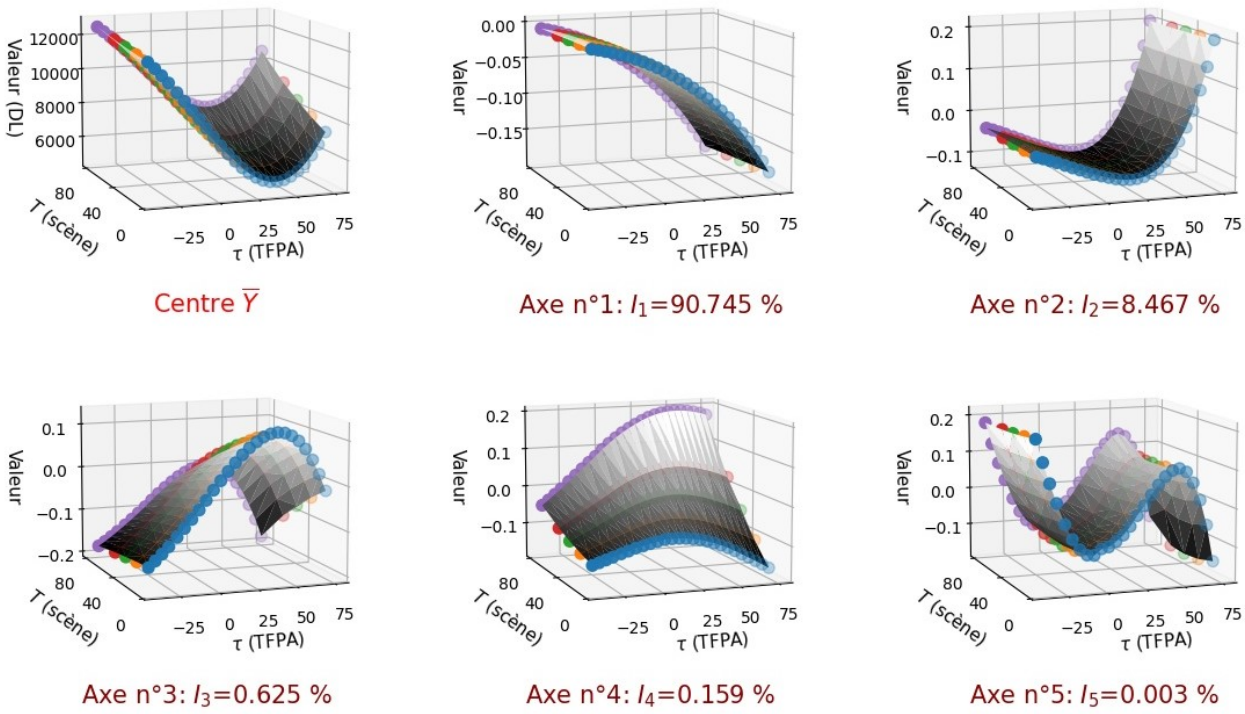


FIGURE 3.19 – **Comportement général du capteur selon la température de scène et la TFPA. Courbes 3D du centre et des cinq axes principaux après ACP**. Les données proviennent de l'ensemble $\mathscr{C}_2$. L'ensemble des situations peut se voir comme un quadrillage de 5 températures de scène T (10°C en bleu, 25°C en orange, 40°C en vert, 55°C en rouge, 80°C en pourpre) par 26 TFPA τ, on dénombre donc ici $5 \times 26 = 130$ situations. L'ACP consiste à étudier la forme du nuage de $M = 307106$ points, chacun de ces points correspond à la réponse d'un pixel face à l'ensemble des situations. L'ACP permet d'établir le centre ainsi que les axes principaux dans lequel évolue le nuage de points. Le centre du nuage est représenté par la surface en haut à gauche. Sa surface associée représente le comportement moyen d'un pixel selon T et τ. Pour les TFPA τ les plus froides, la courbe varie peu selon T à τ fixé. Les autres graphiques représentent les 5 axes principaux (ou atomes). Il leur est associé l'inertie explicative I_k exprimée en pourcentage ($k \in [\![1, 5]\!]$). Ils caractérisent les spécificités comportementales de l'ensemble des pixels par rapport au comportement moyen du capteur. Si on considère le plan passant par le centre et dirigé par les 2 premiers axes, plus de 99.2% du comportement général du capteur est expliqué. Les courbes de ces 2 axes montrent une variabilité en fonction de la scène T peu perceptible (en tout cas moindre que celle induite par la TFPA). Les surfaces associées aux axes 3 et 4 montrent une variabilité en fonction de la scène T qui est d'autant plus perceptible que les TFPA sont grands. La surface de l'axe 5 montre une variabilité selon la scène relativement petite. Quelque soit la dimension du sous-espace affine que l'on considère pour décrire la forme du nuage de points, on déduit que la réponse d'un pixel est beaucoup plus influencée par le facteur TFPA τ que par le facteur scène T et ce d'autant plus que les TFPA sont très froides. Toutes les surfaces observées ici sont surprenamment régulières alors que rien dans le modèle (3.50) ne promeut une telle régularité. Une telle régularité découle probablement des lois physiques et électroniques qui régissent la réponse du capteur. On peut alors envisager, après par exemple un *fit* polynomial 2D (les variables étant τ et T) du centre et des axes principaux, une modélisation de la réponse de chaque pixel selon τ et T (modèle direct) sans a priori d'un modèle affine sous-jacent.

2.2 Analyse de l'espace des biais

2.2.1 Modèle affine avec gain unique et biais

Le modèle affine avec gain et biais dérive du modèle de correction affine avec responsivité et décalage qui pour rappel s'écrit

$$S = \frac{r_\tau - \mathcal{D}_\tau}{\mathcal{R}_\tau}, \tag{3.53}$$

S désigne l'image corrigée, r_τ l'image brute acquise à la TFPA τ, $\mathcal{R}_\tau$ l'image de la responsivité et $\mathcal{D}_\tau$ l'image du décalage en tension.
A l'instar de l'établissement du modèle de correction avec gain et offset présenté dans le chapitre précédent 2.1.2.2, on pose

$$\begin{cases} \mathcal{G}_\tau = \dfrac{\mu(\mathcal{R}_\tau)}{\mathcal{R}_\tau}, & \text{(3.54a)} \\ B_\tau = \mathcal{G}_\tau \mathcal{D}_\tau - \mu(\mathcal{G}_\tau \mathcal{D}_\tau), & \text{(3.54b)} \\ U = \mu(\mathcal{R}_\tau)\, S + \mu(\mathcal{G}_\tau \mathcal{D}_\tau)\,, & \text{(3.54c)} \end{cases}$$

U désigne l'image corrigée, $\mathcal{G}_\tau$ l'image du gain et B_τ l'image du biais. L'image du biais est de par sa définition (3.54b) de moyenne spatiale nulle [10], i.e.

$$\mu(B_\tau) = 0, \tag{3.55}$$

En partant de (3.54c), nous avons

$$\begin{aligned} U &= \mu(\mathcal{R}_\tau)\, S + \mu(\mathcal{G}_\tau \mathcal{D}_\tau) \\ U &= \mu(\mathcal{R}_\tau)\frac{r_\tau - \mathcal{D}_\tau}{\mathcal{R}_\tau} + \mu(\mathcal{G}_\tau \mathcal{D}_\tau) \quad \text{d'après (3.53)} \\ U &= \mu(\mathcal{R}_\tau)\frac{r_\tau}{\mathcal{R}_\tau} - \mu(\mathcal{R}_\tau)\frac{\mathcal{D}_\tau}{\mathcal{R}_\tau} + \mu(\mathcal{G}_\tau \mathcal{D}_\tau) \\ U &= \mathcal{G}_\tau\, r_\tau - \Big(\mathcal{G}_\tau \mathcal{D}_\tau - \mu(\mathcal{G}_\tau \mathcal{D}_\tau)\Big) \quad \text{d'après (3.54a)} \\ U &= \mathcal{G}_\tau\, r_\tau - B_\tau \quad \text{d'après (3.54b)} \end{aligned}$$

Le modèle de correction s'écrit alors

$$U = \mathcal{G}_\tau\, r_\tau - B_\tau. \tag{3.56}$$

Pour déterminer le gain $\mathcal{G}_\tau$ en fonction de la TFPA τ, il faut combiner la définition du gain (3.54a) et le calcul de la responsivité après modélisation polynomiale des images brutes [11] (3.40a). Il est courant dans la littérature de considérer que l'image de gain ne varie pas selon la TFPA [10, 19, 20, 26, 27]. Pour déterminer ce gain unique $\mathcal{G}_0$ nous proposons de considérer le gain obtenu pour la TFPA $\tau_0 = 20°C$ (température moyenne de l'air

10. La moyenne spatiale $\mu(\cdot)$ est établie sur l'ensemble des pixels du masque $\mathbf{M}$.
11. Rien ne nous prouve que la modélisation polynomiale des images brutes avec correction delta nous permette d'obtenir des polynômes plus précis pour la responsivité, bien au contraire a priori. Nous optons donc pour considérer la modélisation polynomiale des images brutes sans delta.

ambiant en intérieur mais également médiane des TFPA dans la plupart des données de calibration).

$$\mathcal{G}_0 = \mathcal{G}_{\tau_0} = \frac{\mu(\mathcal{R}_{\tau_0})}{\mathcal{R}_{\tau_0}}. \tag{3.57}$$

Le modèle de correction que nous allons considérer dans cette partie est le modèle affine à un seul paramètre avec gain unique qui s'écrit

$$U = \mathcal{G}_0\, r_\tau - B_\tau, \tag{3.58}$$

où U désigne l'image corrigée, r_τ l'image brute délivrée par la caméra pour la TFPA τ et B_τ l'image du biais. L'espace des corrections dans le cadre du modèle (3.58) se limite donc à l'espace des biais. Nous allons voir comment construire une famille de biais à partir d'un jeu de données $\mathscr{C}_k$.

2.2.2 Construction d'une famille de biais

Soit $r_\tau^{(T)}$ l'image brute d'une scène uniforme de température de scène T acquise à la TFPA τ. Montrons que le biais B_τ vérifie

$$B_\tau = \mathcal{G}_0\, r_\tau^{(T)} - \mu(\mathcal{G}_0\, r_\tau^{(T)}). \tag{3.59}$$

Preuve :
L'image corrigée U par le modèle (3.58) de cette scène uniforme est une image constante. Par conséquent $\exists K \in \mathbb{R}$ tel que

$$\forall (x,y) \in \mathbf{M}, \quad U(x,y) = K$$

or d'après (3.58) on a

$$\forall (x,y) \in \mathbf{M}, \quad U(x,y) = \mathcal{G}_0(x,y)\, r_\tau^{(T)}(x,y) - B_\tau(x,y).$$

On obtient alors

$$\forall (x,y) \in \mathbf{M}, \quad K = \mathcal{G}_0(x,y)\, r_\tau^{(T)}(x,y) - B_\tau(x,y). \tag{3.60}$$

On déduit de (3.60) que

$$\begin{aligned} \mu(K) &= \mu(\mathcal{G}_0\, r^{(T)} - B_\tau) \\ K &= \mu(\mathcal{G}_0\, r_\tau^{(T)}) - \mu(B_\tau) \\ K &= \mu(\mathcal{G}_0\, r^{(T)}), \quad \text{car } \mu(B_\tau) = 0 \text{ d'après (3.55).} \end{aligned}$$

L'équation (3.60) se réécrit alors en (3.59), ce qui finit la démonstration.

En résumé, pour obtenir le biais à partir d'une image brute de scene uniforme [12], il suffit de multiplier (terme à terme) cette dernière par l'image du gain unique puis de centrer le tout (i.e. retrancher la moyenne spatiale). On peut constater que l'image de biais définie

12. Une caméra thermique avec obturateur mécanique considère l'image brute uniforme de l'obturateur pour fréquemment déterminer une image de biais.

par (3.59) ne dépend pas de la température T de la scène uniforme.

On peut alors transformer toutes les images brutes d'un jeu de calibration $\mathscr{C}$ en une famille de biais $\mathscr{B}$ définie de la manière suivante

$$\begin{cases} \mathscr{B} = \left(B_j^{(i)}\right)_{\substack{i\in[\![1,N_T]\!]\\ j\in[\![1,N_e]\!]}} \\ B_j^{(i)} = \mathcal{G}_0\, r_j^{(i)} - \mu(\mathcal{G}_0\, r_j^{(i)}). \end{cases}$$

Chaque biais $B_j^{(i)}$ de la famille $\mathscr{B}$ est associé à une mesure de TFPA $\tau_{_{i,j}}$. On peut alors construire la famille des TFPA

$$\mathscr{T} = \left(\tau_{_{i,j}}\right)_{\substack{i\in[\![1,N_T]\!]\\ j\in[\![1,N_e]\!]}}$$

associée à $\mathscr{B}$. Pour simplifier l'écriture, nous renoterons les familles de biais et de TFPA sous la forme

$$\begin{cases} \mathscr{B} = (B_1, ..., B_N) \\ \mathscr{T} = (\tau_1, ...\tau_{_N}), \end{cases} \tag{3.61}$$

où $N = N_T\, N_e$, chaque biais B_n ($n \in [\![1, N]\!]$) est considéré comme une image de $\mathbb{R}^{\mathbf{M}}$, où $\mathbf{M}$ désigne le masque de pixel non défectueux.

2.2.3 Analyse de l'espace des biais

Une fois la famille de biais $\mathscr{B}$ constituée (comme illustrée par la figure 3.20a), nous allons nous intéresser à l'espace décrit par l'ensemble des biais. Nous proposons deux approches. La première consiste à considérer chaque biais comme un point de $\mathbb{R}^M$ (on fera l'assimilation $\mathbb{R}^{\mathbf{M}} \approx \mathbb{R}^M$, où M =card($\mathbf{M}$).), l'espace étudié est affine et nous ferons une ACP autour du centre de la famille des biais. La deuxième approche consiste à considérer chaque biais comme un vecteur de $\mathbb{R}^M$, l'espace étudié est vectoriel et nous ferons une ACP non centrée. Dans les deux cas il s'agit de considérer la matrice Z de taille [13] $M \times N$ définie par

$$Z = \begin{pmatrix} \vdots & & \vdots \\ B_1 & \dots & B_N \\ \vdots & & \vdots \end{pmatrix}. \tag{3.62}$$

Chaque colonne de cette matrice est un biais de la famille $\mathscr{B}$. Pour procéder à une étude statistique de l'espace engendré par les biais, nous avons donc envisagé les 2 approches suivantes.

13. On suppose $M \gg N$, le nombre de pixels non défectueux est très supérieur aux nombres d'images de calibration.

— 1-ACP (centrée) : on cherche à représenter l'espace des biais comme un sous-espace affine.
On définit le centre $\overline{Z}$ par

$$\overline{Z} = \frac{\sum_{n=1}^{N} B_n}{N}$$

L'ACP consiste à faire la décomposition SVD de $Z - \overline{Z}$, i.e.

$$\exists U \in \mathbf{O}(M),\ \exists \Sigma \in \mathbf{D}(M,N), \exists V \in \mathbf{O}(N), \quad Z = \overline{Z} + U\,\Sigma\,V^t \tag{3.63}$$

Le sous espace affine $\mathbb{A}_{N-1}$ dans lequel évoluent les biais se définit par

$$\mathbb{A}_{N-1} = \overline{Z} + \mathrm{Vect}(B_1 - \overline{Z}, ..., B_{N-1} - \overline{Z}), \tag{3.64}$$

et on a aussi

$$\mathbb{A}_{N-1} = \overline{Z} + \mathrm{Vect}(U_1, ..., U_{N-1}). \tag{3.65}$$

L'espace affine $\mathbb{A}_{N-1}$ est composé de N éléments : le centre $\overline{Z}$ et les atomes $(U_1, ...U_{N-1})$. Pour quantifier la proportion de variabilité expliquée par chacun d'entre eux, nous voudrions leur associer des pourcentages d'explication que nous noterons $(J_0, ...J_{N-1})$. Il nous faut préalablement définir pour chaque élément $k \in [\![0, N-1]\!]$ la variance explicative E_k. Pour $k \in [\![1, N-1]\!]$, puisqu'il s'agit d'une ACP, la variance explicative associée à l'axe k vaut $E_k = \sigma_k^2$ où σ_k désigne le k^{ème} élément sur la diagonale de la matrice Σ. En ce qui concerne le centre $\overline{Z}$, la variance explicative correspond à l'erreur quadratique résultant du modèle consistant à approximer chaque biais comme étant égal au centre. On définit la variance explicative du centre $\overline{Z}$ par

$$E_0 = \|Z - \overline{Z}\mathbf{1}^*\|^2$$

où $\mathbf{1}^* \in \mathbb{R}^N$ est un vecteur ligne rempli de 1, ce qui correspond à l'erreur quadratique en approximant chaque biais par le centre. Le pourcentage d'explication du $k^{\text{ième}}$ élément est le rapport de sa variance explicative sur la somme de toutes les variances explicatives. On définit alors les pourcentages explicatifs des N éléments de l'espace affine $\mathbb{A}_{N-1}$ par

$$\begin{cases} J_0 = \dfrac{E_0}{E_0 + \sum_{n=1}^{N-1} \sigma_n^2} & \text{pour le centre } \overline{Z} \\ J_k = \dfrac{\sigma_k^2}{E_0 + \sum_{n=1}^{N-1} \sigma_n^2} & \text{pour l'axe } U_k. \end{cases} \tag{3.66}$$

Lorsque l'on cherche à réduire l'espace, nous considérerons $\mathbb{A}_p$ l'espace affine de dimension $p \leq N-1$ défini par

$$\mathbb{A}_p = \overline{Z} + \mathrm{Vect}(U_1, ..., U_p). \tag{3.67}$$

Soit une image $X \in \mathbb{R}^{\mathbf{M}}$, sa projection orthogonale $\widehat{X}$ sur $\mathbb{A}_p$ vérifie

$$\widehat{X} = \underset{\mathbb{A}_p}{\mathrm{proj}}\, X = \overline{Z} + \sum_{n=1}^{p} \langle X - \overline{Z}, U_n \rangle\, U_n. \tag{3.68}$$

— 2-ACP non centrée : on cherche à représenter l'espace des biais comme un sous-espace vectoriel.

L'ACP non centrée consiste à faire la décomposition SVD de Z, ce qui donne

$$\exists U \in \mathbf{O}(M), \exists \Sigma \in \mathbf{D}(M,N), \exists V \in \mathbf{O}(N), \quad Z = U\,\Sigma\,V^t \tag{3.69}$$

Les N premières colonnes U_1,...,U_N de la matrice U forment une base orthonormale de Vect(B_1,...,B_N).
$\forall k \in [\![1, N]\!]$, on définit J_k le pourcentage explicatif de l'axe U_k en tant que rapport de la variance explicative de l'axe U_k sur la somme de toutes les variances explicatives, i.e

$$J_k = \frac{\sigma_k^2}{\sum_{n=1}^{N} \sigma_n^2}. \tag{3.70}$$

Le sous espace vectoriel $\mathbb{V}_N = \text{Vect}(B_1,...,B_N)$ vérifie

$$\mathbb{V}_N = \text{Vect}(U_1, ..., U_N). \tag{3.71}$$

Lorsque l'on cherche à réduire l'espace, nous considérerons $\mathbb{V}_p$ l'espace vectoriel de dimension $p \leq N$ défini par

$$\mathbb{V}_p = \text{Vect}(U_1, ..., U_p). \tag{3.72}$$

Soit une image $X \in \mathbb{R}^{\mathbf{M}}$, sa projection orthogonale $\widehat{X}$ sur $\mathbb{V}_p$ vérifie

$$\widehat{X} = \underset{\mathbb{V}_p}{\text{proj}}\, X = \sum_{n=1}^{p} \langle X, U_n \rangle\, U_n. \tag{3.73}$$

Montrons que

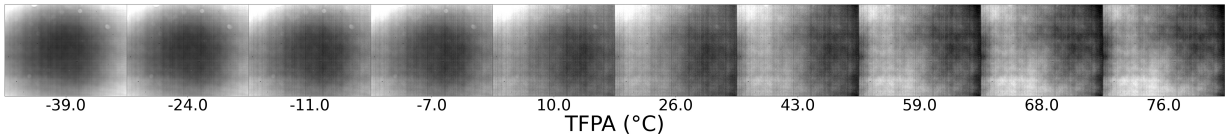
$$\mathbb{A}_{N-1} \subset \mathbb{V}_N. \tag{3.74}$$

Preuve : Soit $Z \in \mathbb{A}_{N-1}$, d'après (3.64), $\exists(\alpha_1, .., \alpha_{N-1}) \in \mathbb{R}^{N-1}$, $Z = \overline{Z} + \sum_{n=1}^{N-1} \alpha_n (B_n - \overline{Z})$. Puisque $\overline{Z} \in \mathbb{V}_N$ et $B_n \in \mathbb{V}_N$, nous avons $Z \in \mathbb{V}_N$ □

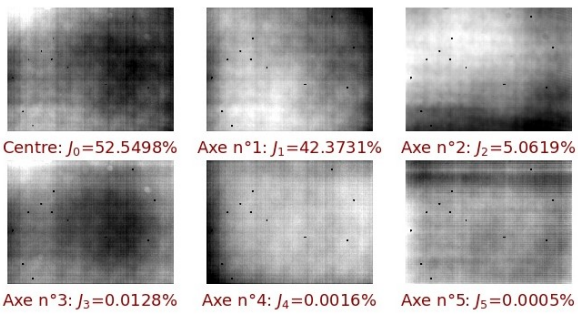
Si on fait l'hypothèse que la dimension de $\mathbb{V}_N$ est N (vérifiée en pratique), la dimension de $\mathbb{A}_{N-1}$ est alors égale à $N-1$ et l'inclusion dans (3.74) est stricte.

Lorsque qu'il ne s'agit pas de réduire la dimension de l'espace des biais, il est donc préférable de considérer l'approche ACP non centrée puisque l'espace décrit est plus grand mais nous ne pouvons pas affirmer la même chose lorsque l'on désire réduire de dimension. Dans la littérature cependant, Tanwar et al. [60], dans le cadre de la gestion massive des données, préconisent également l'emploi de la ACP non centrée plutôt que de l'ACP lorsque l'on cherche à réduire la dimension de l'espace observé. A priori l'approche ACP non centrée est donc préférable mais nous voudrions nous en assurer à partir de nos propres expérimentations en répondant à la problématique suivante.

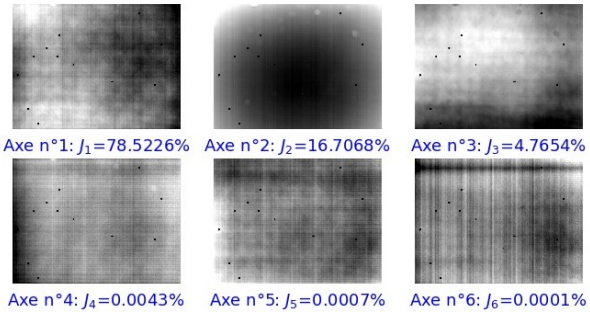
Soit un biais quelconque B_τ où τ désigne la TFPA, est-il plus proche de l'espace affine $\mathbb{A}_{N-1}$ défini par (3.65) ou de l'espace vectoriel $\mathbb{V}_N$ défini par (3.71) ? La figure 3.20a nous



(a) Famille de biais



(b) Approche ACP (centrée)



(c) Approche ACP non centrée

FIGURE 3.20 – **Famille des biais, axes principaux obtenus par les approches ACP centrée et ACP non centrée.** Les données proviennent de l'ensemble $\mathscr{C}_2$. L'affichage en niveau de gris de chacune des images est centré autour de leur moyenne spatiale et se fait sur une étendue propre de 5 écart-types spatiaux. Les points noirs sur les images correspondent aux pixels défectueux. (a) La famille de biais est composée de 130 biais. Nous n'en montrons ici que 10, la mesure de la TFPA associée à chacun de ces biais se lit en abscisse. Lorsque leur TFPA est proche, les images de deux biais se ressemblent assez fortement. Plus la TFPA est froide, plus l'image du biais est spatialement lisse. Plus la TFPA est chaude, plus apparaissent des effets de ligne et de colonne. La famille des 130 biais est fortement corrélée. (b) Le centre et les 5 axes principaux (avec leur pourcentage explicatif J_k, $0 \leq k \leq 5$ défini par (3.66)) obtenus par ACP sont affichés. (c) Les 6 axes principaux obtenus par ACP non centrée sont affichés avec leur pourcentage explicatif J_k, $1 \leq k \leq 6$ défini par (3.70). L'image du premier axe de l'approche ACP non centrée ressemble assez fortement au biais dont la TFPA est de 76°C, on observe en effet des *patterns* de ligne, de colonne et de basse fréquence assez similaires. L'image du deuxième axe de l'approche ACP non centrée ressemble un peu au biais dont la TFPA est -38°C, on observe en effet des images spatialement lisses avec des *patterns* basse fréquence assez similaires ainsi que des taches claires et circulaires disposées aux mêmes endroits (tous ces motifs s'apparentent à des effets engendrés par l'optique de la caméra). On observe quelques correspondances visuelles entre les images des axes principaux obtenus par les deux approches : l'axe n°1 de l'approche ACP non centrée ressemble un peu à l'axe n°1 de l'ACP, l'axe n°3 (ACP non centrée) ressemble assez fortement à l'axe n°2 (ACP), l'axe n°4 (ACP non centrée) un peu à l'axe n°4 (ACP). Pour l'approche ACP le pourcentage explicatif cumulé des 6 éléments affichés est d'environ 99.9997. Pour l'approche ACP non centré le pourcentage explicatif cumulé des 6 éléments affichés est d'environ 99.9999. Lorsque l'on cherche à ne considérer que les $k \leq 6$ premiers éléments, il est préférable de considérer l'approche ACP non centré puisque on constate (après un calcul simple) que le pourcentage explicatif cumulé est supérieur pour cette approche. On distingue souvent sur les images des axes relatifs aux deux approches des effets colonnaires assez prononcés. On peut les expliquer par l'architecture du circuit de lecture électronique qui retranche pour chacun des pixels d'une colonne la même valeur provenant d'un pixel aveugle (i.e. qui ne voit aucune scène).

permet de comprendre que cette notion de proximité est étroitement liée au TFPA. Si il existe un biais dans la famille dont la TFPA est plus ou moins proche de τ, B_τ sera plus ou moins proche de $\mathbb{A}_{N-1}$ (et de $\mathbb{V}_N$). Nous avons donc décidé d'observer l'image de la différence entre un biais et son projeté orthogonal dans différents cas que nous allons détailler.

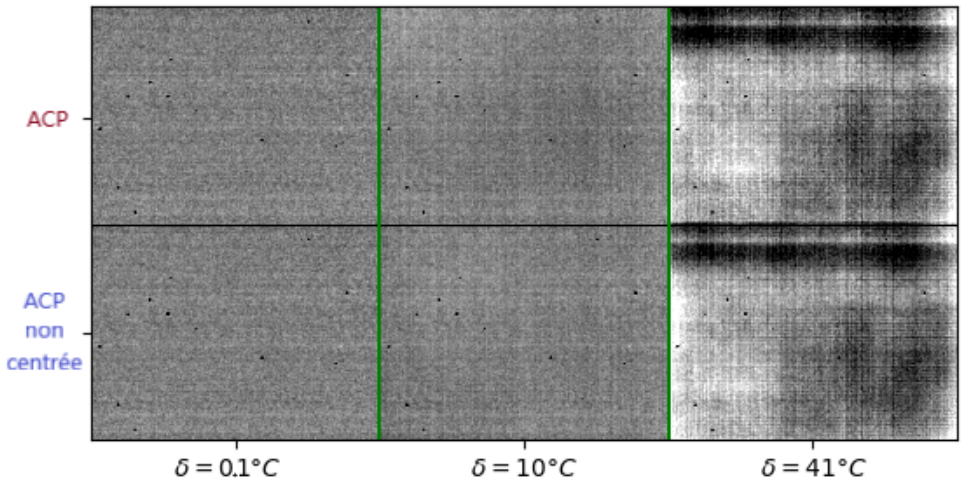
On choisit un biais $B_{n_0} \in \mathscr{B}$ auquel est associé une TFPA [14] τ_{n_0}. On définit un seuil $\delta \geq 0$ et on considère la sous-famille de biais $\mathscr{B}_{(n_0,\delta)} \subset \mathscr{B}$ définie par

$$\mathscr{B}_{(n_0,\delta)} = \{B_n \in \mathscr{B},\, n \in [\![1, N]\!],\, |\tau_n - \tau_{n_0}| \geq \delta\}$$

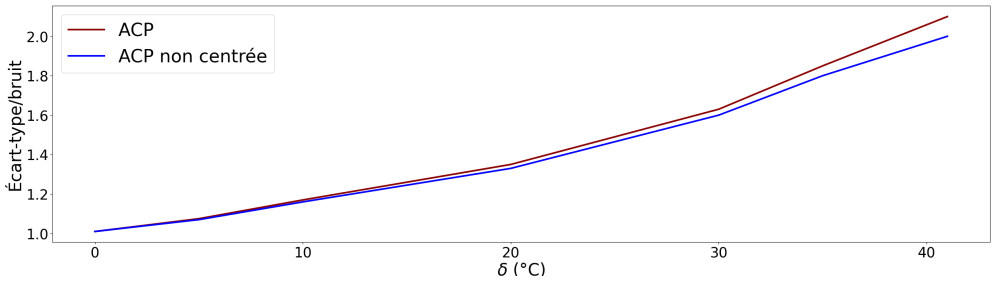
Il s'agit donc de réduire la famille $\mathscr{B}$ à des biais éloignés d'au moins δ en TFPA du biais B_{n_0}.

On applique alors l'approche ACP (respectivement ACP non centrée) à partir de la sous-famille $\mathscr{B}_{(n_0,\delta)}$ et l'on considère l'image de la différence entre B_{n_0} et son projeté orthogonal $\widehat{B_{n_0}}$ sur $\mathbb{A}_{N-1}$ (respectivement sur $\mathbb{V}_N$) selon l'équation (3.68) (respectivement (3.73)). La figure 3.21 permet de visualiser quelques résultats pour différentes valeurs de δ. Plus le seuil δ est grand, plus l'image de $B_{n_0} - \widehat{B_{n_0}}$ présente une norme ℓ^2 élevée, s'écartant d'un *pattern* de bruit aléatoire. Cet effet étant légèrement plus marqué pour l'approche ACP, il est préférable de privilégier l'ACP non centrée. Dit autrement, il est légèrement préférable de considérer l'espace engendré par une famille de biais comme un espace vectoriel plutôt que comme un espace affine. Pour que cette approximation soit valable, il convient de constituer la famille à partir de TFPA relativement proches (par exemple espacées de moins de 10 °C), tout en s'assurant que la plage de calibration soit suffisamment étendue. À titre d'exemple générique, on peut considérer une plage de -40 à 80 °C, la valeur exacte dépendant toutefois du domaine d'application de la caméra.

14. On choisira n_0 de sorte que τ_{n_0} soit au centre des mesures de TFPA présentes dans $\mathscr{T}$.



(a) Différence entre un biais et son projeté orthogonal sur une sous famille selon 2 approches



(b) Distance entre un biais et une sous-famille selon deux approches

FIGURE 3.21 – **Proximité d'un biais avec une sous-famille de biais selon les approches ACP et ACP non centrée.** Les données proviennent de l'ensemble $\mathscr{C}_2$. Un biais B_{n_0} a été choisi avec $\tau_{n_0} = 25$°C. Le seuil δ correspond à une distance minimale entre τ_{n_0} et les TFPA d'une sous famille de biais. Les points noirs sur les images correspondent aux pixels défectueux. L'affichage de chaque image est centré autour de la moyenne spatiale (nulle ici) et se fait sur une étendue commune égale à quatre fois le bruit temporel de mesure (soit une étendue de 6 DL). (a) Différence entre un biais et son projeté orthogonal sur une sous famille en tant qu'espace affine (ACP, première ligne) ou en tant qu'espace vectoriel (ACP non centrée, deuxième ligne). Chaque colonne correspond à un seuil δ différent. Sur la première colonne, $\delta = 0.1$°C, la sous-famille contient au moins un biais avec une TFPA très proche de τ_{n_0}. Dans les deux approches la différence entre B_{n_0} et son projeté s'apparente à une image du bruit temporel avec un très léger effet de ligne. Sur la deuxième colonne, $\delta = 10$°C, les deux images s'apparentent au bruit auquel s'ajoute un léger effet colonnaire. L'image ACP est légèrement plus contrastée que l'image ACP non centrée (plus blanche dans le coin supérieur gauche). Sur la troisième colonne, le seuil est beaucoup plus grand ($\delta = 41$°C), et des *patterns* se distinguant d'une image de bruit apparaissent (effets de ligne, de colonne, de basse fréquence ainsi que de larges barres noires horizontales en haut). Les images ACP et ACP non centrée se ressemblent mais l'image ACP apparait légèrement plus contrastée (e.g. sur la partie gauche de l'image).(b) Écart-type spatial de la différence entre un biais et son projeté orthogonal selon le seuil δ, les écart-types ont été divisés par la valeur moyenne du bruit temporel de mesure. Les images (biais et projeté) étant centrées, l'écart-type est proportionnel à la distance entre le biais et la sous-famille (en tant qu'espace affine ou vectoriel). Plus δ est grand, plus la distance entre B_{n_0} et la sous-famille grandit. C'est davantage le cas pour l'approche ACP. Il vaut mieux considérer une famille de biais comme un espace vectoriel que comme un espace affine, l'approche ACP non centrée est donc à privilégier.

3 Conclusion

Pour chaque caméra, un jeu de données a été constitué afin de calibrer les composantes de responsivité et de décalage du modèle affine en TFPA. Nous avons proposé une méthode reposant sur la modélisation polynomiale des images brutes, qui diffère sensiblement de la méthode de référence *shutterless polynomial* pour la détermination de ces deux composantes. Les expérimentations menées montrent que les performances de correction sont au moins équivalentes, et parfois très légèrement supérieures, à celles obtenues avec la méthode de référence.

L'analyse des erreurs de fit de la modélisation polynomiale des images brutes a révélé des écarts inattendus. En effet, leur distribution pour un même état thermique au sein des pixels ne suivait pas la distribution aléatoire que l'on pouvait attendre. Nous avons formulé l'hypothèse que ces écarts pouvaient résulter d'une imprécision dans la mesure de la TFPA du capteur. Pour tester cette hypothèse, nous avons introduit une modélisation polynomiale des images brutes avec correction conjointe des mesures de TFPA, appelée « correction delta ». Cette approche a permis de réduire très fortement l'erreur quadratique globale d'ajustement. Toutefois, les valeurs estimées pour les imprécisions de mesure de TFPA se sont parfois révélées aberrantes. Par ailleurs, les performances en termes de qualité de correction se sont montrées inférieures, ou au mieux équivalentes, à celles obtenues par la modélisation polynomiale directe des images brutes. Nous en avons conclu que les écarts observés ne peuvent être attribués uniquement à des imprécisions de mesure de TFPA et reflètent également la présence de flux thermiques parasites, difficiles à modéliser paramétriquement.

L'analyse par ACP du comportement général d'un pixel en fonction de la température de scène T et de la TFPA τ met en évidence la prédominance de cette dernière et souligne la difficulté inhérente à la correction d'une image thermique, le rapport entre le signal utile et le signal observé étant très faible. Elle permet également d'envisager un nouveau « modèle direct », généralisant le modèle affine communément considéré, dans lequel la réponse des pixels est représentée comme une combinaison linéaire de fonctions de (T, τ), ces fonctions pouvant être approchées par des polynômes.

Dans le cadre du modèle affine avec gain constant et biais variable, nous avons analysé l'espace des biais. Il apparaît que cet espace peut être considéré comme vectoriel, à condition que la calibration couvre une plage de TFPA adaptée au domaine d'application de la caméra et comporte un nombre suffisant d'acquisitions réalisées à des TFPA distinctes, régulièrement espacées, par exemple avec un pas inférieur à 10 °C.

Le chapitre suivant propose et étudie un modèle de correction d'image brute fondé sur une famille de biais. L'objectif est de déterminer un biais appartenant à l'espace vectoriel engendré par cette famille, en optimisant un critère de qualité d'image. Ce modèle de correction est non paramétrique, ce qui constitue un avantage, puisque l'ensemble des paramètres environnementaux θ, incluant la TFPA, est trop complexe pour être pris en compte explicitement dans un modèle de correction.

Chapitre 4

Modèles de correction à partir d'une famille de biais, débruitage colonnaire

Ce chapitre traite de la correction d'images acquises par une caméra thermique non-refroidie. Il est structuré en deux parties.

La première partie présente un nouveau modèle de correction d'images brutes, basé sur une famille de biais. Ce modèle est dit « mixte », car il combine les approches « *calibration-based* » et « *scene-based* ». La famille de biais est construite à partir de données de calibration, tandis que le modèle de correction exploite également des informations issues de la scène pour déterminer la correction optimale.

La seconde partie est consacrée au débruitage colonnaire. Les effets de colonne, fréquents dans les modèles « *calibration-based* », résultent principalement d'une dérive temporelle des mesures du circuit de lecture électronique. Comme les données de calibration ne tiennent pas compte de cette dérive, elles deviennent mal adaptées par rapport à l'image courante, ce qui peut entraîner l'apparition d'artefacts structurés dans l'image corrigée sous forme de structures verticales aberrantes. Un post-traitement de débruitage colonnaire devient alors nécessaire pour atténuer ces artefacts tout en préservant les détails de l'image.

1 Correction d'une image à partir d'une famille de biais

1.1 Cadre de travail, hypothèses et modèle générique

1.1.1 Correction des pixels défectueux

Avant toute analyse, il est préférable de corriger les pixels défectueux présents dans les images observées. L'ensemble de ces pixels forme un masque que l'on cherche à interpoler. Pour cela, le niveau radiométrique d'un pixel défectueux est remplacé par la moyenne de ses huit voisins adjacents [1]. Lorsqu'un ensemble de pixels défectueux constitue un *cluster* (comme illustré sur la figure 3.8b), on commence par initialiser arbitrairement les valeurs

1. On considère ici le 8-voisinage, ou 8-connexité. Pour un pixel situé en bordure, le nombre de voisins peut être inférieur à huit.

à l'intérieur du *cluster*. En répétant ensuite l'opération consistant à remplacer l'intensité de chaque pixel par la moyenne de ses voisins, la solution converge progressivement vers un état où l'intensité de chaque pixel est égal à la moyenne de ses voisins, ce qui revient à imposer que le Laplacien est nul en chaque pixel défectueux. On approche ainsi la solution de l'équation de Laplace ($\Delta u = 0$), avec des conditions de Dirichlet fixées par les pixels en bordure du masque. Après suffisamment d'itérations, cette méthode produit une interpolation harmonique qui érode peu à peu les *clusters* et restitue des valeurs radiométriques cohérentes.

Remarque 9 (Une autre manière de corriger les pixels défectueux)**.** *Une autre approche possible consiste à utiliser une interpolation fondée sur l'extension lipschitzienne absolument minimale (AMLE, pour absolutely minimizing Lipschitz extension), qui préserve mieux les discontinuités et les structures géométriques [61, 62]. En effet, contrairement à l'interpolation harmonique qui lisse de façon isotrope les valeurs et tend à atténuer les contours, l'AMLE minimise la plus grande pente locale, ce qui permet de conserver des bords nets et d'éviter le sur-lissage. Cette méthode est particulièrement intéressante lorsque des pixels valides sont présents à l'intérieur du masque, car elle gère mieux les structures internes et les discontinuités. Toutefois, dans notre cas, les clusters de pixels défectueux sont en pratique des zones compactes sans pixels valides en leur sein, ce qui fait que l'interpolation harmonique et l'AMLE conduisent à des résultats très similaires. La propagation isotrope de l'interpolation harmonique suffit donc à restituer des niveaux radiométriques cohérents.*

1.1.2 Modèle affine non paramétrique de correction d'une image pseudo-brute

Nous considérons un modèle affine caractérisé par un gain constant et un biais, conformément à l'équation (3.58). Ici, elle est exprimée sous une forme non paramétrique [2]

$$U = \mathcal{G}_0\, r - B, \tag{4.1}$$

où $U, r, \mathcal{G}_0, B \in \mathbb{R}^\Omega$ désignent respectivement l'image corrigée, l'image brute, l'image de gain unique déterminée à partir des données de calibration (cf. figure 4.1a), et l'image du biais.

L'image du biais est construite de manière à satisfaire, conformément à (3.54b),

$$\mu(B) = 0, \tag{4.2}$$

où $\mu(\cdot)$ représente l'opérateur de moyenne spatiale.

Étant donnée une image brute $r \in \mathbb{R}^\Omega$, nous définissons l'image pseudo-brute $v \in \mathbb{R}^\Omega$ par

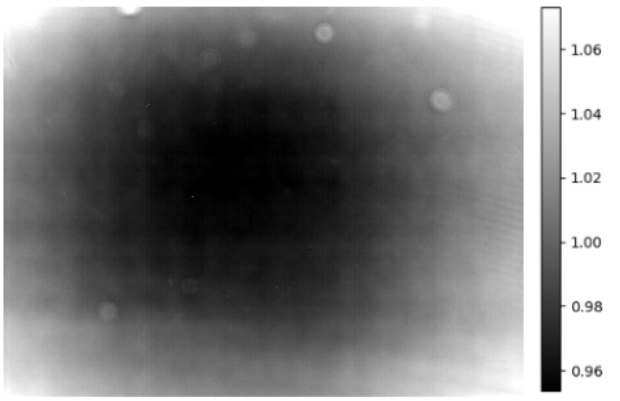
$$v = \mathcal{G}_0\, r. \tag{4.3}$$

Le modèle de correction affine s'écrit alors

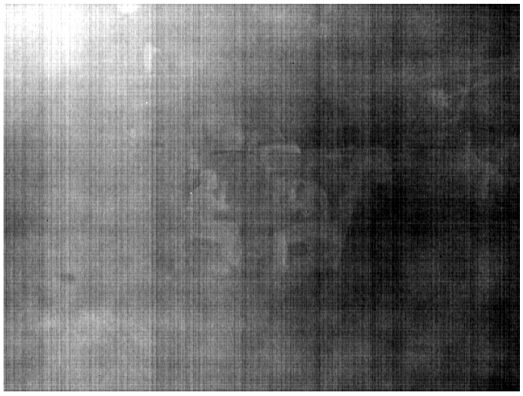
$$U = v - B, . \tag{4.4}$$

2. Cela revient à considérer tous les paramètres, l'ensemble des paramètres environnementaux θ étant trop vaste, voire infini.

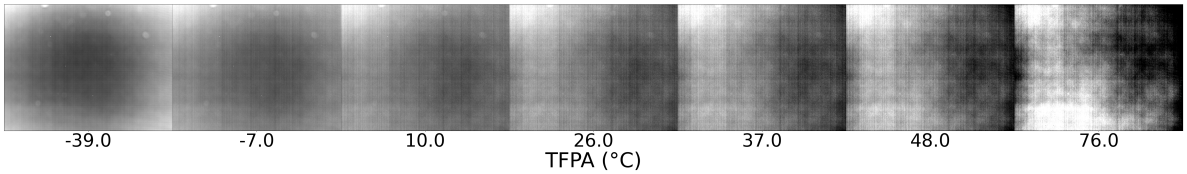
Dans le cadre de ce modèle affine, on peut considérer que l'image pseudo-brute est l'image observée (cf. figure 4.1b) et que le biais s'apparente au *fixed pattern noise* (FPN), un bruit additif qui s'ajoute à la scène, non uniforme spatialement et relativement stationnaire (cf. Chapitre 2). En soustrayant le biais B à l'image observée, l'objectif est d'éliminer ces non-uniformités structurelles propres au capteur et d'obtenir une image plus homogène spatialement.



(a) Gain unique



(b) Image pseudo-brute



(c) Famille de biais

FIGURE 4.1 – **Images du gain, d'une image pseudo-brute et d'une famille de biais.** Données issues de la caméra *ATI-640c*. (a) L'image du gain unique est calculée à partir de données de calibration acquises en exposant la caméra à deux scènes uniformes dans un même environnement, conformément à la méthode de calibration à deux points (2.22). Elle présente une structure spatiale lisse, avec des valeurs radiométriques proches de l'unité. (b) L'image pseudo-brute, considérée ici comme l'image observée, est obtenue par multiplication pixel à pixel de l'image brute par le gain unique. Acquise à une TFPA de 28 °C, elle laisse à peine deviner quelques éléments de scène (notamment deux personnes autour d'une table), mais reste fortement altérée par un bruit additif structuré typique du FPN. (c) La famille de biais est construite à partir d'images de scènes uniformes et du gain unique (seuls 7 des 40 biais constituant la famille sont représentés ici). Les biais correspondant à des TFPA proches (indiquées en abscisse) présentent des motifs similaires (effets de colonne et composantes de basse fréquence localisées aux mêmes endroits), révélant une forte corrélation interne et une dépendance progressive aux conditions environnementales. En particulier, le biais associé à la TFPA de 26 °C présente des similarités frappantes avec l'image pseudo-brute (b), acquise à une TFPA voisine, ce qui confirme que ces biais traduisent une forme structurée de FPN.

1.1.3 Famille de biais

Une famille de biais $\mathscr{B} = (B_1, \dots, B_N)$ est construite, par exemple, à partir d'images de scènes uniformes via la formule (3.59), en imposant (par soustraction d'une constante)

$$\mu(B_n) = 0, \quad \forall n \in [\![1, N]\!]. \tag{4.5}$$

Bien que le cadre soit non paramétrique, chaque biais B_n reste associé à un environnement spécifique, et en particulier à une certaine TFPA.

Dans cette section, nous posons l'hypothèse **H0**, selon laquelle la famille $\mathscr{B}$ est linéairement indépendante (c'est-à-dire libre) [3]. Formellement, cela signifie

$$\textbf{H0}: \quad \forall (\alpha_1, \dots, \alpha_N) \in \mathbb{R}^N, \quad \sum_{n=1}^{N} \alpha_n B_n = 0 \quad \Longrightarrow \quad \forall n \in [\![1, N]\!], \quad \alpha_n = 0. \tag{4.6}$$

La figure 4.1c illustre une famille typique de biais. On observe que les biais associés à des TFPA proches présentent des motifs similaires, rappelant le FPN, ce qui suggère une forte corrélation. Plus généralement, les biais sont d'autant plus corrélés que les conditions environnementales ont peu évolué pendant l'acquisition, c'est-à-dire lorsque la plage temporelle d'acquisition est courte relativement au temps caractéristique de variation de ces conditions. Cette corrélation reflète une dépendance progressive du biais (ou du FPN) aux conditions environnementales.

Nous présentons à présent un modèle générique de correction, fondé sur la recherche d'un biais optimal dans l'espace vectoriel engendré par la famille des biais, selon une métrique $F(\cdot)$.

1.1.4 Modèle générique de correction d'une image brute à partir d'une famille de biais

Le modèle de correction que nous proposons dans cette partie est le suivant.

Soit $F : \mathbb{R}^{\mathbf{\Omega}} \to \mathbb{R}$ une métrique de qualité d'image, supposée décroître avec l'amélioration de la qualité. Étant donnée une image pseudo-brute $v \in \mathbb{R}^{\Omega}$, on cherche un biais $\widehat{B}$ dans l'espace vectoriel engendré par la famille $\mathscr{B}$ tel que l'image corrigée $U = v - \widehat{B}$ soit minimale au sens de la métrique $F(\cdot)$.

Le modèle générique associé à la métrique F s'écrit

$$\widehat{B} \in \underset{B \in \mathrm{Vect}(B_1, \dots, B_N)}{\operatorname{argmin}} F(v - B), \tag{4.7}$$

et l'image corrigée est donnée par :

$$U = v - \widehat{B}. \tag{4.8}$$

Dans la littérature, un modèle relativement similaire a été proposé par Liu et al. [20] (voir chapitre 2, sous-section 2.2.4.2). Le modèle de correction proposé par les auteurs

3. Cette hypothèse est presque toujours satisfaite en raison du bruit de mesure inhérent au processus d'acquisition

consistait initialement à déterminer un décalage comme combinaison affine des images brutes de scènes uniformes. Après multiplication par le gain $\mathcal{G}_0$, cette opération revient, à une image constante près, à déterminer un biais dans l'espace vectoriel engendré par la réunion du gain et de la famille des biais. Nous reviendrons sur la comparaison de notre modèle avec cette méthode en conclusion.

Dans les sous-sections suivantes, nous étudierons trois déclinaisons particulières du modèle générique (4.7). La première est la déclinaison ℓ^2, associée à la métrique de l'écart-type. La deuxième est la déclinaison $\mathscr{H}^1$, associée à la norme euclidienne du gradient spatial. Enfin, la troisième est la déclinaison TV (*total variation*), fondée sur la variation totale.

1.2 Modèle ℓ^2

Le modèle ℓ^2 consiste à déterminer un biais appartenant à l'espace vectoriel engendré par la famille des biais $\mathscr{B}$, de sorte que l'écart-type spatial de l'image corrigée soit minimal. Par commodité d'écriture et de raisonnement, nous allons dans un premier temps définir ce modèle ℓ^2 comme consistant à déterminer un biais appartenant à l'espace vectoriel engendré par $\mathscr{B}$, de sorte que la norme euclidienne de l'image corrigée soit minimale. Écrit de la sorte, le problème devient plus compréhensible, car il s'agit d'une simple projection orthogonale. Nous montrerons ensuite l'équivalence entre cette formulation et celle fondée sur l'écart-type.

1.2.1 Existence, unicité et équivalence des formulations du problème ℓ^2

Le problème consistant à minimiser la norme euclidienne de l'image corrigée s'écrit

$$\widehat{B} \in \underset{B \in \mathrm{Vect}(B_1,\dots,B_N)}{\operatorname{argmin}} \|v - B\|_2, \tag{4.9}$$

où v désigne l'image pseudo-brute, et $(B_1, \dots, B_N)$ la famille des biais. Le problème (4.9) correspond à la projection orthogonale de v sur le convexe fermé non vide $\mathrm{Vect}(B_1, \dots, B_N)$, et admet donc une solution unique.

Nous faisons l'hypothèse **H0** (voir l'énoncé 4.6) que la famille de biais est linéairement indépendante, ce qui implique que la dimension de $\mathrm{Vect}(B_1, \dots, B_N)$ est égale à N. Lorsque nous appliquons un procédé d'orthonormalisation (par exemple Gram-Schmidt ou une SVD), on transforme la famille $(B_1, \dots, B_N)$ en une famille orthonormée $(O_1, \dots, O_N)$ vérifiant

$$\begin{cases} \mathrm{Vect}(B_1, \dots, B_N) = \mathrm{Vect}(O_1, \dots, O_N) & \text{(4.10a)} \\ \langle O_n, O_m \rangle = \delta_{n,m} \quad \forall (n,m) \in [\![1, N]\!]^2 & \text{(4.10b)} \\ \mu(O_n) = 0 \quad \forall n \in [\![1, N]\!], & \text{(4.10c)} \end{cases}$$

où $\delta_{n,m}$ désigne le symbole de Kronecker (valant 1 si $n = m$, et 0 sinon). L'égalité (4.10c) se justifie par le fait que chaque O_n est une combinaison linéaire d'éléments de moyenne spatiale nulle, conformément à (4.5).

La solution de (4.9) s'écrit alors comme la projection orthogonale de v sur $\mathrm{Vect}(O_1, \dots, O_N)$:

$$\widehat{B} = \sum_{n=1}^{N} \langle O_n, v \rangle \, O_n. \tag{4.11}$$

Nous allons désormais montrer que le problème (4.9) revient à minimiser l'écart-type spatial de l'image corrigée. Pour cela, montrons que, pour tout $k \in \mathbb{R}$, le problème (4.9) est invariant par translation d'une image constante, et donc équivalent au problème suivant

$$\widehat{B} \in \operatorname*{argmin}_{B \in \mathrm{Vect}(B_1,\dots,B_N)} \|v - k - B\|_2. \tag{4.12}$$

Preuve : Soit $\widehat{B}$ solution du problème (4.12). Par définition, cela signifie que $\widehat{B}$ est le projeté orthogonal de $v-k$ sur $\mathrm{Vect}(B_1,\dots,B_N)$. Étant donné que la famille $(O_1,\dots,O_N)$ est une base orthonormée de cet espace, on a

$$\begin{aligned}
\widehat{B} &= \sum_{n=1}^{N} \langle O_n, v-k\rangle\, O_n \\
&= \sum_{n=1}^{N} \langle O_n, v\rangle\, O_n - k \sum_{n=1}^{N} \langle O_n, \mathbf{1}_\Omega\rangle\, O_n \\
&= \sum_{n=1}^{N} \langle O_n, v\rangle\, O_n - k \sum_{n=1}^{N} \mathrm{card}(\Omega)\, \mu(O_n)\, O_n \\
&= \sum_{n=1}^{N} \langle O_n, v\rangle\, O_n \quad \text{(d'après (4.10c))}
\end{aligned}$$

On reconnaît alors (4.11), l'unique solution du problème (4.9), ce qui prouve l'équivalence annoncée. □

En choisissant $k = \mu(v)$ dans (4.12), et puisque d'après (4.5), tout $B \in \mathrm{Vect}(B_1,\dots,B_N)$ vérifie $\mu(B) = 0$, le problème (4.9) devient équivalent à

$$\widehat{B} \in \operatorname*{argmin}_{B \in \mathrm{Vect}(B_1,\dots,B_N)} \frac{\|(v-B) - \mu(v-B)\|_2^2}{\mathrm{card}(\Omega)} = \operatorname*{argmin}_{B \in \mathrm{Vect}(B_1,\dots,B_N)} \mu\Big(\big(v - B - \mu(v-B)\big)^2\Big). \tag{4.13}$$

Cette équivalence montre que minimiser la norme euclidienne de l'image corrigée revient à minimiser son écart-type spatial, à condition que les biais considérés soient centrés (c'est-à-dire de moyenne nulle). Ainsi, les deux formulations du modèle ℓ^2, l'une fondée sur la norme euclidienne et l'autre sur l'écart-type, sont rigoureusement équivalentes. Nous adopterons par la suite la formulation basée sur la norme euclidienne (4.9) pour définir le problème ℓ^2.

On peut noter que l'orthonormalisation de la base permet de calculer très simplement les coordonnées de $\widehat{B}$ dans cette base. Nous verrons dans la suite que l'utilisation de cette stratégie d'orthonormalisation de la famille $\mathcal{B}$ s'avérera d'importance majeure quand on considérera des métriques plus complexes.

Nous introduisons à présent une résolution explicite du problème ℓ^2, en recherchant $\widehat{B}$ par ses coefficients dans une base de $\mathrm{Vect}(B_1,\dots,B_N)$, grâce à une formulation matricielle.

1.2.2 Résolution par coefficients dans une base de biais et enjeux numériques

1.2.2.1 Formulation et résolution du problème via les coefficients dans une base initiale

Une manière simple et directe de représenter l'espace vectoriel $\mathrm{Vect}(B_1,\dots,B_N)$ consiste

à l'exprimer à partir des coefficients dans la base $(B_1, \dots, B_N)$

$$B \in \mathrm{Vect}(B_1, \dots, B_N) \iff \exists! \, \alpha \in \mathbb{R}^N, \quad B = \mathbf{B}\,\alpha, \tag{4.14}$$

où $\mathbf{B}$ est une matrice de taille $\mathrm{card}(\Omega) \times N$ dont les vecteurs colonnes sont les biais $B_1, \dots, B_N$ de la famille $\mathscr{B}$. L'unicité de cette décomposition repose sur l'hypothèse que $(B_1, \dots, B_N)$ forme une famille linéairement indépendante, c'est-à-dire que $\mathbf{B}$ est de rang plein. Il en résulte que la solution $\widehat{B}$ du problème (4.9) vérifie

$$\begin{cases} \widehat{B} = \mathbf{B}\,\widehat{\alpha}. & \text{(4.15a)} \\ \widehat{\alpha} \in \underset{\alpha \in \mathbb{R}^N}{\mathrm{argmin}} \, ||v - \mathbf{B}\,\alpha||_2^2 & \text{(4.15b)} \end{cases}$$

Le problème (4.15b) consiste à minimiser une fonction quadratique, convexe et différentiable. Cette fonction est également coercive. En effet, si $\mathbf{B}$ est de rang plein, il existe une constante $c > 0$ telle que $\|\mathbf{B}\alpha\|_2 \geq c\|\alpha\|_2$. On applique alors l'inégalité triangulaire inversée

$$\|v - \mathbf{B}\alpha\|_2 \geq \Big| \, \|\mathbf{B}\alpha\|_2 - \|v\|_2 \, \Big| \geq c\|\alpha\|_2 - \|v\|_2.$$

Ainsi, lorsque $\|\alpha\|_2 \to \infty$, on a $\|v - \mathbf{B}\alpha\|_2 \to \infty$, d'où la coercivité de la fonction objectif. Il suffit donc de déterminer le point critique de cette fonction.

Si on pose

$$J(\alpha) = ||v - \mathbf{B}\,\alpha||_2^2,$$

son gradient est

$$\nabla J(\alpha) = -2\,\mathbf{B}^* \, (v - \mathbf{B}\,\alpha).$$

Le point critique $\widehat{\alpha}$ vérifie alors

$$\nabla J(\widehat{\alpha}) = 0$$

On a alors

$$(\mathbf{B}^*\,\mathbf{B})\,\widehat{\alpha} = \mathbf{B}^*\,v. \tag{4.16}$$

On définit $\mathbf{M}$ la matrice de Gram par

$$\mathbf{M} = \mathbf{B}^*\,\mathbf{B}. \tag{4.17}$$

$\mathbf{M}$ est une matrice $N \times N$, elle est inversible. En effet soit $\alpha \in \mathbb{R}^N$, on a les implications suivantes

$$\begin{aligned} & \mathbf{M}\,\alpha = 0 \\ \Longrightarrow\ & (\mathbf{B}^*\,\mathbf{B})\,\alpha = 0 \\ \Longrightarrow\ & \alpha^*\,\mathbf{B}^*\,\mathbf{B}\,\alpha = 0 \quad \text{en multipliant à gauche par } \alpha^* \\ \Longrightarrow\ & ||\mathbf{B}\,\alpha||_2^2 = 0 \\ \Longrightarrow\ & \mathbf{B}\,\alpha = 0 \\ \Longrightarrow\ & \alpha = 0 \quad \text{car } \mathbf{B} \text{ de rang plein .} \end{aligned}$$

Par suite, d'après (4.16), il existe une unique solution au problème (4.15b), elle s'écrit

$$\widehat{\alpha} = (\mathbf{B}^*\,\mathbf{B})^{-1}\,\mathbf{B}^*\,v \tag{4.18}$$

En combinant (4.15) et (4.18), on obtient la solution du problème ℓ^2 (4.9) qui s'écrit

$$\widehat{B} = \mathbf{B}\,(\mathbf{B}^*\,\mathbf{B})^{-1}\,\mathbf{B}^*\,v \tag{4.19}$$

1.2.2.2 Problèmes pratiques de résolution causés par le mauvais conditionnement

La formule (4.19) met en jeu l'inversion de la matrice de Gram $\mathbf{M}$, ce qui peut engendrer des erreurs numériques selon son conditionnement. Le conditionnement d'une matrice inversible $\mathbf{M}$ relativement à la norme subordonnée euclidienne $|||\cdot|||_2$ est défini par

$$\mathrm{cond}_2(\mathbf{M}) = |||\mathbf{M}|||_2 \, |||\mathbf{M}^{-1}|||_2.$$

La matrice $\mathbf{M} = \mathbf{B}^* \mathbf{B}$ est symétrique à valeur réelle, elle est positive puisque

$$\forall \alpha \in \mathbb{R}^N, \; \alpha^* \mathbf{M} \alpha = ||\mathbf{B} \alpha||_2^2 \geq 0,$$

elle est définie car inversible. $\mathbf{M}$ est donc diagonalisable et ses valeurs propres ordonnées de manière décroissante $\lambda_1, ..., \lambda_N$ sont strictement positives. Le conditionnement de $\mathbf{M}$ par rapport à la norme euclidienne vaut :

$$\mathrm{cond}_2(\mathbf{M}) = \frac{\lambda_1}{\lambda_N}$$

En pratique, le rapport $\frac{\lambda_1}{\lambda_N}$ est souvent très grand Cette situation s'explique par la forte corrélation souvent observée entre les éléments de la famille de biais $\mathscr{B}$. Par exemple, pour la famille $\mathscr{B}$ illustrée en figure 4.1c, le conditionnement de la matrice $\mathbf{M}$ dépasse 10^{20}. Un tel conditionnement engendre une perte d'environ 20 chiffres significatifs. Or, dans un contexte de calcul en précision double, où la précision est d'environ 16 chiffres significatifs, cette perte peut induire des erreurs significatives dans le calcul des coefficients de projection via la formule (4.18).

Pour pallier le problème de mal-conditionnement, on peut effectuer une décomposition en valeurs singulières (SVD) de la matrice $\mathbf{M}$. La méthode de Jacobi unilatérale [63] permet de calculer cette SVD de manière fiable même lorsque la matrice est fortement corrélée,

$$\mathbf{M} = V \, \Sigma \, V^*, \quad V \in \mathbf{O}(N), \; \Sigma \in \mathbf{D}(N, N), \tag{4.20}$$

où $\mathbf{O}(N)$ désigne le groupe des matrices orthogonales et $\mathbf{D}(N, N)$ l'ensemble des matrices diagonales. L'inversion de $\mathbf{M}$ s'effectue alors via

$$\mathbf{M}^{-1} = V \, \Sigma^{-1} \, V^*,$$

et l'on peut régulariser ou tronquer les petites valeurs singulières pour limiter l'amplification du bruit numérique et obtenir une inversion numériquement plus précise. . En pratique, bien que cette approche réduise certaines imprécisions, elle peut conduire à des résultats aberrants, ce qui rend son usage insuffisant. Il devient donc nécessaire d'envisager une méthode alternative pour garantir une correction fiable.

1.2.2.3 Reformulation et résolution dans une base orthonormée

Pour éviter d'avoir à inverser directement la matrice de Gram associée à la base initiale $(B_1, \ldots, B_N)$, on peut reformuler le problème en utilisant une base orthonormée. On construit une famille de vecteurs orthonormés $(O_1, \ldots, O_N)$ à partir de la base initiale. La méthode de Jacobi unilatérale [63] permet de réaliser cette orthonormalisation de manière

fiable, même lorsque les vecteurs sont fortement corrélés, et fournit directement la matrice de passage $\mathbf{P}$ de la base orthonormée $(O_1, \dots, O_N)$ à la base initiale $(B_1, \dots, B_N)$. En notant $\mathbf{O}$ la matrice dont les colonnes sont les vecteurs $(O_1, \dots, O_N)$, on obtient ainsi

$$\mathbf{O} = \mathbf{B}\,\mathbf{P}. \tag{4.21}$$

Dans ce nouveau cadre, la solution $\widehat{B}$ du problème (4.9) s'exprime simplement à partir des coefficients dans la base orthonormée

$$\begin{cases} \widehat{B} = \mathbf{O}\,\widehat{\beta}, & \text{(4.22a)} \\ \widehat{\beta} \in \underset{\beta \in \mathbb{R}^N}{\operatorname{argmin}} \|v - \mathbf{O}\beta\|_2^2. & \text{(4.22b)} \end{cases}$$

En suivant les mêmes calculs que dans le paragraphe 1.2.2.1, et puisque la matrice de Gram $\mathbf{O}^*\,\mathbf{O}$ est l'identité de $\mathscr{M}_N(\mathbb{R})$, le remplacement de $\mathbf{B}$ par $\mathbf{O}$ dans (4.18) fournit directement la solution du problème (4.22b)

$$\widehat{\beta} = \mathbf{O}^*\,v, \tag{4.23}$$

ou encore si on injecte (4.21)

$$\widehat{\beta} = (\mathbf{B}\,\mathbf{P})^*\,v. \tag{4.24}$$

On évite alors le problème d'une matrice possiblement mal conditionnée à inverser.

Par suite la solution du problème ℓ^2 (4.9) s'obtient en combinant (4.22) et (4.24), ce qui donne

$$\widehat{B} = \mathbf{B}\,\mathbf{P}\,(\mathbf{B}\,\mathbf{P})^*\,v. \tag{4.25}$$

Une fois le biais $\widehat{B}$ déterminé par (4.11) (ou par (4.19) ou encore par (4.25)), l'image corrigée U s'obtient par (4.8) qui, pour rappel, s'écrit

$$U = v - \widehat{B}. \tag{4.26}$$

1.2.3 Correction d'images par le modèle ℓ^2

1.2.3.1 Correction d'une scène uniforme

Soit $v^{(T)}$ l'image pseudo-brute d'une scène uniforme de température de scène T. Comme nous l'avons vu dans la section 2.2.2, cette image de scène uniforme permet de définir un biais $\widetilde{B}$ par l'équation (3.59)

$$\widetilde{B} = v^{(T)} - \mu(v^{(T)}). \tag{4.27}$$

En appliquant le modèle affine (4.4), on obtient alors, pour l'image corrigée $\widetilde{U}$

$$\widetilde{U} = v^{(T)} - \widetilde{B} = \mu(v^{(T)}),$$

c'est-à-dire une image parfaitement uniforme. On en déduit que $\widetilde{B}$ constitue le biais « parfait » pour corriger l'image observée $v^{(T)}$.

Appliquer le modèle ℓ^2 à l'image pseudo-brute $v^{(T)}$ revient à résoudre

$$\widehat{B} \in \underset{B \in \operatorname{Vect}(B_1, \dots, B_N)}{\operatorname{argmin}} \|v^{(T)} - B\|_2. \tag{4.28}$$

L'image corrigée U s'obtient ensuite par la formule

$$U = v^{(T)} - \widehat{B}, \tag{4.29}$$

que l'on peut écrire, d'après (4.27), sous la forme

$$U = \widetilde{B} - \widehat{B} + \mu(v^{(T)}). \tag{4.30}$$

En prenant $k = \mu(v^{(T)})$ comme constante dans la formule d'invariance par translation d'une image constante (4.12), et en utilisant (4.27), le problème (4.28) est équivalent à

$$\widehat{B} \in \underset{B \in \mathrm{Vect}(B_1,\dots,B_N)}{\operatorname{argmin}} \|\widetilde{B} - B\|_2. \tag{4.31}$$

Autrement dit, $\widehat{B}$ est le projeté orthogonal du biais « parfait » $\widetilde{B}$ sur l'espace $\mathrm{Vect}(B_1, \dots, B_N)$. D'après (4.30), l'image corrigée U diffère d'une image parfaitement constante uniquement par l'écart entre $\widetilde{B}$ et son projeté orthogonal. Plus cet écart est faible, meilleure est la qualité de la correction. Comme montré dans le chapitre précédent (cf. figure 3.21), cet écart décroît lorsque la TFPA associée à $\widetilde{B}$ est proche de celles des biais constituant la famille $\mathscr{B}$. On en déduit que la correction d'une scène uniforme est d'autant plus efficace que $\mathscr{B}$ contient un ou plusieurs biais acquis à une TFPA proche de celle de la scène à corriger (par exemple 5 °C). Ce constat peut être raisonnablement généralisé aux scènes réelles et à des environnements divers, en étendant l'idée de la TFPA à des conditions environnementales proches entre l'image observée et la famille de biais.

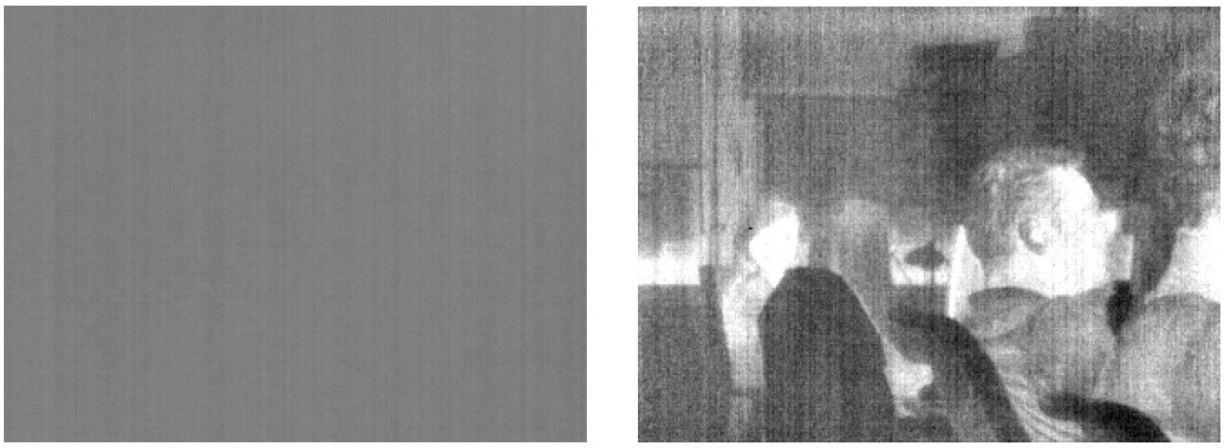
La figure 4.2a illustre la correction d'une scène uniforme par le modèle ℓ^2 : le FPN, bruit structuré lié au capteur, est efficacement supprimé par soustraction entre l'image pseudo-brute et son projeté $\widehat{B}$, produisant une image homogène.

1.2.3.2 Correction d'une scène réelle, insuffisance du modèle ℓ^2

La figure 4.2b montre que la correction de l'image pseudo-brute d'une scène réelle est significativement dégradée. Le modèle ℓ^2 consiste à déterminer un biais $\widehat{B}$ dans l'espace vectoriel $\mathrm{Vect}(B_1, \dots, B_N)$ de manière à minimiser l'écart-type spatial de l'image corrigée. Ce critère est particulièrement adapté à une scène uniforme, où les variations locales sont peu marquées. Cependant, dès lors que la scène présente une structure plus complexe, ce modèle devient inapproprié, car il ne tient pas compte des variations locales nécessaires à une restitution réaliste.

Il existe d'autres métriques plus adaptées pour évaluer la qualité d'une image, en particulier celles qui tiennent compte de la régularité spatiale. Dans une scène réelle, deux pixels voisins reçoivent généralement une irradiance similaire lorsqu'ils appartiennent à l'intérieur d'un même objet. Cependant, ce n'est plus le cas aux contours des objets, bien que ces situations soient moins fréquentes. Ainsi, le gradient spatial de l'image en un pixel est souvent proche de zéro, ce qui traduit la parcimonie du gradient spatial.

Dans ce contexte, nous nous intéressons à deux modèles qui cherchent à minimiser une certaine norme du gradient spatial de l'image reconstruite. Lorsque cette norme est la norme ℓ^1 du gradient spatial, notée $\|\cdot\|_{1,2}$, on parle du modèle TV (*total variation*) dans sa forme isotrope. En revanche, lorsque la norme utilisée est la norme euclidienne (ou ℓ^2) du gradient spatial, notée $\|\cdot\|_{2,2}$, on parle du modèle $\mathscr{H}^1$. Le modèle TV favorise une plus grande parcimonie dans le gradient spatial et pénalise moins les contours par rapport au modèle $\mathscr{H}^1$.



(a) Scène uniforme corrigée par le modèle ℓ^2 (b) Scène réelle corrigée par le modèle ℓ^2

FIGURE 4.2 – **Correction d'images par le modèle ℓ^2.** Données issues de la caméra *Smart IR320*. Une famille de 24 biais a été construite à partir d'acquisitions réalisées lors d'une calibration « en étuve » (12 paires de corps noir chaud à 50 °C et froid à 10 °C, acquises pour des TFPA quasi uniformément réparties entre −20 °C et 60 °C). Les images pseudo-brutes à corriger proviennent d'une séquence vidéo enregistrée 20 jours après la phase de calibration. L'affichage en niveaux de gris est réalisé sur une même étendue dynamique de 700 DL. La scène réelle permet d'évaluer la restitution des détails fins, tandis que la scène uniforme (une planche à température ambiante placée devant la caméra) permet de tester l'efficacité de la suppression du FPN résiduel, un bruit additif fixe lié aux non-uniformités de réponse du capteur. (a) La correction de l'image pseudo-brute d'une scène uniforme (acquise à une TFPA de 27 °C) produit une image globalement homogène et faiblement contrastée. Cependant, quelques effets de colonne persistent. (b) La correction de l'image pseudo-brute d'une scène réelle (acquise à une TFPA de 29 °C) est moins convaincante : certains objets, tels que le dossier du siège en bas au centre ou les placards en haut à droite, ne paraissent pas aussi uniformes qu'ils devraient l'être. En effet, des objets ayant une température et une émissivité relativement homogènes devraient produire une image thermique plus uniforme et moins texturée. Le modèle ℓ^2 cherche à minimiser l'écart-type spatial de l'image corrigée, ce qui le rend particulièrement efficace pour les scènes uniformes. Cependant, cette approche est inadaptée aux scènes réelles, où les différences locales de température et d'émissivité ne sont pas suffisamment préservées pour garantir une restitution fidèle à la réalité.

1.3 Modèle $\mathscr{H}^1$

1.3.1 Définition du gradient spatial d'une image et de son adjoint

Soient une image $u \in \mathbb{R}^\Omega$ et un pixel $(x, y) \in \Omega$, le gradient discret horizontal $\nabla_x(u)$ et le gradient discret vertical $\nabla_y(u)$ sont définis par

$$\nabla_x u(x,y) = \begin{cases} u(x+1,y) - u(x,y) & \text{si } 1 \leq x \leq N_c - 1 \text{ et } 1 \leq y \leq N_l \\ 0 & \text{si } x = N_c \text{ et } 1 \leq y \leq N_l, \end{cases} \tag{4.32a}$$

et

$$\nabla_y u(x,y) = \begin{cases} u(x,y+1) - u(x,y) & \text{si } 1 \leq x \leq N_c \text{ et } 1 \leq y \leq N_l - 1 \\ 0 & \text{si } 1 \leq x \leq N_c \text{ et } y = N_l, \end{cases} \tag{4.32b}$$

où N_l et N_c désignent respectivement le nombre de lignes et de colonnes des images de domaine Ω. Le gradient horizontal et le gradient vertical sont par conséquent des applications linéaires de $\mathbb{R}^\Omega$ vers $\mathbb{R}^\Omega$.

Le gradient spatial discret ∇ est une application linéaire définie par

$$\begin{array}{rccc} \nabla : & \mathbb{R}^\Omega & \mapsto & \mathbb{R}^\Omega \times \mathbb{R}^\Omega \\ & u & \mapsto & (\nabla_x u, \nabla_y u). \end{array}$$

On montre facilement que si $\nabla u = 0$, alors il existe $k \in \mathbb{R}$ tel que $\forall (x, y) \in \Omega,\ u(x, y) = k$. Ce résultat est trivial puisque $\nabla u = 0$ signifie que toutes les différences finies horizontales et verticales sont nulles.

Lemme 1. *L'opérateur ∇ est injectif lorsqu'il est restreint à* $\mathrm{Vect}(B_1, \ldots, B_N)$.

Preuve : Soit $B \in \mathrm{Vect}(B_1, \ldots, B_N)$, avec

$$B = \sum_{i=1}^{N} \alpha_i B_i$$

tel que $\nabla B = 0$. Alors, il existe $k \in \mathbb{R}$ tel que $B(x, y) = k$ pour tout $(x, y) \in \Omega$. Or, les biais B_i satisfont la contrainte $\mu(B_i) = 0$ (cf. (4.5)). Donc,

$$\mu(B) = \sum_{i=1}^{N} \alpha_i \mu(B_i) = 0,$$

ce qui implique $k = 0$. D'où $B = 0$, et par conséquent ∇ est injectif sur $\mathrm{Vect}(B_1, \ldots, B_N)$. □

1.3.2 Problème $\mathscr{H}^1$, existence et unicité de la solution

Soit $v \in \mathbb{R}^\Omega$ une image pseudo-brute. Le problème associé au modèle $\mathscr{H}^1$ consiste à déterminer le biais $\widehat{B} \in \mathrm{Vect}(B_1, \ldots, B_N)$ qui minimise la norme euclidienne du gradient spatial de l'image corrigée, c'est-à-dire

$$\widehat{B} \in \operatorname*{argmin}_{B \in \mathrm{Vect}(B_1, \ldots, B_N)} \|\nabla(v - B)\|_{2,2}. \tag{4.33}$$

La norme mixte $\|p\|_{2,2}$ pour $p = (p_1, p_2) \in \mathbb{R}^\Omega \times \mathbb{R}^\Omega$ est définie par

$$\|p\|_{2,2} = \sqrt{\sum_{(x,y)\in\Omega} \Big(p_1(x, y)^2 + p_2(x, y)^2\Big)}.$$

Comme l'opérateur ∇ est linéaire, un changement de variable montre que le problème (4.33) équivaut à

$$\begin{cases} \nabla \widehat{B} = \widehat{Z}, & \text{(4.34a)} \\ \widehat{Z} \in \operatorname*{argmin}_{Z \in \mathrm{Vect}(\nabla B_1, \ldots, \nabla B_N)} \|\nabla v - Z\|_{2,2}. & \text{(4.34b)} \end{cases}$$

Puisqu'il s'agit de la projection orthogonale de ∇v sur un sous-espace vectoriel fermé de $\mathbb{R}^\Omega \times \mathbb{R}^\Omega$, le problème (4.34b) admet une unique solution $\widehat{Z}$.

De plus, cette solution vérifie

$$\exists! \, \widehat{B} \in \mathrm{Vect}(B_1, \dots, B_N), \quad \widehat{Z} = \nabla \widehat{B}, \tag{4.35}$$

l'existence découlant de la linéarité de ∇, et l'unicité du lemme 1.

Par conséquent, comme le problème (4.34b) admet une solution unique, et que cette dernière correspond à un unique élément $\widehat{B} \in \mathrm{Vect}(B_1, \dots, B_N)$ via (4.35), on conclut d'après (4.34a) que le problème $\mathscr{H}^1$ (4.33) admet une solution unique.

1.3.3 Formulation et résolution du problème selon les coefficients dans une base

1.3.3.1 Formulation et résolution via les coefficients dans la base d'origine

En raison de la linéarité de ∇ et du lemme 1, nous avons l'équivalence suivante

$$(B_1, \dots, B_N) \text{ est une famille libre} \iff (\nabla B_1, \dots, \nabla B_N) \text{ est une famille libre.} \tag{4.36}$$

Puisque nous faisons l'hypothèse **H0** (voir (4.6)) d'indépendance linéaire de la famille $(B_1, \dots, B_N)$, on en déduit que la famille $(\nabla B_1, \dots, \nabla B_N)$ est également linéairement indépendante.

Pour résoudre le problème $\mathscr{H}^{\mathbf{1}}$ (4.33), nous procédons comme pour le modèle ℓ^2, en le reformulant à l'aide d'un vecteur de coefficients (voir sous-section 1.2.2.1). Dans ce cadre, l'opérateur gradient ∇ est considéré comme un opérateur linéaire allant de $\mathbb{R}^\Omega$ vers $\mathbb{R}^{2\,\mathrm{card}(\Omega)}$, obtenu en concaténant les composantes horizontales et verticales. Cette représentation permet de simplifier la notation et de traiter le gradient comme un vecteur dans un espace euclidien de dimension $2\,\mathrm{card}(\Omega)$. Le problème $\mathscr{H}^{\mathbf{1}}$ (4.33) est alors équivalent à

$$\begin{cases} \nabla \widehat{B} = \nabla \mathbf{B}\, \widehat{\alpha}, & \text{(4.37a)} \\ \widehat{\alpha} \in \underset{\alpha \in \mathbb{R}^N}{\mathrm{argmin}} \, \|\nabla v - \nabla \mathbf{B}\, \alpha\|_{2,2}^2, & \text{(4.37b)} \end{cases}$$

où $\nabla \mathbf{B}$ est une matrice de taille $2\,\mathrm{card}(\Omega) \times N$ dont les colonnes sont les gradients des biais $(\nabla B_1, \dots, \nabla B_N)$. Enfin, la norme mixte $\|p\|_{2,2}$, pour un élément $p \in \mathbb{R}^{2\,\mathrm{card}(\Omega)}$, correspond simplement à la norme euclidienne de p.

Soit $\widehat{B}$ la solution de (4.33). Puisque $\widehat{B}$ est combinaison linéaire d'éléments de moyenne nulle, nous avons $\mu(\widehat{B}) = 0$ et pour les mêmes raisons nous avons aussi $\forall \widehat{\alpha} \in \mathbb{R}^N$, $\mu(\mathbf{B}\, \widehat{\alpha}) = 0$. On a alors

$$\begin{aligned} & \nabla \widehat{B} = \nabla \mathbf{B}\, \widehat{\alpha} \\ \Longrightarrow & \nabla \widehat{B} = \sum_{i=1}^{N} \widehat{\alpha}_i \nabla B_i \\ \Longrightarrow & \nabla (\widehat{B} - \sum_{i=1}^{N} \widehat{\alpha}_i \, B_i) = 0 \text{ par linéarité de } \nabla \\ \Longrightarrow & \widehat{B} = \mathbf{B}\, \widehat{\alpha} \text{ d'après le lemme 1,} \end{aligned}$$

et la réciproque est également vraie.

Le problème $\mathscr{H}^{\mathbf{1}}$ (4.37) est alors équivalent à

$$\begin{cases} \widehat{B} = \mathbf{B}\,\widehat{\alpha}, & \text{(4.38a)} \\ \widehat{\alpha} \in \underset{\alpha\in\mathbb{R}^N}{\operatorname{argmin}} ||\nabla v - \nabla\mathbf{B}\,\alpha||_{2,2}^2. & \text{(4.38b)} \end{cases}$$

On applique le procédé déjà vu dans la sous-section 1.2.2.1, en remplaçant v par ∇v et $\mathbf{B}$ par $\nabla\mathbf{B}$ dans le problème (4.15), la solution de (4.38b) s'obtient après remplacement dans l'équation (4.18). Ce qui donne

$$\widehat{\alpha} = (\,\nabla\mathbf{B}^*\,\nabla\mathbf{B})^{-1}\;\nabla\mathbf{B}^*\,\nabla\,v. \tag{4.39}$$

La matrice de Gram $\nabla\mathbf{B}^*\,\nabla\mathbf{B}$ est inversible car $(\nabla\,B_1, ..., \nabla\,B_N)$ est une famille linéairement indépendante. En combinant (4.38a) et (4.39), on obtient la solution au problème $\mathscr{H}^{\mathbf{1}}$ (4.33) qui s'écrit

$$\widehat{B} = \mathbf{B}\,(\,\nabla\mathbf{B}^*\,\nabla\mathbf{B})^{-1}\;\nabla\mathbf{B}^*\,\nabla\,v. \tag{4.40}$$

En pratique, le conditionnement de la matrice de Gram ${}^t\nabla\mathbf{B}\,\nabla\mathbf{B}$ est souvent très élevé, ce qui rend son inversion numériquement instable. Pour éviter ces difficultés, on reformule le problème dans une base orthonormale.

1.3.3.2 Reformulation du problème dans une base orthonormée

On orthonormalise la famille $(\nabla B_1, \ldots, \nabla B_N)$ et l'on note $(W_1, \ldots, W_N)$ la base orthonormale ainsi obtenue.

On considère alors la matrice $\mathbf{W}$ de taille $2\,\mathrm{card}(\Omega) \times N$, dont les vecteurs colonnes sont $(W_1, \ldots, W_N)$. En notant $\mathbf{Q}$ la matrice de passage de la base $(W_1, \ldots, W_N)$ vers la base $(\nabla B_1, \ldots, \nabla B_N)$, on a la relation

$$\mathbf{W} = \nabla\mathbf{B}\,\mathbf{Q}. \tag{4.41}$$

Le problème $\mathscr{H}^{\mathbf{1}}$ (4.33) est équivalent à

$$\begin{cases} \nabla\widehat{B} = \nabla\mathbf{B}\,\mathbf{Q}\,\widehat{\beta}. & \text{(4.42a)} \\ \widehat{\beta} \in \underset{\beta\in\mathbb{R}^N}{\operatorname{argmin}} \|\nabla v - \mathbf{W}\beta\|_{2,2}^2, & \text{(4.42b)} \end{cases}$$

En suivant le même raisonnement que celui ayant permis d'établir l'équivalence entre (4.37a) et (4.38a), on montre que le problème (4.42) est équivalent à

$$\begin{cases} \widehat{B} = \mathbf{B}\,\mathbf{Q}\,\widehat{\beta}. & \text{(4.43a)} \\ \widehat{\beta} \in \underset{\beta\in\mathbb{R}^N}{\operatorname{argmin}} \|\nabla v - \mathbf{W}\beta\|_{2,2}^2, & \text{(4.43b)} \end{cases}$$

Le problème (4.43b) est équivalent à (4.38) en posant simplement $\nabla\mathbf{B} = \mathbf{W}$. Dès lors, à l'instar du calcul ayant conduit à la formule (4.39), on obtient directement que la solution de (4.43b) s'écrit

$$\widehat{\beta} = \left(\mathbf{W}^*\mathbf{W}\right)^{-1}\mathbf{W}^*\nabla v. \tag{4.44}$$

Puisque les vecteurs colonnes de $\mathbf{W}$ forment une base orthonormée, la solution de (4.43b) s'écrit aussi

$$\widehat{\beta} = \mathbf{W}^* \nabla v. \tag{4.45}$$

La solution du problème $\mathscr{H}^{\mathbf{1}}$ (4.33) s'obtient alors en combinant (4.43a), (4.45) et (4.41) :

$$\widehat{B} = \mathbf{B}\,\mathbf{Q}\,(\nabla \mathbf{B}\,\mathbf{Q})^* \nabla v. \tag{4.46}$$

Ce procédé permet d'éviter l'inversion directe d'une matrice possiblement mal conditionnée, comme dans la formule (4.40). Il repose sur l'orthonormalisation d'une famille de vecteurs possiblement fortement corrélés, étape nécessaire dans nos expérimentations et sans problème pratique. Cette orthonormalisation, par exemple via la méthode de Jacobi unilatérale [63], permet de construire une base orthonormée et la matrice de passage correspondante, assurant ainsi la stabilité numérique et limitant la perte de précision. En cas de corrélation extrêmement forte, certaines valeurs singulières peuvent être très proches de zéro. La régularisation ou la troncature de ces valeurs permet alors de limiter l'amplification du bruit et de garantir une bonne précision lors du calcul des projections dans la base initiale.

1.3.4 Correction d'images par le modèle $\mathscr{H}^{\mathbf{1}}$

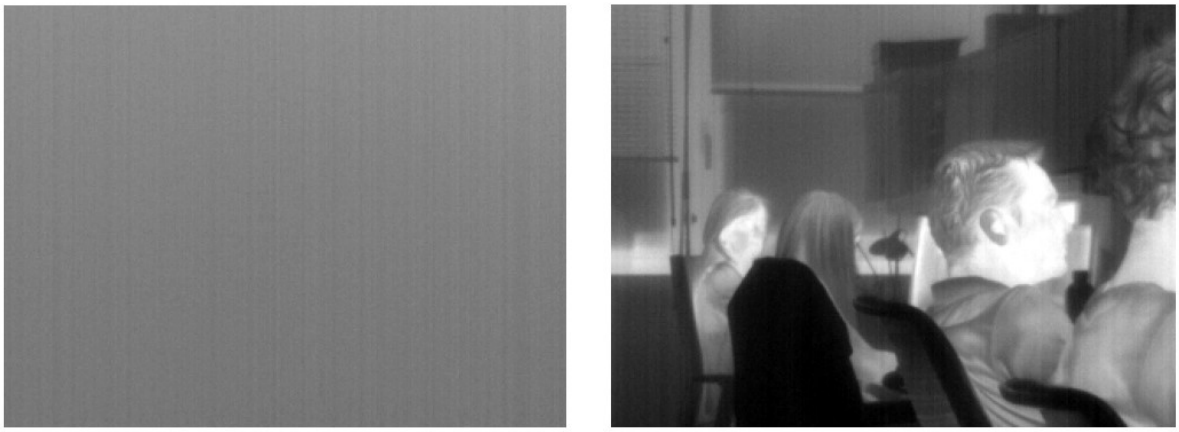
Une fois le problème $\mathscr{H}^{\mathbf{1}}$ (4.33) résolu par la formule (4.40) ou la formule (4.46), nous obtenons l'image corrigée U par la formule (4.4). La figure 4.3 montre que la correction d'une image pseudo-brute est tout à fait convenable. Si l'on compare avec le modèle ℓ^2 (figure 4.2), on remarque que la correction d'une scène uniforme donne des résultats visuellement à peu près semblables, mais que la correction d'une scène réelle est d'une bien meilleure qualité pour le modèle $\mathscr{H}^{\mathbf{1}}$.

Certaines méthodes de débruitage d'image consistent à minimiser une fonction coût composée de deux termes : un terme d'attache aux données et un terme de régularisation. Le premier quantifie la fidélité de l'image reconstruite par rapport à l'image bruitée en utilisant une norme de la différence entre ces deux images. Le second impose des contraintes sur la structure de l'image recherchée, en s'appuyant sur une norme du gradient spatial. Lorsque cette norme est la norme euclidienne (ou ℓ^2) appliquée au gradient (notée $||\nabla \cdot ||_{2,2}$), on parle de régularisation $\mathscr{H}^1$. En revanche, si l'on utilise la norme ℓ^1 du gradient (souvent notée $||\nabla \cdot ||_{1,2}$), on parle de régularisation par variation totale (*Total Variation*, TV). La régularisation $\mathscr{H}^1$ favorise les variations douces, mais tend à lisser les contours nets, ce qui peut entraîner une perte de détails (effet de flou). À l'inverse, TV est capable de préserver les discontinuités, notamment les bords d'objets, mais peut introduire des artefacts visuels « en escaliers », un phénomène connu sous le nom de *staircasing*. En raison de sa capacité à préserver les contours d'objet, TV est souvent préférée à $\mathscr{H}^1$ dans les méthodes de débruitage. Cependant, la nature non différentiable [4] de la norme ℓ^1 rend la minimisation plus délicate. On fait alors appel à des algorithmes de type primaux-duaux, spécialement conçus pour traiter efficacement ce type de fonctionnelles non lisses.

4. La norme ℓ^2 est également non différentiable mais son carré l'est, ce qui fait que le problème $\mathscr{H}^{\mathbf{1}}$ que nous avons développé est un problème d'optimisation différentiable.

Les modèles que nous proposons se distinguent des méthodes classiques de débruitage : les rôles du terme d'attache aux données et du terme de régularisation y sont, en quelque sorte, entrelacés.

Nous allons à présent nous concentrer sur le modèle reposant sur la régularisation TV de l'image corrigée à partir d'une famille de biais.



(a) Scène uniforme (b) Scène réelle

FIGURE 4.3 – **Correction d'images par le modèle $\mathscr{H}^1$.** La famille des biais ainsi que les deux images pseudo-brutes à corriger sont identiques à celles de la figure 4.2. L'affichage en niveaux de gris est réalisé sur une étendue dynamique commune de 700 DL. (a) La correction de l'image pseudo-brute d'une scène uniforme par le modèle $\mathscr{H}^1$ donne une image faiblement contrastée d'apparence uniforme, bien que quelques effets colonnaires persistent. (b) La correction de l'image pseudo-brute d'une scène réelle par le modèle $\mathscr{H}^1$ montre de bons résultats : les détails fins ainsi que la texture de la scène sont bien préservés, bien que quelques effets colonnaires demeurent.

1.4 Modèle TV

1.4.1 Problème TV

Soit une image pseudo-brute $v \in \mathbb{R}^\Omega$, le modèle TV consiste à trouver le biais $\widehat{B}$ dans l'espace vectoriel $\mathrm{Vect}(B_1, ..., B_N)$ qui minimise la TV de l'image corrigée. Il s'écrit

$$\widehat{B} \in \underset{B \in \mathrm{Vect}(B_1,...,B_N)}{\operatorname{argmin}} ||\nabla (v - B)||_{1,2}. \tag{4.47}$$

où la norme $||\cdot||_{1,2} : \mathbb{R}^\Omega \times \mathbb{R}^\Omega \to \mathbb{R}_+$ est définie par

$$\forall p = (p_x, p_y) \in \mathbb{R}^\Omega \times \mathbb{R}^\Omega, \quad ||p||_{1,2} = \sum_{(x,y)\in\Omega} \sqrt{p_x(x,y)^2 + p_y(x,y)^2}. \tag{4.48}$$

À l'instar du modèle $\mathscr{H}^1$ (4.33), le problème TV (4.47) consiste à déterminer le projeté de ∇v sur le convexe fermé non vide $\mathrm{Vect}(\nabla B_1, \ldots, \nabla B_N)$ selon la norme mixte $||\cdot||_{1,2}$.

L'existence d'une solution est assurée, car il s'agit de la projection d'un point sur un fermé convexe dans un espace de Hilbert. En revanche, l'unicité n'est pas garantie, puisque la norme $||\cdot||_{1,2}$, définie comme une somme de normes ℓ^2 par blocs, hérite de la structure non strictement convexe de la composante ℓ^1, ce qui implique que le minimum peut être atteint sur un ensemble non réduit à un point, par exemple un segment.

Remarque 10 (Exemple de non-unicité de projection en norme ℓ^1 dans le plan $\mathbb{R}^2$)**.** *Soit $p = (1,1) \in \mathbb{R}^2$ et l'ensemble convexe fermé*

$$C = \{(x, -x) \mid x \in \mathbb{R}\},$$

qui correspond à la deuxième bissectrice dans $\mathbb{R}^2$.

La distance en norme ℓ^1 entre p et $q = (x, -x) \in C$ est

$$d_1(p, q) = |1 - x| + |1 + x| = 2\max(|x|, 1).$$

Pour tout $x \in [-1, 1]$, cette distance est minimale et constante puisqu'elle vaut 2.

La projection de p sur C n'est donc pas unique : c'est tout le segment $\{(x, -x) \mid x \in [-1, 1]\}$.

Cet exemple illustre que la projection en norme ℓ^1, et par extension la norme mixte $||\cdot||_{1,2}$, sur un convexe fermé peut ne pas être unique.

1.4.2 Formulation du problème selon les coefficients dans une base, résolution avec plusieurs variantes de Chambolle-Pock

1.4.2.1 Formulation et résolution via les coefficients dans la base d'origine

Pour le modèle $\mathscr{H}^{\mathbf{1}}$, nous avons montré l'équivalence entre les problèmes (4.33) et (4.38). De manière analogue, on peut établir que le problème TV (4.47) est équivalent au problème suivant

$$\begin{cases} \widehat{B} = \mathbf{B}\,\widehat{\alpha}, & \text{(4.49a)} \\ \widehat{\alpha} \in \underset{\alpha \in \mathbb{R}^N}{\operatorname{argmin}} \left\| \nabla \mathbf{B}\,\alpha - \nabla\, v \right\|_{1,2}. & \text{(4.49b)} \end{cases}$$

Les colonnes de $\mathbf{B}$ sont les biais $(B_1, \ldots, B_N)$, considérés comme des vecteurs réels de dimension $\operatorname{card}(\Omega)$, tandis que les colonnes de $\nabla\mathbf{B}$ sont les gradients correspondants $(\nabla B_1, \ldots, \nabla B_N)$, vus comme des vecteurs de dimension $2\operatorname{card}(\Omega)$.

La norme mixte $\|\cdot\|_{1,2}$ pour $p = (p_1, \ldots, p_{2\operatorname{card}(\Omega)})$ est définie par

$$\|p\|_{1,2} = \sum_{i=1}^{\operatorname{card}(\Omega)} \sqrt{p_i^2 + p_{i+\operatorname{card}(\Omega)}^2}, \tag{4.50}$$

ce qui constitue une redéfinition adaptée au cadre matriciel.

Nous allons maintenant résoudre le problème (4.49b) de plusieurs manières à l'aide de l'algorithme de Chambolle–Pock.

1.4.2.2 Résolution numérique par l'algorithme de Chambolle–Pock avec choix de pas égaux

Dans leur article « *A first-order primal-dual algorithm for convex problems with applications to imaging* » [28], Chambolle et Pock réécrivent un problème de minimisation non différentiable sous la forme d'un problème de recherche de point selle, en s'appuyant sur la transformée de Legendre–Fenchel d'une fonction.

Les calculs détaillés ainsi que les définitions sont fournis en annexe (voir Annexe 2.6).

Une solution $\widehat{\alpha}$ du problème (4.49b) peut alors être approchée comme la limite d'une suite $(\alpha^{(k)})_{k\in\mathbb{N}}$ construite via un schéma numérique primal-dual itératif, défini par

$$\forall k \in \mathbb{N}, \quad \begin{cases} p^{(k+1)} = \operatorname{proj}_{\mathbf{b}}\left(p^{(k)} - \sigma\,\nabla v + \sigma\,\nabla\mathbf{B}\,\overline{\alpha}^{(k)}\right), \\ \alpha^{(k+1)} = \alpha^{(k)} - \tau\,\nabla\mathbf{B}^*\,p^{(k+1)}, \\ \overline{\alpha}^{(k+1)} = \alpha^{(k+1)} + \theta(\alpha^{(k+1)} - \alpha^{(k)}). \end{cases} \tag{4.51}$$

avec les conditions

$$\begin{cases} \sigma > 0, \quad \tau > 0, \\ \tau\,\sigma\,\rho(\nabla\mathbf{B}^*\,\nabla\mathbf{B}) < 1, \\ \theta = 1. \end{cases} \tag{4.52}$$

On peut choisir dans un premier temps les conditions (4.52) comme suit

$$\begin{cases} \sigma = \dfrac{0.99}{\rho(\nabla\mathbf{B}^*\,\nabla\mathbf{B})} \\ \tau = \dfrac{0.99}{\rho(\nabla\mathbf{B}^*\,\nabla\mathbf{B})} \\ \theta = 1, \end{cases} \tag{4.53}$$

ce qui garantit la convergence du schéma numérique.

On dit qu'un schéma est stable lorsque l'énergie associée diminue de manière régulière après un nombre conséquent d'itérations. Il est instable si, même après de nombreuses itérations, l'énergie peut remonter, contredisant l'objectif de minimisation.

Théoriquement, le schéma défini par (4.51) et (4.53) est garanti convergent. En pratique, cette convergence peut être lente et présenter des fluctuations, notamment lorsque la base $\nabla\mathbf{B}$ est mal conditionnée.

Pour mieux analyser ce phénomène, on définit l'image corrigée à l'itération k par

$$U^{(k)} = v - \mathbf{B}\,\alpha^{(k)}, \tag{4.54}$$

où $\alpha^{(k)} \in \mathbb{R}^N$ est le vecteur des coefficients à la $k^{\text{ème}}$ itération, et l'énergie correspondante par

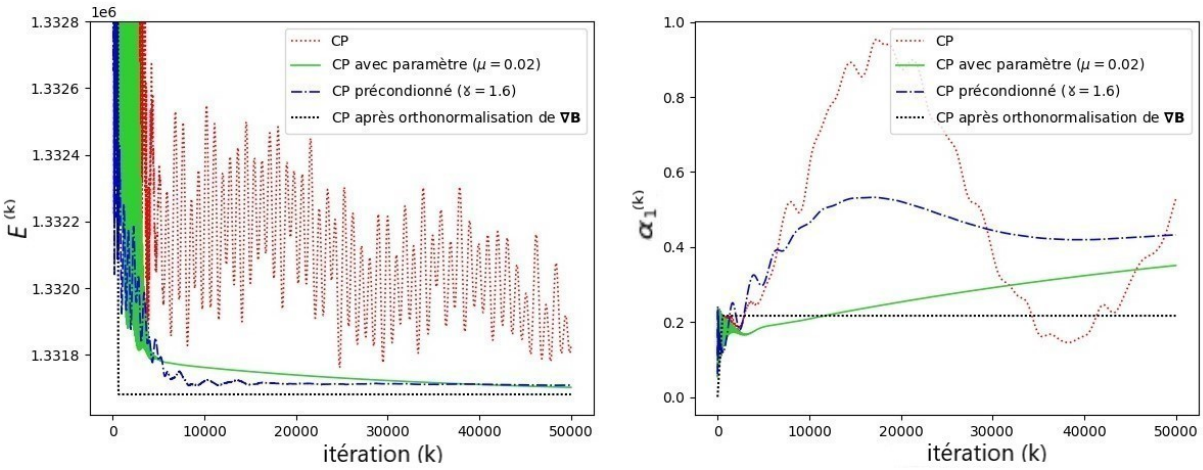
$$E^{(k)} = \left\|\nabla U^{(k)}\right\|_{1,2}, \tag{4.55}$$

quantité que l'on cherche à minimiser dans (4.49b) (et aussi dans (4.47)).

La figure 4.4 montre que le schéma peut présenter de fortes fluctuations malgré un grand nombre d'itérations. Ces instabilités peuvent être dues à un mauvais choix des pas d'itération ou au mauvais conditionnement de la base $\nabla\mathbf{B}$.

Dans les applications classiques du schéma de Chambolle-Pock au débruitage d'image, les variables primale et duale évoluent dans des espaces de dimensions comparables. Dans ce contexte, ce type d'instabilité est rarement observé. Dans notre cas, la variable primale $\alpha \in \mathbb{R}^N$ appartient à un espace de dimension bien inférieure à celle de la variable duale $p \in \mathbb{R}^{2\,\mathrm{card}(\Omega)}$, avec $N \ll 2\,\mathrm{card}(\Omega)$. Le choix standard de pas égaux $\tau = \sigma$ dans (4.53) peut alors provoquer les instabilités observées.

Pour y remédier, nous proposons d'introduire une version du schéma de Chambolle-Pock avec des pas non nécessairement égaux.



(a) Évolution de l'énergie (b) Évolution de la première composante de α

FIGURE 4.4 – **Évolution de l'énergie à minimiser et de la solution du problème TV reformulé selon (4.49b), au fil des itérations, pour différentes variantes de l'algorithme de Chambolle–Pock.** Le vecteur α est de dimension 24 correspondant au cardinal de la famille de biais. (a) L'énergie $E^{(k)}$ décroît de manière très irrégulière avec Chambolle–Pock à pas égaux (courbe pointillés rouges), illustrant un schéma instable et une convergence lente. L'introduction de pas non nécessairement égaux via un paramètre d'ajustement (courbe verte) réduit les fluctuations et stabilise le schéma, mais la convergence reste lente. La version préconditionnée (courbe noire discontinue) accélère légèrement la décroissance initiale mais demeure relativement instable et n'atteint pas le minimum observé avec la base orthonormée (courbe pointillés noirs), qui converge très rapidement après quelques centaines d'itérations. (b) La première composante $\alpha_1^{(k)}$ varie fortement avec des pas égaux, moins avec les versions ajustée et préconditionnée. Dans la version orthonormée, $\alpha_1^{(k)}$ converge rapidement vers une valeur limite autour de 0.2. Travailler dans une base orthonormée suffit à stabiliser le schéma et accélère fortement la vitesse de convergence.

1.4.2.3 Résolution numérique par l'algorithme de Chambolle–Pock avec pas variables

Il s'agit de considérer le schéma (4.51) en prenant les conditions suivantes, qui vérifient

(4.52)

$$\begin{cases} \sigma = \dfrac{0.99}{\mu\,\rho(\nabla\mathbf{B}^*\,\nabla\mathbf{B})} \\ \tau = \dfrac{0.99\,\mu}{\rho(\nabla\mathbf{B}^*\,\nabla\mathbf{B})} \\ \theta = 1 \\ \mu > 0 \quad \text{paramètre à ajuster.} \end{cases} \tag{4.56}$$

La figure 4.4 montre qu'avec l'introduction des pas variables, l'instabilité de l'énergie et du vecteur de coefficients au fil des itérations est nettement atténuée. Toutefois, la convergence reste lente, donc insatisfaisante. C'est pourquoi nous allons chercher une variante algorithmique différente, à partir d'un schéma de Chambolle-Pock avec préconditionnement.

1.4.2.4 Résolution numérique par l'algorithme de Chambolle–Pock préconditionné

Dans un autre article « *Diagonal preconditioning for first order primal-dual algorithms in convex optimization* » [29], Chambolle et Pock ont proposé une variante de leur schéma numérique, spécifiquement conçue pour les cas où l'opérateur linéaire présente une structure interne complexe, susceptible de ralentir la convergence. Dans notre situation, l'opérateur concerné est $\nabla\mathbf{B}$, dont les vecteurs colonnes sont fortement corrélés, ce qui constitue un cas ralentissant la convergence.

Nous désignerons cette variante sous le nom de « schéma de Chambolle-Pock préconditionné ». Dans leur article, les auteurs introduisent une famille de préconditionneurs paramétrés par $\gamma \in [0, 2]$, spécialement conçue pour accélérer la convergence du schéma.

Le schéma associé à Chambolle-Pock préconditionné [29] s'écrit

$$\forall k \in \mathbb{N} \begin{cases} p^{(k+1)} = \underset{\mathbf{b}}{\text{proj}}\left(p^{(k)} - \mathbf{\Sigma}\,\nabla\,v_0 + \mathbf{\Sigma}\;\nabla\mathbf{B}\,\overline{\alpha}^{(k)}\right) \\ \alpha^{(k+1)} = \alpha^{(k)} - \mathbf{T}\,\nabla\mathbf{B}^*\,p^{(k+1)} \\ \overline{\alpha}^{(k+1)} = \alpha^{(k+1)} + \theta(\alpha^{(k+1)} - \alpha^{(k)}), \end{cases} \tag{4.57}$$

où $\mathbf{\Sigma}$ et $\mathbf{T}$ sont des matrices symétriques définies positives de taille respective $2\,\text{card}(\Omega) \times 2\,\text{card}(\Omega)$ et $N \times N$. La convergence du schéma est garantie dès lors que les conditions

$$\begin{cases} ||\mathbf{\Sigma}^{\frac{1}{2}}\,\nabla\mathbf{B}\,\mathbf{T}^{\frac{1}{2}}||_2^2 < 1, \\ \theta = 1, \end{cases} \tag{4.58}$$

sont satisfaites.

Chambolle et Pock proposent de considérer

$$\mathbf{T} := \begin{cases} \mathbf{T}_{i,i} = \dfrac{1}{\sum\limits_{1\le j\le 2\,\text{card}(\Omega)} |\nabla\mathbf{B}_{i,j}|^{2-\gamma}}, \quad i \in [\![1, N]\!] \\ \mathbf{T}_{i,k} = 0 \text{ si } i \neq k, \quad (i,k) \in [\![1, N]\!]^2, \end{cases} \tag{4.59a}$$

$$\mathbf{\Sigma} := \begin{cases} \mathbf{\Sigma}_{j,j} = \dfrac{1}{\sum\limits_{1\leq i\leq N} |\nabla \mathbf{B}_{i,j}|^{\gamma}}, \quad j \in [\![1, 2\,\mathrm{card}(\Omega)]\!] \\ \\ \mathbf{\Sigma}_{j,l} = 0 \text{ si } j \neq l, \quad (j,l) \in [\![1, 2\,\mathrm{card}(\Omega)]\!]^2, \end{cases} \tag{4.59b}$$

$$0 \leq \gamma \leq 2. \tag{4.59c}$$

Ainsi définies, $\mathbf{\Sigma}$ et $\mathbf{T}$ vérifient (4.58) [5].

La figure 4.4 montre que la convergence du schéma de Chambolle-Pock préconditionné reste lente malgré un réglage fin du paramètre $\gamma \in [0, 2]$.

Une recherche bibliographique nous a conduits à nous intéresser aux méthodes dites de « *matching pursuit* ». Ces techniques algorithmiques visent à décomposer un signal comme une combinaison linéaire d'éléments issus d'un dictionnaire. Lorsqu'un dictionnaire redondant (c'est-à-dire composé d'éléments fortement corrélés) est utilisé, ces algorithmes peuvent s'avérer particulièrement lents [6]. Gribonval et al. recommandent, dans ce contexte, de sélectionner en priorité les atomes présentant la plus forte corrélation avec le signal, afin de favoriser une décomposition parcimonieuse [64]. Par ailleurs, Davis et al. [65, 66] ont montré que de meilleures propriétés de convergence pouvaient être obtenues en imposant une orthogonalité entre les atomes sélectionnés.

Il nous semble alors pertinent d'orthonormaliser les vecteurs colonnes de la matrice $\nabla\mathbf{B}$, comme nous l'avions proposé pour résoudre le modèle $\mathscr{H}^{\mathbf{1}}$ (dans la sous-section 1.3.3.2).

1.4.2.5 Résolution numérique dans une base orthonormée

Nous reprenons les matrices introduites dans le cadre du modèle $\mathscr{H}^1$: la matrice $\mathbf{W}$, dont les colonnes forment une famille orthonormale, ainsi que la matrice de passage $\mathbf{Q}$, liées par la relation (4.41). Le problème (4.49b) est alors équivalent à

$$\begin{cases} \widehat{\alpha} = \mathbf{Q}\,\widehat{\beta} & \text{(4.60a)} \\ \mathbf{W} = \nabla\mathbf{B}\,\mathbf{Q} & \text{(4.60b)} \\ \widehat{\beta} \in \underset{\beta\in\mathbb{R}^N}{\mathrm{argmin}} \|\mathbf{W}\,\beta - \nabla\, v\|_{1,2} & \text{(4.60c)} \end{cases}$$

La résolution de (4.60c) s'effectue à l'aide du schéma de Chambolle-Pock, comme cela a été fait précédemment pour (4.49b). Il suffit ici de remplacer l'opérateur $\nabla\mathbf{B}$ par $\mathbf{W}$ dans le schéma (4.51), ce qui conduit au schéma suivant

$$\forall k \in \mathbb{N} \begin{cases} p^{(k+1)} = \underset{\mathbf{b}}{\mathrm{proj}} \left(p^{(k)} - \sigma\,\nabla\, v_0 + \sigma\,\mathbf{W}\,\overline{\beta}^{(k)} \right) \\ \beta^{(k+1)} = \beta^{(k)} - \tau\ \mathbf{W}^*\,p^{(k+1)} \\ \overline{\beta}^{(k+1)} = \beta^{(k+1)} + \theta(\beta^{(k+1)} - \beta^{(k)}), \end{cases} \tag{4.61}$$

5. Pour garantir la stricte inégalité dans (4.58), il est nécessaire que cette inégalité soit strictement satisfaite. Théoriquement, cela peut être assuré en multipliant les paramètres $\mathbf{T}$ et $\mathbf{\Sigma}$ par des constantes strictement positives μ et ν telles que $\mu\nu < 1$. Cependant, les auteurs constatent qu'en pratique, prendre $\mu = \nu = 1$ ne cause aucun problème de convergence.

6. Notamment en raison de la complexité induite par la redondance et du coût des projections successives.

Puisque les vecteurs colonnes $\mathbf{W}$ forment une famille orthonormale nous avons $\rho(\mathbf{W}^*\,\mathbf{W}) = 1$. Les conditions assurant une convergence du schéma sont donc

$$\begin{cases} \sigma > 0,\, \tau > 0 \\ \tau\sigma < 1 \\ \theta = 1. \end{cases} \tag{4.62}$$

Nous proposons alors de considérer les conditions suivantes [7]

$$\begin{cases} \sigma = 0.99 \\ \tau = 0.99 \\ \theta = 1. \end{cases} \tag{4.63}$$

Dans le cadre de cette reformulation du problème (4.49b) dans une base orthonormée, et de sa résolution via le schéma (4.61), il est pertinent, afin de comparer les résultats obtenus avec ceux issus des formulations précédentes (sans orthonormalisation), de définir selon (4.60a) à la $k^{\text{ème}}$ itération le vecteur de coefficients dans la base d'origine par

$$\alpha^{(k)} = \mathbf{Q}\,\beta^{(k)}, \tag{4.64}$$

l'image corrigée par

$$U^{(k)} = v - \mathbf{B}\,\mathbf{Q}\,\beta^{(k)} = v - \mathbf{B}\,\alpha^{(k)}. \tag{4.65}$$

La figure 4.4 met en évidence que la reformulation du problème (4.49b) dans une base orthonormée permet d'obtenir une convergence à la fois plus rapide et plus stable.

On peut cependant considérer le problème différemment. Le modèle TV originel (4.47) consiste à déterminer le biais. Plutôt que de le représenter par ses coefficients dans une base, orthonormale ou non, nous proposons une nouvelle formulation, dite « intrinsèque », qui permet d'estimer le biais sans recours à une base.

1.4.3 Modèle TV par une méthode intrinsèque

Le problème TV (4.47) s'écrit

$$\widehat{B} \in \underset{B\in\mathbb{R}^\Omega}{\operatorname{argmin}} \Big(||\nabla(B - v)||_{1,2} + \delta_{\operatorname{Vect}(B_1,\dots,B_N)}(B) \Big), \tag{4.66}$$

on peut appliquer l'algorithme de Chambolle-Pock. Le calcul détaillé en annexe 2.6 montre que la suite $B^{(k)}$ du schéma suivant converge vers une solution de (4.47)

$$\forall k \in \mathbb{N}, \quad \begin{cases} p^{(k+1)} = \underset{\mathbf{b}}{\operatorname{proj}}\Big(p^{(k)} - \sigma\nabla v + \sigma\nabla\overline{B}^{(k)} \Big), \\ B^{(k+1)} = \underset{\operatorname{Vect}(B_1,\dots,B_N)}{\operatorname{proj}} \Big(B^{(k)} - \tau\,\nabla^* p^{(k+1)} \Big), \\ \overline{B}^{(k+1)} = B^{(k+1)} + \theta\Big(B^{(k+1)} - B^{(k)} \Big), \end{cases} \tag{4.67}$$

7. Le schéma de Chambolle-Pock avec un paramètre μ (cf. conditions (4.56)) a donné des vitesses de convergence très similaires.

avec les conditions

$$\begin{cases} \sigma = \dfrac{0.99}{2\mu\sqrt{2}}, \\ \tau = \dfrac{0.99\mu}{2\sqrt{2}}, \\ \theta = 1, \\ \mu > 0 \quad \text{paramètre à ajuster.} \end{cases} \tag{4.68}$$

Afin de comparer cette méthode algorithmique à celles fondées sur la reformulation du problème selon les coefficients (4.49b), on définit, pour la k-ième itération, l'image corrigée par

$$U^{(k)} = v - B^{(k)}. \tag{4.69}$$

1.4.4 Étude comparative des méthodes algorithmiques de résolution du problème TV

Le principe est le suivant. Pour chacun des schémas algorithmiques vus précédemment, on considère, pour la $k^{\text{ème}}$ itération, l'image corrigée $U^{(k)}$ (définie par (4.54), (4.65) ou (4.69)) et l'énergie $E^{(k)}$ définie par (4.55).

À la suite d'un nombre « infini » d'itérations (en pratique quelques millions), on établit la valeur de l'énergie minimale $\overline{E}$, à laquelle on associe l'image corrigée $\overline{U}$. On définit ensuite les fonctions suivantes

$$f_1(k) = \frac{\left|E^{(k)} - \overline{E}\right|}{\overline{E}}, \tag{4.70}$$

qui mesure la variation relative de l'énergie, et

$$f_2(k) = \frac{\left\|U^{(k)} - \overline{U}\right\|_\infty}{\left\|\overline{U}\right\|_\infty}, \tag{4.71}$$

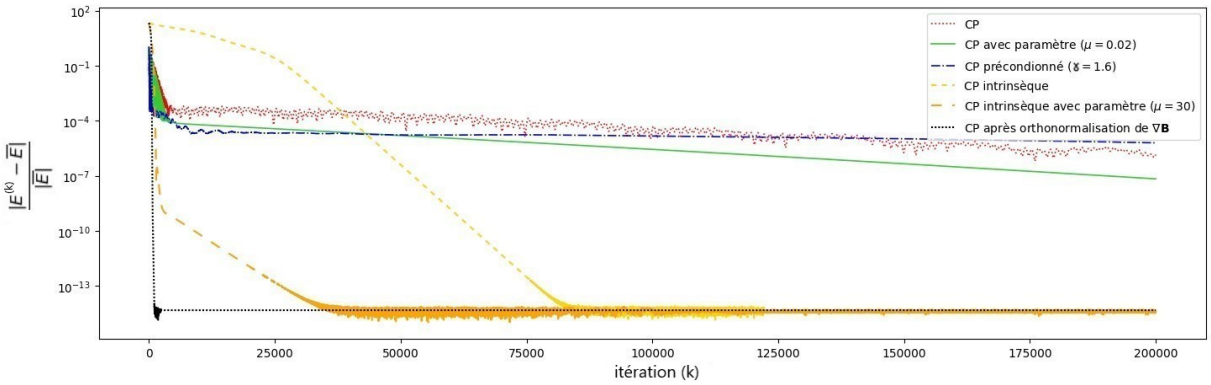
qui évalue la variation relative de l'image corrigée en norme infinie.

La figure 4.5 montre que seule la méthode intrinsèque, ainsi que la méthode de recherche des coefficients après orthonormalisation de la famille des gradients des biais, permettent d'obtenir, après un nombre raisonnable d'itérations, une convergence à la précision double. Lorsque le nombre d'éléments de la famille des biais n'est pas trop petit (en pratique, au moins 10), les méthodes fondées sur la recherche de coefficients nécessitent de travailler dans une base orthonormée, ce qui garantit une convergence plus rapide.

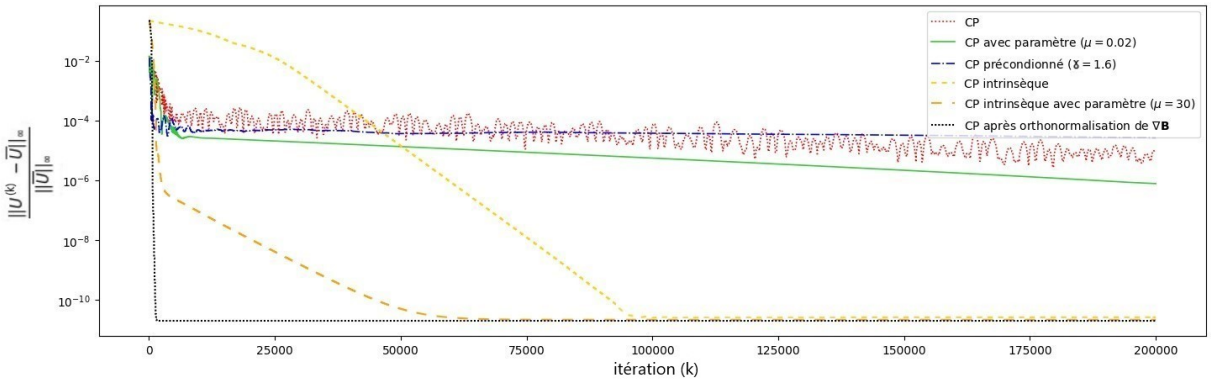
1.4.5 Correction d'images par le modèle TV

La figure 4.6 montre que les corrections obtenues, aussi bien sur la scène uniforme que sur la scène réelle, sont très similaires à celles fournies par le modèle $\mathscr{H}^{\mathbf{1}}$, et donc d'une qualité satisfaisante.

Toutefois, pour évaluer rigoureusement la pertinence des modèles que nous avons développés, il est indispensable de comparer leurs performances à celles des méthodes de correction issues de l'état de l'art.

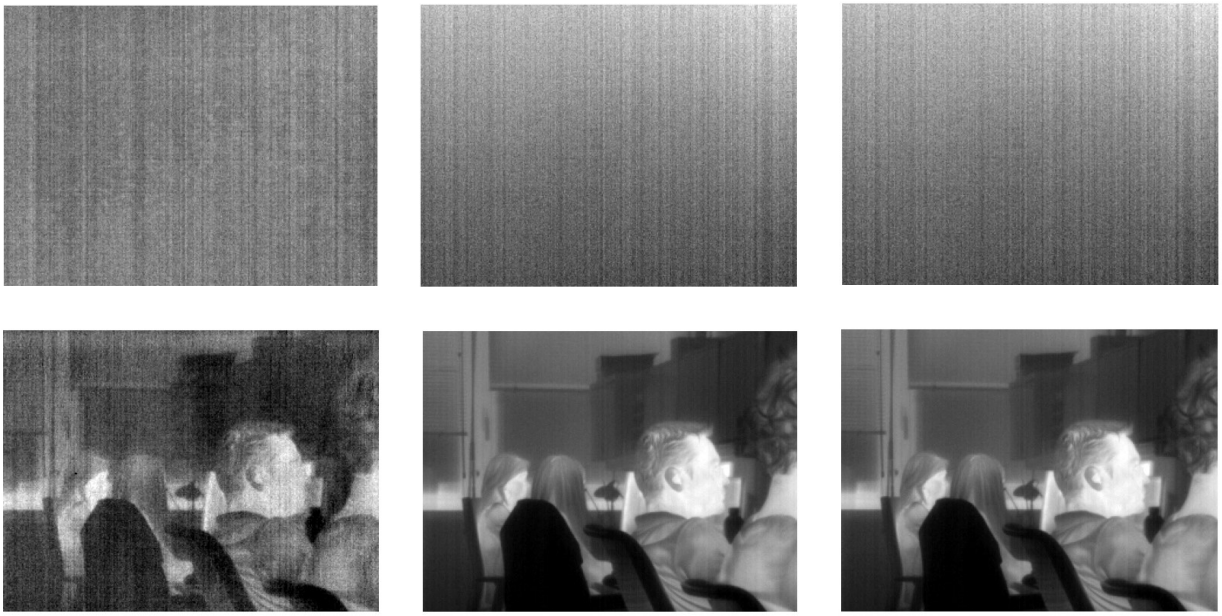


(a) Variation relative de l'énergie



(b) Variation relative de l'image corrigée en norme infinie

FIGURE 4.5 – **Courbes de convergence des différentes méthodes algorithmiques pour résoudre le problème** TV**.** Le cardinal de la famille de biais est de 24. L'énergie minimale $\overline{E}$ est définie comme la plus petite valeur observée parmi toutes les méthodes après un million d'itérations. L'image corrigée correspondante est notée $\overline{U}$. (a) L'énergie relative décroît globalement au fil des itérations k pour l'ensemble des variantes testées, mais une hiérarchie nette se dégage entre elles. La décroissance est très rapide pour la méthode de Chambolle–Pock appliquée après orthonormalisation de la famille des gradients des biais (courbe en pointillés noirs) : la convergence numérique (ici relative à une précision double associée à une valeur avoisinant 10^{-15}) est atteinte en moins d'un millier d'itérations. Elle reste rapide et régulière pour la méthode intrinsèque avec paramètre (courbe discontinue orange), la convergence numérique étant atteinte après environ quarante mille itérations. Elle devient plus lente pour la méthode intrinsèque sans paramètre (courbe discontinue jaune), où la convergence numérique en précision double est atteinte un peu en dessous de cent mille itérations. La décroissance est lente pour les méthodes Chambolle–Pock préconditionnée (courbe discontinue violette) et avec paramètre (courbe verte), mais dans les deux cas, la convergence numérique demeure hors d'atteinte après deux cent mille itérations. Enfin, pour la méthode Chambolle–Pock sans paramètre (donc avec des pas égaux, courbe en pointillés rouges), la décroissance est lente et instable, l'énergie oscillant encore après un nombre conséquent d'itérations, et la convergence numérique reste hors de portée. (b) On observe un comportement similaire pour la variation relative de l'image corrigée. En conclusion, pour résoudre le problème TV à l'aide de l'un des schémas numériques proposés ici dans un nombre raisonnable d'itérations (quelques dizaines de milliers au maximum), il semble nécessaire soit de rechercher les coefficients dans une base orthonormée de la famille des gradients des biais, soit de recourir à la méthode intrinsèque.



(a) Modèle ℓ^2 (b) Modèle $\mathscr{H}^1$ (c) Modèle TV

FIGURE 4.6 – **Correction d'une scène réelle et d'une scène uniforme par les modèles ℓ^2, $\mathscr{H}^1$ et TV.** La famille des biais, ainsi que les deux images pseudo-brutes à corriger, sont identiques à celles des figures 4.2 et 4.3. La première ligne illustre la correction d'une scène uniforme par les trois modèles, avec un affichage en niveaux de gris normalisé sur une étendue commune de 140 DL. Les images obtenues sont faiblement contrastées, d'apparence homogène, avec des effets de colonnes visibles. La correction de l'image produite par le modèle ℓ^2 (a) est plus texturée et moins lisse que celle obtenue pour les modèles $\mathscr{H}^1$ (b) et TV (c), bien que son écart-type soit deux fois inférieur. Les résultats issus des modèles $\mathscr{H}^1$ et TV sont visuellement indistinguables et présentent un effet basse fréquence. La seconde ligne montre la correction d'une scène réelle, avec un affichage en niveaux de gris normalisé sur une étendue commune de 700 DL. La solution obtenue avec le modèle ℓ^2 (a) présente des défauts : certains objets dont la température et l'émissivité sont supposées homogènes (comme les dossiers de sièges ou les placards) apparaissent de manière irrégulière. Les corrections fournies par les modèles $\mathscr{H}^1$ (b) et TV (c), en revanche, sont de très bonne qualité et identiques visuellement. La similarité visuelle entre les résultats des modèles $\mathscr{H}^1$ et TV peut sembler surprenante au premier abord, car la régularisation $\mathscr{H}^1$ produit généralement des images floues, tandis que la régularisation TV génère des images « en escalier », c'est-à-dire composées de zones constantes par morceaux. Cette ressemblance s'explique ici par le faible nombre de degrés de liberté (24) impliqués dans le modèle. En conclusion, les modèles $\mathscr{H}^1$ et TV permettent d'obtenir des corrections fiables et visuellement satisfaisantes, y compris dans le cas de scènes réelles, contrairement au modèle ℓ^2, qui n'est adapté que pour des scènes uniformes.

1.5 Comparaison des performances des modèles ℓ^2, $\mathscr{H}^1$ et TV avec le modèle *shutterless polynomial*

1.5.1 Limites des modèles présents dans la littérature

Comme nous l'avons vu dans le deuxième chapitre (section 2.2), la littérature distingue deux grandes classes de modèles de correction. Les modèles dits *« scene-based »* reposent

sur des hypothèses relativement fortes, telles que la présence d'un mouvement significatif de la scène et une faible variabilité temporelle de l'environnement thermique, ce qui limite leur champ d'application ou engendre des artefacts. Les modèles dits *« calibration-based »* nécessitent quant à eux des données acquises en usine, mais manquent de robustesse lorsque la caméra est utilisée dans un environnement différent de ceux envisagés lors de la calibration, en particulier si la TFPA de l'image à corriger diffère sensiblement de celles observées. Parmi ces approches, le modèle *shutterless polynomial* (section 2.2.2) constitue aujourd'hui une référence, bien qu'il puisse se révéler inadapté dans des conditions critiques, par exemple lorsque l'environnement thermique de la caméra diffère fortement de ceux rencontrés lors de la calibration, ce qui conduit à une inadéquation des coefficients et à une restitution dégradée du signal thermique (voir figure 3.17).

1.5.2 Données de calibration et de test utilisées pour l'étude comparative

Les modèles de correction développés dans ce chapitre s'appuient sur des données de calibration à partir desquelles est construite la famille des biais. En pratique, ces données consistent en des paires d'images de scènes uniformes (corps noir froid et chaud), acquises à différentes valeurs de TFPA. Une image de gain unique, notée $\mathcal{G}_0$, est d'abord estimée à partir d'une ou plusieurs paires [8]. La famille des biais est ensuite construite à partir de $\mathcal{G}_0$ et de l'ensemble des paires, selon la formule (3.59).

Le modèle *shutterless polynomial* associe à chaque paire une image de responsivité et une image de décalage correspondant à la TFPA de cette paire. Un ajustement polynomial est ensuite réalisé sur l'ensemble des paires disponibles (de degré 2 pour la responsivité et de degré 3 pour le décalage), afin de permettre l'interpolation de ces paramètres à toute valeur de TFPA. Les mêmes données de calibration sont utilisées pour tous les modèles afin de garantir une comparaison équitable.

L'étude repose sur des acquisitions réalisées avec la caméra *ATI-640c*, réparties en trois ensembles distincts : deux ensembles de calibration, destinés à l'apprentissage des modèles, et un ensemble de test, utilisé pour évaluer leur performance sur des scènes réelles.

Le premier ensemble de calibration comprend 20 paires d'images de corps noirs à 10°C et 70°C, acquises en « étuve stabilisée », c'est-à-dire dans un environnement thermiquement homogène et maintenu stable avant chaque prise de vue. Ces données couvrent une plage de TFPA répartie quasi uniformément dans l'intervalle $[-38°\text{C}, 74°\text{C}]$. Un sous-ensemble est sélectionné pour les expérimentations, soit pour restreindre l'amplitude de TFPA, soit pour réduire le nombre de points de calibration.

Le second ensemble contient 12 paires d'images similaires, mais acquises cette fois dans un environnement spatialement hétérogène (résistance chauffante placée derrière la caméra). Chaque scène est néanmoins stabilisée thermiquement avant acquisition. Ces données couvrent une plage de TFPA également répartie quasi uniformément, dans l'intervalle $[-23°\text{C}, 44°\text{C}]$.

Enfin, l'ensemble de test comprend 7 images de scènes réelles, acquises en « étuve stabilisée » (la caméra observant la scène au travers d'un hublot), pour des TFPA comprises

8. Nous avons choisi d'estimer ce gain à partir de l'ensemble des paires, conformément à la formule proposée au chapitre 3 (3.57), bien qu'il soit également possible de le calculer à partir d'une seule paire selon la méthode de calibration à deux points (2.22)

entre −39°C et 80°C. Les 7 scènes correspondantes présentent des éléments visuels qui peuvent varier d'une image à l'autre, bien qu'elles partagent le même point de vue fixe.

1.5.3 Étude comparative

Lors de la comparaison des modèles de correction basés sur une famille de biais, que nous avons proposés, avec le modèle *shutterless polynomial*, les expérimentations menées permettent d'établir les constatations suivantes :

— La qualité de correction obtenue avec les modèles $\mathscr{H}^{\mathbf{1}}$ et TV n'est jamais inférieure à celle du modèle de référence (cf. figures 4.7, 4.8, 4.9).

— Les modèles $\mathscr{H}^{\mathbf{1}}$ et TV font preuve d'une meilleure robustesse lorsqu'ils sont utilisés dans un environnement thermique différent de ceux rencontrés lors de la phase de calibration. Cette robustesse se manifeste non seulement en termes de température (comme la variation de la TFPA, cf. figure 4.7), mais aussi en fonction de la nature de l'environnement thermique, qu'il soit homogène ou non (cf. figure 4.8).

— Les modèles $\mathscr{H}^{\mathbf{1}}$ et TV présentent une bonne robustesse même lorsque le nombre de données de calibration est réduit (cf. figure 4.9). En effet, même avec un sous-ensemble plus restreint des données de calibration, ces modèles continuent de délivrer des résultats de correction de haute qualité, ce qui les rend particulièrement adaptés à des situations où les ressources de calibration sont limitées.

— Les modèles ℓ^2, $\mathscr{H}^{\mathbf{1}}$ et TV sont robustes lorsqu'une des données de calibration est corrompue (*outlier*). Cette propriété peut être expliquée par la nature du biais estimé, qui s'apparente au FPN. Par conséquent, il est peu probable qu'une donnée aberrante présente une corrélation spatiale significative avec le FPN ; cette dernière est alors naturellement ignorée par le modèle. Cette robustesse est maintenue tant que la proportion de données corrompues n'est pas excessive.

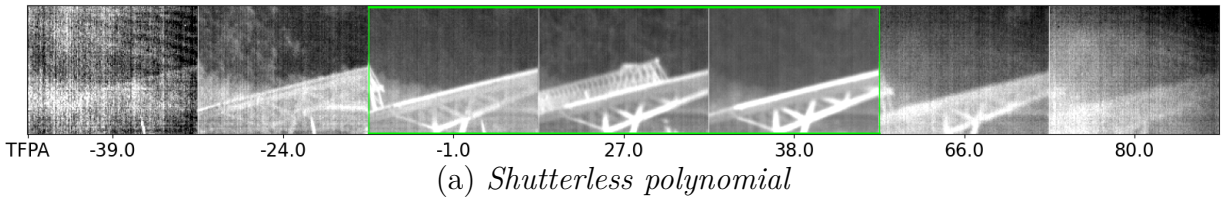
— Le modèle ℓ^2 donne des images corrigées de scènes réelles avec de nombreux artefacts de reconstruction.

— Les modèles $\mathscr{H}^{\mathbf{1}}$ et TV produisent des corrections équivalentes.

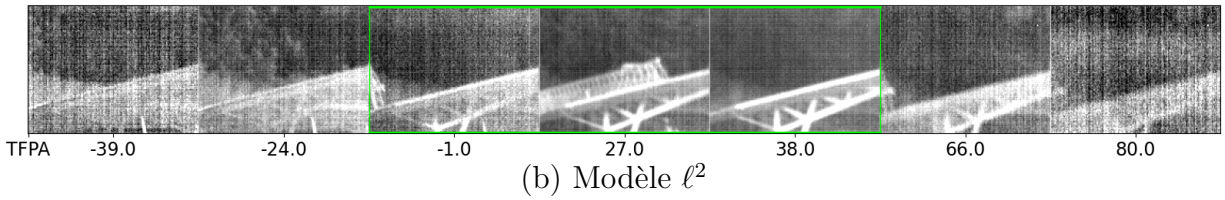
En conclusion, le modèle que nous avons proposé, dans ses déclinaisons $\mathscr{H}^{\mathbf{1}}$ et TV, offre des performances de correction supérieures ou égales à celles du modèle de référence. La déclinaison ℓ^2, par contre, délivre de nombreux artefacts dans ses corrections de scène réelle. Ces résultats ouvrent également la voie à des procédures de calibration moins contraignantes.

Les protocoles actuellement en usage, conçus pour l'estimation du modèle *shutterless polynomial*, visent à modéliser la responsivité et le décalage comme des fonctions paramétriques de la TFPA, tout en s'efforçant de neutraliser l'influence des autres composantes de l'environnement thermique.

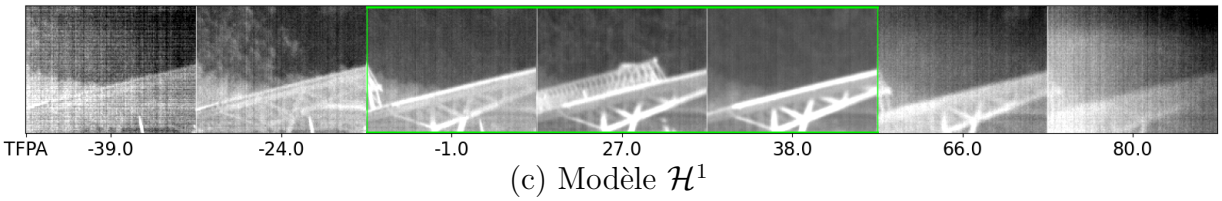
À l'inverse, notre approche suggère qu'une calibration effectuée dans des environnements thermiques variés, incluant notamment des configurations inhomogènes et instables dans le temps, permettrait d'améliorer la qualité de la correction, en particulier lorsque la caméra est exposée à des conditions similaires. La diversité des environnements explorés lors de la calibration apparaît ainsi comme un levier prometteur. De plus, ces procédures



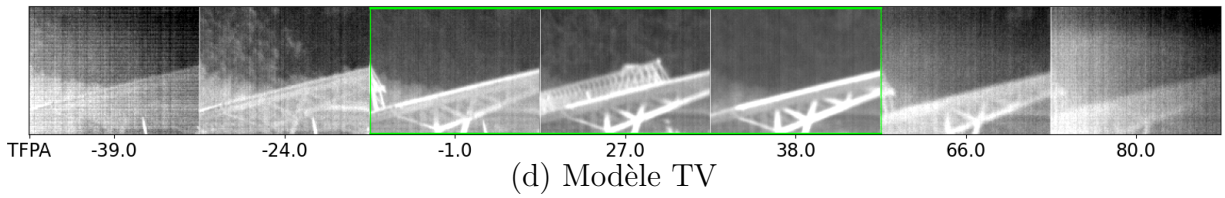
(a) *Shutterless polynomial*



(b) Modèle ℓ^2



(c) Modèle $\mathcal{H}^1$



(d) Modèle TV

FIGURE 4.7 – **Correction d'images délivrées par une caméra placée dans une étuve à sept TFPA différentes, selon quatre méthodes, avec des données de calibration restreintes en TFPA.** Les données utilisées proviennent de la caméra *ATI-640c*. La calibration repose sur 11 paires d'images de corps noir chaud (70 °C) et froid (10 °C), acquises dans une « étuve stabilisée » à 11 TFPA distinctes, réparties quasi-uniformément sur l'intervalle $[-18, 43\ °C]$, qui constitue la plage de calibration (zones encadrées en vert). Les données de test se composent de 7 images brutes de scènes réelles, également acquises dans l'« étuve stabilisée », à travers un hublot. La valeur de TFPA est indiquée en abscisse pour chaque image de test. L'affichage montre un zoom centré sur la moyenne de chaque image reconstruite, couvrant un intervalle égal à quatre fois son écart-type. Pour les images délivrées à une TFPA située à l'intérieur de la plage de calibration (TFPA de −1, 27 et 38 °C), notre modèle basé sur les métriques $\mathscr{H}^1$ et TV (sous-figures c et d) fournit une correction de bonne qualité, comparable à celle de la méthode *shutterless polynomial* (sous-figure a). En revanche, la métrique ℓ^2 (sous-figure b) introduit des artefacts de lignes et de colonnes, dégradant la qualité de la texture. Lorsque la TFPA de l'image à corriger est hors de la plage de calibration (notamment pour −39 et 80 °C), les modèles $\mathscr{H}^1$ (c) et TV (d) produisent une correction dégradée par des effets de lignes, de colonnes et un voile basse fréquence. Toutefois, ces modèles permettent une meilleure distinction des contours et des textures, comparativement à la méthode *shutterless polynomial* (a). Pour −39 °C, le modèle TV conserve mieux la texture que $\mathscr{H}^1$. La correction ℓ^2 (b) permet également de distinguer les contours, mais la restitution des textures (par exemple, celle du ciel) est fortement altérée. En conclusion, pour les cas en interpolation de TFPA (zones encadrées en vert), les modèles $\mathscr{H}^1$ et TV offrent une correction comparable à la méthode *shutterless polynomial*. En revanche, pour les cas en extrapolation, ces mêmes modèles fournissent une correction de meilleure qualité.



(a) *Shutterless polynomial*



(b) Modèle ℓ^2

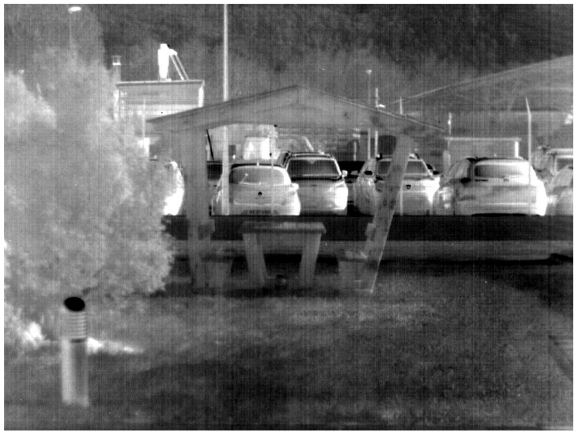


(c) Modèle $\mathcal{H}^1$

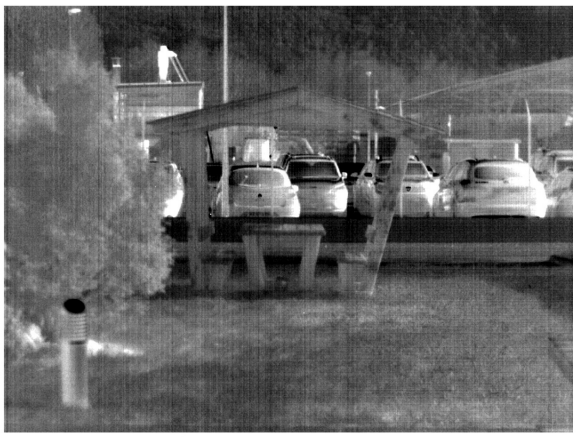


(d) Modèle TV

FIGURE 4.8 – **Correction d'une image acquise en milieu homogène à partir d'une calibration réalisée en milieu inhomogène.** Les données proviennent de la caméra *ATI-640c*. Les données de calibration sont constituées de 12 paires de corps noir chaud (70°C) et froid (10°C), acquises dans un environnement spatialement inhomogène mais stabilisé temporellement. Ces paires couvrent quasi uniformément l'intervalle de TFPA $[-23°\text{C}, 44°\text{C}]$. L'image brute de test a été acquise à une TFPA de $-1°$C, dans une « étuve stabilisée », soit un environnement spatialement homogène, distinct de par sa nature de celui des données de calibration. La scène observée comporte notamment, en arrière-plan en haut, une forêt à l'ombre, donc supposée présenter une irradiance à peu près uniforme. (a) La correction obtenue avec le modèle de référence *shutterless polynomial* est globalement satisfaisante, mais présente des résidus de basse fréquence perceptibles notamment dans l'arrière-plan, qui apparaît légèrement teinté au lieu d'être noir. (b) Le modèle ℓ^2 souffre d'effets de colonne et de fluctuations de basse fréquence plus prononcés, rendant l'arrière-plan gris. (c) et (d) Les modèles $\mathscr{H}^{\mathbf{1}}$ et TV produisent une correction visuellement indiscernable et de très bonne qualité. L'arrière-plan est parfaitement noir, ce qui est cohérent avec sa faible luminance émise dans la scène. Ces résultats suggèrent que les modèles $\mathscr{H}^{\mathbf{1}}$ et TV sont particulièrement robustes au changement de nature de l'environnement thermique entre calibration et test.



(a) *Shutterless polynomial*



(b) Modèle ℓ^2



(c) Modèle $\mathcal{H}^1$



(d) Modèle TV

FIGURE 4.9 – **Correction d'une image à partir d'un nombre de données de calibration restreint.** Les données proviennent de la caméra *ATI-640c*. Les données de calibration sont composées de 4 paires de corps noir chaud (70°C) et froid (10°C), acquises en « étuve stabilisée » pour des TFPA de −38, −7, 26 et 63°C. L'image de test, acquise en « étuve stabilisée » à une TFPA de −24°C, se situe entre les points de calibration de −38 et −7°C. L'affichage des images corrigées est centré sur la moyenne et couvre cinq fois l'écart-type relatif à l'image. (a) L'image corrigée par la méthode *shutterless polynomial* permet de distinguer nettement les contours et la texture de la scène. Un effet latéral de basse fréquence reste visible car l'arbre à gauche apparaît plus clair que le gazon en bas alors que leur luminance devrait être comparable et l'arrière-plan en haut n'est pas assez sombre alors qu'il correspond à une zone de luminance très faible. (b) L'image corrigée par le modèle ℓ^2 est lisible mais présente des artefacts sous forme d'effets de ligne et de colonne, ce qui dégrade la qualité de la correction. (c) et (d) Les images corrigées par les modèles $\mathcal{H}^1$ et TV sont visuellement très proches. La restitution des détails fins et des contours n'est pas affectée par des effets de basse fréquence : l'arbre et le gazon ont des nuances de gris comparables, et l'arrière plan est assez noir. Contrairement à la figure 4.8, des effets de colonne apparaissent pour les corrections $\mathcal{H}^1$ et TV, dus à une dérive temporelle des mesures électroniques, l'écart temporel entre calibration et test étant de 15 jours (contre 5 jours pour la figure 4.8). Dans cette configuration d'interpolation, nos modèles basés sur les métriques $\mathcal{H}^1$ et TV offrent une correction visuellement supérieure à la méthode *shutterless polynomial*, notamment pour la restitution des luminances émises par les objets.

pourraient être plus rapides et moins coûteuses à mettre en œuvre que les protocoles traditionnels, qui visent à stabiliser température et humidité pour minimiser l'impact d'une variabilité des flux parasites avant acquisition.

Nous avons observé, sur certaines images, malgré une correction satisfaisante, des effets résiduels de colonnes plus ou moins prononcés (par exemple, sur les figures 4.2a, 4.3 et 4.9c-d). Le débruitage colonnaire, un dispositif de post-traitement, est essentiel en imagerie thermique non-refroidie. Avant d'aborder cette étape en post-traitement d'une image corrigée, il semble pertinent d'envisager l'intégration du débruitage colonnaire directement dans nos modèles, en décomposant chaque biais de la famille $\mathscr{B}$ en éléments colonnaires.

1.6 Décomposition colonnaire des biais et intégration d'un débruitage colonnaire dans le modèle de correction

1.6.1 Causes des artefacts de colonnes

Les effets de colonnes résiduels sont inhérents aux modèles de correction « *calibration-based* ». Ces artefacts dépendent du contraste de la scène et semblent fortement influencés par un facteur temporel, notamment l'intervalle entre la prise de vue et la calibration. Lorsque cet intervalle est court, comme montré à la figure 4.8d, les artefacts restent peu visibles, tandis qu'un intervalle prolongé accentue les effets de colonne, comme le montre la comparaison entre les figures 4.10a et 4.10b.

Selon des retours d'expérience provenant de l'industrie des imageurs thermiques, l'une des causes principales de ce phénomène serait la dérive temporelle des mesures du circuit de lecture électronique. Cette dérive proviendrait notamment de la manière dont le signal d'un pixel est généré à partir d'un pixel aveugle de la même colonne (cf. figure 2.9).

Avant de mettre en place un débruitage colonnaire en post-traitement, il est pertinent de se demander si ces artefacts pourraient être corrigés conjointement avec la reconstruction de l'image pseudo-brute, au moment de la correction via les modèles de correction à partir d'une famille de biais que nous avons considérés précédemment.

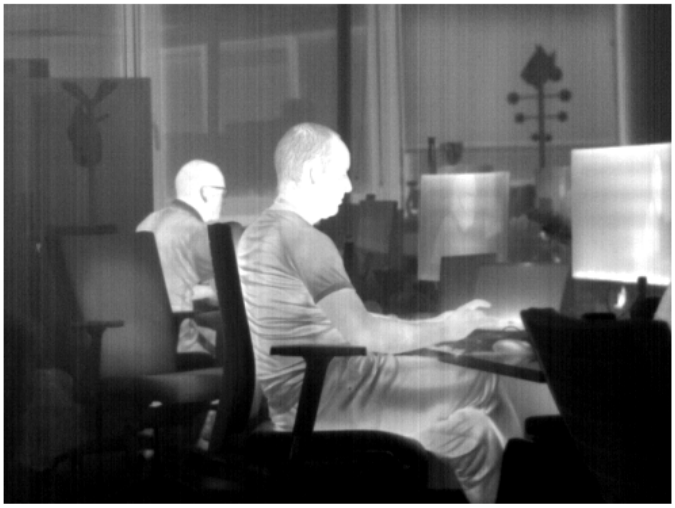
1.6.2 Proposition d'un débruitage colonnaire intégré à nos modèles de correction

L'observation d'un biais (cf. figure 4.1) met en évidence des effets de colonnes, qui pourraient être associés aux mêmes phénomènes responsables des artefacts colonnaires observés après correction. Pour remédier à ces artefacts, nous proposons de décomposer chaque image de biais de la famille $\mathscr{B}$ en un sous-ensemble d'éléments colonnaires. Formulons cette décomposition de la manière suivante : soit un biais $B_n \in \mathscr{B}$, il peut être transformé en un sous-ensemble de la forme

$$\mathcal{B}_n = \{B_n^{(1)}, \ldots, B_n^{(N_c)}\},$$

avec

$$\forall i \in [\![1, N_c]\!], \quad \forall (x,y) \in [\![1, N_c]\!] \times [\![1, N_l]\!], \quad B_n^{(i)}(x,y) = \begin{cases} B_n(x,y) & \text{si } x = i, \\ 0 & \text{sinon.} \end{cases}$$



(a) Image corrigée par le modèle TV à partir d'une calibration non actualisée



(b) Image « vérité terrain »

FIGURE 4.10 – **Image corrigée avec effet de colonne résiduel, image corrigée sans effet colonnaire.** Données issues de la caméra *ATI-640b*. Les deux images sont affichées selon une même dynamique de niveaux de gris. (a) Correction obtenue avec le modèle TV, à partir d'une calibration réalisée en « étuve stabilisée », comprenant 13 paires d'images de corps noir à 10°C et 70°C, couvrant quasi uniformément les TFPA de −38°C à 77°C. L'image test, acquise à la TFPA 25°C, a été produite quatre mois après la phase de calibration. La correction est globalement satisfaisante (textures, structures), mais un effet colonnaire résiduel subsiste, perceptible notamment dans les zones uniformes (stores, placards). (b) Image corrigée par soustraction directe d'un biais estimé à partir d'une scène uniforme acquise quelques secondes auparavant, dans un environnement quasi identique. Cette correction représente la « vérité terrain ». On constate que les effets colonnaires ont complètement disparu. Les effets de colonne observés sur (a) dépendent du contraste de la scène et se manifestent dans tous les modèles « calibration-based », en particulier lorsque la calibration n'est pas actualisée, introduisant ainsi un facteur temporel entre la prise de vue et la calibration.

où N_c et N_l désignent respectivement le nombre de colonnes et de lignes d'une image.

On peut ensuite considérer la famille décomposée $\mathscr{B}'$ comme la réunion de ces sous-ensembles, c'est-à-dire

$$\mathscr{B}' = \bigcup_{1 \leq n \leq N} \mathcal{B}_n.$$

Ainsi, $\mathscr{B}'$ contient $N \times N_c = M$ images, linéairement indépendantes, tout comme leurs gradients spatiaux, en raison du bruit de mesure. Les modèles de correction sont alors appliqués dans leurs variantes ℓ^2, $\mathscr{H}^{\mathbf{1}}$ et TV, en tenant compte de la famille décomposée $\mathscr{B}'$.

Dans le cas ℓ^2, la résolution est plus délicate car les éléments de $\mathscr{B}'$ ne sont pas centrés. On reprend donc le modèle initial, qui consiste à minimiser la variance spatiale de l'image corrigée.

Soit une image $u \in \mathbb{R}^\Omega$ (vue comme un vecteur colonne). Sa variance est donnée par

$$\sigma^2(u) = \mu(u^2) - \Big(\mu(u)\Big)^2.$$

Puisque l'on peut écrire

$$\begin{cases} \mu(u^2) = \dfrac{\|u\|_2^2}{\operatorname{card}(\Omega)} \\ \mu(u) = \dfrac{{}^t\mathbf{1}_\Omega\, u}{\operatorname{card}(\Omega)} \end{cases}$$

où $\mathbf{1}_\Omega$ désigne un vecteur colonne de taille $\operatorname{card}(\Omega)$ dont tous les éléments valent un, on en déduit

$$\sigma^2(u) \propto \|u\|_2^2 - \frac{\Big({}^t\mathbf{1}_\Omega\, u\Big)^2}{\operatorname{card}(\Omega)}.$$

On cherche alors à résoudre le problème suivant

$$\begin{cases} \widehat{\alpha} \in \underset{\alpha \in \mathbb{R}^M}{\operatorname{argmin}} \|v - \mathbf{B}'\alpha\|_2^2 - \dfrac{\Big({}^t\mathbf{1}_\Omega(v - \mathbf{B}'\alpha)\Big)^2}{\operatorname{card}(\Omega)} & \text{(4.72a)} \\ \widehat{B} = \mathbf{B}'\widehat{\alpha}, & \text{(4.72b)} \end{cases}$$

où $\mathbf{B}'$ est une matrice dont les vecteurs colonnes sont les biais de la famille décomposée $\mathscr{B}'$. L'existence et l'unicité de la solution du problème (4.72a) ne sont pas garanties, mais celui-ci est résolu de manière heuristique en recherchant un point critique.

Si l'on pose

$$J(\alpha) = \|v - \mathbf{B}'\,\alpha\|_2^2 - \frac{\Big({}^t\mathbf{1}_\Omega\,(v - \mathbf{B}'\,\alpha)\Big)^2}{\operatorname{card}(\Omega)},$$

alors son gradient s'écrit

$$\nabla J(\alpha) = -2\,{}^t\mathbf{B}'\,(v - \mathbf{B}'\,\alpha) + 2\,\overline{\mathbf{B}'}\;{}^t\mathbf{1}_\Omega\,(v - \mathbf{B}'\,\alpha),$$

où

$$\overline{\mathbf{B}'} = \begin{bmatrix} \mu(B'_1) \\ \vdots \\ \mu(B'_M) \end{bmatrix}.$$

Soit $\widehat{\alpha}$ un point critique de $J(\cdot)$, il vérifie l'équation

$$\nabla J(\widehat{\alpha}) = 0.$$

En supposant que la matrice $\left({}^t\mathbf{B}' - \overline{\mathbf{B}'}\ {}^t\mathbf{1}_{\mathbf{\Omega}}\right)\mathbf{B}'$ est inversible, on obtient explicitement

$$\widehat{\alpha} = \left(\left({}^t\mathbf{B}' - \overline{\mathbf{B}'}\ {}^t\mathbf{1}_{\mathbf{\Omega}}\right)\mathbf{B}'\right)^{-1}\left({}^t\mathbf{B}' - \overline{\mathbf{B}'}\ {}^t\mathbf{1}_{\mathbf{\Omega}}\right)v. \tag{4.73}$$

En pratique, une décomposition en valeurs singulières (SVD) est utilisée pour pallier le mauvais conditionnement de la matrice à inverser et limiter les erreurs numériques.

L'image corrigée U, correspondant à la formulation ℓ^2, s'obtient alors en combinant les expressions (4.73), (4.72b) et (4.8).

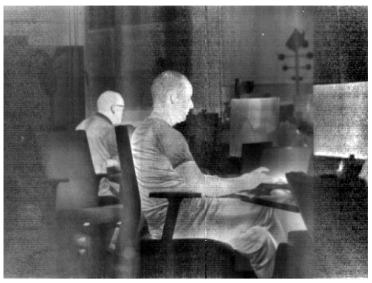
1.6.3 Résultats de correction à partir d'une famille décomposée

La figure 4.11 montre les résultats obtenus avec nos modèles de correction appliqués à une famille de biais décomposée par décomposition colonnaire. Bien que cette approche élimine les effets de colonnes visibles issus des corrections avec la famille initiale (cf. figure 4.10a), elle introduit de nouveaux artefacts latéraux de moyenne et basse fréquence, dégradant ainsi la qualité de l'image corrigée. Le modèle TV se distingue des deux autres par des artefacts latéraux moins importants, faisant de lui la méthode la plus performante dans ce contexte.

Ces résultats suggèrent que les artefacts de colonne observés après correction sont indépendants des structures colonnaires présentes dans les biais de calibration. Par conséquent, l'intégration du débruitage colonnaire à nos modèles de correction par une famille décomposée de la sorte semble inadaptée. Une stratégie plus efficace consisterait à traiter ce débruitage en post-traitement, après la correction initiale par un modèle *« calibration-based »*, afin d'isoler les sources d'erreur et de mieux cibler les artefacts sans affecter la structure de l'image.

1.7 Conclusion sur les modèles de correction reposant sur une famille de biais

Le modèle de correction que nous avons proposé se rapproche de celui de Liu et al. [20] (voir chapitre 2, sous-section 2.2.4.2). Dans ce dernier, le décalage est déterminé comme combinaison affine des images brutes de scènes uniformes. Après multiplication pixel par pixel par l'image de gain constant, cette opération revient, à une image constante près, à déterminer un biais dans l'espace vectoriel engendré par la réunion du gain et de la famille des biais. Les coefficients sont estimés en minimisant la norme euclidienne entre l'image corrigée V et sa version lissée $\widetilde{V}$, obtenue par moyenne locale sur une fenêtre carrée. Pour interprétation, si cette fenêtre est réduite aux 8 voisins immédiats d'un pixel, la différence $\widetilde{V} - V$ peut être vue comme une discrétisation du laplacien. Le critère des auteurs revient ainsi à minimiser la norme euclidienne du laplacien discret fondé sur la 8-connexité de l'image corrigée, favorisant une image spatialement lisse. La méthode proposée par Liu et al. présente cependant trois limites. Premièrement, la minimisation est restreinte à une zone uniforme « marquée », ce qui constitue une contrainte pratique. Deuxièmement, si la zone contient un voile affine (intensité variant linéairement selon une



(a) Famille décomposée (ℓ^2)



(b) Famille décomposée ($\mathscr{H}^{\mathbf{1}}$)



(c) Famille décomposée (TV)



(d) Famille d'origine (ℓ^2)



(e) Famille d'origine ($\mathscr{H}^{\mathbf{1}}$)



(f) Famille d'origine (TV)

FIGURE 4.11 – **Correction d'une image à partir d'une famille de biais décomposée, comparaison avec la correction à partir de la famille d'origine.** Caméra, données de calibration, image de test et affichage en niveaux de gris identiques à ceux de la figure 4.10a. (a) L'image corrigée par le modèle ℓ^2 à partir d'une famille décomposée présente une texture dégradée et des artefacts latéraux de moyenne et basse fréquence. (b) Le modèle $\mathscr{H}^{\mathbf{1}}$ améliore la qualité de la texture, mais des artefacts latéraux persistent, notamment une bande blanche verticale au premier tiers de l'image. (c) Le modèle TV est affecté mais les artefacts sont visuellement plus discrets, fournissant une correction globalement plus satisfaisante que $\mathscr{H}^{\mathbf{1}}$. (d), (e) et (f) correspondent aux corrections à partir de la famille d'origine non décomposée, pour ℓ^2, $\mathscr{H}^{\mathbf{1}}$ et TV, respectivement. Ces corrections présentent des artefacts bien moins marqués. Pour $\mathscr{H}^{\mathbf{1}}$ et TV, des effets de colonnes apparaissent, mais les artefacts latéraux de moyenne et basse fréquence observés avec la famille décomposée sont absents. En conclusion, l'intégration d'un débruitage colonnaire en enrichissant la famille de biais par éléments colonnaires ne donne pas les résultats attendus : les effets de colonne haute fréquence visibles avec la famille non décomposée disparaissent, mais de nouveaux artefacts latéraux de moyenne et basse fréquence apparaissent, compromettant la qualité globale de la correction.

direction), le laplacien nul ne détecte pas cette variation et le modèle ne la conserve pas, ce qui peut conduire à une correction incorrecte. Troisièmement, la résolution numérique ne considère que quatre images de calibration dont la TFPA est la plus proche de celle de l'image à corriger, éludant ainsi les problèmes posés par l'utilisation d'une famille plus large d'images fortement corrélées. Dans notre approche, la recherche des coefficients se fait sur l'ensemble de l'image sans nécessité de zone « marquée ». Nous n'utilisons pas la norme euclidienne du laplacien comme métrique mais une mesure telle que TV, moins pénalisante sur les contours, et les corrections peuvent différer d'une image constante près plutôt que d'un voile affine. Enfin, nous proposons une résolution adaptée qui traite correctement la forte corrélation entre les éléments de la famille et permet d'utiliser un nombre arbitraire d'images de calibration.

Nos modèles fondés sur les métriques $\mathscr{H}^{\mathbf{1}}$ et TV permettent, dans de nombreux cas expérimentaux, une restitution de bonne qualité des luminances, des textures et des contours des scènes observées. Lorsque l'environnement de la caméra est « proche » de ceux rencontrés lors de la calibration (interpolation), la correction obtenue est d'une qualité comparable ou supérieure à celle du modèle *shutterless polynomial*. Lorsque l'environnement diffère significativement de ceux rencontrés lors de la phase de calibration, la correction reste correcte mais moins optimale, tout en surpassant nettement celle du modèle de référence *shutterless polynomial*. Cette robustesse face aux variations de l'environnement thermique entre calibration et utilisation de la caméra permet à ces modèles de maintenir une performance supérieure, en particulier lorsque les conditions de test s'écartent fortement de celles de calibration. Par contre, la déclinaison ℓ^2, plus sensible aux artefacts, ne permet pas une correction satisfaisante dans les configurations testées. Elle engendre des artefacts qui altèrent les détails fins et la texture.

Même lorsque l'environnement thermique de l'image test est proche de ceux rencontrés lors de la phase de calibration, des motifs colonnaires peuvent apparaître, en particulier si l'écart temporel entre la calibration et l'acquisition est important. Ces artefacts, présents dans tous les modèles reposant sur des données de calibration, peuvent dégrader la qualité visuelle. La tentative d'intégrer le débruitage colonnaire directement dans le modèle de correction, via une famille de biais décomposée, ne s'est pas révélée concluante.

Cela justifie l'introduction, dans la section suivante, d'un débruitage colonnaire en post-traitement visant à éliminer ces effets de colonne sur une image corrigée par un modèle pour lequel ces artefacts sont les seuls présents.

2 Débruitage colonnaire en imagerie thermique

La correction par des méthodes dites « *calibration-based* » ou « mixtes » peut laisser apparaître un bruit colonnaire résiduel, appelé « *stripes* » (cf. figure 4.10a). En revanche, les méthodes « *scene-based* », qui exploitent directement les informations issues de la scène observée, ne génèrent pas ce type d'artefact.

Ce bruit colonnaire est d'amplitude généralement inférieure à celle du bruit affectant l'image brute avant correction (*fixed pattern noise*, FPN) mais peut néanmoins dégrader la qualité radiométrique des images corrigées. Il résulte principalement d'une dérive temporelle des caractéristiques électroniques entre la calibration et l'acquisition. Le débruitage colonnaire est alors envisagé comme un post-traitement appliqué après la correction FPN.

Le bruit colonnaire, qui se manifeste sous forme de bandes verticales appelées *stripes*, est présent dans plusieurs domaines d'imagerie, notamment thermique, satellitaire, hyperspectrale et médicale. En imagerie thermique, la figure 4.12a illustre ce motif dont l'amplitude peut atteindre plusieurs degrés Celsius, compromettant l'interprétation radiométrique. L'effet colonnaire observé pour une colonne x apparaît indépendant de l'indice colonnaire, ce qui suggère que la perturbation colonnaire peut être considérée comme indépendante de x.

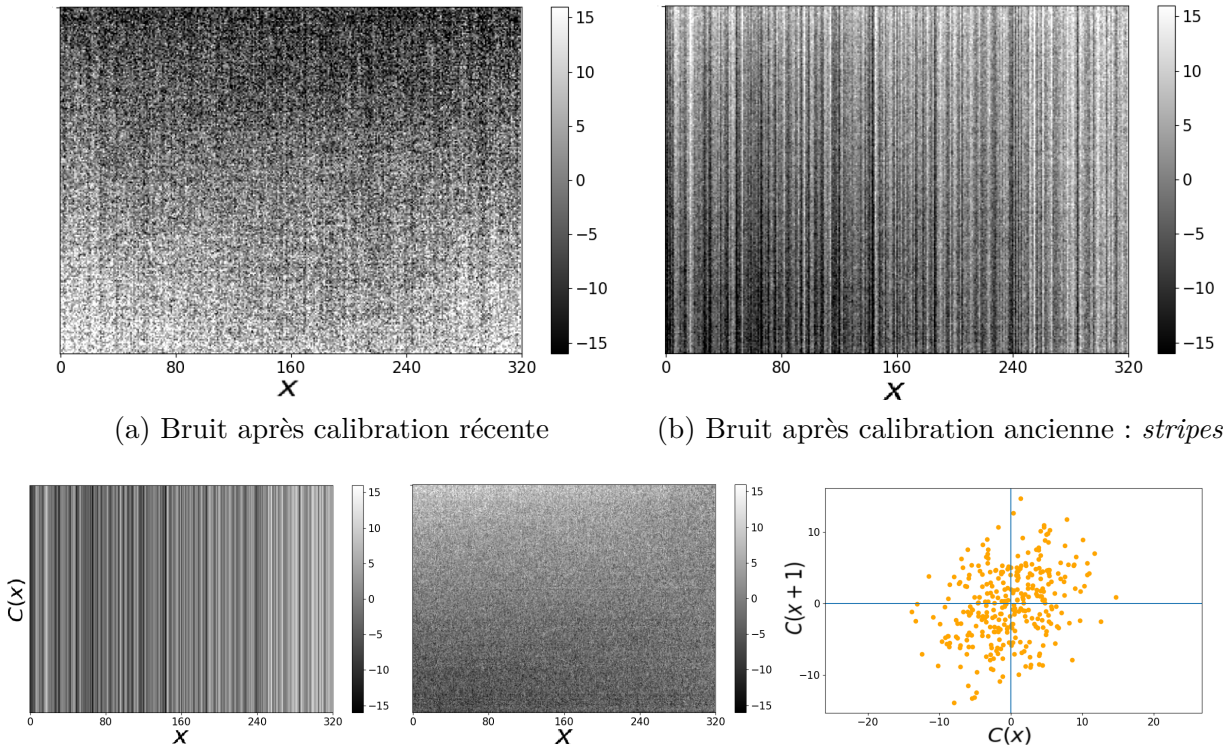
2.1 État de l'art des modèles de débruitage colonnaire

Les méthodes existantes pour supprimer le bruit colonnaire dans les images thermiques se répartissent généralement en quatre grandes familles. La première regroupe les approches statistiques, qui exploitent les propriétés radiométriques des colonnes pour corriger les écarts, par exemple via un ajustement des moments statistiques (*moment matching*). La seconde catégorie inclut les méthodes de filtrage spatial ou fréquentiel, qui traitent localement l'image ou agissent dans le domaine des fréquences pour atténuer les artefacts colonnaires. Une troisième famille rassemble les méthodes variationnelles, qui formulent la suppression du bruit comme un problème d'optimisation régularisé, permettant de réduire les oscillations verticales tout en préservant les détails de la scène. Enfin, des approches récentes exploitent l'apprentissage par réseaux de neurones, apprenant directement à partir d'exemples de paires images bruitées/propres la structure du bruit colonnaire et celle des détails de l'image, ce qui permet de traiter efficacement des scènes complexes ou peu texturées.

Il convient de noter que certaines approches combinent ces catégories, notamment des méthodes variationnelles intégrées à des schémas de filtrage spatial, tirant ainsi parti à la fois des avantages du filtrage et de la modélisation variationnelle.

2.1.1 Méthodes de débruitage colonnaire fondées sur les statistiques

On distingue deux sous-familles parmi les approches fondées sur des propriétés statistiques, les méthodes « *moment matching* » qui ajustent les moments statistiques (moyenne et écart-type) de chaque colonne, et les méthodes de correction par harmonisation des histogrammes colonnaires qui ajustent les distributions radiométriques entre colonnes voisines afin d'atténuer le bruit colonnaire.



(a) Bruit après calibration récente

(b) Bruit après calibration ancienne : *stripes*

(c) Médianes colonnaires *stripes*

(d) Résidu *stripes*

(e) Corrélation colonnes adjacentes

FIGURE 4.12 – **Analyse d'un bruit structuré de type *stripes* dominé par une structure colonnaire et une faible corrélation spatiale.** Données issues de la caméra *Smart IR320*. (a) Image corrigée d'une scène uniforme à l'aide d'un modèle « calibration-based » L'image courante a été acquise cinq jours après la phase de calibration. Le bruit observé est globalement faible, composé d'une composante isotrope et d'un bruit colonnaire très atténué mais encore perceptible. (b) Image d'un bruit structuré de type *stripes* obtenue avec exactement les mêmes données de calibration que dans (a), appliquées à une scène uniforme acquise cette fois six mois après la phase de calibration. On observe une forte structuration colonnaire, nettement plus marquée que dans (a), due à une dérive temporelle des mesures issues du circuit électronique de lecture de la caméra. (c) À partir du motif de *stripes* présenté en (b), on calcule pour chaque colonne $x \in \{1, \ldots, 320\}$ la médiane des intensités le long de l'axe vertical. Cela permet de reconstruire une image synthétique, purement colonnaire, notée $C(x)$, qui correspond à la composante colonnaire dominante du bruit structuré. Cette image présente une amplitude d'environ 30 DL (soit 1,2 °C), avec des variations irrégulières selon x. (d) L'image résiduelle, obtenue après soustraction de la composante colonnaire $C(x)$ (correspondant à l'image (c)) à chaque colonne de (b), révèle un bruit plus isotrope, avec une légère structure verticale de basse fréquence. Ce bruit présente une intensité inférieure à celle de la composante colonnaire extraite en (c). (e) Nuage de points représentant la corrélation entre colonnes adjacentes. Pour chaque $x \leq N_c - 1$, on trace le point $(C(x), C(x+1))$. Le nuage apparaît distribué de manière quasi uniforme dans un disque de rayon 15 et le coefficient de corrélation de Pearson vaut $4 \cdot 10^{-5}$, suggérant une quasi-indépendance de la composante colonnaire par rapport à l'indice de colonne x. Le bruit structuré de type *stripes* peut ainsi être assimilé à une perturbation purement colonnaire, indépendante le long des colonnes.

2.1.1.1 Méthodes « *moment matching* »

Les méthodes « moment matching » reposent sur l'hypothèse que le bruit colonnaire

se traduit par un décalage et une mise à l'échelle des niveaux radiométriques de chaque colonne [30, 31, 67]. Chaque colonne possède ainsi sa propre moyenne et son écart-type, différents de ceux des autres colonnes. Ces méthodes cherchent à corriger ces variations de deux manières principales. La première approche, dite globale, ajuste la moyenne et l'écart type de chaque colonne pour les faire correspondre à des valeurs de référence calculées sur l'ensemble de l'image [30]. La seconde approche, dite locale, lisse spatialement ces moments entre colonnes voisines afin de conserver des transitions plus naturelles [67].

Cette stratégie permet de réduire efficacement les bandes verticales indésirables. Néanmoins, elle présente des limites dans les zones à fort contraste local. Dans ces régions, qu'il s'agisse d'un ajustement global ou d'un lissage local, la correction peut générer des artefacts sous forme de bandes sombres autour des objets. Ces défauts sont dus à une surcorrection provoquée par des statistiques qui ne reflètent pas la variabilité locale de la scène et ne sont pas corrigés efficacement par les méthodes globales ou semi-locales. Cela souligne le besoin d'approches plus fines ou adaptatives pour gérer correctement les variations locales tout en supprimant le bruit colonnaire.

2.1.1.2 Méthodes de correction par harmonisation des histogrammes colonnaires

Une autre sous-famille de méthodes statistiques pour la correction du bruit colonnaire, développée et proposée par Tendero et al., consiste à ajuster les histogrammes des colonnes afin de les rendre plus proches les uns des autres. Cette approche, connue sous le nom de « MIRE » (*midway infrared equalization*), réalise ce rapprochement en lissant ou en alignant les distributions, dans le but de rendre l'image plus homogène et d'atténuer les bandes verticales indésirables [32, 33].

Pour une colonne x, l'histogramme cumulé, aussi appelé fonction de répartition cumulative (*cumulative distribution function*, CDF), est une fonction $H_x : [0, 2^{N_{\text{bits}}} - 1] \to [0, 1]$ définie par

$$H_x(l) = \frac{1}{N_l} \sum_{y=1}^{N_l} \mathbf{1}_{\{U_0(x,y) \leq l\}},$$

où N_l désigne le nombre de lignes de l'image, N_{bits} la quantification en bits de l'image observée U_0, $U_0(x, y)$ la valeur du pixel en position $(x, y) \in \Omega$, et $\mathbf{1}_{\{\mathbf{c}\}}$ désigne la fonction indicatrice valant 1 si la condition $\mathbf{c}$ est vraie, 0 sinon. Cette fonction $H_x(l)$ correspond à la proportion de pixels de la colonne x dont la valeur est inférieure ou égale à l.

La correction par « spécification d'histogramme » consiste à transformer les valeurs radiométriques de chaque pixel afin que la distribution de la colonne corresponde à une distribution cible, appelée « spécification » [68]. Dans une image affectée par du bruit colonnaire, la distribution originale d'une colonne peut être considérée comme décalée et étirée par les *stripes* inhérentes à la colonne, c'est-à-dire que ses niveaux radiométriques subissent un « shift » et une mise à l'échelle par rapport à la distribution réelle de la scène.

Pour une colonne x, la spécification $\widetilde{H_x}$ est définie comme une moyenne lissée des CDF des colonnes voisines

$$\widetilde{H_x}(l) = \sum_{x'=x-n}^{x+n} G_\sigma(x - x')\, H_{x'}(l),$$

où G_σ est un noyau gaussien de largeur σ appliqué dans la direction des colonnes, et $2n + 1$ la largeur de la fenêtre de lissage. Cette approche, dite « midway » [69], définit la

distribution cible comme un compromis entre la colonne considérée et ses voisines, reflétant la continuité de l'image et la faible variation radiométrique entre colonnes adjacentes. Le bruit de *stripes* est ainsi réduit dans la spécification, car le lissage atténue les décalages locaux.

La correction consiste ensuite à déplacer les valeurs radiométriques de chaque pixel pour que la distribution de la colonne suive la spécification. Ce déplacement est effectué selon un transport optimal unidimensionnel [70], c'est-à-dire que les pixels sont réarrangés de manière monotone pour atteindre la distribution cible tout en minimisant le déplacement global. La transformation appliquée aux pixels est donnée par

$$V(x,y) = T_x(U_0(x,y)) = \widetilde{H_x}^{-1}\Big(H_x(U_0(x,y))\Big),$$

où $\widetilde{H_x}^{-1}$ est la fonction quantile inverse de la spécification. Cette approche permet d'homogénéiser les niveaux radiométriques entre colonnes tout en respectant la variabilité locale de la scène et en réduisant le risque de surcorrection.

La spécification $\widetilde{H_x}$, obtenue par lissage gaussien des fonctions de répartition cumulées des colonnes voisines, intègre des caractéristiques radiométriques propres à la scène elle-même. Par conséquent, la correction appliquée à chaque colonne dépend à la fois du bruit colonnaire et du contenu de la scène dans les colonnes voisines. Cette dépendance peut engendrer des modifications non désirées des détails radiométriques, notamment dans les zones à fort contraste, ainsi que l'apparition d'artefacts. La méthode de correction « MIRE » [32] constitue une référence solide et sera utilisée comme méthode de comparaison pour évaluer les performances des méthodes que nous proposerons.

2.1.2 Méthodes de filtrage spatial et fréquentiel

Le filtrage des images pour réduire le bruit colonnaire peut être effectué dans le domaine spatial ou fréquentiel.

Le filtrage spatial constitue une méthode efficace pour atténuer le bruit colonnaire, ou *stripes*, directement dans l'espace image. Un filtre guidé 1-D peut être appliqué à l'image bruitée en utilisant l'image elle-même comme guide [34]. Cette approche lisse les variations horizontales correspondant aux bandes de bruit colonnaire tout en préservant les détails verticaux et la structure réelle de la scène.

Les méthodes fréquentielles exploitent des transformées en ondelettes discrètes (*discrete wavelet transform*, DWT) ou des filtres fréquentiels pour isoler les composantes à haute fréquence associées aux variations verticales indésirables des *stripes* [35]. En supprimant ces composantes, le bruit colonnaire est efficacement atténué, mais cette approche peut également supprimer des détails réels (comme des structures verticales) lorsque leurs fréquences se superposent à celles du bruit.

2.1.3 Méthodes variationnelles

Les méthodes variationnelles estiment les composants d'une image en minimisant une fonctionnelle d'énergie combinant un terme d'attache aux données et des termes de régularisation. Dans ce cadre, l'image observée U_0 est modélisée comme la somme de l'image réelle V et du bruit colonnaire C

$$U_0 = V + C. \tag{4.74}$$

Chen et al. [36] proposent une régularisation de type variation totale sur l'image corrigée V pour favoriser la régularité spatiale et préserver les détails de la scène. Pour le bruit de *stripes*, ils imposent deux régularisations complémentaires. La première consiste en la norme L_1 du gradient vertical de C,

$$\|\nabla_y C\|_1, \tag{4.75}$$

ce qui favorise des *stripes* quasi-constantes le long des colonnes tout en autorisant certaines variations locales importantes. La deuxième régularisation, de type $\ell^{2,1}$, est définie par

$$\sum_{i=1}^{N_c} \|C_i\|_2, \tag{4.76}$$

où C_i représente la $i^{\text{ième}}$ colonne de C et N_c le nombre total de colonnes. Cette norme encourage les valeurs de chaque colonne à rester globalement proches de zéro, mais autorise certaines colonnes à présenter des amplitudes plus importantes. L'hypothèse sous-jacente est que le bruit colonnaire peut être concentré dans quelques colonnes et faible ailleurs, ce qui n'est pas toujours réaliste lorsque des *stripes* sont présentes sur la majorité des colonnes.

De nombreuses méthodes combinent plusieurs approches pour améliorer la correction du bruit, en particulier en associant des méthodes variationnelles et des techniques de filtrage spatial. Un exemple concret de ce type d'approche combinée est présenté dans ce qui suit.

2.1.4 Un exemple d'approche combinant méthode variationnelle et filtrage fréquentiel

Wang et al. [71] proposent par exemple une approche hybride associant décomposition fréquentielle et méthode variationnelle. Cette méthode se décompose en quatre étapes et combine à la fois la décomposition en ondelettes et une régularisation variationnelle, offrant un cadre pratique pour la correction du bruit colonnaire.

i. Décomposition fréquentielle de l'image

L'image observée U_0 est modélisée comme la somme de deux composantes :

— X : image de basse et moyenne fréquence (approximation globale), contenant peu ou pas de détails fins,

— F : image haute fréquence, incluant détails fins, bords, textures et bruit.

On a alors

$$U_0 = X + F. \tag{4.77}$$

La composante haute fréquence F, qui contient à la fois les détails et le bruit, est ensuite décomposée en trois sous-bandes directionnelles via une transformée en ondelettes discrètes (DWT) [9]

$$F = (V_1, H_1, D_1), \tag{4.78}$$

où

9. La DWT utilisée correspond ici aux ondelettes de Haar, permettant de séparer les composantes verticale, horizontale et diagonale de l'image.

— V_1 représente la composante verticale, contenant à la fois le bruit colonnaire et les contours verticaux d'objets de la scène,

— H_1 la composante horizontale,

— D_1 la composante diagonale.

Le bruit de *stripes* se manifeste essentiellement dans V_1, provoquant des fluctuations rapides et régulières entre colonnes tout en partageant l'information des objets verticaux, ce qui complique sa séparation des structures réelles de l'image.

ii. Lissage de l'image par variation totale

L'image U_0 est débruitée via un problème variationnel en supposant que les oscillations dues au bruit de *stripes* sont de faible amplitude comparées aux discontinuités des objets. L'image débruitée $\widehat{G}$ est obtenue en résolvant le problème

$$\widehat{G} \in \underset{G\in\mathbb{R}^{\Omega}}{\operatorname{argmin}} \lambda_1\|\nabla_x G\|_1 + \lambda_2\|\nabla_y(U_0 - G)\|_1 + \|U_0 - G\|_2^2, \tag{4.79}$$

où ∇_x et ∇_y désignent les gradients discrets selon les axes horizontal et vertical, et $\lambda_1, \lambda_2 > 0$ sont des coefficients de régularisation.

Les rôles des termes sont les suivants

— $\|\nabla_x G\|_1$ favorise la régularité dans la direction horizontale de l'image corrigée tout en préservant les discontinuités verticales correspondant aux contours des objets.

— $\|\nabla_y(U_0 - G)\|_1$ cible le bruit de *stripes* en favorisant une quasi-constance le long des colonnes tout en autorisant des variations ponctuellement importantes grâce à la norme ℓ^1.

— $\|U_0 - G\|_2^2$ assure la fidélité de l'image débruitée par rapport à l'image observée.

iii. Débruitage de la composante verticale par filtrage guidé

Le filtrage guidé [72] utilise une image guide pour orienter le débruitage, de sorte que l'image filtrée reste proche de l'entrée tout en suivant les structures locales du guide.

Dans notre approche, la composante verticale V_1 est débruitée en s'inspirant des structures locales de $\widehat{G}$, image guide presque exempte de bruit de *stripes*.

iv. Reconstruction finale

La composante verticale filtrée est réinjectée dans la décomposition en ondelettes, puis l'image finale est reconstruite par transformée en ondelettes inverse.

Cette méthode multi-étapes, combinant analyse multi-échelle, problème variationnel et filtrage guidé, supprime efficacement le bruit colonnaire tout en préservant les structures fines et les contours verticaux. Cependant, elle peut s'avérer coûteuse en temps de calcul et sensible au choix des paramètres, ce qui peut limiter son application en temps réel. De plus, l'hypothèse selon laquelle le bruit de *stripes* a une amplitude inférieure aux variations liées aux contours verticaux n'est pas toujours vérifiée (par exemple, dans des scènes faiblement contrastées), ce qui peut affecter la qualité du débruitage colonnaire dans certains cas.

2.1.5 Méthodes d'apprentissage profond pour la correction du bruit colonnaire

Récemment, plusieurs méthodes fondées sur l'apprentissage profond ont été proposées pour corriger le bruit colonnaire sur les images infrarouges. Bien qu'elles soient prometteuses, il nous est impossible de les reproduire dans le cadre de ce travail, faute de moyens et de compétences pour entraîner et manipuler les réseaux de neurones.

Comme présenté au chapitre 2 (section 2.2.5), la correction des *stripes* peut être réalisée efficacement grâce à des réseaux de neurones convolutifs (CNN, pour *convolutional neural network*) supervisés. Cao et al. proposent un modèle résiduel entraîné sur des images simulées, où des motifs de *stripes* artificiels sont ajoutés à des images thermiques de référence [55]. Le réseau apprend à estimer le résidu correspondant aux *stripes* verticales, facilitant ainsi la suppression des artefacts. L'apprentissage résiduel permet de concentrer la capacité du réseau sur la correction des bandes de bruit, tout en préservant les structures naturelles de l'image. Une fonction de perte de type ℓ^1 est utilisée pour mieux gérer la faible amplitude du bruit par rapport aux variations naturelles de la scène, et l'extraction multi-échelle avec *pooling* étendu permet de distinguer les *stripes* des structures verticales.

L'*asymmetric sampling correction network* (ASCNet) proposé par Yuan et al. [73] est un réseau de neurones entraîné pour corriger les images infrarouges affectées par un bruit colonnaire. L'entraînement consiste à présenter au réseau de nombreux exemples d'images bruitées et de leurs correspondantes propres, afin qu'il apprenne à prédire le bruit présent dans chaque image et à le retirer.

Le traitement commence par un réseau de neurones convolutionnel (CNN) qui décompose l'image bruitée en un ensemble de cartes de représentation.

La transformée en ondelettes de Haar résiduelle (RHDWT, pour *residual Haar discrete wavelet transform*) est ensuite appliquée sur ces cartes pour séparer le bruit vertical des détails horizontaux et des textures de l'image. Cette étape comporte deux branches. La branche modèle applique des filtres spécifiques pour estimer le bruit de type *stripes* dans chaque colonne. Certaines colonnes contiennent cependant à la fois du bruit et des structures réelles, comme des objets verticaux. Pour gérer cela, la branche résiduelle apprend à partir d'exemples d'images bruitées et propres pour compléter et corriger l'estimation de la branche modèle. En combinant ces deux branches, le réseau obtient des cartes de représentation enrichies, permettant de distinguer efficacement le bruit des détails réels de la scène.

Après la RHDWT, le module de correction de non-uniformité des colonnes prend ces cartes enrichies et ajuste les différences d'intensité entre colonnes causées par le bruit. Il homogénéise chaque colonne tout en préservant les détails horizontaux et la structure de l'image, en s'appuyant sur des informations locales et globales pour distinguer le bruit des variations réelles de la scène.

Une fois cette correction appliquée, les cartes de représentation sont débruitées et cohérentes verticalement. Le *pixel shuffle* réorganise ensuite ces cartes pour reconstruire l'image à sa résolution initiale, en assemblant toutes les informations corrigées sans perte de détails ni introduction de biais vertical. Cette étape corrige notamment l'effet de réduction de résolution induit par les premières couches CNN.

Enfin, le réseau prédit le bruit vertical et le soustrait de l'image d'entrée pour produire l'image finale corrigée. Pendant l'entraînement, cette image finale est comparée à l'image

propre correspondante à l'aide d'une fonction de perte ℓ^2, c'est-à-dire l'erreur quadratique moyenne entre les pixels des deux images. Cette mesure sert à ajuster les paramètres du réseau afin d'améliorer progressivement ses prédictions et obtenir une correction plus précise du bruit colonnaire.

Ainsi, ASCNet combine des outils mathématiques et des modules de deep learning adaptés aux images infrarouges : la transformée en ondelettes pour isoler le bruit directionnel, les modules de correction des non-uniformités colonnaires pour homogénéiser les colonnes en tenant compte des dépendances locales et globales, et le *pixel shuffle* pour reconstruire l'image sans perte de détails. Il convient de noter avec précaution qu'il nous est impossible de tester directement cette méthode dans le cadre de ce travail.

Les méthodes classiques de suppression du bruit colonnaire, qu'elles soient fondées sur les statistiques, le filtrage ou les approches variationnelles, reposent souvent sur l'hypothèse que les *stripes* ont une amplitude plus faible que les structures naturelles de la scène. Cette hypothèse limite leur efficacité dans les scènes faiblement contrastées, où le bruit colonnaire peut avoir une amplitude comparable aux variations radiométriques de la scène. Notre démarche ne fait pas cette hypothèse restrictive, ce qui la rend plus flexible et robuste face à des situations variées, notamment lorsque l'amplitude des *stripes* est importante par rapport au contenu de la scène. Cette absence d'hypothèse ouvre la voie à des solutions plus générales et adaptatives. Dans ce contexte, nous présentons différentes approches expérimentales pour la correction du bruit colonnaire.

2.2 Hypothèses de travail et modèle de perturbation

Pour dépasser les limites des méthodes classiques, notre démarche s'appuie sur des hypothèses de travail plus souples et mieux adaptées aux différents contextes d'imagerie thermique. Nous présentons ci-après ces hypothèses ainsi que le modèle de perturbation que nous utilisons pour représenter les *stripes* dans les images.

2.2.1 Hypothèses de travail

Dans cette démarche, nous proposons plusieurs modèles de correction du bruit colonnaire reposant sur les hypothèses suivantes :

— **H1 Offset colonnaire constant.** Le bruit de stripes est modélisé comme un offset constant au sein de chaque colonne de l'image, identique sur toutes les lignes pour une colonne donnée (cf. comparaison entre les sous-figures 4.12b et 4.12c). Cette hypothèse repose sur l'architecture du détecteur (cf. figure 2.9). Pour chaque colonne, la tension mesurée sur un pixel actif sensible à la scène est soustraite à celle d'un pixel aveugle insensible à la scène situé dans la même colonne. La dérive temporelle des caractéristiques du pixel aveugle entraîne une variation de la tension mesurée, ce qui explique que chaque colonne possède un offset constant propre.

— **H2 Indépendance spatiale.** Les offsets colonnaires sont supposés indépendants entre eux, ce qui est illustré par la sous-figure 4.12e.

2.2.2 Modèle de perturbation

Parmi les hypothèses présentées, c'est principalement l'hypothèse **H1** d'un offset colonnaire constant qui motive le modèle additif

$$U_0(x,y) = V(x,y) + C(x), \quad \forall (x,y) \in \Omega, \tag{4.80}$$

où $U_0 \in \mathbb{R}^\Omega$ représente l'image observée bruitée, $V \in \mathbb{R}^\Omega$ l'image réelle de la scène, et $C \in \mathbb{R}^{N_c}$ le vecteur des offsets colonnaires, dépendant uniquement de la coordonnée horizontale x, avec N_c le nombre de colonnes de l'image.

Le modèle de correction s'écrit alors

$$V(x,y) = U_0(x,y) - C(x). \tag{4.81}$$

Ce modèle additif permet de poser le problème de correction de manière claire.

Par la suite, nous réécrirons le modèle de correction (4.81) de manière plus compacte

$$V = U_0 - C, \tag{4.82}$$

en assimilant le vecteur C à une image dont les valeurs sont constantes dans chaque colonne.

Le débruitage colonnaire peut être un problème mal posé. La figure 4.13 illustre ce point. L'image observée, affectée de bruit colonnaire, peut correspondre à plusieurs scènes originales. L'ambiguïté apparaît lorsqu'une structure verticale réelle est présente.

Nous nous appuyons désormais sur cette formulation pour proposer une méthode explicite de correction des offsets colonnaires.

2.3 Débruitage colonnaire par le modèle TVx

L'hypothèse **H2** suppose que les offsets colonnaires sont indépendants entre eux, sans corrélation spatiale horizontale. Cette indépendance génère un signal d'offsets possiblement très irrégulier le long des colonnes, ce qui se traduit par un gradient horizontal $\nabla_x U_0$ fortement bruité et dispersé dans l'image observée U_0.

En revanche, la perturbation étant constante le long de chaque colonne (hypothèse **H1**), les variations verticales inter-lignes ne sont pas modifiées, ce qui préserve le gradient vertical $\nabla_y U_0$.

Ainsi, pour obtenir une image corrigée [10] $V = U_0 - \widehat{C}$ présentant une régularité adaptée à une image naturelle, on cherche un vecteur d'offsets $\widehat{C}$ tel que le gradient horizontal de V soit relativement parcimonieux : les variations horizontales doivent être rares et localisées, correspondant aux contours verticaux réels de la scène. Cette idée se traduit par un problème variationnel que l'on appelle TVx, qui consiste à minimiser la norme ℓ^1 du gradient horizontal (défini selon (4.32a)) de l'image corrigée. Le problème TVx s'écrit

$$\widehat{C} \in \underset{C \in \mathbb{R}^{N_c},\, C(1)=0}{\operatorname{argmin}} \left\| \nabla_x \left(U_0 - C \right) \right\|_1 . \tag{4.83}$$

Cette formulation permet de retrouver un vecteur $\widehat{C}$ qui réduit les variations horizontales artificielles de l'image observée U_0. Le problème TVx permet de déterminer le vecteur

10. Par coordonnées, cela signifie que pour tout pixel $(x,y) \in \Omega$, $V(x,y) = U_0(x,y) - \widehat{C}(x)$.

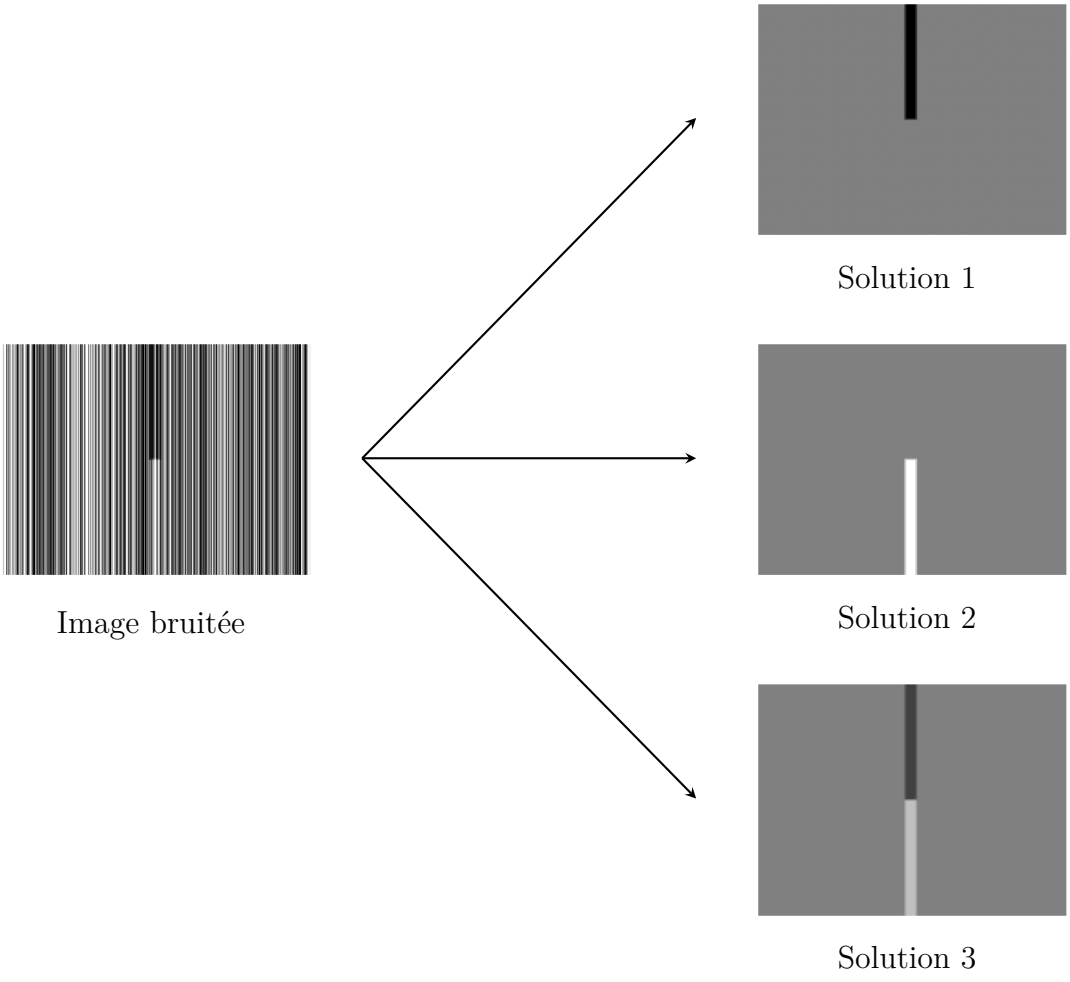


FIGURE 4.13 – **Illustration que le problème de débruitage peut être mal posé.** L'image à gauche est affectée de bruit colonnaire. Selon la perception visuelle, notre œil peut tenter de reconstituer la scène à partir de cette image, ce qui permet de considérer les trois solutions corrigées présentées à droite comme valides. La première montre une structure verticale noire en haut au milieu. La deuxième montre une structure verticale blanche au milieu en bas. La troisième est la moyenne des deux premières solutions. Ces trois interprétations sont toutes compatibles avec ce que l'on observe sur l'image bruitée, illustrant ainsi le caractère intrinsèquement mal posé du problème. D'autres interprétations sont également possibles.

$\nabla_x \widehat{C}$, mais le vecteur $\widehat{C}$ reste défini à une constante près. La contrainte $\widehat{C}(1) = 0$ fixe cette constante, ce qui permet de reconstruire de manière unique $\widehat{C}$ à partir de son gradient horizontal.

Soit $\widehat{C} \in \mathbb{R}^{N_c}$ solution de (4.83), on a

$$\widehat{C} \in \underset{C \in \mathbb{R}^{N_c},\, C(1)=0}{\operatorname{argmin}} \left\| \nabla_x \left(U_0 - C \right) \right\|_1 .$$

En posant $W = \nabla_x C$, puisque $\nabla_x C(N_c) = 0$ (condition de bord de Neumann) il vient

$$\nabla_x \widehat{C} \in \underset{W \in \mathbb{R}^{N_c},\, W(N_c)=0}{\operatorname{argmin}} \left\| \nabla_x U_0 - W \right\|_1 , \quad \widehat{C}(1) = 0,$$

ce qui s'écrit de manière explicite

$$(\nabla_x \widehat{C}(1), \ldots, \nabla_x \widehat{C}(N_c - 1)) \in \underset{(w_1, \ldots, w_{N_c-1}) \in \mathbb{R}^{N_c-1}}{\operatorname{argmin}} \sum_{x=1}^{N_c-1} \sum_{y=1}^{N_l} \left| \nabla_x U_0(x, y) - w_x \right| , \quad \widehat{C}(1) = 0.$$

Par séparabilité des variables, cela conduit à minimiser indépendamment chaque colonne

$$\forall x \in [\![1, N_c - 1]\!], \quad \nabla_x \widehat{C}(x) \in \underset{w \in \mathbb{R}}{\operatorname{argmin}} \sum_{y=1}^{N_l} \left| \nabla_x U_0(x, y) - w \right| , \quad \widehat{C}(1) = 0,$$

et donc

$$\forall x \in [\![1, N_c - 1]\!], \quad \nabla_x \widehat{C}(x) = \text{médiane}\Big(\nabla_x U_0(x, 1), \ldots, \nabla_x U_0(x, N_l) \Big), \quad \widehat{C}(1) = 0, \tag{4.84}$$

où médiane(X) désigne la médiane [11] de l'ensemble X, et N_l le nombre de lignes de l'image.

Puisqu'il est possible de reconstruire $\widehat{C}$ selon

$$\begin{cases} \widehat{C}(1) = 0, \\ \widehat{C}(x+1) = \widehat{C}(x) + \nabla_x \widehat{C}(x), \quad 1 \leq x \leq N_c - 1, \end{cases} \tag{4.85}$$

l'injection de (4.84) conduit à la solution $\widehat{C}$ du problème TVx (4.83)

$$\begin{cases} \widehat{C}(1) = 0, \\ \widehat{C}(x+1) = \widehat{C}(x) + \text{médiane}\Big(\nabla_x U_0(x, 1), \ldots, \nabla_x U_0(x, N_l) \Big), \quad 1 \leq x \leq N_c - 1. \end{cases} \tag{4.86}$$

Ainsi déterminée, cette solution existe et est unique.

L'image corrigée par TVx, $V \in \mathbb{R}^{\Omega}$, s'écrit

$$V = U_0 - \widehat{C}, \tag{4.87}$$

ce qui implique que

$$\nabla_x V(x, y) = \nabla_x U_0(x, y) - \nabla_x \widehat{C}(x), \quad \forall (x, y) \in \Omega. \tag{4.88}$$

11. Pour calculer la médiane, on commence par ordonner toutes les valeurs de la série. Si le nombre de valeurs N_l est impair, la médiane est la valeur centrale. Si N_l est pair, la médiane est la moyenne des deux valeurs centrales.

D'après (4.84), on en déduit que

$$\forall x \in [\![1, N_c]\!], \quad \text{médiane}\Big(\nabla_x V(x,1), \ldots, \nabla_x V(x,N_l)\Big) = 0. \tag{4.89}$$

Autrement dit, dans chaque colonne, la médiane du gradient horizontal de l'image corrigée est nulle. La correction obtenue est donc strictement la même, quelle que soit l'intensité ou la répartition spatiale du bruit colonnaire, dès lors que celui-ci est constant le long de chaque colonne.

Nous présentons à présent les résultats obtenus pour la correction du bruit colonnaire à l'aide du modèle TVx sur une image de test, ainsi qu'une discussion des principales difficultés rencontrées. L'image de test a été générée à partir d'une image de référence parfaite, dite « vérité terrain », en y ajoutant un bruit colonnaire simulé $C \in \mathbb{R}^{N_c}$. Conformément à l'hypothèse **H1**, ce bruit est constant au sein de chaque colonne, c'est-à-dire identique pour toutes les lignes d'une colonne donnée. Selon l'hypothèse **H2**, les offsets colonnaires ont été tirés de manière indépendante pour chaque colonne, sans corrélation spatiale. Par construction, la première colonne est exempte de perturbation, avec $C(1) = 0$, et sert de référence pour l'analyse.

2.3.1 Étude de la correction TVx sur scènes avec ou sans structures verticales

Les figures 4.14 et 4.15 présentent les résultats du débruitage colonnaire obtenu avec le modèle TVx pour deux types de scènes. Dans la figure 4.14, la scène ne comporte pas de structure verticale marquée et la correction est très satisfaisante. Ce cas simple ne pose sans doute pas de difficulté particulière pour les méthodes de suppression de bruit colonnaire. En revanche, la figure 4.15 montre une scène contenant des structures verticales. Après correction, on observe des artefacts latéraux sous forme de discontinuités horizontales. Ce type de configuration rend difficile la distinction entre les structures de la scène et le bruit colonnaire, ce qui peut entraîner une correction incorrecte.

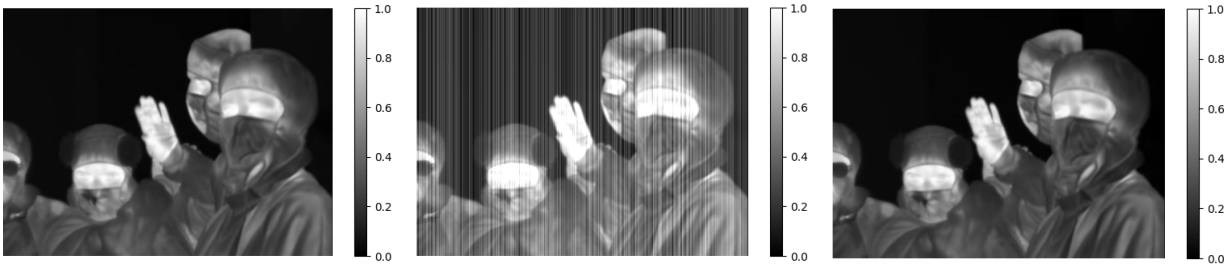
C'est ce second cas que nous retenons pour la suite de l'analyse. La figure 4.16 compare le gradient horizontal de l'image corrigée à celui de la vérité terrain. Le modèle TVx impose une médiane nulle colonne par colonne, ce qui ne reflète pas toujours la réalité, en particulier en présence de structures verticales.

Pour mieux comprendre cette différence, il est pertinent d'analyser le gradient horizontal colonne par colonne.

2.4 Débruitage colonnaire par analyse du gradient horizontal colonne par colonne

Dans l'image vérité, certaines zones à l'intérieur d'objets présentent une irradiance à peu près uniforme, avec de légères variations dues à la texture ou au bruit. Lorsqu'on examine le gradient horizontal colonne par colonne, ces zones apparaissent comme des segments verticaux dans lesquels le signal oscille faiblement autour de zéro. Dans l'image bruitée, on observe un comportement similaire, mais les oscillations du gradient horizontal sont centrées autour d'une valeur décalée, reflétant le biais induit par le bruit colonnaire.

En supposant que le segment identifié correspond effectivement à une zone uniforme dans la scène, la moyenne locale des oscillations du gradient horizontal fournit alors une



(a) Vérité terrain (b) Image bruitée Image corrigée par TVx

FIGURE 4.14 – **Correction d'une image sans structure verticale par le modèle TVx.** Image issue d'une caméra thermique refroidie. L'échelle radiométrique de toutes les images a été réduite à l'intervalle $[0, 1]$. (a) Cette image représente la vérité terrain. L'échelle radiométrique a été réduite à l'intervalle $[0, 1]$. Elle ne présente aucune structure verticale manifeste. (b) L'image bruitée est obtenue par addition de la vérité terrain (a) et d'un bruit purement colonnaire, indépendant selon les colonnes et suivant une loi normale $\mathcal{N}(0, 0{,}1^2)$, constitue une simulation bruitée respectant les hypothèses H1 et H2 (cf. sous-section 2.2.1). L'image est alors dégradée par un bruit structuré verticalement. (c) L'image corrigée à l'aide du modèle TVx appliqué à l'image dégradée (b) est d'une qualité très proche de la vérité terrain (a), les différences sont peu perceptibles. Lorsque la scène ne possède pas de structure verticale, l'image corrigée par le modèle TVx est de très bonne qualité.

estimation du biais recherché. Il devient possible de recentrer le signal de ce segment autour de zéro, ce qui revient à corriger localement le biais induit par le bruit colonnaire. Dans la suite, nous appelons segment « homogène » un intervalle discret d'au moins deux lignes dans le signal formé par le gradient horizontal d'une colonne, où les oscillations sont faibles autour d'une valeur constante.

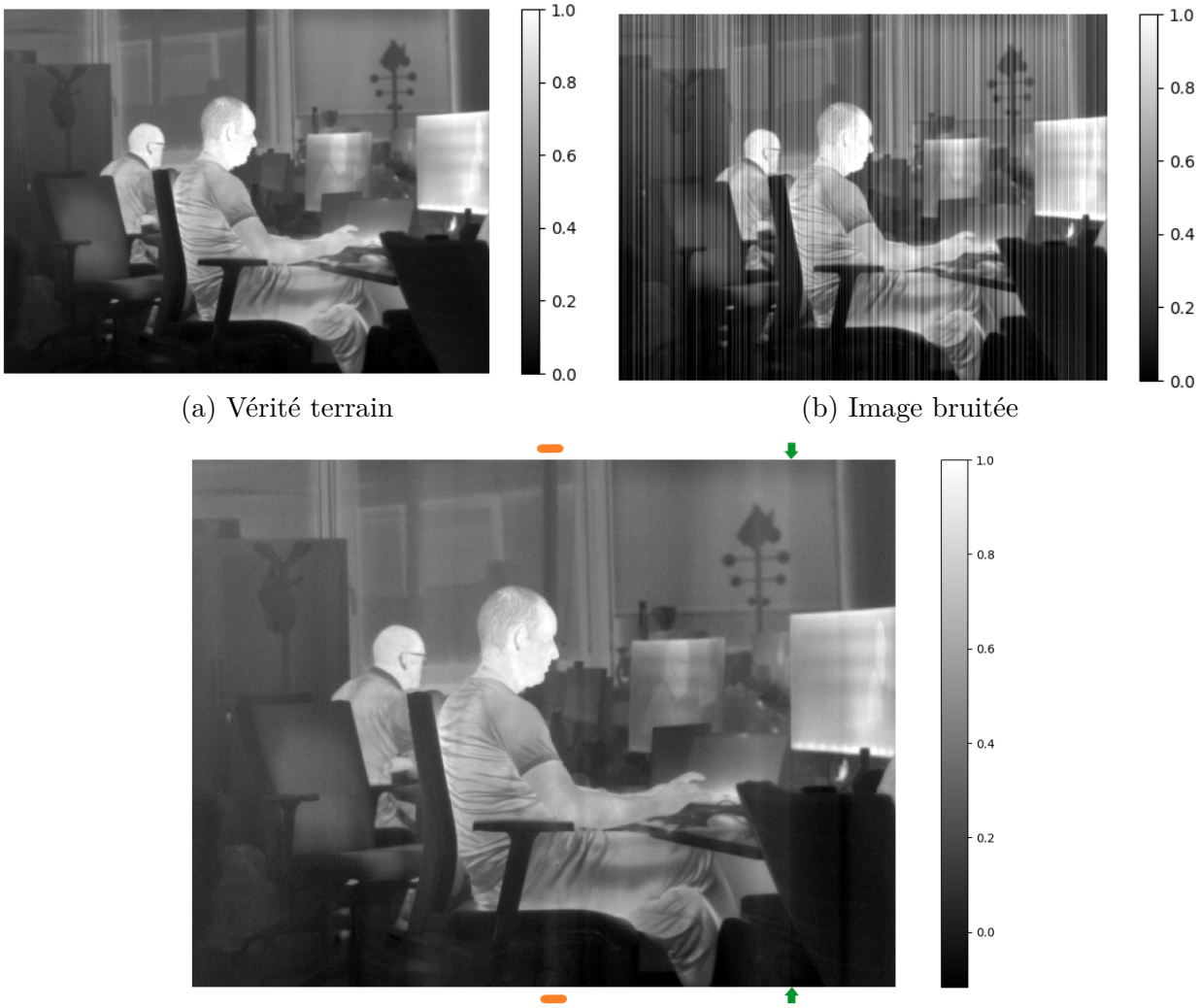
Lorsqu'une structure verticale est présente dans la scène, le signal associé à la vérité terrain peut osciller autour d'une valeur non nulle, propre à cette structure. Un segment extrait dans une telle zone ne représente donc pas une région uniforme, mais une variation réelle de la scène. Dans ce cas, la valeur moyenne d'oscillation du signal associé à l'image bruitée ne reflète pas uniquement le biais induit par le bruit colonnaire, mais combine ce biais avec la composante due à la structure. L'estimation du biais devient alors incorrecte si on recentre les oscillations autour de zéro. La présence de structures verticales peut rendre le problème de correction particulièrement délicat et potentiellement mal posé.

La fiabilité de la méthode que nous envisageons ici dépendra donc de la capacité à identifier les segments homogènes correspondant effectivement à des zones d'irradiance uniforme. Une détection imprécise conduira à des estimations biaisées et réduira l'efficacité de la correction.

2.4.1 Détection des segments homogènes au sein d'une colonne du gradient horizontal

2.4.1.1 Rôle de l'identification des segments homogènes dans l'estimation du bruit colonnaire

Soit $V_0 \in \mathbb{R}^\Omega$ une image thermique de la scène (vérité terrain), et $x_0 \in [\![1, N_c - 1]\!]$ une colonne fixée. Considérons un intervalle discret de lignes $[\![a, b]\!]$, avec $1 \leq a < b \leq N_l$. Si la région rectangulaire délimitée horizontalement par les colonnes x_0 et $x_0 + 1$, et

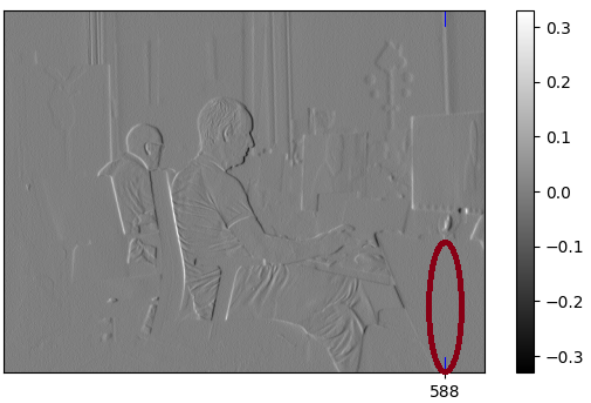


(a) Vérité terrain

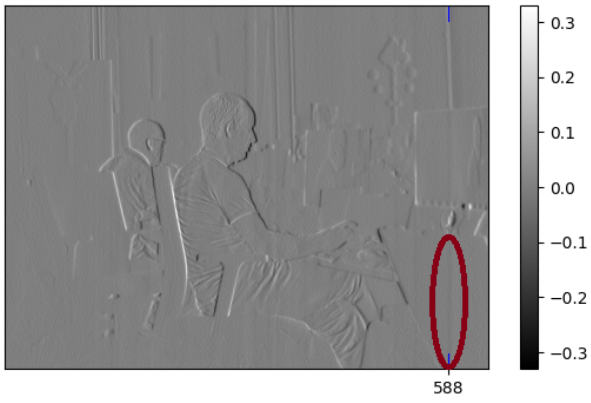
(b) Image bruitée

(c) Image corrigée par le modèle TVx

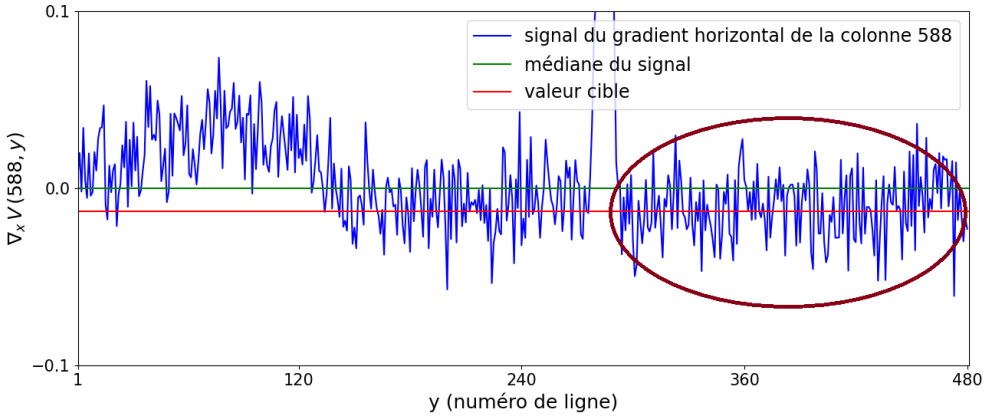
FIGURE 4.15 – **Correction d'une image avec structures verticales par le modèle TVx.**Données issues de la caméra *ATI-640b*. L'échelle radiométrique de toutes les images a été réduite à l'intervalle $[0, 1]$ (a) Elle représente la « vérité terrain ». Cette image présente un nombre de structures verticales assez important. (b) Image obtenue par addition de la vérité terrain (a) et d'un bruit purement colonnaire, indépendant selon les colonnes et suivant une loi normale $\mathcal{N}(0, 0{,}1^2)$. L'image est alors dégradée par des *stripes* verticales. (c) Image (b) corrigée par le modèle TVx. Les *stripes* verticales ont disparu, mais des artefacts latéraux apparaissent, caractérisés par des discontinuités horizontales. Par ailleurs, l'échelle radiométrique globale a augmenté d'environ 10 % en raison d'un voile horizontal basse fréquence. Ce voile résulte de l'intégration colonne par colonne du gradient horizontal estimé, où de petites imprécisions s'accumulent et provoquent une dérive de l'échelle. Des imprécisions plus marquées entraînent localement des discontinuités, visibles dans les zones délimitées par les barres orange et les flèches vertes. Dans la zone orange, une bande agit comme un voile éclaircissant la structure verticale en haut et réduisant l'uniformité sous le siège en bas. Le long de la ligne verticale indiquée par les flèches vertes, en haut, une discontinuité anormale est visible sur le store de la fenêtre, alors que cette région est uniforme dans (a). Plus bas dans cette zone, le dossier du siège présente une irrégularité spatiale : la partie gauche est plus claire que la droite. La présence de structures verticales dans ces zones peut biaiser l'estimation de la correction.



(a) Gradient horizontal vérité terrain



(b) Gradient horizontal correction par TVx



(c) Signal du gradient horizontal de l'image corrigée par TVx, colonne 588

FIGURE 4.16 – **Analyse du gradient horizontal.** (a) Le gradient horizontal de l'image vérité terrain est parcimonieux puisque le niveau radiométrique des pixels est souvent proche de zéro. La zone à l'intérieur de l'ovale bordeaux, située en bas de la colonne 588, présente un gradient proche de zéro, témoignant d'une forte uniformité. D'après la figure 4.15a, cette zone correspond au dossier d'un objet à irradiance relativement uniforme (dos d'un siège), ce qui explique la parcimonie du gradient horizontal dans cette zone. (b) Le gradient horizontal de l'image corrigée par le modèle TVx reste globalement parcimonieux, mais la zone à l'intérieur de l'ovale bordeaux est moins uniforme que dans (a), en raison d'un petit bloc de lignes verticales légèrement plus sombres, où le signal est inférieur à zéro. (c) Le signal du gradient horizontal de l'image corrigée par TVx, pour la colonne 588, est tracé en bleu. Par construction, sa médiane est nulle ; elle est indiquée par la ligne verte. En revanche, la médiane du gradient de la vérité terrain pour cette même colonne est non nulle, et son opposé est représenté par la ligne rouge, qui constitue la valeur cible. La zone délimitée par l'ovale bordeaux correspond à la colonne 588 de la même zone sur la figure 4.15b. À cet endroit, le signal corrigé n'oscille pas autour de zéro, contrairement à celui de la vérité terrain, ce qui rompt la parcimonie attendue. Pour corriger cela, il faudrait translater verticalement le signal corrigé de manière à faire coïncider la ligne verte (médiane actuelle) avec la ligne rouge (valeur cible), ce qui recentrerait les oscillations autour de zéro dans cette zone. La contrainte implicite du modèle TVx imposant une médiane nulle à chaque colonne du gradient horizontal n'est pas toujours pertinente.

verticalement par les lignes a à b, correspond à une zone uniforme de la scène, alors les intensités $V_0(x_0, y)$ et $V_0(x_0 + 1, y)$, pour tout $y \in [\![a, b]\!]$, peuvent être modélisées comme des réalisations d'une variable aléatoire suivant une loi normale $\mathcal{N}(\mu, \rho^2)$, où μ désigne la moyenne radiométrique de la zone uniforme, et ρ^2 la variance due au bruit et/ou à la texture locale.

Dans ce cas, le gradient horizontal défini par

$$\nabla_x V_0(x_0, y) = V_0(x_0 + 1, y) - V_0(x_0, y)$$

suit une loi normale centrée

$$\nabla_x V_0(x_0, y) \sim \mathcal{N}(0, \sigma^2), \tag{4.90}$$

avec [12] $\sigma = \sqrt{2}\rho$. Cette relation fournit un modèle statistique simple du gradient horizontal dans une zone uniforme, utile pour la détection de segments.

Soit U_0 une image dégradée par un bruit colonnaire C selon le modèle (4.80), issue de la scène réelle V_0. En appliquant le gradient horizontal, on obtient, pour une colonne x_0 et tout $y \in [\![1, N_l]\!]$

$$\nabla_x U_0(x_0, y) = \nabla_x V_0(x_0, y) + \nabla_x C(x_0). \tag{4.91}$$

Si un segment $[\![a, b]\!]$ correspond à une zone localement uniforme de la scène, alors le gradient horizontal de la scène $\nabla_x V(x_0, y)$ suit la loi (4.90), ce qui implique

$$\forall y \in [\![a, b]\!], \quad \nabla_x U_0(x_0, y) \sim \mathcal{N}\Big(\nabla_x C(x_0), \sigma^2\Big). \tag{4.92}$$

Dans ce cas, la moyenne empirique du signal $\nabla_x U_0(x_0, \cdot)$ sur l'intervalle $[\![a, b]\!]$,

$$\nabla_x \widehat{C}(x_0) := \frac{1}{b - a + 1} \sum_{y=a}^{b} \nabla_x U_0(x_0, y), \tag{4.93}$$

fournit une estimation naturelle de $\nabla_x C(x_0)$. Si on pose $\widehat{C}(1) = 0$, il est alors possible de déterminer $\widehat{C}$ selon (4.85). L'image corrigée V s'obtient alors selon la formule

$$V = U_0 - \widehat{C}. \tag{4.94}$$

2.4.1.2 Détection de segments homogènes dans un signal par un test du χ^2

L'objectif est d'identifier, le long d'une colonne d'image, des segments où le gradient horizontal présente peu de variation, autrement dit des zones localement homogènes. Pour cela, on se ramène à l'étude du signal unidimensionnel $(X_1, \ldots, X_{N_l})$ représentant les valeurs de $\nabla_x U_0(x_0, \cdot)$ pour une colonne x_0 fixée.

Étant donné un intervalle discret $[\![a, b]\!] \subset [\![1, N_l]\!]$, on souhaite tester si le signal est localement stationnaire autour de sa moyenne. On définit alors la moyenne empirique

$$\overline{X}_{a,b} = \frac{1}{b - a + 1} \sum_{y=a}^{b} X_y,$$

12. L'écart-type ρ peut être estimé à partir d'une zone uniforme de l'image V, permettant ainsi d'obtenir une estimation de σ.

et on introduit les variables normalisées

$$T_y = \frac{X_y - \overline{X}_{a,b}}{\sigma}, \quad \text{pour } y \in [\![a, b]\!],$$

où σ est un écart-type fixé, supposé connu (cf. figure 4.17).

On définit l'hypothèse nulle $\mathbf{H_0}$: le segment $[\![a, b]\!]$ est dit homogène, c'est-à-dire que les variables X_y, pour $y \in [\![a, b]\!]$, sont indépendantes et distribuées selon une même loi normale $\mathcal{N}(\mu, \sigma^2)$, avec une moyenne μ constante (mais inconnue) et une variance σ^2 supposée connue. Sous $\mathbf{H_0}$, la statistique de test

$$S_{a,b} = \sum_{y=a}^{b} T_y^2 \;=\; \sum_{y=a}^{b} \left(\frac{X_y - \overline{X}_{a,b}}{\sigma} \right)^2$$

suit une loi du χ^2 à $b - a$ degrés de liberté [13].

$$\mathbf{H_0} : \; S_{a,b} \sim \chi^2(b - a). \tag{4.95}$$

Lorsque l'hypothèse $\mathbf{H_0}$ n'est pas rejetée, le segment $[\![a, b]\!]$ peut être considéré comme homogène. Afin de quantifier ce degré d'homogénéité et de comparer différents segments candidats, on utilise la p-valeur associée à la statistique $S_{a,b}$. Celle-ci fournit un critère de régularité permettant de hiérarchiser les segments selon leur homogénéité locale.

On note $p_{a,b}$ la p-valeur associée à la statistique $S_{a,b}$, définie comme la fonction de survie de la loi $\chi^2(b - a)$, évaluée en $S_{(a,b)}$, elle s'écrit

$$p_{a,b} = \mathbb{P}\Big(\chi^2(b - a) \geq S_{a,b}\Big). \tag{4.96}$$

Lorsque les variations du signal sur l'intervalle $[\![a, b]\!]$ sont importantes, la p-valeur est proche de zéro, ce qui suggère que le segment est non-homogène, c'est-à-dire que les données s'écartent significativement d'un comportement uniforme. À l'inverse, une p-valeur supérieure à un seuil (par exemple 5%) indique que l'hypothèse d'homogénéité locale ne peut pas être rejetée, ce qui traduit une faible dispersion autour de la moyenne et donc une homogénéité du segment considéré.

Afin de hiérarchiser les segments détectés selon leur niveau d'homogénéité, on introduit le score

$$\gamma_{a,b} = (b - a) \, \log\left(\Big(1 - p_{(a,b)}\Big)^{-1} \right). \tag{4.97}$$

Cette définition repose sur l'idée que la confiance accordée à un segment homogène doit croître avec sa longueur, afin de pénaliser les segments trop courts. En effet, dans une zone supposée uniforme (comme l'intérieur d'un objet), on s'attend à observer un segment homogène relativement long sur une colonne. Un segment trop court risque de ne pas représenter fidèlement cette uniformité, car il peut correspondre à une irrégularité locale, un contour ou du bruit. La longueur joue donc un rôle crucial : un segment long est plus susceptible d'être représentatif d'une zone réellement homogène. La fonction logarithme

13. On perd un degré de liberté car la moyenne μ est inconnue et remplacée par la moyenne empirique $\overline{X}_{a,b}$. Cette estimation impose la contrainte $\sum_{y=a}^{b}(X_y - \overline{X}_{a,b}) = 0$, ce qui réduit à $b - a$ le nombre de composantes indépendantes de T_y, même si σ est supposée connue.

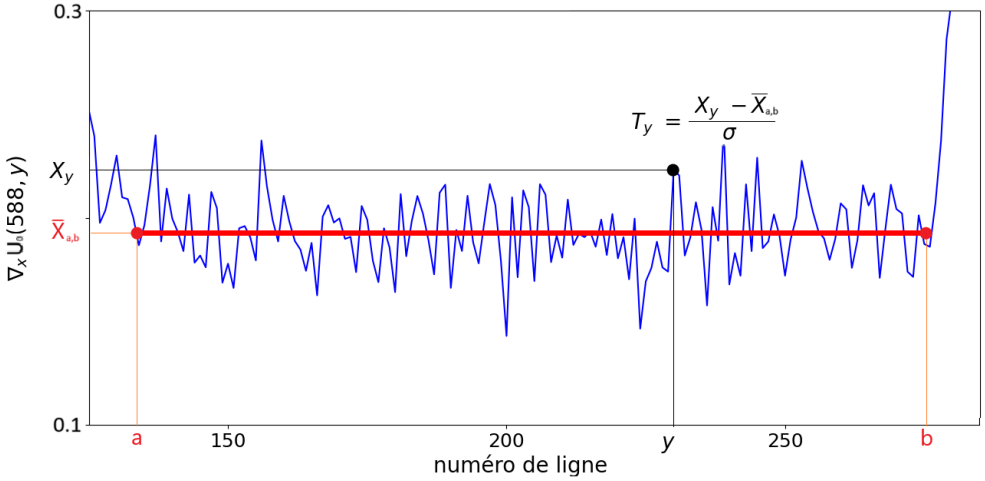


FIGURE 4.17 – **Détection de segment par un test du χ^2.** On considère un signal $(X_1, \ldots, X_{N_l})$ correspondant ici au gradient horizontal de l'image observée U_0 sur la colonne $x_0 = 588$. On suppose l'écart-type σ déterminé ou estimé. Pour tester l'homogénéité d'un intervalle discret $[\![a, b]\!]$, on définit, pour tout $y \in [\![a, b]\!]$, la statistique $T_y = \frac{X_y - \overline{X}_{a,b}}{\sigma}$, où $\overline{X}_{a,b}$ désigne la moyenne locale. Si l'intervalle est homogène, alors $T_y \sim \mathcal{N}(0, 1)$, et la statistique $\dfrac{1}{b-a+1} \displaystyle\sum_{y=a}^{b} T_y^2$ suit une loi du χ^2 à $b - a$ degrés de liberté.

appliquée à l'inverse de $1 - p_{(a,b)}$ permet quant à elle de mieux discriminer les segments non-homogènes, c'est-à-dire ceux pour lesquels $p_{(a,b)}$ est proche de 0.

On en déduit que plus le score $\gamma_{a,b}$ est élevé, plus le segment $[\![a, b]\!]$ est susceptible d'être homogène et représentatif d'une zone uniforme et ainsi servir d'estimation du gradient colonnaire $\nabla_x \widehat{C}(x_0)$.

Ainsi, pour chaque colonne $x_0 \in [\![1, N_c - 1]\!]$, on sélectionne dans un premier temps le segment $[\![a_1, b_1]\!]$ ayant obtenu le meilleur score, auquel on associe la moyenne $\overline{X}_{a_1,b_1}$ et le score γ_{a_1,b_1}. On recherche ensuite, s'il en existe, un second segment $[\![a_2, b_2]\!]$ disjoint du premier, c'est-à-dire tel que $[\![a_2, b_2]\!] \cap [\![a_1, b_1]\!] = \emptyset$, qui maximise le score parmi les segments restants. Ce processus est répété jusqu'à obtenir au plus n_0 segments disjoints, avec n_0 fixé par exemple à 4, ce qui correspond à la considération d'au plus quatre segments homogènes d'intérêt dans le signal colonnaire. Il est possible que moins de segments soient détectés par cette méthode.

D'après (4.93), pour chaque colonne x, on dispose ainsi de $n_x \leq n_0$ candidats pour estimer le gradient du bruit colonnaire, noté $\nabla_x \widehat{C}(x)$, à savoir

$$\nabla_x \widehat{C}(x) \in \left\{ \overline{X}_{a_1,b_1}, \ldots, \overline{X}_{a_{n_x},b_{n_x}} \right\}.$$

La figure 4.18 illustre le résultat de la méthode de détection des segments homogènes développée précédemment. La figure 4.19 montre la correction obtenue lorsque, pour chaque colonne, on sélectionne le segment ayant obtenu le meilleur score. Cette correction s'avère cependant insatisfaisante : une structure verticale présente dans la scène peut être interprétée à tort comme une zone uniforme, ce qui introduit des discontinuités

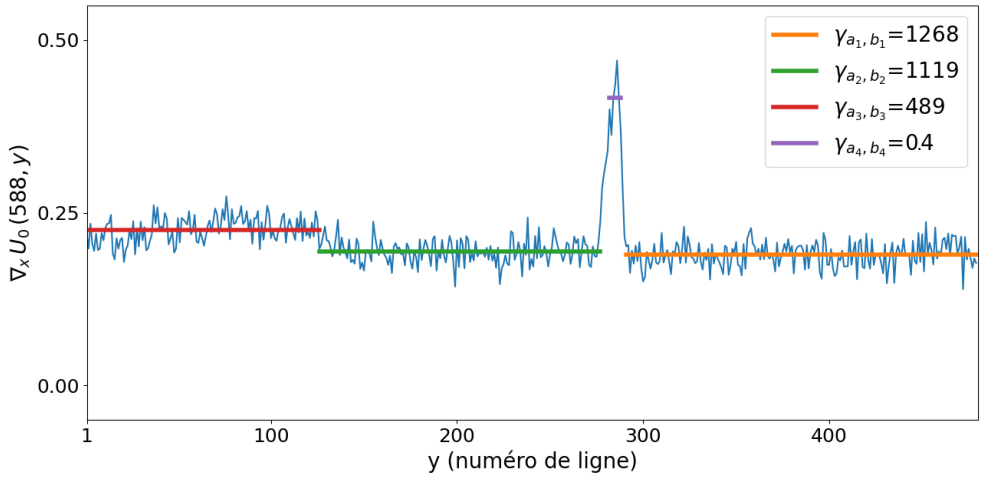


FIGURE 4.18 – **Détection de segments sur un signal.** La méthode de détection est appliquée au gradient horizontal de l'image dégradée U_0, ici pour la colonne $x_0 = 588$. Quatre segments disjoints sont détectés automatiquement, chacun étant caractérisé par la moyenne locale du signal, utilisée comme candidat possible pour estimer le gradient de correction des *stripes* $\nabla_x \widehat{C}(x_0)$, et par un score $\gamma_{a,b}$. Ce score dépend à la fois de la p-valeur du test du χ^2, qui quantifie l'homogénéité statistique du segment, et de sa longueur, de sorte à privilégier les zones à faible variance et suffisamment étendues. Les deux premiers segments (en vert et orange) présentent des scores élevés et proches, ce qui suggère qu'ils sont de bons candidats pour représenter une zone homogène. Leur moyenne est également relativement similaire, renforçant cette impression. Le troisième segment (en rouge), en revanche, présente un score plus modéré. À l'observation du signal, on constate que les oscillations ne sont pas stables sur l'ensemble de l'intervalle : le signal oscille d'abord autour d'une certaine valeur, puis autour d'une autre vers la fin du segment. Cette rupture interne remet en question sa pertinence pour représenter une zone uniforme. Enfin, le quatrième segment (en violet), situé sur un pic du signal, semble correspondre à un contour de l'image réelle. Son score très faible indique clairement qu'il constitue un mauvais candidat pour l'estimation du gradient colonnaire. Ici c'est le segment orange est celui qui se rapproche le plus de la vérité.

horizontales qui dégradent la correction. Le candidat avec le meilleur score n'est donc pas toujours le plus pertinent. Pour éviter ces corrections erronées, nous allons maintenant voir comment combiner la détection de ces segments homogènes avec un critère de régularité relatif à l'image reconstruite.

2.4.2 Correction colonnaire par sélection de segments homogènes via un problème variationnel

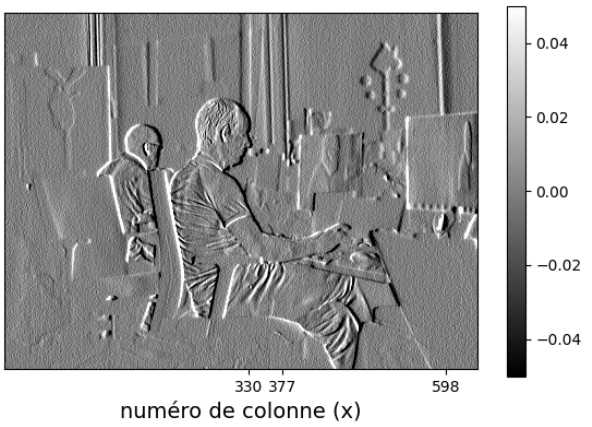
Tout d'abord, nous allons définir la dérivée seconde horizontale d'une image u en un pixel (x, y)

$$\nabla_{xx} u(x,y) = \begin{cases} u(x+2,y) + u(x,y) - 2\,u(x+1,y) & \text{si } 1 \le x \le N_c - 2 \text{ et } 1 \le y \le N_l \\ 0 & \text{si } x \ge N_c - 1 \text{ et } 1 \le y \le N_l. \end{cases} \tag{4.98}$$

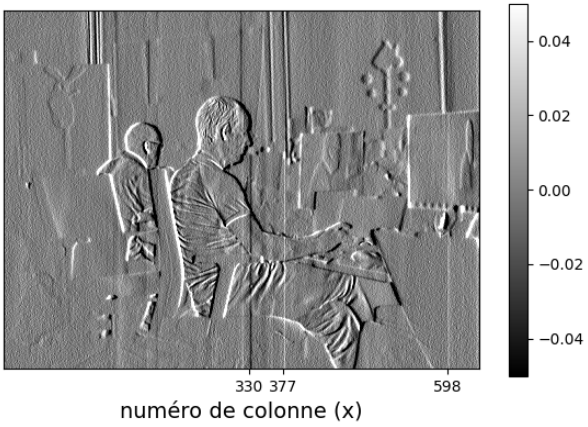
L'objectif est d'estimer $\nabla_x C(x)$, le gradient horizontal du bruit colonnaire pour chaque



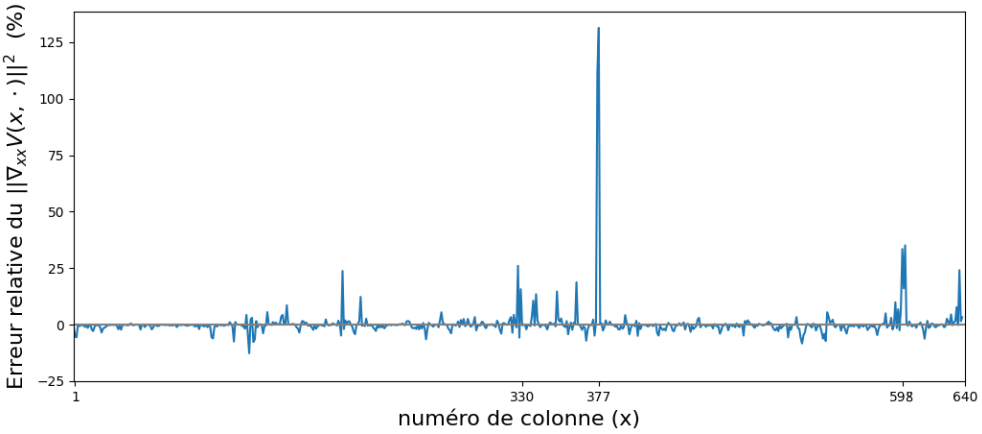
FIGURE 4.19 – **Correction appliquée en sélectionnant pour chaque colonne le segment réalisant le meilleur score.** L'estimation du gradient du bruit colonnaire $\nabla_x \widehat{C}$ a été réalisée colonne par colonne, en sélectionnant systématiquement le segment présentant le meilleur score γ. L'image corrigée présente un voile horizontal irrégulier, avec des discontinuités nettes et marquées dans la direction horizontale, notamment autour des colonnes 330, 377 et 598. Ces artefacts sont causés par la sélection de segments inadaptés, correspondant à des structures verticales ou à des zones non homogènes. L'erreur se propage horizontalement lors de l'intégration du gradient estimé. En conséquence, la correction provoque des décalages et des variations locales de l'intensité radiométrique : deux zones de la scène, qui ont la même irradiance dans la réalité, apparaissent avec des niveaux de gris différents après correction, créant ainsi des contrastes artificiels. Cette perte de cohérence radiométrique constitue un défaut majeur. Ce résultat illustre que le score local γ, pris isolément, est insuffisant pour garantir une correction cohérente : il peut privilégier des structures verticales, entraînant des irrégularités visibles dans l'image corrigée.



(a) Gradient horizontal vérité terrain



(b) Gradient horizontal meilleurs scores



(c) Différences locales sur la dérivée seconde horizontale

FIGURE 4.20 – **Analyse du gradient horizontal de la correction par sélection du meilleur score.** (a) Le gradient horizontal de l'image vérité terrain $\nabla_x V_0$ présente des discontinuités locales clairement liées aux contours réels des objets dans la scène.(b) En comparaison, le gradient horizontal $\nabla_x V$ de l'image corrigée, obtenu en sélectionnant pour chaque colonne le candidat homogène ayant obtenu le meilleur score local γ, présente des discontinuités verticales marquées. Celles-ci se manifestent par des lignes verticales continues dont l'intensité diffère de celle des colonnes adjacentes, donnant l'impression de fils se détachant du fond, en particulier autour des colonnes 330, 377 et 598. Cette sélection peut engendrer des erreurs de détection et donc des artefacts locaux. Par ailleurs, certains contours verticaux d'objets, notamment en haut entre les colonnes 330 et 377, y apparaissent moins nets ou moins marqués. (c) L'erreur relative en pourcentage entre les normes euclidiennes au carré des dérivées secondes horizontales, $\nabla_{xx}V$ (voir définition (4.98)) pour l'image corrigée et $\nabla_{xx}V_0$ pour la vérité terrain, en fonction de la colonne x, est définie par

$$\text{Erreur}(x) = \frac{\sum_{y=1}^{N_l} \left(\nabla_{xx}V(x,y)\right)^2 - \sum_{y=1}^{N_l} \left(\nabla_{xx}V_0(x,y)\right)^2}{\sum_{y=1}^{N_l} \left(\nabla_{xx}V_0(x,y)\right)^2} \times 100.$$

La courbe révèle des pics d'environ 25% aux colonnes 330 et 598, ainsi qu'un pic de 130% à la colonne 377, indiquant que l'image corrigée présente une dérivée seconde horizontale localement bien plus élevée que la vérité terrain. Ce phénomène traduit une instabilité locale du gradient horizontal, liée à la sélection de mauvais candidats. Ce constat souligne que le score seul ne suffit pas à guider la correction, justifiant l'introduction d'une pénalisation explicite sur le carré de $\nabla_{xx}V$ pour favoriser une correction plus régulière horizontalement.

colonne x, à partir des n_x segments homogènes détectés. Pour le $i^{\text{ème}}$ segment, avec $i \in [\![1, n_x]\!]$, $\overline{X}_i^{(x)}$ constitue une estimation possible de $\nabla_x C(x)$ et $\gamma_i^{(x)}$ correspond à son score d'homogénéité. La figure 4.20 montre le gradient horizontal de l'image corrigée $V = U_0 - \widehat{C}$ lorsque, pour chaque colonne, on sélectionne le candidat ayant obtenu le meilleur score. On observe que ce gradient présente des discontinuités horizontales d'une colonne à l'autre, traduisant une instabilité latérale. Pour pallier ce problème, il pourrait être pertinent d'introduire une pénalisation du carré de la norme euclidienne de la dérivée seconde horizontale $\nabla_{xx} V$, afin de limiter les discontinuités et de rendre le gradient horizontal de l'image corrigée le plus lisse possible. Cette approche motive ensuite la formalisation du modèle permettant de combiner la fidélité aux segments détectés et la régularité horizontale.

Pour formaliser, on définit une sélection d'indices

$$I = \Big(i(1), \dots, i(N_c - 1)\Big), \quad I \in \mathbb{I},$$

où $\mathbb{I} = \prod_{x=1}^{N_c-1} [\![1, n_x]\!]$ est l'ensemble de toutes les sélections possibles. À une sélection I, on associe la correction des *stripes* $C^{(I)}$ définie par

$$\begin{cases} C^{(I)}(1) = 0, \\ C^{(I)}(x+1) = C^{(I)}(x) + \overline{X}_{i(x)}^{(x)}, \quad 1 \leq x \leq N_c - 1, \end{cases}$$

et l'image corrigée

$$V^{(I)} = U_0 - C^{(I)}. \tag{4.99}$$

Le problème variationnel consiste à déterminer la sélection optimale solution de

$$\widehat{I} \in \operatorname*{argmin}_{I \in \mathbb{I}} \left[- \sum_{x=1}^{N_c-1} \gamma_{i(x)}^{(x)} + \lambda \sum_{x=1}^{N_c-2} \sum_{y=1}^{N_l} \Big(\nabla_{xx} V^{(I)}(x, y) \Big)^2 \right]. \tag{4.100}$$

Cette formulation favorise les segments candidats présentant un bon score d'homogénéité, sélectionnés selon la procédure décrite précédemment, et impose une régularité horizontale du gradient de l'image corrigée, mesurée par $\nabla_{xx} V^{(I)}$. La pénalisation de ce terme limite les variations brusques entre colonnes, de sorte à obtenir un gradient horizontal plus lisse. Le réglage du paramètre λ présente un comportement non linéaire : sur un large intervalle de valeurs, la solution demeure inchangée, puis une transition brutale apparaît au-delà d'un certain seuil, conduisant à une correction sensiblement différente. Ce comportement par sauts découle de la nature discrète du problème : la sélection des segments homogènes reste stable tant que la pondération du terme de régularité ne modifie pas l'équilibre global. Dans la pratique, pour des images normalisées dans $[0, 1]$, un réglage de λ autour de 10^3 permet d'obtenir une solution stable, correspondant au régime où le compromis entre fidélité aux segments détectés et régularité horizontale est atteint.

La résolution du problème variationnel (4.100) peut se faire efficacement par programmation dynamique. L'algorithme parcourt les colonnes de gauche à droite en mémorisant pour chaque candidat le meilleur coût cumulatif ainsi que l'antécédent permettant de reconstruire la solution optimale. Cette approche garantit une solution optimale globale avec une complexité linéaire en le nombre de colonnes et quadratique en le nombre maximal de candidats par colonne. Afin de formaliser cette démarche nous introduisons les

notions d'énergie partielle, de fonction de coût entre deux colonnes voisines, de fonction valeur partielle et de fonction d'antécédence qui structurent la construction progressive de la solution.

On définit l'énergie partielle pour $x \in [\![2, N_c - 1]\!]$ par

$$E\Big(i(1), \ldots, i(x)\Big) = -\gamma^{(1)}_{i(1)} - \sum_{s=2}^{x} \gamma^{(s)}_{i(s)} + \lambda \sum_{s=1}^{x-1} \sum_{y=1}^{N_l} \Big[(\nabla_x U_0(s+1, y) - \overline{X}^{(s+1)}_{i(s+1)}) - (\nabla_x U_0(s, y) - \overline{X}^{(s)}_{i(s)}) \Big]^2. \tag{4.101}$$

avec $E(i(1)) = -\gamma^{(1)}_{i(1)}$ pour la première colonne $x = 1$ et pour tout $s \in [\![1, x]\!]$, $i(s) \in [\![1, n_s]\!]$, où n_s est le nombre de candidats possibles dans la colonne s.

On remarque que le problème variationnel (4.100) s'écrit de manière équivalente comme la minimisation de l'énergie partielle associée à la dernière colonne $N_c - 1$

$$\widehat{I} \in \underset{I=(i(1),\ldots,i(N_c-1))}{\operatorname{argmin}} E\Big(i(1), \ldots, i(N_c - 1)\Big). \tag{4.102}$$

Pour $x \in [\![1, N_c - 2]\!]$ on définit la fonction de coût entre deux candidats consécutifs $i(x) \in [\![1, n_x]\!]$ et $i(x+1) \in [\![1, n_{x+1}]\!]$ comme la contribution supplémentaire apportée par la colonne $x + 1$ à l'énergie partielle

$$G_x(i(x), i(x+1)) = -\gamma^{(x+1)}_{i(x+1)} + \lambda \sum_{y=1}^{N_l} \Big[(\nabla_x U_0(x+1, y) - \overline{X}^{(x+1)}_{i(x+1)}) - (\nabla_x U_0(x, y) - \overline{X}^{(x)}_{i(x)}) \Big]^2. \tag{4.103}$$

On définit la fonction valeur partielle $J_x(i)$ comme le coût minimal permettant d'atteindre la colonne $x \in [\![1, N_c - 1]\!]$ en choisissant le candidat $i \in [\![1, n_x]\!]$

$$J_x(i) = \min_{(i(1),\ldots,i(x-1)) \in \prod_{w=1}^{x-1} [\![1, n_w]\!]} E\Big(i(1), \ldots, i(x-1), i\Big). \tag{4.104}$$

Ainsi, $J_x(i)$ représente le coût cumulé optimal associé à une trajectoire de sélection se terminant en colonne x par le candidat i.

La fonction valeur partielle $J_x(i)$ satisfait une relation récursive. En effet, en développant l'ajout du dernier terme dans la définition (4.104), on obtient $\forall x \in [\![2, N_c - 1]\!]$, $i \in [\![1, n_x]\!]$

$$\begin{aligned} J_x(i) &= \min_{i(1),\ldots,i(x-1)} E\Big(i(1), \ldots, i(x-1), i\Big) \\ &= \min_{i(1),\ldots,i(x-1)} \Big[E\Big(i(1), \ldots, i(x-2), i(x-1)\Big) + G_x\Big(i(x-1), i\Big) \Big] \\ &= \min_{j \in [\![1, n_{x-1}]\!]} \Big[J_{x-1}(j) + G_x(j, i) \Big]. \end{aligned}$$

On reconnaît alors la relation récursive suivante, appelée équation de Bellman

$$\forall x \in [\![2, N_c - 1]\!],\ \forall i \in [\![1, n_x]\!], \quad J_x(i) = \min_{j \in [\![1, n_{x-1}]\!]} \Big[J_{x-1}(j) + G_x(j, i) \Big]. \tag{4.105}$$

En parallèle, on définit la fonction d'antécédence optimale

$$\forall x \in [\![2, N_c - 1]\!],\ \forall i \in [\![1, n_x]\!], \quad H_x(i) \in \operatorname*{argmin}_{j \in [\![1, n_{x-1}]\!]} \left[J_{x-1}(j) + G_x(j, i) \right], \tag{4.106}$$

qui mémorise, pour chaque position (x, i) (i.e. le $i^{\text{ème}}$ candidat de la colonne x), l'indice j de la colonne précédente ayant permis d'atteindre le coût minimal [14].

Le processus commence à $x = 1$, avec l'initialisation

$$J_1(i) = -\gamma_i^{(1)}, \quad \text{pour tout } i \in [\![1, n_1]\!].$$

Ensuite, on calcule récursivement toutes les valeurs $J_x(i)$ et $H_x(i)$ pour $x = 2$ jusqu'à $x = N_c - 1$, en appliquant les deux équations (4.105) et (4.106) (cf. figure 4.21).

Une fois ces valeurs obtenues, on détermine l'indice final optimal

$$\hat{i}(N_c - 1) \in \operatorname*{argmin}_i J_{N_c - 1}(i),$$

puis on reconstruit la chaîne optimale par rétropropagation

$$\hat{i}(x - 1) = H_x\big(\hat{i}(x)\big), \quad \text{pour } x = N_c - 1, N_c - 2, \ldots, 2.$$

La solution de (4.100) s'écrit alors

$$\widehat{I} = \big(\widehat{i}(1), \ldots, \widehat{i}(N_c - 1)\big),$$

et l'image corrigée $V^{(\widehat{I})}$ s'en déduit conformément aux équations (2.4.2) et (4.99).

Cette stratégie permet de résoudre efficacement le problème global d'optimisation avec une complexité algorithmique en $\mathscr{O}(N_c \cdot n^2)$, où

$$n = \max_{1 \le x \le N_c - 1} n_x.$$

2.4.3 Évaluation visuelle des corrections obtenues

La méthode détecte des segments homogènes censés représenter des zones d'irradiance uniforme. Une fois un segment choisi, l'estimation du gradient horizontal du bruit colonnaire à partir de ce segment reste indépendante de l'intensité du bruit et se déplace simplement si celui-ci est plus fort. La figure 4.22 illustre une correction globalement correcte, bien que subsistent quelques artefacts horizontaux de basse ou moyenne fréquence. Ces artefacts proviennent de petites erreurs locales sur l'estimation de $\nabla_x C$ qui, après intégration pour reconstruire C, se traduisent par des variations douces et étendues le long des colonnes. La pénalisation de l'irrégularité de la dérivée seconde horizontale ($\|\nabla_{xx} V\|^2$) permet d'éviter la sélection de segments correspondant à des structures verticales, mais ne supprime pas entièrement ces petites erreurs résiduelles.

14. Dans le cas où plusieurs indices minimisent exactement la même quantité, on en choisit simplement un arbitrairement. Toutefois cette situation ne s'est jamais produite sur les expériences que nous avons réalisées.

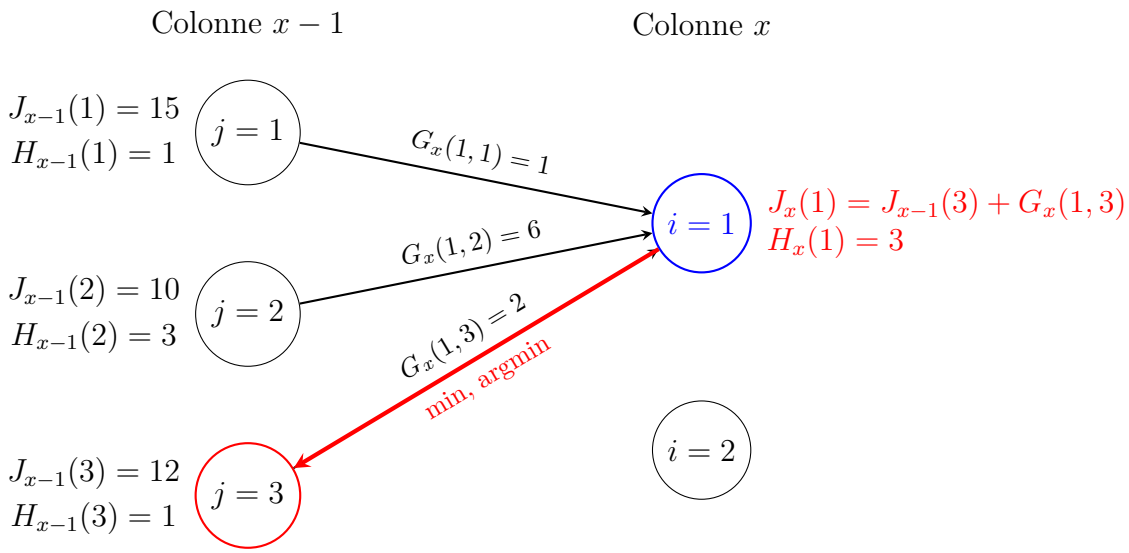


FIGURE 4.21 – **Calcul des fonctions valeur partielle et d'antécédence optimale par l'équation de Bellman.** On suppose que les valeurs partielles et les antécédences optimales sont connues à l'instant $x-1$. La colonne x contient deux candidats, notés $i = 1$ (cercle bleu) et $i = 2$. Ici, on détaille le calcul pour $i = 1$. Pour déterminer les fonctions valeur partielle $J_x(1)$ et d'antécédence optimale $H_x(1)$ définies par les équations de Bellman (4.105) et (4.106), on compare trois trajectoires possibles issues des candidats précédents $j = 1, 2, 3$. Chacune a un coût calculé, correspondant au passage de j vers $i = 1$. Les deux premières trajectoires ont un coût de 16, tandis que la troisième est plus avantageuse avec un coût de 14. Cette dernière est donc sélectionnée, ce qui détermine la meilleure valeur partielle $J_x(1) = 14$ et l'antécédence optimale $H_x(1) = 3$, indiquée par une flèche rouge. La procédure est identique pour $i = 2$, non illustrée ici.

2.5 Comparaison des méthodes de correction proposées avec une méthode existante

Nous comparons nos méthodes de correction du bruit colonnaire à la méthode de référence MIRE [32]. La figure 4.23 illustre cette comparaison sur une image contenant des structures verticales naturelles. Dans cet exemple, la méthode MIRE pose des problèmes, car l'intensité de la perturbation colonnaire est relativement élevée par rapport au contraste de la scène, entraînant des artefacts horizontaux de moyenne fréquence et une restitution imparfaite des détails fins. Les trois méthodes que nous proposons, à savoir TVx, la sélection du candidat au meilleur score et la résolution du problème variationnel par programmation dynamique pour détecter un segment homogène et obtenir une image avec des discontinuités horizontales minimales, fournissent des corrections qui, pour chacune, ne dépendent pas de l'intensité du bruit affectant chaque colonne. Sur cet exemple, la méthode variationnelle combinant score local et régularité horizontale donne des résultats légèrement supérieurs. Sur les exemples précédemment présentés (cf. figures 4.15c, 4.19 et 4.22), elle offre globalement les meilleurs résultats.

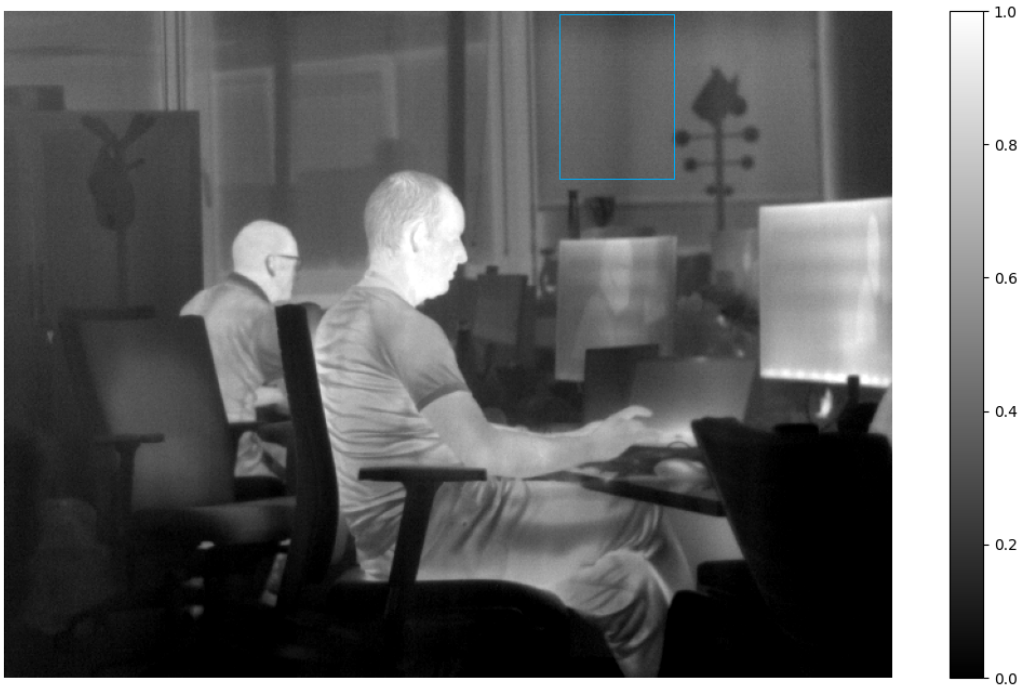
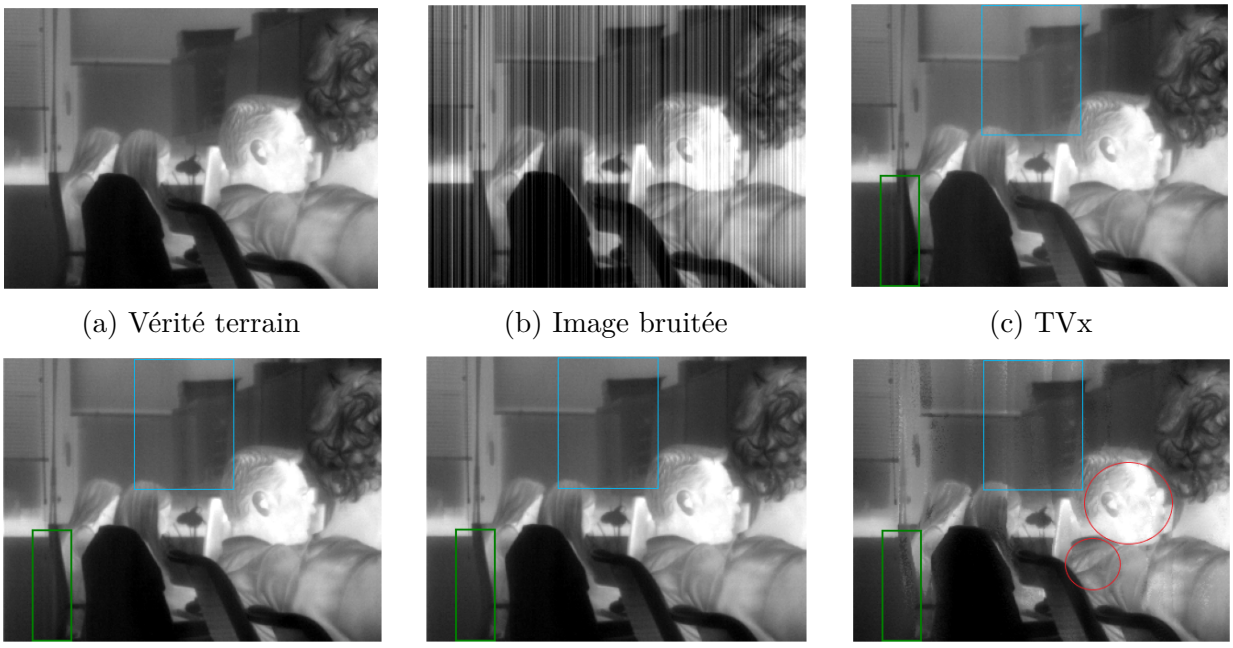


FIGURE 4.22 – **Correction obtenue à partir d'un problème variationnel favorisant la détection de segments homogènes et visant à limiter les discontinuités horizontales.** La méthode proposée prolonge l'approche par sélection de segments homogènes ayant obtenu le meilleur score de représentativité (figure 4.19). Elle corrige les limites de cette dernière, en particulier la présence de discontinuités horizontales marquées, grâce à l'ajout d'un terme de pénalisation fondé sur le carré de la norme euclidienne de la dérivée seconde horizontale de l'image corrigée. La correction obtenue présente une homogénéité globale satisfaisante et une amélioration nette par rapport aux résultats issus de la simple sélection par score. On observe toutefois quelques artefacts résiduels de basse ou moyenne fréquence dans la direction horizontale, notamment dans la zone encadrée en bleu correspondant à une région d'irradiance uniforme. Ces défauts proviennent probablement de la sélection ponctuelle de candidats légèrement inexacts. L'intégration du gradient estimé $\nabla_x \widehat{C}$, qui correspond à la correction du bruit colonnaire, accumule ces petites erreurs et engendre ainsi des artefacts de basse fréquence horizontale dans l'image corrigée.



(a) Vérité terrain (b) Image bruitée (c) TVx

(d) Meilleur score (e) Modèle variationnel (f) MIRE

FIGURE 4.23 – **Comparaison de différentes méthodes de correction du bruit colonnaire.** Données issues de la caméra *Smart IR320*. (a) La vérité terrain correspond à l'image originale de la scène avec ses structures verticales naturelles. (b) L'image bruitée est obtenue en ajoutant à la vérité terrain un bruit purement colonnaire ou *stripes*, indépendant selon les colonnes et distribué selon une loi normale centrée. (c) La correction par le modèle TVx supprime ces stripes mais introduit des artefacts de reconstruction. Dans le rectangle vert, uniforme dans la vérité terrain (a), on distingue une ligne verticale claire. Cette erreur provient d'une structure verticale située au-dessus, qui biaise l'estimation de la correction. Dans le rectangle bleu, le placard apparaît moins net que dans (a), la partie haute ayant perdu de son homogénéité. (d) La correction obtenue en sélectionnant systématiquement le segment au meilleur score γ donne une image légèrement meilleure que TVx (c), mais conserve des défauts similaires. Dans le rectangle vert, une ligne verticale claire reste visible. (e) La correction obtenue à partir du modèle variationnel favorisant la régularité horizontale du gradient se rapproche de la vérité terrain. Dans le rectangle bleu, on distingue clairement le placard, mais la partie haute présente une homogénéité moindre, comme un léger voile horizontal. Dans le rectangle vert, la zone reste relativement uniforme, sans artefacts verticaux marqués. (f) La méthode de référence MIRE [32], utilisant un lissage gaussien des histogrammes colonnaires avec un écart type spatial de 6, fournit une correction globalement dégradée. Un lissage horizontal aussi important était nécessaire en raison de l'intensité du bruit colonnaire dans l'image bruitée (b). Le bas du rectangle vert reste relativement uniforme, bien que quelques artefacts apparaissent à proximité du siège à droite. Le rectangle bleu présente des ondulations horizontales de moyenne fréquence. Certains détails fins sont mal restitués, comme ceux visibles dans les cercles rouges : le visage de l'homme et le pli du vêtement diffèrent légèrement de la réalité (a). Dans l'ensemble, les corrections (c–e) suppriment efficacement les *stripes* et restent insensibles à l'intensité de la perturbation colonnaire, contrairement à la méthode MIRE (f) qui introduit des artefacts plus visibles pour des intensités élevées et restitue moins fidèlement les détails fins.

2.6 Conclusion sur le débruitage colonnaire

Nous avons proposé trois approches de correction du bruit colonnaire, toutes fondées sur l'analyse du gradient horizontal de l'image corrigée. Quelle que soit l'intensité du bruit affectant chaque colonne, les corrections obtenues par ces modèles restent identiques.

La première méthode, notée TVx, repose sur un principe simple : elle impose que, pour chaque colonne, la médiane du gradient horizontal de l'image reconstruite soit nulle. Cette contrainte permet une réduction efficace du bruit colonnaire, mais elle peut générer des artefacts basse ou moyenne fréquence, ainsi que des discontinuités horizontales, en particulier dans les zones où la scène comporte des structures verticales marquées. Dans ces situations, la contrainte de médiane nulle devient inadaptée, car elle interfère avec les gradients réels de la scène.

Pour pallier cette limite, la seconde méthode remplace la contrainte de médiane nulle par une hypothèse locale de constance du signal. Elle repose sur l'idée que, si une portion de colonne correspond à une zone homogène de la scène, alors le gradient horizontal de l'image corrigée doit y être nul. Pour chaque colonne, on détecte des segments homogènes dans le gradient horizontal de l'image bruitée, auxquels un score de confiance est attribué en fonction de leur taille et de la variabilité du niveau radiométrique sur le segment. La correction obtenue en sélectionnant, pour chaque colonne, le segment présentant le meilleur score peut toutefois correspondre à une structure verticale réelle de la scène plutôt qu'à une zone homogène, ce qui introduit des discontinuités horizontales parfois marquées et dégrade la qualité de la correction.

Pour corriger ce second défaut, la troisième méthode prolonge la précédente en imposant une cohérence spatiale entre colonnes adjacentes. Un problème variationnel est alors formulé pour sélectionner simultanément des segments ayant obtenu un bon score tout en limitant les variations brusques du gradient horizontal entre colonnes voisines. La résolution est assurée par programmation dynamique, permettant d'identifier la combinaison optimale en équilibrant la confiance locale et la régularité globale du gradient reconstruit. Cette approche produit une correction généralement satisfaisante, bien que des artefacts latéraux basse ou moyenne fréquence persistent, dus à l'intégration du gradient corrigé qui tend à accumuler des erreurs locales.

Conclusion

Dans cette thèse, nous nous sommes intéressés aux caméras thermiques non-refroidies sans obturateur. Ces capteurs présentent un intérêt industriel majeur grâce à leur faible coût et à leur taille réduite par rapport aux caméras refroidies, et ouvrent la voie à des applications telles que la lutte contre les incendies, l'équipement de drones militaires ou la conduite assistée. Les images brutes issues de ces caméras sont fortement dégradées par un bruit fixe non uniforme spatialement, ce qui rend nécessaire le développement de modèles de correction. Le défi principal réside dans la conception de modèles capables de restituer une qualité radiométrique suffisante, tout en respectant les contraintes d'embarquabilité et en assurant une correction robuste dans toutes les conditions d'utilisation.

La correction des images repose généralement sur un modèle affine dont les composantes doivent être estimées. Les approches *calibration-based* s'appuient sur des données de calibration et combinent fiabilité et faible coût de calcul, ce qui les rend adaptées aux calculs embarqués, tandis que les approches *scene-based* exploitent les informations issues de la scène observée. Ces composantes varient selon l'environnement thermique, phénomène connu sous le nom de *thermal drift*. Dans le modèle paramétrique dominant utilisé par l'entreprise *Lynred*, le « *shutterless polynomial* », cette variabilité est supposée dépendre uniquement de la TFPA, la température mesurée au niveau du plan focal, tandis que les autres facteurs environnementaux sont ignorés. Ce modèle nécessite donc des calibrations réalisées dans des conditions thermiques fortement contrôlées afin de limiter l'influence des autres composantes de l'environnement et demeure peu robuste lorsque la TFPA ou l'environnement diffèrent des conditions explorées lors de la calibration. Cette phase de calibration est longue et coûteuse, constituant un enjeu industriel majeur et motivant la recherche de méthodes permettant de réduire ces contraintes.

L'objectif de cette thèse a été de développer des modèles mathématiques de correction capables de limiter la dépendance aux calibrations contraignantes, tout en assurant une robustesse face aux variations de l'environnement thermique et une qualité radiométrique satisfaisante. L'adaptabilité de ces modèles à des systèmes embarqués, impliquant un faible usage de mémoire et de puissance de calcul, n'a pas constitué un objectif central de la thèse, mais sera évoquée afin d'évaluer leur faisabilité d'intégration dans de tels contextes. Il s'agit à présent de présenter les principales contributions de cette thèse, en illustrer les implications et en mettre en perspective l'intérêt pour les applications industrielles et la réduction des contraintes expérimentales.

Contribution 1 : Modélisation polynomiale des images brutes de calibration

La première contribution porte sur la modélisation polynomiale des images brutes issues des acquisitions de calibration. Comparé au modèle « *shutterless polynomial* », le modèle proposé est très proche. Chaque composante du modèle affine est représentée par un polynôme en fonction de la TFPA pour chaque pixel. Cette approche fournit des résultats très similaires avec un gain de précision très modeste. Les deux méthodes reposent sur des cartes de coefficients pour les composantes du modèle affine et sont équivalentes en termes de mémoire et de puissance de calcul, la principale différence résidant dans le traitement en amont des données.

Par ailleurs, une méthode a été développée pour déterminer le degré optimal d'un polynôme avant de procéder à sa régression sur les données. Cette approche, applicable au-delà du domaine de l'imagerie thermique, permet de limiter significativement le risque d'*overfitting*, comme le confirment les résultats obtenus sur plusieurs jeux de données.

L'analyse des erreurs de fit révèle les limites d'un modèle paramétrique utilisant uniquement la TFPA. Même lorsque la calibration est réalisée dans une étuve où la température et l'humidité sont stabilisées dans le temps et homogènes spatialement, des variations parasites persistent d'une acquisition à l'autre. Si l'on construit un modèle affine en considérant uniquement la TFPA à partir de ces données, certaines fluctuations réduisent la précision de l'estimation des composantes du modèle affine. Comme il n'est pas possible pour l'instant de concevoir un modèle multi-paramétrique intégrant tous les flux parasites difficiles à mesurer, ces observations montrent que la TFPA seule ne suffit pas à caractériser l'ensemble de l'environnement thermique et justifient le recours à des approches non paramétriques capables de prendre en compte ces flux parasites.

Contribution 2 : Analyse statistique des données de calibration

La deuxième contribution porte sur l'analyse statistique des données de calibration et s'articule en deux points principaux.

Le premier point concerne l'étude, par ACP, du comportement du capteur en fonction de la température de la scène T et de la TFPA τ. Cette analyse suggère qu'il est possible de définir un modèle direct, dans lequel la réponse du capteur est exprimée comme une fonction de ces deux paramètres. L'examen des surfaces associées aux axes principaux indique que ces composantes peuvent être approchées par des polynômes en (T, τ). Dans ce cas, le modèle *shutterless polynomial* apparaît comme un cas particulier de ce modèle direct. Plus généralement, ce modèle peut être considéré comme une extension du modèle affine paramétrique en TFPA, indépendamment du choix de la représentation des axes principaux par des fonctions polynomiales. Bien que ce modèle direct n'ait pas été développé en détail, il constitue une alternative plus complète au modèle affine pour caractériser la réponse du capteur en fonction de la scène et de la TFPA.

Le second point concerne l'analyse de l'espace des biais dans le cadre du modèle affine à gain constant et biais variable. Il montre que la famille de biais calculée à partir des données de calibration peut être approchée par un espace vectoriel. Cette approximation reste valable dès lors que la calibration couvre une plage de TFPA adaptée au domaine d'application de la caméra et comporte un nombre suffisant de points distincts, régulièrement espacés, par exemple d'environ 10 °C. En pratique, les éléments de cette famille

de biais présentent une forte corrélation spatiale.

Contribution 3 : Correction fondée sur une famille de biais

La troisième contribution présente un modèle de correction « mixte » non paramétrique, fondé sur une famille de biais construite à partir des données de calibration, et actuellement en cours de dépôt de brevet.

Les résultats obtenus montrent que les déclinaisons $\mathscr{H}^{\mathbf{1}}$ et TV offrent une robustesse supérieure face aux changements d'environnement thermique par rapport à la méthode de référence *shutterless polynomial*, ouvrant la voie à des procédures de calibration moins contraignantes. Dans le modèle *shutterless polynomial*, les composantes du modèle affine sont ajustées sur les données de calibration à l'aide de polynômes en fonction de la TFPA, limitant la dépendance environnementale à une seule variable. Pour que l'estimation des coefficients soit fiable, la calibration doit être réalisée dans un environnement thermiquement homogène afin de réduire la variabilité des flux parasites. Le modèle proposé s'affranchit de cette contrainte en s'appuyant sur une famille de biais construite à partir d'images de calibration acquises dans des environnements thermiques variés. Cette approche non paramétrique exploite la diversité des situations pour couvrir l'ensemble des variations possibles du capteur, y compris celles liées à des flux parasites difficiles à mesurer, offrant ainsi une correction plus robuste. En élargissant le spectre des conditions de calibration, on réduit le coût et la complexité des campagnes tout en améliorant la qualité de la correction, ouvrant la voie à des procédures plus rapides, moins coûteuses et mieux adaptées aux usages opérationnels.

La déclinaison $\mathscr{H}^{\mathbf{1}}$ de ce modèle semble particulièrement adaptée à un embarquement industriel. La réduction de dimension de la famille de biais, réalisée par décomposition en valeurs singulières, permet d'isoler l'information essentielle pour la correction en ne conservant que les composantes traduisant les principales variabilités des images de biais correcteurs, éliminant les éléments redondants ou peu significatifs. Cette démarche réduit l'utilisation des ressources lors du calcul du biais correcteur tout en conservant l'essentiel de l'information de calibration. La formule de correction reste directe et peu exigeante en puissance de calcul. Comparée à *shutterless polynomial*, cette approche n'impose pas de contrainte supplémentaire en termes de ressources, ce qui facilite son intégration dans des systèmes embarqués.

Contribution 4 : Débruitage colonnaire

La quatrième contribution concerne le débruitage des artefacts colonnaires (*stripes*), susceptibles d'apparaître après l'application d'un modèle fondé sur des données de calibration. Ce débruitage constitue un post-traitement destiné à corriger ces artefacts et deux approches ont été développées.

Le premier modèle, TVx, est un modèle variationnel relativement simple et constitue une étape préalable au second modèle, plus élaboré. Il élimine efficacement le bruit colonnaire, mais peut introduire des artefacts lorsque la scène contient des structures verticales, se traduisant par des discontinuités horizontales. La puissance de calcul et la mémoire nécessaires restent faibles, ce qui ouvre la voie à une application industrielle.

La seconde méthode détecte d'abord des segments homogènes dans le gradient horizontal de l'image bruitée et leur attribue un score de confiance. Un problème variationnel est ensuite résolu afin de reconstruire l'image en favorisant les segments les plus fiables tout en produisant un gradient horizontal lisse. Cette approche corrige efficacement le bruit colonnaire tout en limitant l'impact des structures verticales, mais sa complexité computationnelle reste trop élevée pour une intégration immédiate dans des systèmes embarqués.

Les méthodes proposées se sont concentrées sur le débruitage colonnaire d'une seule image. Contrairement au bruit structuré spatial affectant l'image brute, qui peut varier dans le temps en raison du *thermal drift*, le bruit colonnaire reste relativement constant sur une séquence de plusieurs minutes. Lorsqu'on souhaite corriger le bruit colonnaire sur une séquence comportant du mouvement, cette stationnarité peut être exploitée pour améliorer la correction. La diversité des textures et structures à travers les images successives permet d'obtenir plusieurs estimations du bruit colonnaire, chaque mouvement générant des variations spatiales distinctes. Ces estimations peuvent ensuite être combinées, par exemple par moyennage temporel, afin de raffiner la correction.

Dans l'ensemble, ces travaux soulignent l'intérêt des approches non paramétriques et « mixtes » ainsi que la diversité des environnements de calibration pour assurer une correction de qualité et une robustesse face aux variations d'environnement thermique.

Annexe : algorithme de Chambolle-Pock

Quelques définitions

On définit la transformée de Legendre-Fenchel de F par

$$\forall y \in Y, \quad F^*(y) = \sup_{p \in Y} \langle y, p \rangle - F(p), \tag{107}$$

où $F : Y \mapsto \mathbb{R} \cup \{+\infty\}$ est convexe, propre et semi-continue inférieurement (sci).

On définit l'opérateur proximal de f par

$$\forall v \in X, \quad \operatorname*{prox}_{\lambda f}(v) = \operatorname*{argmin}_{x \in X} \frac{1}{2}\|x - v\|^2 + \lambda f(x), \tag{108}$$

où $f : X \mapsto \mathbb{R} \cup \{+\infty\}$ est convexe, propre et sci, et $\lambda > 0$.

Problème et résolution par Chambolle-Pock

On cherche à résoudre le problème d'optimisation convexe suivant

$$\operatorname*{argmin}_{x \in X} F(K\,x) + G(x), \tag{109}$$

où

- X désigne l'espace primal et $x \in X$ la variable primale,
- Y désigne l'espace dual et $y \in Y$ la variable duale,
- $F : Y \mapsto \mathbb{R} \cup \{+\infty\}$ est convexe, propre et sci,
- $G : X \mapsto \mathbb{R} \cup \{+\infty\}$ est convexe, propre et sci,
- $K : X \mapsto Y$ est un opérateur linéaire continu, avec son adjoint noté K^*.

Le problème (109) peut se reformuler comme la recherche d'un point selle

$$\operatorname*{argmin}_{x \in X} \max_{y \in Y} \langle Kx, y \rangle + G(x) - F^*(y), \tag{110}$$

où F^* est la transformée de Fenchel–Legendre définie en (107).

D'après l'article « A first-order primal-dual algorithm for convex problems with applications to imaging » de Chambolle et Pock [28], la suite $\left(x^{(k)}\right)$ définie par le schéma primal-dual

$$\forall k \in \mathbb{N}, \quad \begin{cases} y^{(k+1)} = \underset{\sigma F^*}{\operatorname{prox}}\left(y^{(k)} + \sigma K\, \overline{x}^{(k)}\right), \\ x^{(k+1)} = \underset{\tau G}{\operatorname{prox}}\left(x^{(k)} - \tau K^*\, \overline{y}^{(k)}\right), \\ \overline{x}^{(k+1)} = x^{(k+1)} + \theta\left(x^{(k+1)} - x^{(k)}\right) \end{cases} \tag{111}$$

converge vers une solution $x^\star$ du problème (109) sous les conditions suivantes

$$\begin{cases} \sigma > 0, \quad \tau > 0, \\ \tau\sigma\|K\|^2 < 1, \\ \theta = 1. \end{cases} \tag{112}$$

En général, la solution $x^\star$ n'est pas nécessairement unique, puisque la fonction objectif $x \mapsto F(Kx) + G(x)$ peut être convexe sans être strictement convexe.

L'unicité est garantie si, par exemple, F est strictement convexe et K est injectif, ou si G est strictement convexe. Dans ce cas, la fonction totale est strictement convexe, ce qui assure une solution unique.

Premier cas : recherche d'un vecteur de coefficients

— $X = \mathbb{R}^N$, la variable primale est $\alpha \in \mathbb{R}^N$,
— $Y = \mathbb{R}^{2\Omega}$, la variable duale est $p \in \mathbb{R}^{2\Omega}$,
— $F : \mathbb{R}^{2\Omega} \mapsto \mathbb{R}$, définie par $F(p) = \|p\|_{1,2}$ selon (4.50),
— $K = \nabla\mathbf{B} : \mathbb{R}^N \mapsto \mathbb{R}^{2\Omega}$, $\nabla\mathbf{B}^*$ son opérateur adjoint,
— $G : \mathbb{R}^N \mapsto \mathbb{R}$, définie par $G(\alpha) = 0$.

On rappelle que la fonction indicatrice de l'ensemble $\mathbf{c} \subset E$ (où E désigne un espace vectoriel), notée $\delta_{\mathbf{c}} : E \mapsto \{0, +\infty\}$, est définie par

$$\forall x \in E, \quad \delta_{\mathbf{c}}(x) = \begin{cases} 0 & \text{si } x \in \mathbf{c}, \\ +\infty & \text{sinon.} \end{cases} \tag{113}$$

La transformée de Fenchel–Legendre F^* est la fonction indicatrice de la boule unité [15]

$$\mathbf{b} = \left\{ p = (p_1, \ldots, p_{2\,\mathrm{card}(\Omega)}) \in \mathbb{R}^{2\Omega} \,\middle|\, \forall i \in [\![1, \mathrm{card}(\Omega)]\!],\ \sqrt{p_i^2 + p_{i+\mathrm{card}(\Omega)}^2} \leq 1 \right\}, \tag{114}$$

c'est-à-dire :

$$F^*(p) = \delta_{\mathbf{b}}(p). \tag{115}$$

Le schéma devient alors :

$$\forall k \in \mathbb{N}, \quad \begin{cases} p^{(k+1)} = \underset{p \in \mathbb{R}^{2\Omega}}{\operatorname{argmin}} \left\| p - \left(p^{(k)} + \sigma\, \nabla\mathbf{B}\, \overline{\alpha}^{(k)}\right)\right\|_{2,2}^2 + \sigma\, \delta_{\mathbf{b}}(p), \\ \alpha^{(k+1)} = \alpha^{(k)} - \tau\, \nabla\mathbf{B}^*\, \overline{p}^{(k)}, \\ \overline{\alpha}^{(k+1)} = \alpha^{(k+1)} + \theta\left(\alpha^{(k+1)} - \alpha^{(k)}\right). \end{cases} \tag{116}$$

15. Formule utilisée dans l'article [28].

ou, sous forme projetée

$$\forall k \in \mathbb{N}, \quad \begin{cases} p^{(k+1)} = \underset{\mathbf{b}}{\mathrm{proj}} \left(p^{(k)} + \sigma \, \nabla \mathbf{B} \, \overline{\alpha}^{(k)} \right), \\ \alpha^{(k+1)} = \alpha^{(k)} - \tau \, \nabla \mathbf{B}^* \, \overline{p}^{(k)}, \\ \overline{\alpha}^{(k+1)} = \alpha^{(k+1)} + \theta \left(\alpha^{(k+1)} - \alpha^{(k)} \right). \end{cases} \tag{117}$$

Puisque la norme de l'opérateur $\nabla \mathbf{B}$ est égale à la racine carrée du rayon spectral $\rho(\nabla \mathbf{B}^* \nabla \mathbf{B})$, on a

$$\|\nabla \mathbf{B}\| = \sqrt{\max_{i \in [\![1,N]\!]} \lambda_i} = \sqrt{\rho(\nabla \mathbf{B}^* \nabla \mathbf{B})}, \tag{118}$$

où $\lambda_1, \ldots, \lambda_N$ désignent les valeurs propres de la matrice de Gram $\nabla \mathbf{B}^* \nabla \mathbf{B}$. On en déduit que la condition $\tau \, \sigma \, \rho(\nabla \mathbf{B}^* \nabla \mathbf{B}) < 1$ dans (112) garantit la convergence théorique du schéma (117).

Dans ce cas, on ne peut pas prouver la stricte convexité de la fonction objectif $F(\nabla \mathbf{B} \, \alpha) + G(\alpha)$, car G est nul et F est une norme ℓ^1 appliquée au gradient, donc tous deux non strictement convexes. L'unicité de la solution n'est donc pas garantie.

Deuxième cas : recherche d'une image

— $X = \mathbb{R}^\Omega$, la variable primale est $B \in \mathbb{R}^\Omega$,
— $Y = \mathbb{R}^\Omega \times \mathbb{R}^\Omega$, la variable duale est $p \in \mathbb{R}^\Omega \times \mathbb{R}^\Omega$,
— $F : \mathbb{R}^\Omega \times \mathbb{R}^\Omega \mapsto \mathbb{R}$ définie par $F(p) = \|p\|_{1,2}$ selon (4.48),
— $K = \nabla : \mathbb{R}^\Omega \mapsto \mathbb{R}^\Omega \times \mathbb{R}^\Omega$, opérateur gradient spatial, ∇^* son opérateur adjoint.
— $G : \mathbb{R}^\Omega \mapsto \mathbb{R}$ définie par $G(u) = \delta_{\mathrm{Vect}(B_1,\ldots,B_N)}(u)$ (fonction indicatrice de l'espace vectoriel engendré par les B_i).

Ici la boule unité $\mathbf{b}$ est défini par

$$\mathbf{b} = \left\{ p = (p_1, p_2) \in \mathbb{R}^\Omega \times \mathbb{R}^\Omega \;\middle|\; \forall (x,y) \in \Omega, \; \sqrt{p_1(x,y)^2 + p_2(x,y)^2} \leq 1 \right\}. \tag{119}$$

Puisque F^* est l'indicatrice de la boule $\mathbf{b}$, le schéma s'écrit

$$\forall k \in \mathbb{N}, \quad \begin{cases} p^{(k+1)} = \underset{p \in \mathbb{R}^\Omega \times \mathbb{R}^\Omega}{\mathrm{argmin}} \, \frac{1}{2} \left\| p - \left(p^{(k)} + \sigma \nabla \, \overline{B}^{(k)} \right) \right\|_{2,2}^2 + \sigma \, \delta_{\mathbf{b}}(p), \\ B^{(k+1)} = \underset{B \in \mathbb{R}^\Omega}{\mathrm{argmin}} \, \frac{1}{2} \left\| B - \left(B^{(k)} - \tau \nabla^* \, \overline{p}^{(k)} \right) \right\|_2^2 + \delta_{\mathrm{Vect}(B_1,\ldots,B_N)}(B), \\ \overline{B}^{(k+1)} = B^{(k+1)} + \theta \left(B^{(k+1)} - B^{(k)} \right). \end{cases} \tag{120}$$

ou sous forme projetée

$$\forall k \in \mathbb{N}, \quad \begin{cases} p^{(k+1)} = \underset{\mathbf{b}}{\mathrm{proj}} \left(p^{(k)} + \sigma \nabla \, \overline{B}^{(k)} \right), \\ B^{(k+1)} = \underset{\mathrm{Vect}(B_1,\ldots,B_N)}{\mathrm{proj}} \left(B^{(k)} - \tau \nabla^* \, \overline{p}^{(k)} \right), \\ \overline{B}^{(k+1)} = B^{(k+1)} + \theta \left(B^{(k+1)} - B^{(k)} \right). \end{cases} \tag{121}$$

Puisque

$$|||\nabla||| \leq 2\sqrt{2}, \tag{122}$$

on en déduit que la condition $8\,\tau\sigma < 1$ dans (112) garantit la convergence théorique du schéma (121).

Dans ce cas, on ne peut garantir l'unicité de la solution car la fonction objectif $F(\nabla u) + G(u)$ est convexe mais pas strictement convexe. En effet F est une norme ℓ^1 appliquée au gradient, donc non strictement convexe, et G est la fonction indicatrice d'un sous-espace vectoriel, donc également non strictement convexe. L'unicité n'est donc pas assurée.

Bibliographie

[1] D. L. Perry and E. L. Dereniak, "Linear theory of nonuniformity correction in infrared staring sensors," *Optical Engineering*, vol. 32, no. 8, pp. 1854–1859, 1993.

[2] D. L. Perry, "Linear theory of nonuniformity correction in platinum silicide focal plane arrays," in *Infrared Technology XVIII*, vol. 1762, pp. 60–69, SPIE, 1993.

[3] G. Bieszczad, T. Orżanowski, T. Sosnowski, and M. Kastek, "Method of detectors offset correction in thermovision camera with uncooled microbolometric focal plane array," in *Electro-Optical and Infrared Systems : Technology and Applications VI*, vol. 7481, pp. 200–207, SPIE, 2009.

[4] S. Chang and Z. Li, "Single-reference-based solution for two-point nonuniformity correction of infrared focal plane arrays," *Infrared Physics & Technology*, vol. 101, pp. 96–104, 2019.

[5] D. Zhang, H. Sun, D. Wang, J. Liu, and C. Chen, "Modified two-point correction method for wide-spectrum lwir detection system," *Sensors*, vol. 23, no. 4, p. 2054, 2023.

[6] M. Sheng, J. Xie, and Z. Fu, "Calibration-based nuc method in real-time based on irfpa," *Physics Procedia*, vol. 22, pp. 372–380, 2011.

[7] P. W. Nugent and N. J. Pust, "Correcting for focal-plane-array temperature dependence in microbolometer infrared cameras lacking thermal stabilization," 2013.

[8] A. Tempelhahn, H. Budzier, V. Krause, and G. Gerlach, "Improving the shutterless compensation method for tec-less microbolometer-based infrared cameras," in *Infrared Technology and Applications XLI*, vol. 9451, pp. 378–387, SPIE, 2015.

[9] A. Tempelhahn, H. Budzier, V. Krause, and G. Gerlach, "Shutter-less calibration of uncooled infrared cameras," *Journal of Sensors and Sensor Systems*, vol. 5, no. 1, pp. 9–16, 2016.

[10] Y. Cao, Tisse, and Christel-Loic, "Shutterless solution for simultaneous focal plane array temperature estimation and nonuniformity correction in uncooled long-wave infrared camera," *Applied optics*, vol. 52, no. 25, pp. 6266–6271, 2013.

[11] J. G. Harris and Y.-M. Chiang, "Nonuniformity correction of infrared image sequences using the constant-statistics constraint," *IEEE Transactions on Image Processing*, vol. 8, no. 8, pp. 1148–1151, 1999.

[12] S. N. Torres and M. M. Hayat, "Kalman filtering for adaptive nonuniformity correction in infrared focal-plane arrays," *JOSA A*, vol. 20, no. 3, pp. 470–480, 2003.

[13] R. C. Hardie, F. Baxley, B. Brys, and P. Hytla, "Scene-based nonuniformity correction with reduced ghosting using a gated lms algorithm," *Optics express*, vol. 17, no. 17, pp. 14918–14933, 2009.

[14] L. Geng, Q. Chen, and W. Qian, "An adjacent differential statistics method for irfpa nonuniformity correction," *IEEE Photonics Journal*, vol. 5, no. 6, pp. 6801615–6801615, 2013.

[15] Q. Wang, J. Dai, X. Sun, and Y. Wang, "A new scene-based non-uniformity correction algorithm for infrared focal plane array," in *Journal of Physics : Conference Series*, vol. 48, p. 403, IOP Publishing, 2006.

[16] C. Zuo, Q. Chen, G. Gu, and X. Sui, "Scene-based nonuniformity correction algorithm based on interframe registration," *JOSA A*, vol. 28, no. 6, pp. 1164–1176, 2011.

[17] P. Toulis, T. Horel, and E. M. Airoldi, "The proximal robbins–monro method," *Journal of the Royal Statistical Society Series B : Statistical Methodology*, vol. 83, no. 1, pp. 188–212, 2021.

[18] B. Lv, S. Tong, Q. Liu, and H. Sun, "Statistical scene-based non-uniformity correction method with interframe registration," *Sensors*, vol. 19, no. 24, p. 5395, 2019.

[19] A. SARAGAGLIA and A. Durand, "Method of infrared image processing for non-uniformity correction."

[20] C. Liu, X. Sui, G. Gu, and Q. Chen, "Shutterless non-uniformity correction for the long-term stability of an uncooled long-wave infrared camera," *Measurement Science and Technology*, vol. 29, no. 2, p. 025402, 2018.

[21] H. Akaike, "A new look at the statistical model identification," *IEEE transactions on automatic control*, vol. 19, no. 6, pp. 716–723, 1974.

[22] G. Schwarz, "Estimating the dimension of a model," *The annals of statistics*, pp. 461–464, 1978.

[23] J. Ding, V. Tarokh, and Y. Yang, "Bridging AIC and BIC : a new criterion for autoregression," *IEEE Transactions on Information Theory*, vol. 64, no. 6, pp. 4024–4043, 2017.

[24] K. P. Burnham and D. R. Anderson, "Multimodel inference : understanding AIC and BIC in model selection," *Sociological methods & research*, vol. 33, no. 2, pp. 261–304, 2004.

[25] D. M. Blei, A. Kucukelbir, and J. D. McAuliffe, "Variational inference : A review for statisticians," *Journal of the American statistical Association*, vol. 112, no. 518, pp. 859–877, 2017.

[26] R. Olbrycht, B. Więcek, and G. De Mey, "Thermal drift compensation method for microbolometer thermal cameras," *Applied Optics*, vol. 51, no. 11, pp. 1788–1794, 2012.

[27] M. Krupiński, T. Sosnowski, H. Madura, and S. Gogler, "Temperature influence compensation in microbolometer detector for image quality enhancement," *QIRT Journal, May*, 2016.

[28] A. Chambolle and T. Pock, "A first-order primal-dual algorithm for convex problems with applications to imaging," *Journal of Mathematical Imaging and Vision*, vol. 40, no. 1, pp. 120–145, 2011.

[29] T. Pock and A. Chambolle, "Diagonal preconditioning for first order primal-dual algorithms in convex optimization," in *2011 International Conference on Computer Vision*, pp. 1762–1769, IEEE, 2011.

[30] Z. Gong, H. Leng, J. Cao, J. Wang, Q. Wu, X. Fan, B. Zhao, L. Yang, and J. Liao, "A new destriping method combining moment matching and interpolation in infrared images," in *2014 4th IEEE International Conference on Information Science and Technology*, pp. 410–413, IEEE, 2014.

[31] Y.-S. Xie, J.-N. Wang, and K. Shang, "An improved approach based on moment matching to destriping for hyperion data," *Procedia Environmental Sciences*, vol. 10, pp. 319–324, 2011.

[32] Y. Tendero, S. Landeau, and J. Gilles, "Non-uniformity correction of infrared images by midway equalization," *Image Processing On Line*, vol. 2, pp. 134–146, 2012.

[33] Y. Tendero and J. Gilles, "Admire : a locally adaptive single-image, non-uniformity correction and denoising algorithm : application to uncooled ir camera," in *Infrared Technology and Applications XXXVIII*, vol. 8353, pp. 580–595, SPIE, 2012.

[34] Y. Cao, M. Y. Yang, and C.-L. Tisse, "Effective strip noise removal for low-textured infrared images based on 1-d guided filtering," *IEEE transactions on circuits and systems for video technology*, vol. 26, no. 12, pp. 2176–2188, 2015.

[35] W. Chang, Y. ZHANG, W. Xu, and J. Song, "An effective strip noise removal method for remote sensing image.," *Journal of Geodesy & Geoinformation Science*, vol. 5, no. 4, 2022.

[36] Y. Chen, T.-Z. Huang, X.-L. Zhao, L.-J. Deng, and J. Huang, "Stripe noise removal of remote sensing images by total variation regularization and group sparsity constraint," *Remote sensing*, vol. 9, no. 6, p. 559, 2017.

[37] Q. Wang, Z. Zhang, and Z. Li, "Infrared stripe correction algorithm based on wavelet decomposition and total variation-guided filtering," *Journal of the European Optical Society-Rapid Publications*, vol. 15, no. 1, 2019.

[38] F. Sakuma, A. Ono, and S. Hattori, "A new characteristic equation for radiation thermometers," *Transactions of the Society of Instrument and Control Engineers*, vol. 23, no. 8, pp. 862–863, 1987.

[39] F. Sakuma, H. Sakate, B. C. Johnson, C. Gibson, G. Machin, T. Ricolfi, M. Battuello, J. Fischer, and H. Jung, "International comparison of radiation temperature scales among five national metrological laboratories using a transfer radiation thermometer," *Metrologia*, vol. 33, no. 3, p. 241, 1996.

[40] F. Liebmann and T. Kolat, "Exploration of alternative methods for fitting radiation thermometer signal to blackbody temperature," *Fluke Calibration, American Fork, Utah, USA. Accessed : Jan*, vol. 10, 2021.

[41] I. Ribet, M. Caes, S. Derelle, S. Bernhardt, and J. Jaeck, "Détecteurs matriciels pour l'infrarouge," *Techniques de l'Ingénieur*, pp. E4060–v2, 2021.

[42] E. M. Smith, J. C. Ginn, A. P. Warren, C. J. Long, D. Panjwani, R. E. Peale, and D. J. Shelton, "Linear bolometer array using a high tcr vox-au film," in *Infrared Technology and Applications XL*, vol. 9070, pp. 589–596, SPIE, 2014.

[43] A. Wolf, J. E. Pezoa, and M. Figueroa, "Modeling and compensating temperature-dependent non-uniformity noise in ir microbolometer cameras," 2016.

[44] M. Krupiński, G. Bieszczad, T. Sosnowski, H. Madura, and S. Gogler, "Non-uniformity correction in microbolometer array with temperature influence compensation," *Metrology and Measurement Systems*, no. 4, 2014.

[45] G. D. Boreman and C. Costanzo, "Compensation for gain nonuniformity and nonlinearity in hgcdte infrared charge-coupled-device focal planes," *Optical Engineering*, vol. 26, no. 10, pp. 981–984, 1987.

[46] R. Olbrycht and B. Wiecek, "New approach to thermal drift correction and gain determination in microbolometer thermal cameras," in *12th International Conference on Quantitative Infrared Thermography*, pp. 7–11, 2014.

[47] H. Budzier and G. Gerlach, "Calibration of uncooled thermal infrared cameras," *Journal of Sensors and Sensor Systems*, vol. 4, no. 1, pp. 187–197, 2015.

[48] J. E. Pezoa, S. N. Torres, J. P. Córdova, and R. A. Reeves, "An enhancement to the constant range method for nonuniformity correction of infrared image sequences," in *Progress in Pattern Recognition, Image Analysis and Applications : 9th Iberoamerican Congress on Pattern Recognition, CIARP 2004, Puebla, Mexico, October 26-29, 2004. Proceedings 9*, pp. 525–532, Springer, 2004.

[49] D. A. Scribner, K. A. Sarkady, M. R. Kruer, J. T. Caulfield, J. Hunt, and C. Herman, "Adaptive nonuniformity correction for ir focal-plane arrays using neural networks," in *Infrared Sensors : Detectors, Electronics, and Signal Processing*, vol. 1541, pp. 100–109, SPIE, 1991.

[50] B. Bercu and P. Fraysse, "A robbins–monro procedure for estimation in semiparametric regression models," 2012.

[51] S. N. Torres, E. M. Vera, R. A. Reeves, and S. K. Sobarzo, "Adaptive scene-based nonuniformity correction method for infrared-focal plane arrays," in *Infrared Imaging Systems : Design, Analysis, Modeling, and Testing XIV*, vol. 5076, pp. 130–139, SPIE, 2003.

[52] M. Guizar-Sicairos, S. T. Thurman, and J. R. Fienup, "Efficient subpixel image registration algorithms," *Optics letters*, vol. 33, no. 2, pp. 156–158, 2008.

[53] S. N. Torres, C. S. Martin, D. G. Sbarbaro, and J. E. Pezoa, "A neural network for nonuniformity and ghosting correction of infrared image sequences," in *Image Analysis and Recognition : Second International Conference, ICIAR 2005, Toronto, Canada, September 28-30, 2005. Proceedings 2*, pp. 1208–1216, Springer, 2005.

[54] E. Lielamurs, A. Cvetkovs, R. Novickis, and K. Ozols, "Mathematical model and synthetic data generation for infra-red sensors," *Sensors*, vol. 22, no. 23, p. 9458, 2022.

[55] Z. He, Y. Cao, Y. Dong, J. Yang, Y. Cao, and C.-L. Tisse, "Single-image-based nonuniformity correction of uncooled long-wave infrared detectors : A deep-learning approach," *Applied optics*, vol. 57, no. 18, pp. D155–D164, 2018.

[56] J. Guan, R. Lai, A. Xiong, Z. Liu, and L. Gu, "Fixed pattern noise reduction for infrared images based on cascade residual attention cnn," *Neurocomputing*, vol. 377, pp. 301–313, 2020.

[57] Y. Li, N. Liu, and J. Xu, "Infrared scene-based non-uniformity correction based on deep learning model," *Optik*, vol. 227, p. 165899, 2021.

[58] R. Shibata, "Selection of the order of an autoregressive model by Akaike's information criterion," *Biometrika*, vol. 63, no. 1, pp. 117–126, 1976.

[59] M. Pakdemirli, "A new perturbation approach to optimal polynomial regression," *Mathematical and Computational Applications*, vol. 21, no. 1, p. 1, 2016.

[60] S. Tanwar, T. Ramani, and S. Tyagi, "Dimensionality reduction using pca and svd in big data : A comparative case study," in *Future Internet Technologies and Trends : First International Conference, ICFITT 2017, Surat, India, August 31-September 2, 2017, Proceedings 1*, pp. 116–125, Springer, 2018.

[61] F. Cao, "Absolutely minimizing lipschitz extension with discontinuous boundary data," *Comptes Rendus de l'Académie des Sciences-Series I-Mathematics*, vol. 327, no. 6, pp. 563–568, 1998.

[62] A. Almansa, F. Cao, Y. Gousseau, and B. Rougé, "Interpolation of digital elevation models using amle and related methods," *IEEE transactions on geoscience and remote sensing*, vol. 40, no. 2, pp. 314–325, 2002.

[63] Z. Drmač and K. Veselić, "New fast and accurate jacobi svd algorithm. i," *SIAM Journal on matrix analysis and applications*, vol. 29, no. 4, pp. 1322–1342, 2008.

[64] R. Gribonval and E. Bacry, "Harmonic decomposition of audio signals with matching pursuit," *IEEE Transactions on signal processing*, vol. 51, no. 1, pp. 101–111, 2003.

[65] G. Davis, S. Mallat, and Z. Zhang, "Adaptive time-frequency approximations with matching pursuits," in *Wavelet Analysis and Its Applications*, vol. 5, pp. 271–293, Elsevier, 1994.

[66] G. Davis, S. Mallat, and M. Avellaneda, "Adaptive greedy approximations," *Constructive approximation*, vol. 13, pp. 57–98, 1997.

[67] Y. Kang, L. Pan, M. Sun, X. Liu, and Q. Chen, "Destriping high-resolution satellite imagery by improved moment matching," *International Journal of Remote Sensing*, vol. 38, no. 22, pp. 6346–6365, 2017.

[68] D. Coltuc, P. Bolon, and J.-M. Chassery, "Exact histogram specification," *IEEE Transactions on Image processing*, vol. 15, no. 5, pp. 1143–1152, 2006.

[69] J. Delon, "Midway image equalization," *Journal of Mathematical Imaging and Vision*, vol. 21, no. 2, pp. 119–134, 2004.

[70] J. Rabin, J. Delon, and Y. Gousseau, "Transportation distances on the circle," *Journal of Mathematical Imaging and Vision*, vol. 41, no. 1, pp. 147–167, 2011.

[71] E. Wang, P. Jiang, X. Li, and H. Cao, "Infrared stripe correction algorithm based on wavelet decomposition and total variation-guided filtering," *journal of the european optical society-rapid publications*, vol. 16, no. 1, p. 1, 2020.

[72] K. He, J. Sun, and X. Tang, "Guided image filtering," *IEEE transactions on pattern analysis and machine intelligence*, vol. 35, no. 6, pp. 1397–1409, 2012.

[73] S. Yuan, H. Qin, X. Yan, S. Yang, S. Yang, N. Akhtar, and H. Zhou, "Ascnet : Asymmetric sampling correction network for infrared image destriping," *IEEE Transactions on Geoscience and Remote Sensing*, 2025.

Table des matières

Résumé **vii**

Abstract **ix**

1 Introduction **1**
1 Imagerie thermique, principes et contexte industriel . . . 1
1.1 Rayonnement thermique émis par un objet . . . 1
1.2 Image brute issue d'une caméra thermique non-refroidie . . . 2
1.3 Contexte industriel et domaines d'application . . . 3
2 État de l'art des modèles de correction existants . . . 5
2.1 Modélisation affine de la réponse du capteur . . . 5
2.2 Méthodes « *calibration-based* » . . . 6
2.2.1 Calibration à deux points . . . 7
2.2.2 Modélisation polynomiale en TFPA . . . 7
2.3 Méthodes *scene-based* . . . 8
2.3.1 Méthodes fondées sur l'analyse statistique . . . 8
2.3.2 Méthodes fondées sur le recalage d'images . . . 9
2.4 Méthodes « mixtes » . . . 9
2.5 Limites des modèles de correction . . . 10
3 Problématique et objectifs de la thèse . . . 11
3.1 Enjeux industriels . . . 11
3.2 Objectifs de la thèse . . . 11
3.3 Démarche . . . 12
4 Résumé des travaux et contributions . . . 12
4.1 Modélisation polynomiale des images brutes . . . 13
4.1.1 Protocole de calibration et données acquises . . . 13
4.1.2 Méthode de la littérature et motivation de notre approche 13
4.1.3 Modélisation polynomiale des images brutes . . . 14
4.1.3.1 Choix du degré du polynôme à partir d'un test d'*overfitting* . . . 14
4.1.3.2 Masque de pixels défectueux . . . 15
4.1.4 Observation des erreurs de *fit* et correction delta . . . 15
4.1.4.1 Analyse des erreurs de *fit* . . . 15
4.1.5 Évaluation des performances sur un sous-ensemble de test 16
4.2 Analyse statistique des données de calibration . . . 17
4.2.1 Étude du comportement des pixels à partir d'une ACP . . 17
4.2.2 Analyse de l'espace des biais du modèle affine . . . 18

4.2.2.1 Modèle affine avec gain constant et biais variable 18
4.2.2.2 Construction de la famille de biais 19
4.2.2.3 Analyse de l'espace des biais 19
4.3 Correction d'une image à partir d'une famille de biais 20
4.3.1 Modèle variationnel générique 20
4.3.2 Modèle ℓ^2 21
4.3.3 Modèle $\mathscr{H}^1$ 22
4.3.4 Modèle TV 22
4.3.5 Étude comparative 23
4.4 Débruitage colonnaire 24
4.4.1 État de l'art des modèles de débruitage colonnaire 24
4.4.1.1 Approches statistiques 24
4.4.1.2 Méthodes de filtrage spatial ou fréquentiel 24
4.4.1.3 Méthodes variationnelles 25
4.4.2 Modèle de perturbation 25
4.4.3 Modèle TVx 25
4.4.4 De l'analyse du gradient horizontal à la formulation d'un modèle variationnel 26
4.4.4.1 Détection des segments homogènes au sein d'une colonne du gradient horizontal 26
4.4.4.2 Problème variationnel 28
5 Organisation de la thèse 28

2 Principes physiques, modèles de correction existants 31
1 Principes physiques et électroniques de l'imagerie thermique 31
1.1 Rayonnement du corps noir 32
1.1.1 Définition d'un corps noir 32
1.1.2 Lois physiques 33
1.2 Corps gris, émissivité 35
1.3 Rayonnement infrarouge, irradiance reçue par un pixel 36
1.4 Technologie des caméras thermiques non-refroidies et génération de l'image brute 38
1.4.1 Capteur d'une caméra thermique non-refroidie 39
2 Modélisation du capteur et état de l'art des modèles de correction 45
2.1 Modélisation affine du capteur 45
2.1.1 Paramètres environnementaux qui influent sur la réponse du capteur 45
2.1.2 Modèle affine multi-paramétrique 45
2.1.2.1 Modèle affine direct avec responsivité et décalage 45
2.1.2.2 Reformulation du modèle affine multi-paramétrique avec gain et offset 47
2.1.2.3 Dérive de la réponse du capteur selon l'environnement (*thermal drift*) 47
2.1.3 Modélisation affine à un seul paramètre (TFPA) 49
2.1.4 Modèle affine avec gain unique 49
2.2 État de l'art des modèles de correction 49

2.2.1 Évolution des corrections thermiques, des capteurs refroidis aux méthodes *less shutter* et *shutterless* 50
2.2.1.1 Méthode de la calibration à deux points 50
2.2.1.2 Modèles de correction pour une caméra équipée d'un *shutter* 51
2.2.1.3 Vers des stratégies *less shutter* et *shutterless* . . . 51
2.2.2 Méthodes principalement « *calibration-based* » 51
2.2.2.1 Modélisation polynomiale en TFPA 52
2.2.2.2 Modèle polynomial avec ré-estimation de la mesure de TFPA 53
2.2.3 Méthodes « scene-based » 54
2.2.3.1 Méthodes fondées sur les statistiques 54
2.2.3.2 Méthodes fondées sur le recalage d'images 58
2.2.4 Méthodes « mixtes » 58
2.2.4.1 Modèle de correction avec une ou deux références 58
2.2.4.1.1 Modèle de correction avec une seule référence 59
2.2.4.1.2 Modèle de correction avec deux références 60
2.2.4.2 Correction d'une image à partir d'une famille de scenes uniformes 61
2.2.5 Méthodes de correction par apprentissage avec réseaux de neurones 63
3 Conclusion 63

3 Modélisation du capteur, étude des données de calibration 65
1 Modèles de correction fondé sur une modélisation polynomiale des images brutes 67
1.1 Établissement d'un jeu de données de calibration 67
1.1.1 Calibration en « étuve stabilisée » 67
1.1.2 Jeu de données de calibration 68
1.2 Rappel de la méthode *shutterless polynomial* utilisée à titre comparatif 69
1.3 Modélisation polynomiale des images brutes de scène uniforme . . . 70
1.3.1 Présentation du problème 71
1.3.2 Détermination du degré polynomial 72
1.3.2.1 Revue des pratiques sur le choix du degré polynômial 72
1.3.2.2 Détermination du degré propre à un pixel à partir d'un test d'*overfitting* 72
1.3.2.3 Carte des degrés 79
1.3.2.4 Détermination d'un degré global 79
1.4 Détection des pixels défectueux 81
1.4.1 Définition d'un pixel défectueux 81
1.4.2 Construction du masque de pixels non défectueux 81
1.4.3 Étude du détecteur par analyse du comportement de l'ensemble des pixels 83

1.5 Modèle de correction par prise en compte des erreurs de mesure de la TFPA . . . 84
1.5.1 Modèle de la correction delta . . . 86
1.5.2 Résolution algorithmique de la correction delta . . . 86
1.5.3 Résultats expérimentaux de la correction delta, insuffisance du modèle paramétrique. . . . 89
1.6 Correction d'une scène uniforme, comparaison des différentes méthodes . . . 92
1.6.1 Détermination de la responsivité et décalage selon la TFPA 92
1.6.2 Correction de l'image brute d'une scène uniforme . . . 93
1.6.3 Étude des performances sur un jeu de calibration restreint, extrapolation en TFPA . . . 96
2 Analyse statistique des données de calibration . . . 100
2.1 Étude du comportement général d'un pixel à partir d'une ACP . . . 100
2.1.1 Réponse du capteur face à un ensemble de situations . . . 100
2.1.2 Analyse par composantes principales . . . 100
2.1.3 Observations, interprétations et perspectives d'étude du modèle direct . . . 102
2.2 Analyse de l'espace des biais . . . 104
2.2.1 Modèle affine avec gain unique et biais . . . 104
2.2.2 Construction d'une famille de biais . . . 105
2.2.3 Analyse de l'espace des biais . . . 106
3 Conclusion . . . 112

4 Modèles de correction, débruitage colonnaire 113
1 Correction d'une image à partir d'une famille de biais . . . 113
1.1 Cadre de travail, hypothèses et modèle générique . . . 113
1.1.1 Correction des pixels défectueux . . . 113
1.1.2 Modèle affine non paramétrique de correction d'une image pseudo-brute . . . 114
1.1.3 Famille de biais . . . 116
1.1.4 Modèle générique de correction d'une image brute à partir d'une famille de biais . . . 116
1.2 Modèle ℓ^2 . . . 117
1.2.1 Existence, unicité et équivalence des formulations du problème ℓ^2 . . . 117
1.2.2 Résolution par coefficients dans une base de biais et enjeux numériques . . . 118
1.2.2.1 Formulation et résolution du problème via les coefficients dans une base initiale . . . 118
1.2.2.2 Problèmes pratiques de résolution causés par le mauvais conditionnement . . . 120
1.2.2.3 Reformulation et résolution dans une base orthonormée . . . 120
1.2.3 Correction d'images par le modèle ℓ^2 . . . 121
1.2.3.1 Correction d'une scène uniforme . . . 121

1.2.3.2 Correction d'une scène réelle, insuffisance du modèle ℓ^2 . . . 122
1.3 Modèle $\mathscr{H}^1$ . . . 123
1.3.1 Définition du gradient spatial d'une image et de son adjoint 123
1.3.2 Problème $\mathscr{H}^1$, existence et unicité de la solution . . . 124
1.3.3 Formulation et résolution du problème selon les coefficients dans une base . . . 125
1.3.3.1 Formulation et résolution via les coefficients dans la base d'origine . . . 125
1.3.3.2 Reformulation du problème dans une base orthonormée . . . 126
1.3.4 Correction d'images par le modèle $\mathscr{H}^1$ . . . 127
1.4 Modèle TV . . . 128
1.4.1 Problème TV . . . 128
1.4.2 Formulation du problème selon les coefficients dans une base, résolution avec plusieurs variantes de Chambolle-Pock 129
1.4.2.1 Formulation et résolution via les coefficients dans la base d'origine . . . 129
1.4.2.2 Résolution numérique par l'algorithme de Chambolle–Pock avec choix de pas égaux . . . 130
1.4.2.3 Résolution numérique par l'algorithme de Chambolle–Pock avec pas variables . . . 131
1.4.2.4 Résolution numérique par l'algorithme de Chambolle–Pock préconditionné . . . 132
1.4.2.5 Résolution numérique dans une base orthonormée 133
1.4.3 Modèle TV par une méthode intrinsèque . . . 134
1.4.4 Étude comparative des méthodes algorithmiques de résolution du problème TV . . . 135
1.4.5 Correction d'images par le modèle TV . . . 135
1.5 Comparaison des performances des modèles ℓ^2, $\mathscr{H}^1$ et TV avec le modèle *shutterless polynomial* . . . 137
1.5.1 Limites des modèles présents dans la littérature . . . 137
1.5.2 Données de calibration et de test utilisées pour l'étude comparative . . . 138
1.5.3 Étude comparative . . . 139
1.6 Décomposition colonnaire des biais et intégration d'un débruitage colonnaire dans le modèle de correction . . . 143
1.6.1 Causes des artefacts de colonnes . . . 143
1.6.2 Proposition d'un débruitage colonnaire intégré à nos modèles de correction . . . 143
1.6.3 Résultats de correction à partir d'une famille décomposée 146
1.7 Conclusion sur les modèles de correction reposant sur une famille de biais . . . 146
2 Débruitage colonnaire en imagerie thermique . . . 149
2.1 État de l'art des modèles de débruitage colonnaire . . . 149

2.1.1 Méthodes de débruitage colonnaire fondées sur les statistiques . . . 149
2.1.1.1 Méthodes « *moment matching* » . . . 150
2.1.1.2 Méthodes de correction par harmonisation des histogrammes colonnaires . . . 151
2.1.2 Méthodes de filtrage spatial et fréquentiel . . . 152
2.1.3 Méthodes variationnelles . . . 152
2.1.4 Un exemple d'approche combinant méthode variationnelle et filtrage fréquentiel . . . 153
2.1.5 Méthodes d'apprentissage profond pour la correction du bruit colonnaire . . . 155
2.2 Hypothèses de travail et modèle de perturbation . . . 156
2.2.1 Hypothèses de travail . . . 156
2.2.2 Modèle de perturbation . . . 157
2.3 Débruitage colonnaire par le modèle TVx . . . 157
2.3.1 Étude de la correction TVx sur scènes avec ou sans structures verticales . . . 160
2.4 Débruitage colonnaire par analyse du gradient horizontal colonne par colonne . . . 160
2.4.1 Détection des segments homogènes au sein d'une colonne du gradient horizontal . . . 161
2.4.1.1 Rôle de l'identification des segments homogènes dans l'estimation du bruit colonnaire . . . 161
2.4.1.2 Détection de segments homogènes dans un signal par un test du χ^2 . . . 164
2.4.2 Correction colonnaire par sélection de segments homogènes via un problème variationnel . . . 167
2.4.3 Évaluation visuelle des corrections obtenues . . . 172
2.5 Comparaison des méthodes de correction proposées avec une méthode existante . . . 173
2.6 Conclusion sur le débruitage colonnaire . . . 176

Conclusion **177**

Annexe **181**

Bibliographie **185**

Table des matières **190**

Liste des figures **196**

Liste des figures

1.1 Image brute issue de caméras thermiques non-refroidies. 3
1.2 Domaines d'application 5
1.3 Images de responsivité et de décalage 6
1.4 Analyse de l'erreur de *fit* d'un état. 15
1.5 Correction de scènes uniformes en extrapolation. 16
1.6 Comportement général du capteur, ACP. 18
1.7 Famille des biais 19
1.8 Axes principaux obtenus par ACP centrée et ACP non centrée. 20
1.9 Correction d'une image à partir d'une famille de biais, comparaison avec méthode de référence 23
1.10 Caractère mal posé du débruitage colonnaire 26
1.11 Correction d'une image par débruitage colonnaire 27

2.1 Spectre électromagnétique. 33
2.2 Catastrophe ultraviolette. 34
2.3 Loi de Planck pour différents éléments. 36
2.4 Absorption spectrale du détecteur. 37
2.5 Lien entre température de scène et irradiance 39
2.6 Schéma d'un microbolomètre associé à un pixel 40
2.7 Bilan des flux thermiques. 40
2.8 Temps de réponse thermique d'un microbolomètre 41
2.9 Circuit de lecture. 42
2.10 Comparaison entre les caméras thermiques refroidies et non-refroidies . . . 43
2.11 Image brute, image corrigée, FPN 44
2.12 Images de responsivité et de décalage 46
2.13 Illustration du *thermal drift* 48
2.14 Correction par le modèle *shutterless polynomial* 52
2.15 Comparatif de quelques modèles de correction de l'état de l'art. 59

3.1 Variabilités induites par des changements de TFPA et de scène. 66
3.2 Schéma fonctionnel comparatif pour établir responsivité et décalage 71
3.3 Illustration de l'*underfitting* et de l'*overfitting*. 73
3.4 Conditionnement de la matrice de Vandermonde 75
3.5 Explication du test d'*overfitting* selon une première approche naïve. 77
3.6 Explication du test d'*overfitting*, rectification de l'approche naïve. 78
3.7 Illustration du test d'*overfitting*, comparaison avec BIC 80
3.8 Images comparatives carte des degrés et masque 82
3.9 Création du masque de pixels non défectueux, histogramme des RMSE. . . 83

3.10 Comportement des pixels selon le degré obtenu par le test d'*overfitting* . . 85
3.11 Analyse de l'erreur de *fit* avant correction delta. 87
3.12 Analyse de l'erreur de *fit* après correction delta. 90
3.13 Valeurs de la correction delta. 91
3.14 Comparaisons des responsivités et décalages pour le jeu $\mathscr{C}_1$. 94
3.15 Comparaisons des responsivités et décalages pour le jeu $\mathscr{C}_2$. 95
3.16 Correction d'une scène uniforme, comparatif. 97
3.17 Correction de scènes uniformes à partir d'un jeu de calibration restreint. . 99
3.18 Comportement des pixels dans quatre situations. 101
3.19 Comportement général du capteur, ACP, courbes 3D du centre et des atomes 103
3.20 Famille des biais, axes principaux obtenus par les approches ACP centrée et ACP non centrée. 109
3.21 Proximité d'un biais avec une sous-famille, comparatif ACP ACP non centrée. 111

4.1 Images du gain, d'une image pseudo-brute et d'une famille de biais 115
4.2 Correction d'images par le modèle ℓ^2 123
4.3 Correction d'images par le modèle $\mathscr{H}^1$ 128
4.4 Évolution de l'énergie et des coefficients au fil des itérations. 131
4.5 Courbes de convergence des méthodes TV 136
4.6 Correction d'images par les modèles ℓ^2, $\mathscr{H}^1$ et TV 137
4.7 Correction d'images avec des données de calibration restreintes en TFPA, comparaison de quatre méthodes 140
4.8 Correction croisée entre milieux homogène et inhomogène 141
4.9 Correction d'une image à partir d'un nombre de données de calibration restreint 142
4.10 Effet colonnaire résiduel et « vérité terrain » 144
4.11 Correction d'une image à partir d'une famille de biais décomposée 147
4.12 Analyse d'un bruit structuré de type *stripes* 150
4.13 Illustration que le problème de débruitage est mal posé 158
4.14 Correction d'une image sans structure verticale par le modèle TVx 161
4.15 Correction d'une image avec structures verticales par le modèle TVx 162
4.16 Analyse du gradient horizontal 163
4.17 Détection de segment par un test du χ^2 166
4.18 Détection de segments, résultats sur un signal 167
4.19 Correction avec le meilleur score 168
4.20 Analyse du gradient de la correction par meilleur score 169
4.21 Calcul des fonctions valeur partielle et d'antécédence optimale 173
4.22 Correction obtenue par une méthode variationnelle 174
4.23 Comparaison de différentes méthodes de correction du bruit colonnaire . . 175