



HAL
open science

Apprentissage sur des données multispectrales et multiplateformes au profit des systèmes de mission de surveillance

Christophe Caubet

► **To cite this version:**

Christophe Caubet. Apprentissage sur des données multispectrales et multiplateformes au profit des systèmes de mission de surveillance. Vision par ordinateur et reconnaissance de formes [cs.CV]. Université de Bordeaux, 2023. Français. NNT : 2023BORD0217 . tel-04390283

HAL Id: tel-04390283

<https://theses.hal.science/tel-04390283>

Submitted on 29 Jan 2024

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

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

THÈSE

PRÉSENTÉE À

L'UNIVERSITÉ DE BORDEAUX

ÉCOLE DOCTORALE DE MATHÉMATIQUES ET
D'INFORMATIQUE

par **Christophe CAUBET**

POUR OBTENIR LE GRADE DE

DOCTEUR

SPÉCIALITÉ : INFORMATIQUE

**Apprentissage sur des données multispectrales et
multiplateformes au profit des systèmes de mission
de surveillance**

Date de soutenance : 20 septembre 2023

Devant la commission d'examen composée de :

M. Pascal DESBARATS	Professeur, Université de Bordeaux	Directeur de thèse
M. Gilles GUERRINI	Docteur, Thales DMS	Co-directeur de thèse
M. Philippe JOLY	Professeur, Université Toulouse III	Rapporteur
M. Bertrand KERAUTRET	Professeur, Université Lyon 2	Rapporteur
Mme. Nadine COUTURE	Professeure, ESTIA	Présidente du Jury
Mme. Petra GOMEZ	Maîtresse de conférences, Université La Rochelle	Examinatrice
M. Jean-Philippe DOMENGER	Professeur, Université de Bordeaux	Invité, Co-directeur de thèse

À ma famille

Remerciements

Pour commencer, je tiens à remercier mes directeurs de thèse Pascal Desbarats et Jean-Philippe Domenger du LaBRI et Gilles Guerrini de Thales DMS. Pascal, pour m'avoir présenté ce sujet, pour votre soutien technique, administratif et surtout moral tout au long de cette thèse sans jamais me bâcher, merci. Jean-Philippe, pour vos conseils éclairés qui m'ont poussé à trouver de nouvelles perspectives et pour nos longues discussions qui m'ont fait passer littéralement des mauvaises réponses aux bonnes questions lorsque j'en avais besoin, merci. Gilles (en dernier mais c'est parce qu'on garde le meilleur pour la fin!), pour avoir proposé cette thèse même si on n'a finalement pas pu étudier le quart de la niche qu'est l'imagerie hyperspectrale, et pour votre aide sur tant d'aspects malgré la période difficile du (post-)covid, merci.

Je remercie également les membres du jury Nadine Couture, Petra Gomez, Philippe Joly et Bertrand Kerautret d'avoir accepté d'évaluer mon travail ainsi que pour tous les conseils, remarques et questions précieux qu'ils m'ont partagé.

Merci aussi à tous mes collègues de Thales : Antoine, Benoît, Didier, Emmanuel, Jonathan, Léa, Luc, Romain, Raphaël, Quentin, Thibault, Thomas, Titien et Vincent. Sans oublier bien sûr David, Léo, Maëlle et Thomak dont la présence m'a tant aidé et particulièrement lors de cette troisième année de thèse qui n'était pas la plus facile. Enfin, je ne peux pas oublier Ema et Olivier ; pour tout ce qu'ils ont fait et représenté pour moi lors de ces presque 4 années passées à leurs côtés, merci.

De même, je remercie les gens du LaBRI, Amaury, Colleen, Imed, Sam, Vincent et Yanis, pour tous les bons moments passés ensemble malgré mon manque de présence flagrant au labo (désolé!).

Merci à mes collègues de bureau Adrien, Emile et Marie avec qui nous nous sommes mutuellement soutenus lors de nos thèses respectives, ainsi qu'à Romain, Teiki, Valentin et toute la bande de Wololo.

Je remercie Alex, Jason et Kev, Alan et Kévin, Loïc qui m'ont tant aidé à décompresser ainsi que pour tous ces bons moments passés et à venir avec vous. Merci beaucoup également aux "Anciens" Corentin, Florian, Florent, Jean-Pascal, Sylvain, Thomas, Thomas, Tom et Yoann qui sont là depuis toutes ces années. Sans vous je sais pas si j'aurais eu le courage de me lancer ni de continuer cette thèse, donc merci. Pour vous remercier, je m'en vais vous attraper des hyperfantômes! (non)

Enfin, merci infiniment à toute ma famille, et particulièrement mes parents et ma soeur, pour leur humour(?) douteux et surtout pour tout leur soutien. Merci.

Table des matières

Introduction	1
1 Contexte	4
1.1 Introduction	4
1.2 Contexte technique	4
1.3 Contexte scientifique	6
1.3.1 Dimensions spectrale et spatiales	6
1.3.2 Conditions d'acquisition	8
1.3.3 Application des images hyperspectrales	9
1.3.4 Imagerie hyperspectrale pour la surveillance	10
1.4 Contexte opérationnel	12
1.5 Conclusion	12
2 Études spectrales	13
2.1 Introduction	13
2.2 Bases de données	14
2.3 Transformations des spectres	14
2.4 Analyse visuelle de spectres de plastique	18
2.5 Mesures de similarité	19
2.5.1 Définition	19
2.5.2 Analyse des spectres de plastique	21
2.5.3 Variabilité spectrale	21
2.5.4 Mélanges	23
2.6 Extraction des pics	25
2.7 Conclusion	26
3 Acquisition d'images hyperspectrales	27
3.1 Introduction	27
3.2 Caractéristiques des caméras hyperspectrales	28
3.3 Caméras sélectionnées	30
3.3.1 Cubert FirefEYE S185	30
3.3.2 Headwall SWIR-384	31
3.4 Acquisition d'images de véhicules	32
3.5 Acquisition d'images de matériaux	33
3.5.1 Enjeux	34
3.5.2 Échantillons	34
3.5.3 Acquisitions extérieures	34
3.5.4 Acquisitions intérieures	35

3.6	Conclusion	37
4	Calibration	38
4.1	Introduction	38
4.2	Définition de la calibration	39
4.3	Dérive temporelle des valeurs brutes	41
4.4	Origine de la dérive temporelle	43
4.5	Impact sur la radiance	45
4.6	Méthodes de sélection d'image <i>DC</i> pour le calcul de la radiance	46
4.6.1	« <i>Single</i> »	46
4.6.2	« <i>Closest</i> »	47
4.6.3	« <i>Previous</i> »	47
4.6.4	« <i>Next</i> »	48
4.6.5	« <i>Linear session</i> »	49
4.6.6	« <i>Linear closest</i> »	49
4.6.7	Conclusion de la sélection de l'image <i>DC</i>	50
4.7	Comparaison des références de blanc	51
4.8	Calibration par référence de gris	52
4.9	Limitations de la calibration	56
4.10	Conclusion	57
5	Analyse d'images hyperspectrales	58
5.1	Introduction	58
5.2	Pré-traitements	59
5.2.1	Reconstruction d'images couleur	59
5.2.2	Segmentation et boîtes englobantes de véhicules	59
5.2.3	Segmentation d'objets	61
5.2.4	Correction atmosphérique	63
5.2.5	Normalisation des spectres	65
5.3	Analyse visuelle	65
5.3.1	Analyse d'objets de même couleur	65
5.3.2	Analyse d'objets de même matériau	69
5.3.3	Observations générales	70
5.4	Mesures de similarité	71
5.4.1	Classification de spectres	71
5.4.2	Détection de matériaux	72
5.4.3	Discrimination de pixels	74
5.4.4	Réduction de dimension	77
5.4.5	Détection de matériaux avec ACP	79
5.4.6	Discrimination de pixels avec ACP	80
5.5	Acquisitions nocturnes/sans lumière	81
5.6	Conclusion	82

6 Réseaux de neurones	83
6.1 Introduction	83
6.2 Détection de véhicules	84
6.2.1 Réseau spectral	84
6.2.2 Réseau colorimétrique	90
6.2.3 Réseau à convolutions 3D	91
6.2.4 Conclusion	91
6.3 Réseaux existants pour la détection d’objets	92
6.3.1 YOLOv3	92
6.3.2 Mask-RCNN	99
6.4 Réseaux pour la classification d’images hyperspectrales	102
6.4.1 Terminologie	102
6.4.2 <i>DeepHyperX</i> (Audebert, 2019)	103
6.4.3 <i>hyperspectral_deeplearning_review</i> (Paoletti, 2019)	105
6.4.4 Conclusion	105
6.5 Conclusion	106
Conclusion	108
Annexes	113

Introduction

L'imagerie hyperspectrale est une technologie permettant d'acquérir des images dans plusieurs dizaines ou centaines de bandes du spectre électromagnétique (Landgrebe [2002]). Ces images, combinant l'imagerie et la spectrométrie, permettent d'estimer la présence et la position de compositions chimiques dans des images et permettent ainsi la discrimination de matériaux à longue distance (Mou *et al.* [2017]). L'amélioration récente des capacités de calcul des machines a permis de traiter les quantités importantes de données contenues dans les images hyperspectrales.

Les systèmes de mission de surveillance permettent de gérer une mission à travers plusieurs aspects : la trajectoire des plateformes, la programmation des capteurs, l'acquisition des données par les capteurs et le traitement de ces données. L'acquisition de données peut être conduite par différents types de capteurs (visible, radar, ...) embarqués sur des plate-formes telles que les avions de surveillance ou les drones. Le traitement de ces données permet alors d'en extraire des informations pour établir une situation dans différents domaines : aérien, maritime et terrestre. Ces informations sont de cinq types :

- *Détection* de la présence ou absence d'un élément dans un ensemble de données.
- *Discrimination* des éléments en plusieurs catégories en fonction de leurs caractéristiques respectives.
- *Reconnaissance* d'un élément en associant les caractéristiques de cet élément aux caractéristiques de classes déjà existantes.
- *Identification* d'un élément précis, à travers des caractéristiques qui permettent de le différencier d'autres éléments appartenant à la même classe.
- *Localisation* d'un élément connu afin de connaître sa position.

Pour les systèmes de mission de surveillance, le traitement des données doit être fonctionnel dans autant de situations que possible. Un système de détection trop sensible détecte des éléments non-prioritaires qui noient les informations les plus importantes, et un système peu sensible risque de rater une partie cruciale de cette information prioritaire. De plus, il est nécessaire de développer des systèmes de surveillance de plus en plus sophistiqués afin de réduire les limitations des systèmes existants et de répondre à des besoins grandissants. Dans le cadre de l'exploitation de certaines données acquises par les capteurs en contexte hostile ou militaire, il est également important de considérer la présence d'éléments tels des leurres ou des éléments camouflés, dont la fonction est de tromper l'interprétation des informations dans les systèmes de

surveillance. Il est alors nécessaire de développer de nouvelles méthodes permettant de détecter ces éléments et de les reconnaître comme tels, par exemple en discriminant les leurres des objets qu'ils imitent.

Les drones sont des véhicules à faible altitude permettant de faire de la surveillance dans des contextes variés et pendant un temps long. De plus, ils bénéficient d'une forte mobilité sur de grandes distances et peuvent se déplacer dans des zones qui ne sont pas accessibles aux humains (zone polluée, zone de danger...). Les drones sont donc un outil primordial pour les systèmes de mission de surveillance.

Nous nous intéressons alors à l'utilisation de l'imagerie hyperspectrale pour la discrimination et la localisation de matériaux dans le cadre de missions de surveillance par drones. La discrimination de matériaux possède de nombreuses applications pour la surveillance telles que la détection de véhicules ennemis ou la reconnaissance des types de blindages de ces véhicules.

Les travaux menés jusqu'à présent sur ce sujet considèrent soit des données privées, soit des bases de données publiques limitées sur plusieurs aspects. Tout d'abord, les conditions dans lesquelles ces bases de données sont acquises diffèrent entre les bases de données, et un même objet ou matériau ne présente pas les mêmes caractéristiques en fonction de la base de données. De plus, le faible nombre de bases de données publiques ne permet pas de généraliser les caractéristiques de ces matériaux. Enfin, ces données publiques sont satellitaires ou aériennes à haute altitude, et donc très différentes d'images acquises depuis des systèmes terrestres ou depuis des systèmes aériens tels que les drones.

Cette thèse, sous dispositif Cifre (Convention industrielle de formation par la recherche) validé par l'ANRT (Association Nationale Recherche Technologie), se déroule entre l'entreprise Thales DMS (Defense Mission Systems) France SAS et le Laboratoire Bordelais de Recherche en Informatique (LaBRI). Cette thèse s'intéresse à des problématiques de détection, de discrimination et de reconnaissance de matériaux par utilisation de l'imagerie hyperspectrale dans des conditions d'acquisition variées au sol ou faible altitude.

Nos travaux sont les suivants. Dans un premier temps, nous analyserons l'état de l'art de l'imagerie hyperspectrale, les analyses conduites sur ces images ainsi que les méthodes utilisées pour leur traitement (Chapitre 1). Cette analyse se poursuivra au Chapitre 2 avec l'analyse de bases de données spectrales pour l'extraction des caractéristiques spectrales de matériaux d'intérêt. Dans le Chapitre 3, nous présenterons l'acquisition de deux grandes bases de données hyperspectrales terrestres, de matériaux d'intérêt pour les systèmes de mission de surveillance, acquises dans de multiples conditions d'illumination. La calibration d'images hyperspectrales et ses enjeux et difficultés en conditions opérationnelles seront présentées en Chapitre 4. Ces données seront ensuite pré-traitées et analysées avec des méthodes sans apprentissage dans le Chapitre 5 et des limitations de ces méthodes seront présentées. Enfin, ces analyses

se poursuivront dans le Chapitre 6 avec l'utilisation de réseaux de neurones.

Contributions

Deux articles ont été publiés au cours de cette thèse :

1. Caubet Christophe, Guerrini Gilles, Desbarats Pascal et Domenger Jean-Philippe, 2022a. Case Study of a Calibration Problem in Acquired Hyperspectral Images. Dans 2022 IEEE International Conference on Image Processing (ICIP), pages 2261–2265. IEEE. ISBN 978-1-6654-9620-9. doi : 10.1109/ICIP46576.2022.9897839.
2. Caubet Christophe, Guerrini Gilles, Desbarats Pascal et Domenger Jean-Philippe, 2022b. Evaluation of the Operational Contribution of the Use of Neural Networks on Hyperspectral Images for the Benefit of Airborne Surveillance. Dans SET-312-RSM. Bled, Slovenia. URL <https://hal.science/hal-03963144v1>

Chapitre 1

Contexte

Sommaire

1.1	Introduction	4
1.2	Contexte technique	4
1.3	Contexte scientifique	6
1.3.1	Dimensions spectrale et spatiales	6
1.3.2	Conditions d'acquisition	8
1.3.3	Application des images hyperspectrales	9
1.3.4	Imagerie hyperspectrale pour la surveillance	10
1.4	Contexte opérationnel	12
1.5	Conclusion	12

1.1 Introduction

Dans ce chapitre, nous présenterons les contextes dans lesquels se place cette thèse. Nous présentons d'abord le principe et le fonctionnement de l'imagerie hyperspectrale. Cette présentation technique mène à une description des travaux en rapport avec ce type d'imagerie et les méthodes existantes pour leur traitement. Enfin, nous ferons le lien entre ces travaux existants et les systèmes de mission de surveillance.

1.2 Contexte technique

Ce que l'on appelle communément « lumière » est en réalité l'ensemble des ondes du spectre électromagnétique appartenant au domaine visible, de longueur entre 380 et 740 nanomètres (nm). Le spectre électromagnétique est constitué de plusieurs domaines, différenciés par la longueur des ondes, tels que l'ultraviolet, le visible et l'infrarouge, voir Fig. 1.1.

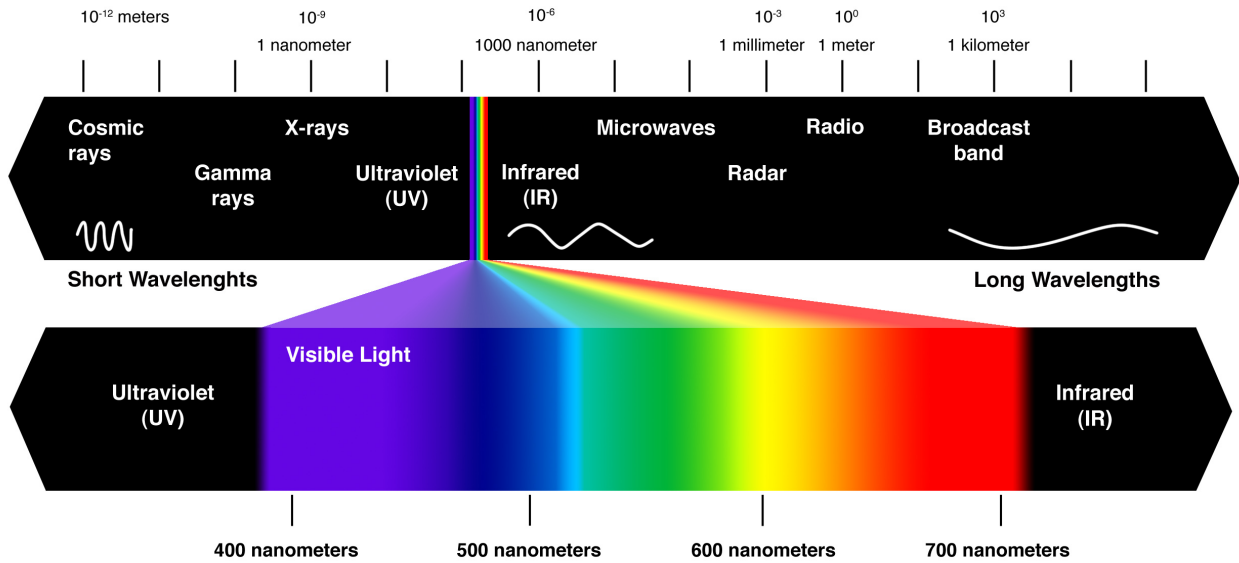


FIGURE 1.1 – Domaines du spectre électromagnétique. Figure extraite de <https://chrisparkins.wordpress.com/2013/09/01/the-colour-of-light/>, dernière visite le 21/03/2023.

Les images couleur classiques visent à imiter la vision humaine de la lumière. Ces images sont constituées de trois canaux de couleur, rouge, vert et bleu, construits à partir d'interpolations de la lumière du domaine visible. Cela signifie que ces images, tout comme la vision humaine, sont limitées à une faible portion du spectre électromagnétique. De plus, ces images sont également limitées dans la quantité d'information qu'elles représentent, car seuls trois canaux représentent toutes les ondes électromagnétiques mesurant entre 380 et 740 nanomètres.

Il est alors possible de constituer des images avec un nombre plus important de canaux (ou « bandes spectrales »), échantillonnant le spectre électromagnétique. En fonction du nombre de bandes spectrales, nous appelons alors ces images multispectrales ou hyperspectrales, mais la manière de différencier ces deux types d'image varie en fonction des capteurs et de l'état de l'art (Aasen *et al.* [2018]). Dans cette thèse, nous considérons qu'une image multispectrale possède au plus 40 bandes spectrales et une image hyperspectrale au moins 41 bandes spectrales et nous précisons le nombre de bandes lorsque nécessaire. L'augmentation du nombre de bandes implique cependant une réduction des dimensions spatiales en raison de contraintes liées aux capteurs d'acquisition (voir Sec. 3.2). Il est alors nécessaire de faire un compromis sur les dimensions de l'image. Une échelle du type de donnée en fonction des résolutions spatiales et de la résolution spatiale est représentée en Figure 1.2.

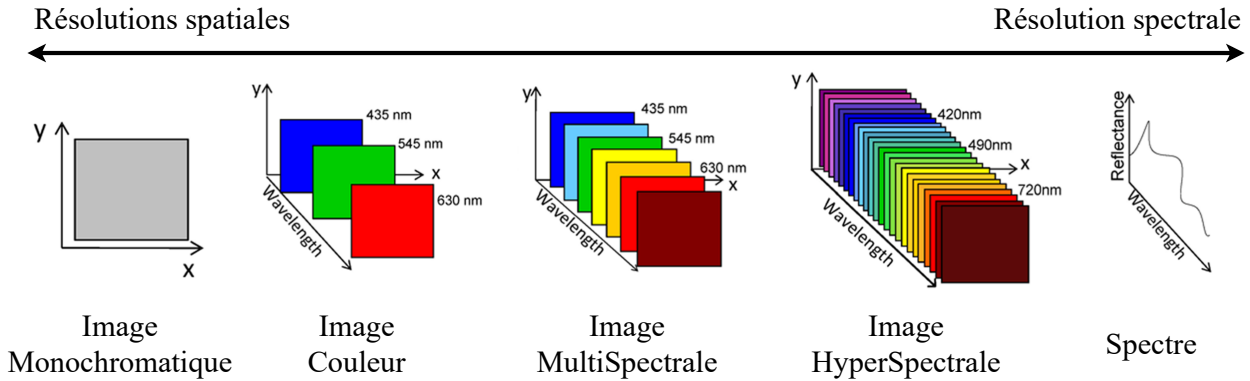


FIGURE 1.2 – Types de données de l’imagerie monochrome à la spectrométrie, en fonction du compromis entre les résolutions spatiales et la résolution spectrale. Figure extraite de [Li *et al.* \[2013\]](#).

1.3 Contexte scientifique

1.3.1 Dimensions spectrale et spatiales

L’apport de l’imagerie hyperspectrale et de l’imagerie multispectrale sont comparés par [Govender *et al.* \[2007\]](#) et [Schlerf *et al.* \[2005\]](#). [Schlerf *et al.* \[2005\]](#) montrent, à travers une analyse de végétation, que l’imagerie hyperspectrale apporte une information plus pertinente que l’imagerie multispectrale, mais pas systématiquement. [Govender *et al.* \[2007\]](#) concluent que l’imagerie hyperspectrale offre davantage de possibilités grâce à sa quantité d’information plus détaillée. Toutefois, les auteurs concluent également que les dimensions des images hyperspectrales sont très limitées par les systèmes d’acquisition existants et que ces images demandent un temps de traitement beaucoup plus important que pour les images multispectrales. En termes d’utilisation, [Shafiee *et al.* \[2016\]](#) recommandent de privilégier l’imagerie multispectrale lorsque les bandes spectrales les plus pertinentes pour une certaine application sont connues.

Afin de bénéficier des différents aspects des images hyperspectrales et multispectrales, il est possible d’effectuer une analyse sur des images hyperspectrales à basse résolution spatiale, de supprimer les bandes les moins pertinentes, puis d’utiliser une caméra multispectrale à plus haute résolution spatiale avec les bandes sélectionnées pour diminuer les contraintes de temps de calcul ou de dimensions spatiales. Dans [Maire et Lefebvre \[2015\]](#), un coefficient est associé à chaque bande spectrale et pondéré à l’aide d’un algorithme génétique afin de ne conserver que les 10 bandes les plus pertinentes. Il est également possible d’éliminer des bandes avec une méthode d’Information Mutuelle (MI) afin de ne conserver que les bandes apportant une information nouvelle. Avec cette

méthode, [Nhaila et al. \[2019\]](#) suppriment jusqu'à 35% des bandes spectrales sans perte des performances de classification, et [Guo et al. \[2006\]](#) montrent qu'il est possible de supprimer près de 60% des bandes de la base de données *Indian Pines* ([Baumgardner et al. \[2015\]](#)) sans perte de précision pour la classification de matériaux. Dans [Khan et al. \[2013\]](#), seules les bandes permettant d'améliorer la précision d'une tâche de détection sont conservées. [Jia et al. \[2012\]](#) présentent une suppression de bandes manuelle complétée par une sélection de bandes par *Affinity Propagation*. Les auteurs démontrent également que la présence de certaines bandes réduit la précision de la classification et que supprimer ces bandes permet d'améliorer cette précision.

Plutôt que supprimer des bandes spectrales et l'information contenue par ces bandes, il est également possible de conserver cette information tout en réduisant le temps de calcul à l'aide de méthodes de réduction de dimension. Après application de ces méthodes, l'information des bandes spécifiques n'est pas préservée, mais une nouvelle information émerge. L'une des méthodes de réduction de dimension les plus utilisées est l'ACP (Analyse en Composantes Principales) ([Farrell et Mersereau \[2005\]](#)). Dans [Bergman \[1996\]](#), seules les 15 premières composantes de l'ACP sont conservées et analysées. [Liu et al. \[2017\]](#) utilisent une ACP pour améliorer les performances de 5 des 7 méthodes de classification proposées en comparaison aux mêmes méthodes appliquées aux images hyperspectrales. Il existe aussi d'autres méthodes de réduction de dimension, telles que le SSFC (*Spectral-Spatial Feature-based Classification*) proposé par [Zhao et Du \[2016\]](#), conçu pour des tâches de classification et qui permet d'atteindre de meilleurs résultats que l'ACP pour ces tâches. Enfin, [Uzair et al. \[2015\]](#) proposent un algorithme de fusion de bandes spectrales dont les performances de classification surpassent celles des autres méthodes proposées, dont l'ACP.

En termes de dimensions spatiales, il est possible d'appliquer une super-résolution (SR) ([Dong et al. \[2015\]](#)) afin de bénéficier d'une plus haute résolution spatiale avec les images hyperspectrales. [Mei et al. \[2017\]](#) proposent un réseau de neurones à convolutions 3D basé sur un mécanisme d'apprentissage (*self-learning*) pour la super-résolution, et ce réseau atteint une meilleure précision que les autres méthodes évaluées. [Sidorov et Hardeberg \[2019\]](#) proposent deux réseaux auto-encodeurs à convolutions respectivement 2D et 3D, mais la précision de ces réseaux est inférieure à celle du réseau proposé par [Mei et al. \[2017\]](#). Une autre approche, basée sur des données acquises depuis plusieurs capteurs, consiste à fusionner une image hyperspectrale à faible résolution spatiale avec une image multispectrale à haute résolution spatiale avec un réseau à auto-encodeurs ([Yao et al. \[2020\]](#)). [Uezato et al. \[2020a\]](#) proposent une approche similaire pour la fusion d'images couleur (RGB) haute-résolution et hyperspectrales faible-résolution. Les auteurs montrent que cette approche est applicable à d'autres types d'images, telles qu'une image panchromatique haute-résolution et une image multispectrale

faible-résolution, ce qui ouvre la voie à des traitements de super-résolution en plusieurs étapes.

1.3.2 Conditions d'acquisition

Lors du traitement d'images hyperspectrales, il est nécessaire de prendre en compte les conditions d'acquisition (capteur utilisé, temps d'exposition, angle et intensité lumineuse...). En effet, les images brutes produites par les capteurs ne sont pas reproductibles (Burger et Geladi [2005]) et sont sujettes à des biais liés au capteur et aux conditions d'illumination. D'après Wendel et Underwood [2017], la calibration d'image rendue possible par la présence d'une cible de calibration dans la scène, c'est-à-dire un objet de réflectance connue et constante dans le domaine spectral observé, est la seule méthode pour calibrer parfaitement les spectres. Les auteurs mentionnent cependant qu'il n'est pas toujours possible d'utiliser une telle cible lors d'acquisitions hyperspectrales. De plus, l'utilisation de plusieurs cibles de calibration ayant différents niveaux de réflectance (par exemple 2%, 50%, 75% et 99%) permet une meilleure calibration des spectres (Geladi *et al.* [2004]; Burger et Geladi [2005]). En raison du prix élevé des standards de cibles de calibration, Shaikh *et al.* [2021] montrent qu'avec un protocole adapté, une cible de calibration de gamme moyenne peut s'avérer suffisante pour la calibration de données hyperspectrales. En revanche, lorsque l'illumination varie fortement entre deux acquisitions de la cible de calibration, la réflectance d'un échantillon varie tout aussi fortement et peut dépasser 100% (Zaman *et al.* [2014]), ce qui ne devrait pas être possible dans un système correctement calibré. Miura et Huete [2009] constatent que ces biais de calibration sont particulièrement importants lorsque l'écart de temps entre l'acquisition de la cible de calibration et celle d'un échantillon est de plus de 15 minutes. Il est alors nécessaire de faire des acquisitions aussi fréquentes que possible de la cible de calibration, ou même d'avoir la cible constamment dans le champ de vision de la caméra.

Ces variations de conditions d'acquisition ou d'illumination peuvent également se présenter au sein d'une même scène observée, principalement en raison de la présence d'ombres (Bondi *et al.* [2016]). Wendel et Underwood [2017] montrent qu'il est possible de corriger partiellement l'illumination lorsque suffisamment de données sont connues à-priori sur celle-ci, mais que l'utilisation de cibles de calibration offre une meilleure correction des variations d'illumination. Dans Uezato *et al.* [2020b], une approche multimodale est proposée, dans laquelle un LiDAR est utilisé conjointement à une caméra hyperspectrale. Ce LiDAR est utilisé pour calculer la hauteur des différents pixels dans la scène et estimer la position des zones d'ombre afin de les corriger. Le problème de la calibration dans des conditions variables d'illumination n'est cependant pas résolu (Aasen *et al.* [2018]).

L'illumination peut aussi être perturbée de plusieurs manières par les parti-

cules contenues dans l’atmosphère (Adler-Golden *et al.* [1999]). Tout d’abord, les particules contenues dans l’atmosphère dispersent la lumière. Cette dispersion est particulièrement observable sur des images acquises par des systèmes aériens ou satellitaires car elle crée un « effet d’adjacence », c’est-à-dire que la lumière d’un objet au sol est dispersée dans les pixels adjacents à ceux qui visent cet objet. Les auteurs utilisent un modèle physique pour proposer une correction de cet effet d’adjacence à partir d’une estimation de celui-ci. L’autre effet mis en valeur par Adler-Golden *et al.* [1999] est celui de l’« absorption atmosphérique » : certaines longueurs d’onde du spectre électromagnétique sont absorbées par l’atmosphère, et une partie de la lumière émise par le soleil n’atteint donc pas la surface de la terre. De nombreux modèles physiques permettant de simuler l’absorption atmosphérique ont été créés (Healey *et al.* [1999]; Poutier *et al.* [2002]; Miesch *et al.* [2005]; Thomas *et al.* [2011]) afin de réduire l’impact de l’atmosphère sur les spectres. Plutôt que de modéliser cet effet et de corriger les spectres, Adler-Golden *et al.* [2001] proposent une méthode de discrimination de spectres qui prend en compte les variations spectrales liées à l’absorption atmosphérique et atteint des performances satisfaisantes sur des échantillons acquis dans des conditions différentes.

1.3.3 Application des images hyperspectrales

Une fois traitées, ces données hyperspectrales, généralement satellitaires, sont principalement utilisées pour la classification de matériaux. Landgrebe [2002] classe les pixels en utilisant l’algorithme ECHO (Extraction and Classification of Homogeneous Objects). Villa *et al.* [2013] projettent les pixels spectraux, représentés dans un espace à 216 dimensions, vers un espace à 2 dimensions afin d’utiliser les regroupements (ou *clusters*) de pixels pour la classification. D’autres utilisent les Machines à Vecteurs de Support (SVM) pour la classification de matériaux (Camps-Valls et Bruzzone [2004]; Melgani et Bruzzone [2004]; Fauvel *et al.* [2007]). Ouerghemmi *et al.* [2017] proposent d’utiliser deux SVM sur des images hyperspectrale et multispectrale d’une même scène et de fusionner les décisions des deux SVM, améliorant ainsi les performances globales. Cependant, depuis près d’une décennie, l’émergence des réseaux de neurones pour la classification a permis de surpasser les résultats des méthodes classiques utilisées jusque là (Paoletti *et al.* [2019]). Parmi les approches se basant sur des réseaux de neurones, certaines utilisent des réseaux à convolution (CNN) (Zhao et Du [2016]) afin de bénéficier des caractéristiques spatiales, et d’autres utilisent des réseaux de neurones récurrents (RNN) (Guo *et al.* [2018]; Mou *et al.* [2017]) afin de bénéficier des caractéristiques spectrales. D’autres méthodes combinent ces deux approches dans des « réseaux récurrents à convolution » (RCNN) (Mou *et al.* [2019]) ou « réseaux à convolution récurrents » (CRNN) (Wu et Prasad [2017]) afin de bénéficier à la fois des caractéristiques spatiales et spectrales des images hyperspectrales.

D'autres types de réseaux sont également utilisés, par exemple [Zhu et al. \[2018\]](#) utilisent un réseau antagoniste génératif (GAN) avec convolutions 3D pour la classification de matériaux.

Cependant, ces travaux de classification utilisent pratiquement tous les mêmes données, des images hyperspectrales issues de challenges publics ([Ma et al. \[2019\]](#)). Ainsi, les résultats de l'état de l'art sont saturés en raison du manque de données hyperspectrales publiques ([Ghamisi et al. \[2017\]](#)).

D'autres publications présentent des méthodes de classification sur de nouvelles données. [Lagrange et al. \[2017\]](#) utilisent un modèle bayésien hiérarchique pour la classification et le démelange de spectres de végétation dans leurs propres images aériennes. [Parsons et al. \[2018\]](#) détectent et classifient des échantillons de corail avec une SVM dans leurs propres images aériennes. [Amigo et al. \[2015\]](#) utilisent un PLS-DA (Partial Least Squares-Discriminant Analysis) pour la classification de pixels et pour la discrimination d'échantillons de plastique après plusieurs étapes de pré-traitements. Dans [Shafiee et al. \[2016\]](#), un réseau de neurones non-profond est utilisé pour classifier des échantillons de miel purs et frelatés avec 95% de précision. [Khan et al. \[2013\]](#) montrent l'apport de l'imagerie hyperspectrale en classifiant des échantillons d'encre qui apparaissent identiques pour les caméras couleur avec une méthode de regroupement (ou *clustering*). Dans [Shaikh et al. \[2021\]](#), une ACP et une SVM sont utilisées pour classifier quelques échantillons de neige et d'asphalte acquis dans des conditions variables d'illumination pour prouver la fiabilité de leur protocole de calibration. Toutefois, ces nouvelles données sont rarement rendues publiques, et ne peuvent donc pas être exploitées par l'ensemble de la communauté scientifique.

1.3.4 Imagerie hyperspectrale pour la surveillance

Dans cette thèse, ce sont plutôt ces dernières applications de la classification qui nous intéressent, mais appliquées au domaine de la surveillance. Pour la caractérisation de matériaux artificiels liés à des contextes urbains, [Herold \[2004\]](#); [Herold et al. \[2004\]](#); [Roberts et Herold \[2004\]](#) proposent une base de données d'échantillons d'asphalte, mais ces données sont purement spectrales et ne contiennent aucun contexte spatial.

En termes d'application dans la surveillance terrestre, d'autres travaux considèrent la tâche du pistage. La méthode de pistage hyperspectrale proposée par [Xiong et al. \[2019\]](#) obtient de meilleures performances que les méthodes de pistage couleur et que la plupart des autres méthodes hyperspectrales. [Xiong et al. \[2020\]](#) proposent une base de données de vidéos acquises simultanément en couleur et en multispectral (16 bandes) avec les mêmes caractéristiques. Les auteurs montrent que la plupart des méthodes de pistages sur des vidéos couleur sous-performent sur la vidéo multispectrale. Cependant, la méthode proposée ne permet pas d'améliorer la performance du pistage. [Li et al. \[2020\]](#)

proposent d'intégrer de nouvelles vidéos à la base de données précédente. Une nouvelle méthode est proposée, mais ses performances sont en-deçà de celles de l'état de l'art. Dans [Nguyen et al. \[2010\]](#), l'utilisation de l'imagerie hyperspectrale permet d'améliorer la qualité du pistage et de l'identification de visages par rapport à des images couleur, en se concentrant sur le pistage de visages ou de vêtements.

Mais en plus des limitations mentionnées précédemment des capteurs multispectraux et hyperspectraux, l'acquisition de vidéos est d'autant plus limitée par leur nombre d'images par seconde. Afin de réduire ces limitations, [Mian et Hartley \[2012\]](#) proposent un système conjoint d'acquisition et traitement d'images permettant de restaurer des vidéos multispectrales et hyperspectrales de meilleure résolution spatiale, spectrale et temporelle. [Meng et al. \[2020\]](#) construisent et présentent un capteur personnalisé pour l'acquisition de vidéos multispectrales (28 bandes) par filtrage des données acquises par le capteur et par une reconstruction du signal avec un réseau de neurones 3D à « *self-attention* ».

D'autres travaux proposent des applications diverses liées à la surveillance. [Uzair et al. \[2014\]](#) et [Uzair et al. \[2015\]](#) proposent des méthodes de reconnaissance de visage dans des images hyperspectrales et leurs résultats surpassent de loin ceux des autres méthodes analysées. [Yuen et Richardson \[2010\]](#) présentent l'imagerie hyperspectrale pour la détection d'objets artificiels et de mines en milieux naturels ainsi que pour de la discrimination de matériaux (vêtements contre véhicules). Le Groupe de Synthèse Hyperspectrale a établi un rapport ([Groupe de Synthèse Hyperspectral \(GSH\) \[2008\]](#)) présentant des applications de l'imagerie hyperspectrale en contexte urbain et pour la défense, telles que le ciblage, la détection de camouflage et l'analyse de terrain. Les auteurs recommandent le domaine spectral 400-2500nm pour ces types d'applications, et recommandent de privilégier des systèmes multi-modaux (hyperspectral avec RADAR) pour certains besoins de défense tels que la détection de camouflage. Dans [Goldberg et al. \[2003\]](#), l'utilisation de bandes spécifiques (capteur multispectral) dans l'infrarouge moyen et long (MWIR-LWIR) est proposé pour la détection de véhicules tactiques, de mines et de missiles ballistiques, y compris en conditions nocturnes. [Bergman \[1996\]](#) et [Sanders et al. \[1998\]](#) effectuent respectivement de la discrimination et de la détection de matériaux en simulant des conditions opérationnelles, avec des images aériennes acquises à plusieurs altitudes différentes. Dans [Sanders et al. \[1998\]](#), en raison de ressemblances importantes entre les cibles et l'arrière-plan, la détection de matériaux échoue, et l'amélioration des dimensions spatiales des images par réduction de l'altitude ne permet pas d'améliorer les résultats de cette détection. Dans [Bergman \[1996\]](#), les acquisitions sont conduites à plusieurs moments de la journée afin d'obtenir une variabilité liée à l'illumination. Les matériaux représentés sur les données (tissus, plastique, métal) sont correctement discriminés entre eux et c'est également le cas pour la discrimination entre des véhicules réels et

des leurres. En revanche, la discrimination des différents types de véhicules n'apporte pas de résultats concluants.

1.4 Contexte opérationnel

Dans ce cadre, l'utilisation de l'imagerie hyperspectrale pourrait permettre d'améliorer les systèmes existants grâce à l'apport de nouvelles informations telles que les matériaux, ou la détection d'objets non perçus par les types de capteurs actuellement utilisés. Pour cela, il nous est nécessaire de connaître les caractéristiques spectrales de matériaux pouvant apparaître dans des scénarios de missions de surveillance : acier, verre, béton, bois, sols...

1.5 Conclusion

Dans ce chapitre, nous avons présenté l'imagerie hyperspectrale. Nous avons analysé et décrits les travaux existants afin de comprendre les problématiques liées à l'utilisation de l'imagerie hyperspectrale et afin d'en extraire des méthodes et applications pertinentes dans le contexte des systèmes de mission de surveillance.

Chapitre 2

Études spectrales

Sommaire

2.1	Introduction	13
2.2	Bases de données	14
2.3	Transformations des spectres	14
2.4	Analyse visuelle de spectres de plastique	18
2.5	Mesures de similarité	19
2.5.1	Définition	19
2.5.2	Analyse des spectres de plastique	21
2.5.3	Variabilité spectrale	21
2.5.4	Mélanges	23
2.6	Extraction des pics	25
2.7	Conclusion	26

2.1 Introduction

L'imagerie hyperspectrale contient des données définies à la fois selon des dimensions spatiales et spectrale. Une exploitation préalable de la dimension spectrale permet de définir son utilité ainsi que les enjeux de son utilisation dans des contextes de surveillance. Dans ce chapitre, nous présentons d'abord une liste de bases de données spectrales existantes. Nous concentrons ensuite l'étude sur deux de ces bases de données pour l'extraction et l'analyse de caractéristiques spectrales. Des méthodes d'analyse et de traitement sont appliquées à ces spectres afin de définir les domaines spectraux d'intérêt pour la suite de nos travaux.

2.2 Bases de données

Nous présentons en Table 2.1 une liste des bases de données spectrales existantes, le type d’objets ou de matériaux qu’elles contiennent ainsi que les domaines spectraux couverts par chaque base de données. Lorsque deux domaines spectraux sont listés pour une même base, cela signifie que les spectres de la base ont des domaines spectraux différents entre eux, situés entre les deux domaines spectraux présentés.

Base de données	Contenu	Domaine spectral
CAVE (Yasuma <i>et al.</i> [2010])	Objets divers	400-700nm
CSIRO (CSIRO [2020])	Minéraux	300-2500nm 6000-15000nm
ECOSTRESS (Baldrige <i>et al.</i> [2009] (Meerdink <i>et al.</i> [2019])	Végétation Matériaux artificiels	300-2500nm 2500nm-25000nm
GHISA (Thenkabail <i>et Aneece</i> [2019])	Végétation	350-2500nm
RELab (Pieters [1983])	Minéraux	220-450nm 16673-200000nm
Santa Barbara Urban Spectral Library (Herold <i>et al.</i> [2004])	Asphalte	440-2370nm
USGS7 (Kokaly <i>et al.</i> [2017])	Objets et matériaux divers Matériaux artificiels	205-1123nm 2202-216006nm

TABLE 2.1 – Liste des bases de données spectrales

Pour la suite de ce chapitre, nous avons sélectionné les bases de données *RELab* et *USGS7* en raison du grand nombre de spectres qu’elles contiennent. Plus précisément, *RELab* est utilisée afin de décrire certains aspects fondamentaux de la variabilité spectrale des matériaux et des mélanges de matériaux, et *USGS7* est utilisée dans l’objectif de classifier des objets ou matériaux artificiels en fonction de leurs caractéristiques spectrales.

2.3 Transformations des spectres

Spectres

Un spectre prend la forme d’un signal 1D exprimé par une intensité en fonction de la longueur d’onde, voir Figure 2.1. Cette intensité est généralement exprimée en AREF (*Absolute REFlectance* ou « réflectance absolue »),

c'est-à-dire la proportion de la lumière qui est réfléchiée par l'objet analysé, indépendamment pour chaque longueur d'onde.

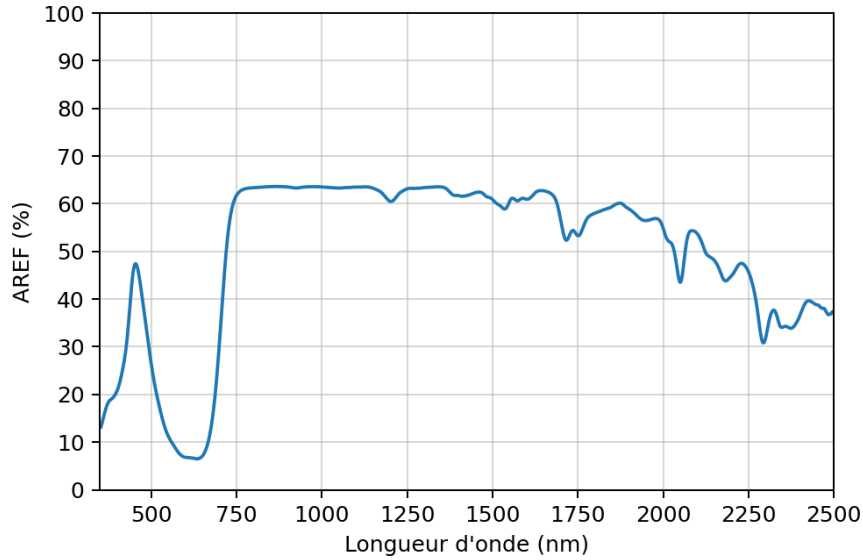


FIGURE 2.1 – Exemple de spectre d'un objet bleu en nylon, par AREF en fonction de la longueur d'onde de la lumière, de 350 à 2500nm.

Normalisation linéaire

La normalisation linéaire consiste à ne considérer que l'évolution du spectre et s'opère en convertissant l'ensemble des valeurs d'intensité d'un spectre vers l'intervalle de valeurs $[0, 1]$. Cette opération peut permettre de trouver des corrélations entre des spectres présentant une forme similaire mais une AREF moyenne différente. En revanche, la normalisation ne permet pas de conserver les informations de quantité de AREF des spectres. Un exemple de normalisation est présenté en Figure 2.2.

Distribution

La distribution est une autre méthode de normalisation des spectres, qui permet d'éviter la perte de l'information de l'écart-type. Chaque valeur du spectre est divisée par la somme des valeurs de ce spectre, résultant en un vecteur de probabilités. L'inconvénient de cette méthode est que la moyenne de la distribution dépend directement du nombre de valeurs contenu dans le spectre. Ainsi, il n'est pas possible, avec cette méthode, de comparer des spectres définis sur des domaines spectraux différents ou possédant un nombre de bandes différents. Un exemple de distribution est présenté en Figure 2.3.

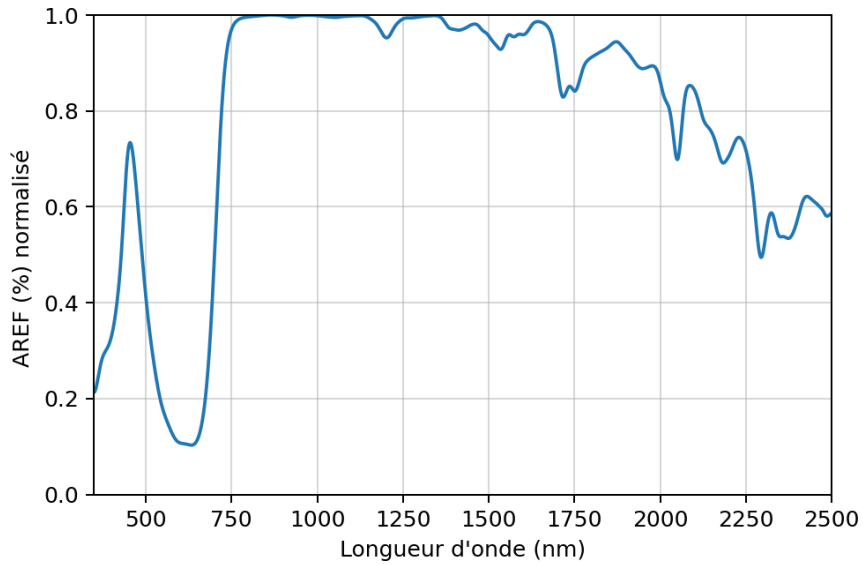


FIGURE 2.2 – Spectre de la Figure 2.1 après normalisation, de 350 à 2500nm.

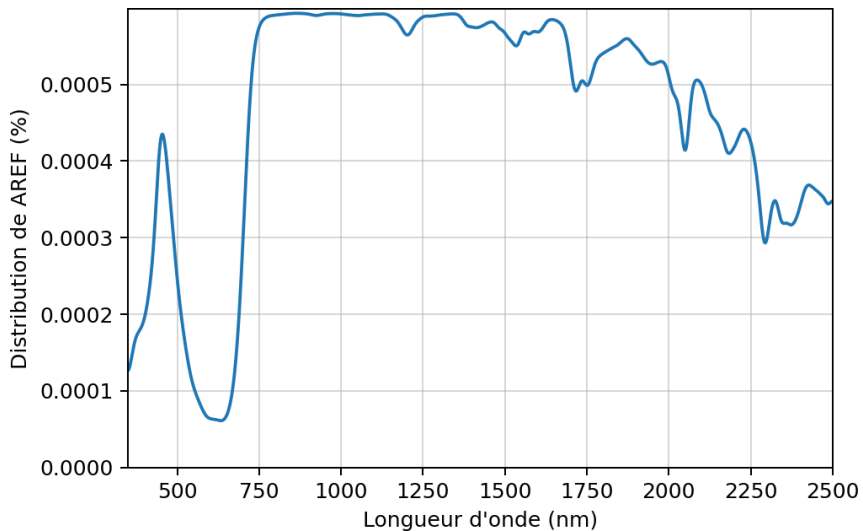


FIGURE 2.3 – Spectre de la Figure 2.1 après distribution, de 350 à 2500nm.

La normalisation et la distribution sont deux méthodes permettant de visualiser un même spectre à des échelles différentes. Nous allons ensuite présenter des méthodes appliquant des transformations au spectre lui-même.

Dérivées première et seconde

La dérivée d'un spectre est une mesure de l'évolution locale du signal. Nous avons extrait les dérivées première et seconde aux spectres. La dérivée première d'un spectre correspond à la variation ou la pente du signal en chaque point de

2. Études spectrales

celui-ci. La dérivée seconde correspond aux variations de la dérivée première, et peut permettre de trouver certains pics d'intensité faibles masqués par des pics plus importants. Les dérivées évaluant les variations présentes dans un spectre, une normalisation linéaire ou une distribution du spectre n'ont d'impact que sur la magnitude des dérivées. Nous présentons en Figures 2.4a et 2.4b des exemples de dérivée première et seconde, respectivement.

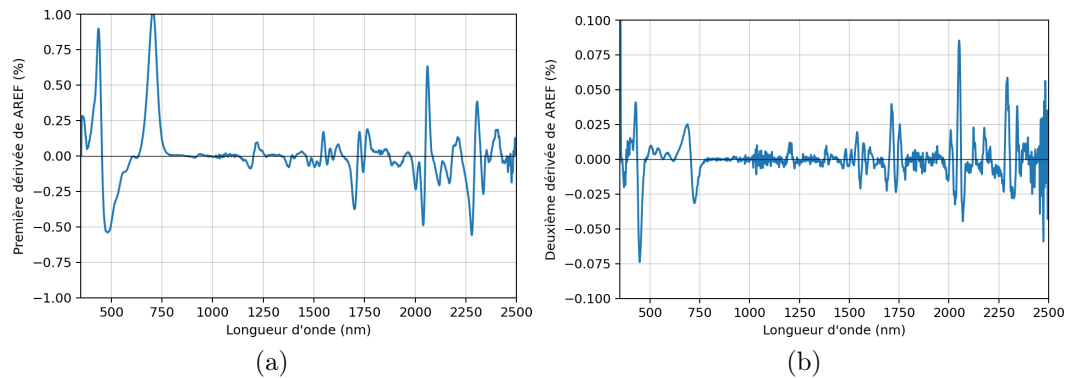


FIGURE 2.4 – Dérivée (a) première et (b) seconde du spectre de la Figure 2.1, de 350 à 2500nm.

Lissage

Lisser un spectre permet de réduire son bruit tout en préservant le signal qu'il contient. Certains spectres possèdent un niveau de bruit important, sous la forme de fortes variations locales, par exemple en raison d'imprécisions du capteur. Il existe plusieurs fonctions permettant de lisser les spectres, nous avons choisi d'appliquer un filtre gaussien sur le voisinage de chaque point du spectre en raison de sa qualité de préservation du signal. L'effet du lissage spectral a un effet particulièrement important sur les dérivées du spectre lissé, car celles-ci représentent les variations de ce spectre.

Nous présentons en Figure 2.5 des exemples de dérivée première et seconde après lissage. Nous pouvons constater que les dérivées ont des variations locales beaucoup moins importantes, ce qui permet de mieux identifier certaines variations du signal, par exemple entre 1000nm et 1250nm sur la deuxième dérivée. Cependant, ce lissage peut également détruire une partie de l'information spectrale, et ne doit pas être utilisé systématiquement.

Histogramme

L'histogramme est une représentation de la répartition des valeurs d'un spectre. L'histogramme dépend fortement de l'échelle du spectre dont il est issu. À l'inverse des dérivées, l'histogramme préserve les informations globales

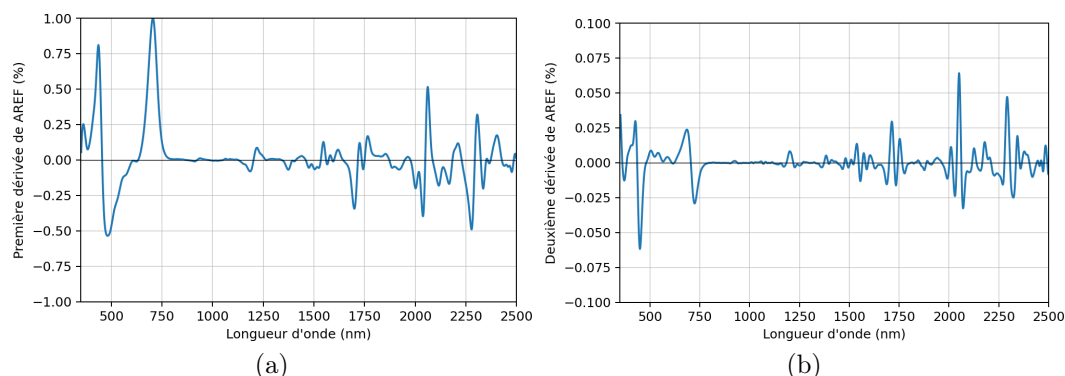


FIGURE 2.5 – Dérivée (a) première et (b) seconde du spectre de la Figure 2.1 après lissage du spectre avec un filtre gaussien, de 350 à 2500nm.

des spectres au détriment des variations locales. Un exemple d’histogramme est présenté en Figure 2.6. Nous pouvons observer que pour le spectre exemple (voir Figure 2.1), la majorité des valeurs ont une AREF comprise entre 60% et 65%.

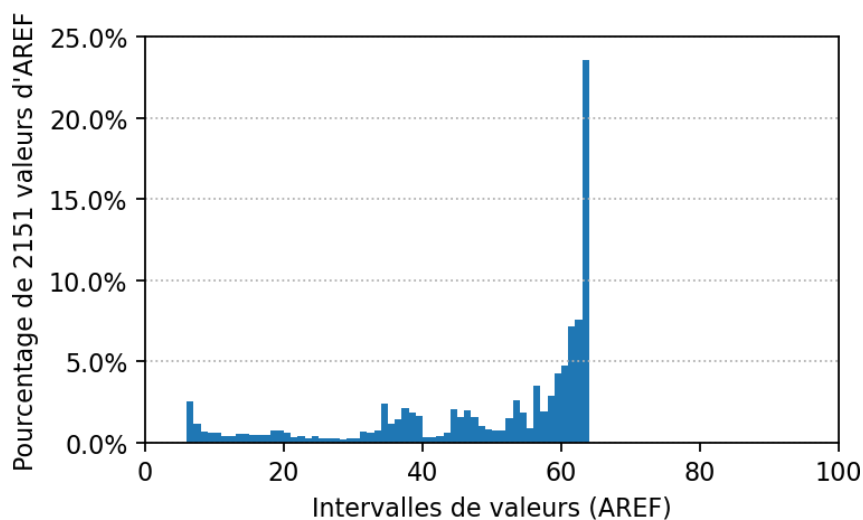


FIGURE 2.6 – Histogramme des valeurs de AREF du spectre de la Figure 2.1, de 350 à 2500nm.

2.4 Analyse visuelle de spectres de plastique

Afin de proposer des perspectives à l’utilisation de spectres pour la caractérisation de matériaux, nous proposons une analyse visuelle d’objets en plastique (nylon et vinyle), dont les spectres sont présentés en Figure 2.7. En-dessous de 1200nm, les spectres n’existent que peu de caractéristiques

communes et sont très différents pour un même type de matériau, particulièrement lorsque les objets observés sont visuellement de couleurs différentes. Cependant, entre 1200 et 2300nm, des caractéristiques spectrales similaires apparaissent entre les échantillons de même matériau et différentes des caractéristiques d'autres matériaux.

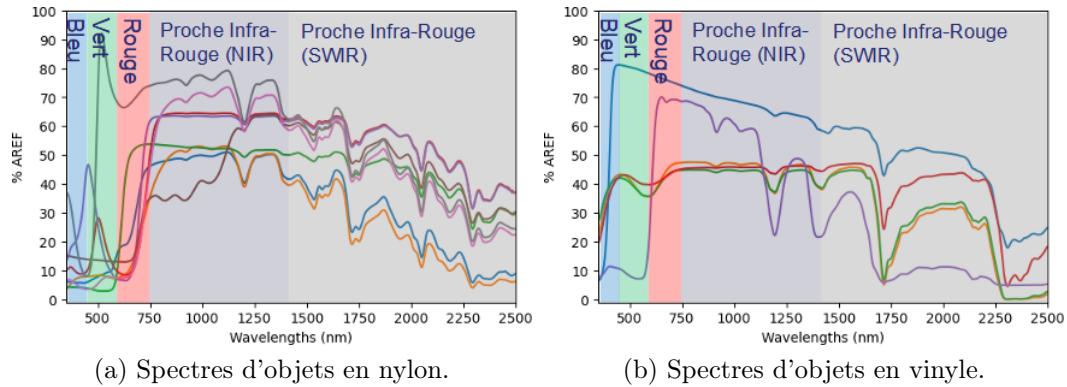


FIGURE 2.7 – Spectres d'objets en (a) nylon et (b) vinyle en fonction du domaine spectral, de 350 à 2500nm.

De plus, cette analyse ne se limite pas aux échantillons présentés, car des caractéristiques semblables sont trouvées entre les échantillons des autres classes de matériaux (plastiques, béton...). Cela nous amène à conclure que le domaine spectral 1200-2300nm est le plus adapté à la caractérisation de matériaux artificiels.

2.5 Mesures de similarité

2.5.1 Définition

En fonction de ses propriétés, une fonction $f(s, t)$ permettant de calculer une différence quantitative entre deux données s et t de même type est appelée une « mesure de similarité » (lorsque $f(s, s) = 0$) ou une « mesure de dissimilarité » (lorsque $f(s, s) \neq 0$). Ces fonctions sont particulièrement adaptées pour comparer des spectres entre eux, à condition que les spectres s et t comparés soient de la même taille L .

Une mesure de similarité $f(s, t)$ est appelée « mesure de distance » lorsqu'elle respecte 3 propriétés : la réflexivité ($f(s, s) = 0$), la non-négativité ($f(s, t) \geq 0$) et la symétrie ($f(s, t) = f(t, s)$).

Les mesures de (dis)similarité ou de distance comparent chaque valeur du vecteur s à la valeur correspondante du vecteur t . Chaque valeur n'est alors considérée qu'une fois et individuellement afin d'évaluer sa différence avec la valeur correspondante de l'autre spectre.

Spectral Angle Mapper (SAM)

Le Spectral Angle Mapper (SAM) (Yuhua *et al.* [1992]) est l'une des mesures de similarité les plus utilisées pour comparer deux spectres (Deborah *et al.* [2015]). Cette méthode consiste à considérer les spectres s et t , tous deux de taille L , comme des points dans un espace à L dimensions et à calculer l'angle entre ces deux points. Le SAM est exprimé à partir de l'expression suivante :

$$SAM(s, t) = \cos^{-1} \left(\frac{\sum_{\lambda}^L s_{\lambda} t_{\lambda}}{\sqrt{\sum_{\lambda}^L s_{\lambda}^2} \sqrt{\sum_{\lambda}^L t_{\lambda}^2}} \right). \quad (2.1)$$

Distance Euclidienne (ED)

La Distance Euclidienne considère également les vecteurs comme des points dans un espace à L dimensions, mais calcule la distance entre ces points. Elle est exprimée par :

$$ED(s, t) = \sqrt{\sum_{\lambda}^L (s_{\lambda} - t_{\lambda})^2}. \quad (2.2)$$

Divergence Kullback-Leibler (KL)

L'expression de la divergence Kullback-Leibler (KL) (Kullback et Leibler [1951]) entre deux densités de probabilité $p(\cdot)$ et $q(\cdot)$ est donnée par :

$$KL(p, q) = \int_{-\infty}^{+\infty} p(x) \log \left(\frac{p(x)}{q(x)} \right) dx. \quad (2.3)$$

Dans notre cas, nous considérerons des probabilités discrètes telles que la distribution d'un spectre ou la distribution de l'histogramme d'un spectre. On peut alors calculer la KL discrète entre les densités de probabilité discrètes P et Q telle que :

$$KL(P, Q) = \sum_i P(i) \log \left(\frac{P(i)}{Q(i)} \right). \quad (2.4)$$

Nous notons que la divergence Kullback-Leibler n'est pas une mesure de distance car elle ne respecte pas la symétrie ($KL(P, Q) \neq KL(Q, P)$). Afin de respecter la symétrie, il est possible de considérer la somme $KL(P, Q) + KL(Q, P)$ comme mesure de distance.

Tanimoto Coefficient (TC)

Le Coefficient de Tanimoto (TC) (Tanimoto [1957]) est une autre manière d'évaluer la dissimilarité entre deux spectres. Il est défini par :

$$TC(s, t) = \frac{\sum_{\lambda}^L s_{\lambda} t_{\lambda}}{\sum_{\lambda}^L s_{\lambda}^2 + \sum_{\lambda}^L t_{\lambda}^2 - \sum_{\lambda}^L s_{\lambda} t_{\lambda}}. \quad (2.5)$$

Le TC n'est pas une mesure de distance car il ne respecte pas la réflexivité ($TC(x, x) \neq 0$, car $TC(x, x) = 1$).

2.5.2 Analyse des spectres de plastique

Pour corroborer les résultats de notre analyse visuelle, nous proposons de calculer les valeurs de SAM entre des échantillons de plastique (nylon, vinyl) et un échantillon de goudron sur les domaines spectraux 350-2500nm et 1200-2300nm. Ces valeurs sont affichées en Figure 2.8. Selon l'analyse conduite en Section 2.4, le domaine 350-1200nm contient des informations indépendantes du type de matériau. Les échantillons d'un même type de matériau devraient alors avoir des valeurs de SAM beaucoup plus faibles sur le domaine 1200-2300nm (voir Fig. 2.8b) que sur le domaine 350-2500nm (voir Fig. 2.8a).

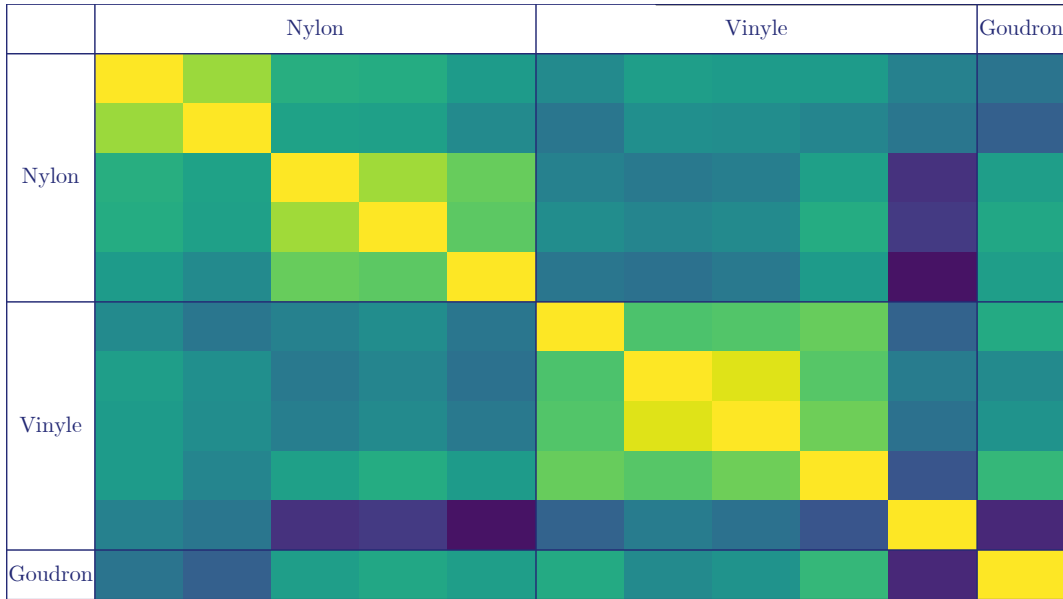
Comme attendu, à part quelques exceptions, les valeurs de SAM sont plus faibles sur le domaine 1200-2300nm (Fig. 2.8b). En revanche, les échantillons dont la valeur de SAM a été la plus réduite ne sont pas ceux appartenant à une même classe. Par exemple, alors que le spectre de l'échantillon de goudron est supposément très différent des autres échantillons, ses valeurs de SAM sont du même ordre de grandeur. De plus, l'un des échantillons de vinyle présente des valeurs de SAM importantes dans les deux domaines spectraux étudiés. Il s'agit du spectre ayant un pic important dans le rouge, présenté en Figure 2.7b. Ce spectre présente des variations importantes, particulièrement dans l'infrarouge, ce qui peut expliquer ces valeurs extrêmes de SAM.

Cette analyse signifie que, pour les matériaux étudiés, la sélection du domaine spectral 1200-2300nm permet de réduire la distance évaluée entre les spectres. Cependant, cette réduction se fait de manière indiscriminée et ne permet pas de rapprocher entre eux les échantillons appartenant pourtant à une même classe.

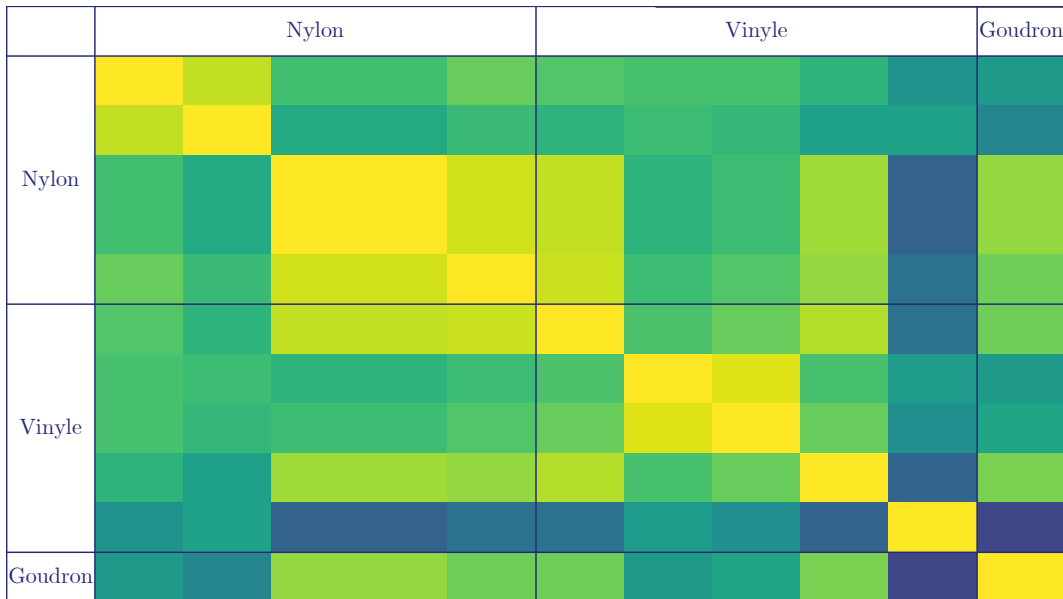
2.5.3 Variabilité spectrale

La variabilité spectrale désigne le fait que plusieurs spectres, produits par une même méthode d'acquisition, dans les mêmes conditions et sur un même type de matériau, ne sont pas nécessairement identiques. Ces différences peuvent venir de plusieurs facteurs et produire des effets différents sur les spectres. Par exemple, Schneider *et al.* [2009] analysent l'impact de la luminosité et de l'humidité sur des échantillons de roche et observent que la luminosité impacte la magnitude du spectre et peu la forme de celui-ci, tandis que l'humidité affecte fortement la forme du spectre mais peu sa magnitude.

Nous présentons en Figure 2.9 les valeurs de SAM (*Spectral Angle Mapper*) d'échantillons de quartz de différentes tailles jusqu'à $180\mu m$, humides ou secs. Dans un premier temps, nous pouvons constater que la taille des échantillons a un impact très important sur la valeur du SAM, et les spectres les plus



(a) Valeurs de SAM sur le domaine spectral 350-2500nm.



(b) Valeurs de SAM sur le domaine spectral 1200-2300nm.

FIGURE 2.8 – Valeurs de SAM entre des spectres de nylon, de vinyl et de goudron sur deux domaines spectraux. Les valeurs de SAM s’étendent de 0 (jaune, échantillons similaires) à 0.8 (bleu, échantillons différents).

différents sont ceux acquis sur des échantillons de taille différente. L’humidité des échantillons a également un impact sur la valeur du SAM, mais celui-ci est beaucoup plus faible que l’impact de la taille des échantillons.

La variabilité spectrale, analysée ici sous les aspects de l’humidité et surtout

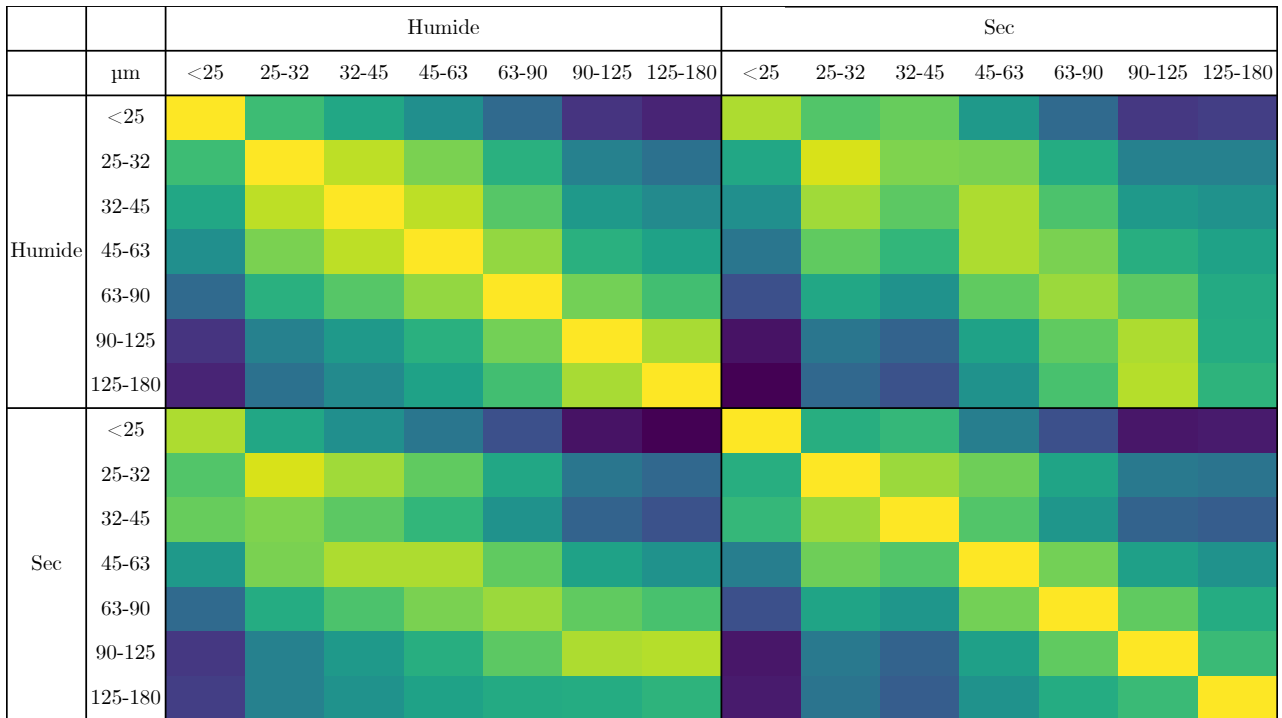


FIGURE 2.9 – Valeurs de SAM entre les spectres d’amas de quartz de différentes tailles. Les valeurs de SAM s’étendent de 0 (jaune, échantillons similaires) à 0.253 (bleu, échantillons différents).

de la taille des échantillons, peut donc impacter de manière importante des échantillons qui sont pourtant constitués du même type de matériau. Lors du traitement de spectres, il est alors nécessaire de prendre en compte cette variabilité, par exemple à travers l’acquisition de plusieurs échantillons d’un même type de matériau.

2.5.4 Mélanges

En plus de la variabilité spectrale, qui affecte les spectres des matériaux purs, les spectres de mélanges de matériaux peuvent être affectés de deux manières différentes : on parle alors de mélanges linéaires et non-linéaires.

Un mélange linéaire se produit typiquement lorsque les échantillons de matériaux différents sont séparés, mais spatialement proches. La lumière perçue par le capteur sur la délimitation entre ces deux échantillons contient alors une partie de la lumière réfléchiée par chacun des matériaux. On parle de mélange linéaire car le spectre composite acquis correspond à l’addition des spectres des différents échantillons purs.

Un mélange non-linéaire se produit lorsque l’échantillon lui-même est un composite de plusieurs matériaux. Le spectre produit a alors des caractéris-

tiques différentes de celles des spectres des matériaux purs qui le composent. Nous présentons en Figure 2.10 des spectres de mélanges de diaspore et de quartz. Vers 4000nm, le mélange constitué de 25% de diaspore et 75% de quartz a un spectre beaucoup plus proche de celui du quartz que de celui du quartz, malgré sa faible composition en diaspore. Sur le pic d'AREF autour de 6800nm, la mélange 25/75 présente une réflectance absolue plus élevée que chacun de ses composants, montrant l'apparition de caractéristiques nouvelles dans ce mélange.

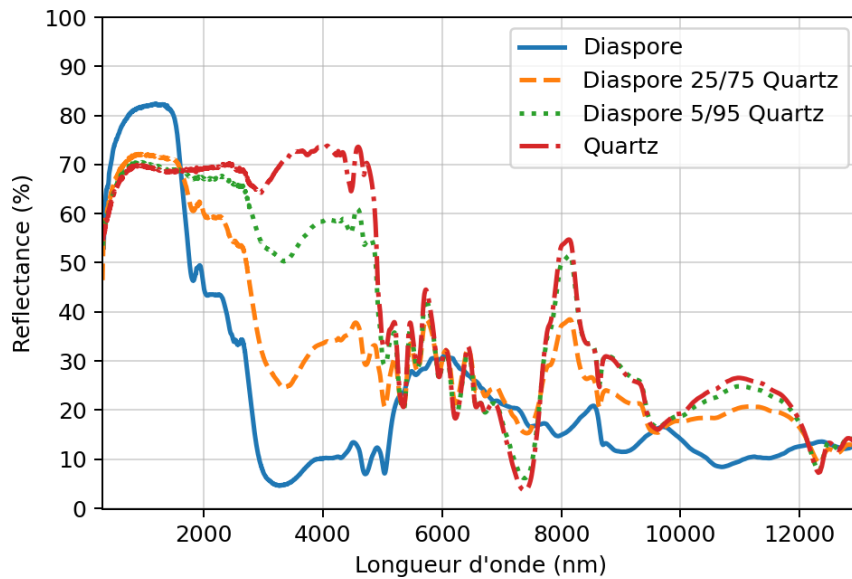


FIGURE 2.10 – Spectres de mélanges de diaspore et de quartz, de 300 à 13000nm.

Il existe un deuxième type de mélange non-linéaire, qui ne se produit pas en mélangeant physiquement des matériaux, mais lorsque la lumière reçue par un échantillon est celle qui a été réfléchié ou transmise par un premier objet, car celui-ci absorbe une partie de la lumière qu'il reçoit. Ce type de mélange est en réalité très fréquent lors d'acquisitions avec éclairage solaire, avec l'atmosphère absorbant une partie de la lumière qui la traverse.

Afin de retrouver les matériaux purs dont sont issus certains mélanges, il est possible d'appliquer un « démélange spectral » (Bioucas-Dias *et al.* [2012]; Farah *et al.* [2010]; Zare et Ho [2014]). Le démélange est utile dans de nombreux domaines, mais particulièrement dans les cas où la résolution au sol est élevée (par ex. pour l'imagerie satellitaire, avec des pixels $> 10\text{cm}$) ou dans les scènes avec un relief important et donc des réflexions importantes entre les différents échantillons. Le démélange est une tâche complexe, sur laquelle de nombreux travaux sont menés y compris à ce jour (Janiczek *et al.* [2020]; Borsoi *et al.* [2021]; Zhao *et al.* [2019]). Par la suite, nous prendrons en compte les problématiques de mélange linéaire et non-linéaire lorsqu'elles sont applicables,

sans utiliser de méthode de démixage sauf indication contraire.

2.6 Extraction des pics

Une autre méthode d'analyse spectrale est celle de l'extraction des pics. Contrairement aux mesures de similarité, cette méthode ne permet pas d'évaluer une différence entre deux spectres, mais d'évaluer et de localiser les variations présentes dans un spectre. De plus, les pics extraits sont des caractéristiques qui pourraient permettre d'établir des caractéristiques communes aux échantillons d'une même classe d'objet et qui diffèrent avec les caractéristiques d'autres classes d'objets.

Un pic correspond à un point du spectre dont l'intensité est supérieure à celle de son voisinage. Un pic est caractérisé par 3 informations :

- sa localisation, ou la longueur d'onde la plus élevée de la région ;
- sa hauteur, ou différence d'intensité avec la longueur d'onde la plus faible de la même région ;
- sa largeur à mi-hauteur, ou plage de longueurs d'ondes impactées par ce pic.

En plus des pics, nous considérons également les creux, c'est-à-dire les points du spectre dont l'intensité est inférieure à celle de leur voisinage, avec leur hauteur et leur largeur à mi-hauteur.

Nous présentons en Figure 2.11 l'extraction des pics et des creux sur deux spectres de tissu nylon. Sur ces deux spectres représentant un même type de matériau, les pics les plus proéminents sont distincts, tant en localisation qu'en hauteur et en largeur à mi-hauteur, car l'AREF du spectre 2.11a est plus élevée autour de 1000nm.

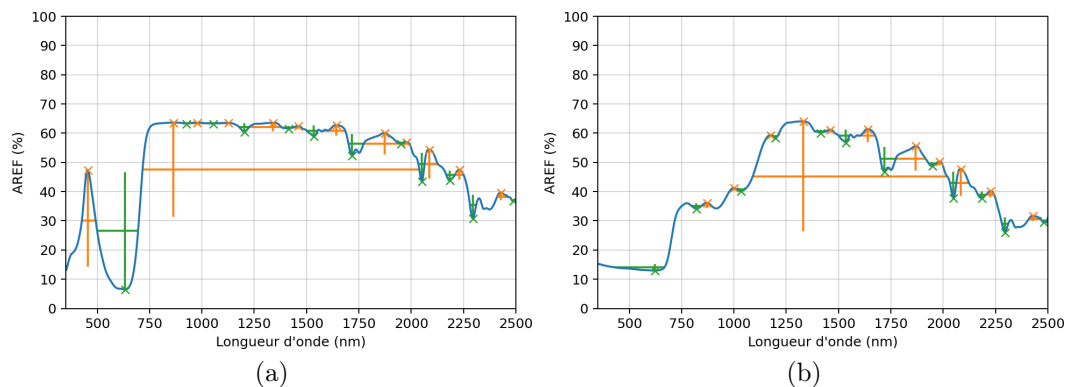


FIGURE 2.11 – Extraction des pics (jaune) et des creux (vert) ainsi que de leur hauteur et largeur à mi-hauteur, sur le spectre de deux objets en tissu nylon.

De plus, les pics et creux qui apparaissent similaires présentent des légères variations entre les spectres, en raison de l'aspect local de l'information extraite par cette méthode. Ces variations compliquent l'extraction de caractéristiques communes à des classes : plus le nombre de spectres représentant cette classe est élevé, plus le nombre de corrélations entre les échantillons d'une même classe est faible. Une autre limitation de cette méthode est, à nouveau en raison de l'aspect local de la méthode, que les paramètres utilisés pour l'extraction des pics a un impact très important sur leur nombre, ce qui nécessite un affinage de ces paramètres pour chaque classe ou pour chaque spectre étudié.

2.7 Conclusion

Dans ce chapitre, nous avons présenté des bases de données spectrales de matériaux. À partir de ces bases de données, nous avons présenté certains enjeux du traitement des spectres, tels que la variabilité spectrale et les mélanges linéaires et non-linéaires. Nous avons utilisé plusieurs méthodes pour extraire des informations depuis ces spectres. Cependant, les bases de données analysées sont limitées en termes d'applications et de variations des conditions d'acquisition. Les résultats obtenus dans ce chapitre permettent notamment de définir le domaine spectral d'intérêt pour l'acquisition de données hyperspectrales adaptées à notre besoin, 1200-2300nm.

Chapitre 3

Acquisition d'images hyperspectrales

Sommaire

3.1	Introduction	27
3.2	Caractéristiques des caméras hyperspectrales	28
3.3	Caméras sélectionnées	30
3.3.1	Cubert FireflEYE S185	30
3.3.2	Headwall SWIR-384	31
3.4	Acquisition d'images de véhicules	32
3.5	Acquisition d'images de matériaux	33
3.5.1	Enjeux	34
3.5.2	Échantillons	34
3.5.3	Acquisitions extérieures	34
3.5.4	Acquisitions intérieures	35
3.6	Conclusion	37

3.1 Introduction

Dans le chapitre précédent, nous avons constaté l'intérêt de la dimension spectrale par l'analyse de plusieurs bases de données existantes. Cependant, les bases de données spectrales mais également hyperspectrales sont soumises à de nombreuses limitations, telles que le faible nombre de données extraites dans des conditions d'illumination et d'acquisition spécifiques. De plus, les bases de données hyperspectrales publiques sont sur-exploitées dans l'état de l'art, et les résultats sur ces données sont saturés (Ghamisi *et al.* [2017]). La définition de nos propres bases de données est alors cruciale.

Dans ce chapitre, nous présentons une liste des caractéristiques principales des caméras hyperspectrales, ainsi que deux caméras sélectionnées pour l'acquisition d'images hyperspectrales. Nous présentons ensuite la création de deux bases de données hyperspectrales, une de véhicules et une de matériaux urbains.

3.2 Caractéristiques des caméras hyperspectrales

Le terme de « caméra hyperspectrale » désigne une très grande diversité de caméras, fabriquées pour correspondre à différents besoins. Des travaux antérieurs ont présenté des listes de capteurs pour des acquisitions satellitaires (Aasen *et al.* [2018]), pour des acquisitions terrestres (Hagen et Kudenov [2013]) ou pour des applications médicales (Vemuri *et al.* [2019]). Dans cette section, nous présentons les caractéristiques nécessaires au choix d'une caméra hyperspectrale adaptée.

Le premier critère, et généralement le plus important lors de la sélection d'une caméra hyperspectrale, est le type de caméra. Les principaux types de caméra sont les suivants :

- « Instantanée » (ou *snapshot*) : La caméra acquiert simultanément toutes les données d'une image hyperspectrale 3D, de taille $X \times Y \times L$. Cette acquisition de données 3D avec un capteur 2D implique des limitations dans chacune des trois dimensions de l'image hyperspectrale.
- « À balayage » (ou *pushbroom* ou *spatial scanning*) : La caméra acquiert simultanément une dimension spatiale et la dimension spectrale, produisant des lignes de signatures spectrales soit des données de taille $X \times L$. L'optique de la caméra balaie la scène perpendiculairement à l'axe X , ce qui permet de reconstituer un axe Y en concaténant les lignes spectrales obtenues. La dimension temporelle est donc transformée en deuxième dimension spatiale.
- « À balayage spectral » (ou *staring* ou *spectral scanning*) : La caméra acquiert simultanément les deux dimensions spatiales X et Y pour une longueur d'onde λ donnée. La caméra change progressivement de λ jusqu'à avoir reconstruit une image de taille $X \times Y \times L$. La dimension temporelle est donc transformée en dimension spectrale.

Afin de définir les dimensions spatiales, deux critères sont utilisés :

- La résolution spatiale, X ou $X \times Y$ selon le type de caméra, décrit la qualité spatiale des données.
- L'angle du champ de vision, dépendant de la lentille de la caméra, permet de connaître les objets de la scène qui sont perçus par la caméra pour une distance donnée. Il permet également de calculer quelle taille représentent ces objets dans l'image. Le champ de vision est particulièrement

important à prendre en compte dans le cas où la caméra est embarquée sur un drone, donc à longue distance des objets observés.

Les critères d'intérêt liés à l'aspect spectral sont la portée spectrale, le nombre de bandes, l'échantillonnage spectral et la résolution spectrale :

- Le nombre de bandes L décrit la qualité spectrale des données.
- Le domaine spectral acquis par la caméra, représenté par deux longueurs d'onde minimale et maximale exprimées en nanomètres (nm).
- L'échantillonnage spectral : Il s'agit de l'écart de longueur d'onde entre deux bandes consécutives, exprimé en nm. Cet échantillonnage est souvent considéré unique et constant pour une même caméra, mais peut changer entre plusieurs domaines spectraux.
- La résolution spectrale (ou « largeur à mi-hauteur » ou *Full-Width at Half Maximum*) de chaque bande : Elle indique la superposition entre les bandes consécutives ainsi que la précision de chaque bande. Tout comme l'échantillonnage, la résolution est considérée constante mais peut changer entre plusieurs domaines spectraux pour une même caméra.

Enfin, les autres critères que nous avons considérés sont :

- La fréquence maximale d'images (ou *max frame rate*) : Définit le nombre d'images qu'il est possible d'acquérir par seconde. Dans le cas d'une caméra *snapshot*, la connaissance de cette fréquence est nécessaire pour des applications telles que le suivi de cible. Pour des caméras *pushbroom* ou *staring*, il s'agit du nombre de lignes par seconde, et la vitesse de construction d'un cube hyperspectral dépend directement de cette fréquence d'acquisition.
- Le temps d'intégration ou d'exposition : Complémentaire à la fréquence d'image, il s'agit du temps nécessaire à l'acquisition d'une image ou d'une ligne.
- *Signal to Noise Ratio* (SNR) : Un indicateur de la qualité du signal dans les images.
- La taille d'encodage : Espace mémoire consommé par le stockage d'une image.
- La taille de la caméra : Nécessaire pour déterminer sa portabilité, notamment sur drone.
- Le poids de la caméra : Nécessaire pour déterminer sa portabilité, notamment sur drone.
- La consommation énergétique de la caméra : Nécessaire pour déterminer son temps d'utilisation, notamment sur drone.
- La résolution spatiale au sol (à 1000m d'altitude) : Calculé à partir du champ de vision et de la résolution spatiale, ce critère permet de prédire la taille des objets dans l'image.

- Le pas temporel au sol (à 120km/h) : Calculé à partir de la fréquence maximale d'images. Ce critère permet de connaître la distance au sol entre deux acquisitions. Dans le cas des caméras *pushbroom*, ce critère permet de calculer la résolution spatiale au sol dans l'axe temporel.

3.3 Caméras sélectionnées

Les caractéristiques mentionnées en Section 3.2 nous ont permis de sélectionner deux caméras, répondant à des besoins opérationnels différents, à partir desquelles conduire nos acquisitions.

3.3.1 Cubert FireflEYE S185

Besoin

Dans l'objectif d'acquérir des images d'objets dynamiques devant un arrière-plan statique, tels que des véhicules en mouvement, il est nécessaire d'avoir une caméra *snapshot*. Toutefois, cet avantage opérationnel a un coût technologique : une résolution spatiale plus faible et des domaines spectraux limités au visible et au proche infrarouge.

Caméra

Pour cela, nous avons choisi la caméra hyperspectrale *snapshot* Cubert FireflEYE S185. Cette caméra acquiert des images de 50x50 pixels dans 138 bandes spectrales de 450 à 998 nanomètres (nm) à raison de 3 images par seconde. Elle a également l'avantage de permettre une étape de *pansharpening*.

Pansharpening

En plus de l'image hyperspectrale de taille 50x50x138, la caméra acquiert des images de taille 1000x1000x1 des mêmes scènes.

Le *pansharpening* consiste à fusionner une image panchromatique (1 bande) à haute résolution spatiale avec une image hyperspectrale à faible résolution spatiale pour obtenir une image hyperspectrale à haute résolution spatiale.

Cependant, cette transformation augmente quadratiquement la taille des données. Le *pansharpening* duplique également de nombreuses informations dans l'image, ce qui peut y introduire des distorsions. Afin de bénéficier de l'image panchromatique tout en limitant les effets de cette transformation, nous avons augmenté la résolution des images en 200x200x138.

Matériel complémentaire

Pour des acquisitions depuis le sol, la caméra est fournie avec un trépied permettant de la stabiliser, la déplacer et la tourner. Elle est contrôlée par un ordinateur portable auquel elle est connectée. Les données sont enregistrées sur un disque dur externe. Cette caméra est alimentée en électricité par deux batteries externes permettant des sessions d'acquisition de plusieurs heures. Une plaque de référence de blanc est également fournie et permet de calibrer les images (voir Section 4.2).

3.3.2 Headwall SWIR-384

Besoin

Opérationnellement, un cas d'emploi type est d'acquérir des images du sol depuis un drone. Pour cela, la caméra doit répondre à des contraintes de poids, taille et consommation d'énergie, mais aussi à des contraintes de résolution spatiale au sol. Les caméras à balayage sont idéales dans le cadre d'acquisitions aériennes à haute altitude, où le capteur balaie la scène grâce à son capteur. Ce type de caméras offre une meilleure résolution spatiale, permettant d'obtenir une résolution au sol plus fine, et spectrale tout en ayant une fréquence d'acquisition d'images plus rapide.

Caméra

Nous avons sélectionné la caméra à balayage Headwall SWIR-384. Sa résolution spatiale de 384 pixels et son champ de vision offrent une résolution spatiale au sol de moins de 30cm pour une altitude de 1000m. Sa fréquence d'acquisition, atteignant 450 lignes par seconde, permet l'obtention d'une résolution spatiale au sol inférieure à 10cm pour une vitesse de 120km/h.

La caméra est capable de reconstituer la deuxième dimension spatiale automatiquement en balayant la scène, en étant montée sur un drone pour un point de vue aérien ou sur un moteur rotatif pour un point de vue terrestre.

Les images sont acquises dans 169 bandes spectrales dans le domaine spectral de 900 à 2500nm, avec un échantillonnage spectral d'environ 9.55nm et une résolution spectrale de 10nm. De plus, la caméra a des caractéristiques de poids (2000g) et de taille (<20x10x20cm) en-dessous de la limite transportable, en plus d'avoir une consommation d'énergie (15W) dans la moyenne par rapport à ce type de caméras.

L'ensemble de ces caractéristiques fait de la Headwall SWIR-384 caméra la plus adaptée sur le marché à notre besoin actuel.

Matériel complémentaire

Pour des acquisitions au sol, la caméra est fournie avec un moteur rotatif. Ils sont tous deux connectés à une tour d'ordinateur permettant de contrôler les paramètres de l'acquisition, la tour étant elle-même contrôlée via un ordinateur portable. Le matériel complémentaire inclut :

- un luxmètre, afin de mesurer la quantité de lumière présente dans la scène des acquisitions ;
- deux plaques de référence de blanc, une en plastique PTFE (Polytétrafluoroéthylène) et une en Spectralon, pour permettre la calibration des matériaux (voir Section 4.2) ;
- une lampe halogène, permettant de fournir une illumination dans l'ensemble du domaine spectral 900-2500nm, pour éclairer les scènes d'acquisitions intérieures.

Un schéma de la configuration des acquisitions est présenté en Figure 3.1.

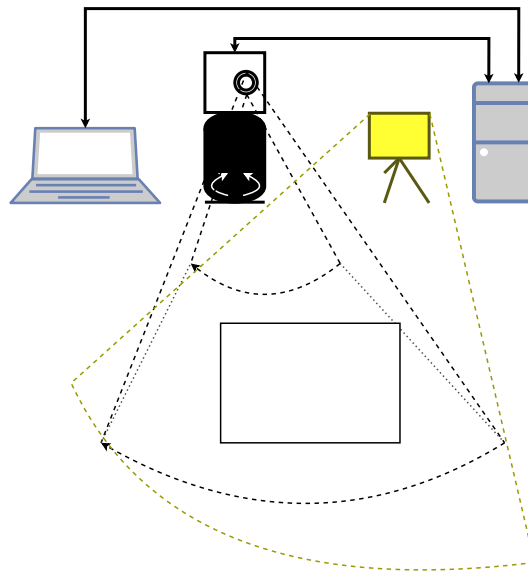


FIGURE 3.1 – Illustration de la configuration des acquisitions. De gauche à droite : l'ordinateur portable, la caméra sur la tête rotative, la lampe halogène, la tour d'ordinateur. La caméra et la lampe sont orientées vers un échantillon.

3.4 Acquisition d'images de véhicules

Notre première base de données a été construite dans le but de fournir une grande variété d'images pour l'utilisation de réseaux de neurones. Afin d'obtenir un nombre élevé d'images avec une grande variabilité et en conditions proches des conditions opérationnelles, nous avons choisi d'acquérir des

images de véhicules en mouvement avec la caméra Cubert FirefEYE S185. Les véhicules sont des objets emblématiques de l'analyse d'image pour la surveillance, et sont également constitués de plusieurs matériaux, ce qui les rend intéressants à analyser spectralement.

Les images de véhicules ont été acquises sous différents angles de vue et à différentes heures de la journée. Nous nous sommes placés sur un pont surplombant une autoroute pour simuler une prise de vue aérienne, avec des angles de vue plongeants allant de 0 à 70 degrés d'inclinaison. Des déplacements sur ce pont nous ont fourni des angles azimutaux variés, permettant d'obtenir des images de véhicules représentatives de la réalité. Des exemples de prises de vues sont présentées en Figure 3.2.

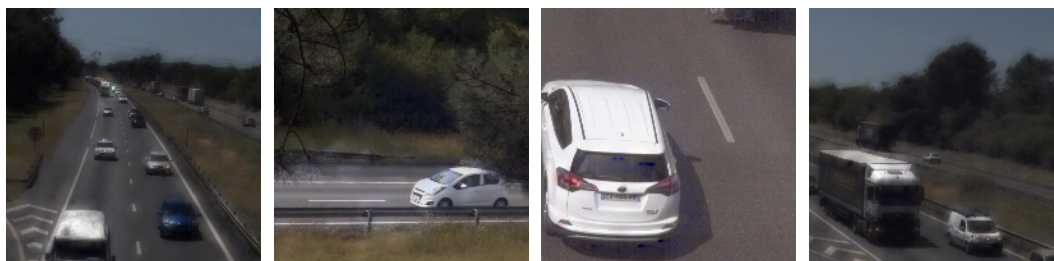


FIGURE 3.2 – Exemple d'images de véhicules lors de différentes prises de vues depuis un pont.

Afin d'avoir une variabilité des conditions d'illumination, les acquisitions ont été réparties sur trois jours, avec une session le matin et une l'après-midi, pour un total de six sessions d'environ 1h30. Les images hyperspectrales ont été acquises à une fréquence de 3 par seconde. Une partie importante de ces images ne contenant cependant pas de véhicules, environ 15000 images de véhicules ont été obtenues au total.

3.5 Acquisition d'images de matériaux

La deuxième base de données a été construite avec plusieurs objectifs. Le premier objectif est d'analyser les spectres de plusieurs objets et matériaux à partir d'images hyperspectrales. Le deuxième objectif d'évaluer la qualité de la calibration d'images hyperspectrales (voir Chapitre 4) en environnement contrôlé. Enfin, le troisième objectif est d'étendre cette calibration à des conditions non contrôlées.

Ces données ont également pour but d'être utilisées pour des tâches de détection et de discrimination de matériaux.

3.5.1 Enjeux

Pour atteindre les objectifs définis précédemment, nous récoltons des images hyperspectrales de matériaux urbains et de surveillance, dans plusieurs conditions d'illumination et d'observation. Nous considérons désormais le domaine spectral NIR-SWIR, comme préconisé en conclusion du Chapitre 2. Nous souhaitons également pouvoir répliquer ces analyses en conditions opérationnelles sur drone à environ 1000m d'altitude. Pour ces raisons, nous utilisons la caméra à balayage Headwall SWIR-384 pour la suite des acquisitions de matériaux.

D'après [Borsoi *et al.* \[2021\]](#), il existe des incohérences entre les données acquises en conditions de laboratoire et celles acquises en extérieur avec éclairage solaire. Afin de prendre en compte cette problématique ainsi que celle de la variabilité spectrale, les acquisitions ont été conduites dans deux environnements distincts : l'un en environnement non-contrôlé en utilisant comme éclairage la lumière du soleil, et l'autre en environnement contrôlé et fermé à la lumière naturelle.

3.5.2 Échantillons

Les bâtiments constituant une part importante des images acquises depuis une vue aérienne, les premiers types d'échantillons à analyser étaient des matériaux de construction. Nous avons ainsi préparé plusieurs blocs de ciment, blocs de béton, des briques, du bitume et des tuiles.

D'autres matériaux d'intérêt choisis étaient ceux pouvant servir de leurres. Pour cela, nous avons sélectionné plusieurs échantillons d'acier pour représenter des vrais véhicules et de bois pour des leurres de véhicules.

3.5.3 Acquisitions extérieures

Les acquisitions en extérieur ont été effectuées pour prendre en compte l'influence et les variations de l'environnement extérieur sur les échantillons mesurés : la météo, l'illumination solaire et une importante absorption atmosphérique. Deux sessions d'acquisition ont été effectuées avec des illuminations solaires différentes : une par conditions nuageuses, une par conditions ensoleillées. Une troisième session a été conduite en conditions nocturnes, sous illumination de la lune, sur un sous-ensemble d'échantillons.

La caméra était éloignée des échantillons de 60cm, surélevée de 35cm par rapport aux échantillons et inclinée à 24 degrés. Ces sessions en extérieur ont résulté en l'acquisition d'environ 100 images hyperspectrales.

Dans l'objectif de chercher des caractéristiques spectrales permettant de différencier un véhicule civil et un véhicule militaire à partir de leur composition métallique ou de leur blindage, nos acquisitions extérieures ont été complétées par l'acquisition d'images d'un véhicule civil sous plusieurs angles.

3.5.4 Acquisitions intérieures

Des acquisitions en environnement intérieur, avec une lampe halogène comme seule source lumineuse, nous ont permis d'obtenir des spectres d'échantillons non soumis aux différentes contraintes des acquisitions en extérieur (luminosité variable, absorption atmosphérique...). Il est notable que dans de telles conditions, l'utilisation de plusieurs lampes est généralement nécessaire pour obtenir des conditions quasi-uniformes d'illumination des échantillons. Dans notre cas, nous avons utilisé une unique lampe halogène comme source d'illumination afin de contrôler l'intensité et l'angle d'incidence de la lumière. Au cours de ces acquisitions, la récolte des images de références de blanc a été conduite séparément des échantillons, en les positionnant à la place des échantillons en début et en fin de chaque session.

Une image de la configuration des acquisitions est présentée en Figure 3.3a. La caméra était éloignée des échantillons de 80cm, surélevée de 35cm par rapport aux échantillons et inclinée à 24 degrés. Ces paramètres sont les mêmes que pour les acquisitions extérieures, à l'exception de l'éloignement de la caméra qui est dû à des limitations d'espace dans les environnements d'acquisition. L'angle de balayage de la tête rotative a été réglé à 30 degrés avec une vitesse de 3 degrés par seconde et une ligne toutes les 22.17ms. Ces valeurs produisent des images de 452 à 454 pixels de largeur, l'incertitude de 2 pixels étant due à des imprécisions du capteur.

Nous avons conduit cinq sessions intérieures, illustrées en Figure 3.3b. Les sessions sont différenciées par l'orientation des objets, posés sur une surface (sessions 1 à 3) ou face à la caméra (sessions 4 et 5), et par l'orientation de la lampe (voir Figure 3.3b). L'ensemble des cinq sessions a résulté en l'acquisition d'environ 300 images.

Cependant, la disposition de la caméra et de la lampe au cours de la session 1 résulte en des réflexions spéculaires. Ces réflexions spéculaires saturent le capteur (Geladi *et al.* [2004]) ou produisent une réflectance supérieure à 100% (Zaman *et al.* [2014]) et ne peuvent donc pas être exploitées spectralement. Les images issues de cette session seront donc ignorées pour l'analyse des données de ces acquisitions.

Acquisitions sans lumière

Des acquisitions complémentaires ont été conduites dans le même environnement sans éclairage de la lampe halogène afin d'identifier de potentielles sources de bruit. Ces nouvelles conditions incluent :

- des acquisitions d'échantillons avec un néon LED comme unique source d'illumination ;
- des acquisitions d'échantillons sans source de lumière ;

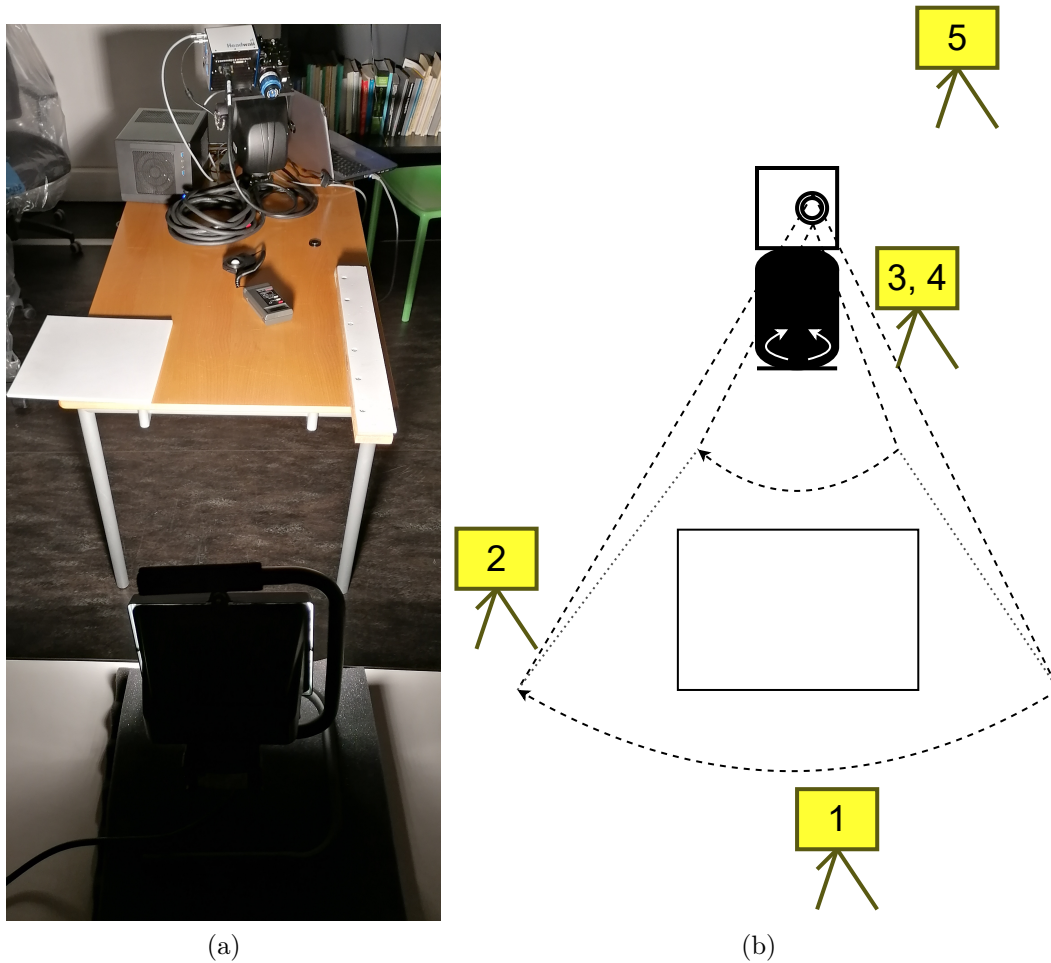


FIGURE 3.3 – Configuration des acquisitions en intérieur. (a) Photo de la configuration de la session 1. (b) Illustration de la configuration des cinq sessions avec la caméra à balayage. La position de la lampe pendant chacune des cinq sessions est indiquée par son numéro.

- des acquisitions, sans source de lumière, d'un écran d'ordinateur présentant des images d'échantillons acquises lors des sessions précédentes.

Par simplicité et en raison du nombre faible d'images acquises dans ces conditions, nous considérons l'ensemble de ces acquisitions complémentaires comme appartenant à une même session 6, et ferons la distinction lorsque nécessaire.

Acquisitions dans le domaine visible

En complément aux acquisitions précédentes, nous avons également conduit des acquisitions en intérieur avec la caméra Cubert FirefEYE S185 dans le domaine spectral 400-1000nm. Pour ces acquisitions, les sources lumineuses

consistent de néons lumineux ainsi que du soleil éclairant à travers une fenêtre en verre absorbant une partie du rayonnement infrarouge.

Au cours de ces acquisitions, des objets de couleur ou de matériau similaires ont été acquis afin de comparer leurs spectres et d'extraire des caractéristiques utiles à la caractérisation de matériaux.

3.6 Conclusion

Dans ce chapitre, les caractéristiques des caméras hyperspectrales ont été présentées, deux caméras ont été sélectionnées et deux bases de données hyperspectrales ont été créées. La première base de données est constituée d'images de véhicules acquises instantanément dans les domaines du visible et du très proche infrarouge. La deuxième contient des images hyperspectrales d'objets et de matériaux acquis dans plusieurs conditions d'illumination dans les domaines du proche infrarouge et infrarouge à ondes courtes. Ces deux bases de données vont ensuite permettre d'effectuer une diversité d'analyses et de traitement des données, spectrales et spatio-spectrales, issues d'objets et de matériaux en rapport avec la surveillance.

Chapitre 4

Calibration

Sommaire

4.1	Introduction	38
4.2	Définition de la calibration	39
4.3	Dérive temporelle des valeurs brutes	41
4.4	Origine de la dérive temporelle	43
4.5	Impact sur la radiance	45
4.6	Méthodes de sélection d'image <i>DC</i> pour le calcul de la radiance	46
4.6.1	« <i>Single</i> »	46
4.6.2	« <i>Closest</i> »	47
4.6.3	« <i>Previous</i> »	47
4.6.4	« <i>Next</i> »	48
4.6.5	« <i>Linear session</i> »	49
4.6.6	« <i>Linear closest</i> »	49
4.6.7	Conclusion de la sélection de l'image <i>DC</i>	50
4.7	Comparaison des références de blanc	51
4.8	Calibration par référence de gris	52
4.9	Limitations de la calibration	56
4.10	Conclusion	57

4.1 Introduction

Nous avons vu au chapitre précédent l'acquisition de deux bases de données hyperspectrales ainsi que des descriptions de ces données. Dans un premier temps, nous définissons la calibration d'images hyperspectrales et ses enjeux. Nous identifions une dérive temporelle des valeurs des images pré-calibration, l'origine de cette dérive et son impact sur les images calibrées, puis nous présentons plusieurs méthodes de calcul de la calibration pour réduire l'impact de la

dérive temporelle, travaux qui ont donné lieu à une publication (Caubet *et al.* [2022a]). Nous comparons l'utilisation de plusieurs références de blanc pour la calibration et nous étudions l'utilisation de spectres intermédiaires, nommés « références de gris ». Enfin, nous présentons les limitations de la calibration en conditions opérationnelles de surveillance.

4.2 Définition de la calibration

Comme mentionné en Section 3.2, les caméras hyperspectrales produisent des images de taille $X \times Y \times L$, avec X et Y les dimensions spatiales et L la dimension spectrale. Les images ainsi produites sont des images brutes, dont les valeurs correspondent aux impulsions électriques détectées par le capteur.

Les données brutes (voir Figs. 4.1a et 4.2a) contiennent de nombreux biais tels que le bruit du capteur de la caméra ou la quantité de lumière dans la scène. Ces biais impactent les données et empêchent de pouvoir comparer entre elles des images acquises dans des conditions d'acquisition différentes (Burger et Geladi [2005]).

Il est alors possible d'effectuer une calibration. La calibration est une méthode permettant d'estimer et d'éliminer ces biais dans les données spectrales ou hyperspectrales. Les images produites par la calibration sont des images de « réflectance », dont les valeurs sont exprimées en pourcentage de lumière réfléchie par l'objet. Après calibration, il devient ainsi possible de comparer entre eux des spectres acquis dans des conditions d'acquisition différentes. Nous présentons en Figures 4.1 et 4.2 des exemples de calibration d'image et de spectre, respectivement. La calibration se décompose en plusieurs étapes.

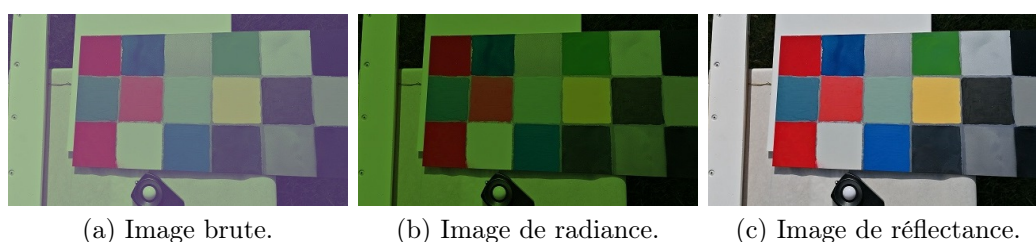


FIGURE 4.1 – Exemple de calibration d'image. L'image présente une plaque de métal peinte de multiples couleurs acquise en conditions extérieures.

La première étape de la calibration consiste à mesurer la quantité de courant d'obscurité (DC ou *Dark Current*). Le DC est le courant détecté par le capteur sans aucune source d'illumination.

Les images DC sont acquises en obstruant l'ouverture de la caméra avec son obturateur afin d'empêcher toute lumière d'atteindre le capteur. Une image du courant d'obscurité DC a donc les mêmes dimensions qu'une image hyperspec-

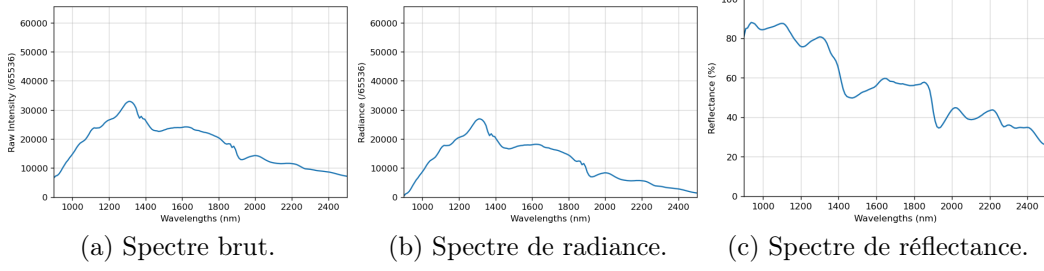


FIGURE 4.2 – Exemple de calibration de spectre. Le spectre représenté est celui d’un échantillon de bois.

trale classique, et une valeur DC est obtenue pour chaque triplet de nombres entiers (x, y, λ) , avec $0 \leq x < X$, $0 \leq y < Y$ et $0 \leq \lambda < L$.

La deuxième étape de la calibration est la conversion des images brutes I_{raw} en images de radiance I_{rad} . Cette transformation est définie pour chaque triplet (x, y, λ) par l’équation :

$$I_{\text{rad}}(x, y, \lambda) = I_{\text{raw}}(x, y, \lambda) - DC(x, y, \lambda). \quad (4.1)$$

L’image de radiance (voir Figs. 4.1b et 4.2b) obtenue correspond à la quantité de lumière perçue par le capteur. Cette image est toujours dépendante de la sensibilité de la caméra, du temps d’exposition et de l’intensité lumineuse de la scène, mais plus du bruit du capteur.

L’étape suivante est la mesure de l’illumination présente dans la scène. Pour cela, il est nécessaire d’avoir une « référence de blanc ». Une référence de blanc est un objet réfléchissant de manière diffuse et à près de 100% la lumière de toutes les longueurs d’onde du domaine spectral considéré. Comme elle représente la quantité de lumière présente dans la scène, la référence de blanc doit être acquise dans les mêmes conditions que les objets d’intérêt. Contrairement au DC , la référence de blanc ne couvre généralement pas toute l’image. Nous définissons $I_{w_{\text{rad}}}$ l’image de radiance de la référence de blanc et $S_w \subset (X \times Y)$ son sous-ensemble de pixels $s = (x, y)$ dans lesquels se trouve la référence de blanc. Nous extrayons un spectre représentatif de la référence de blanc, défini tel que :

$$\bar{w}_{\text{rad}}(\lambda) = \frac{\sum_s^{S_w} I_{w_{\text{rad}}}(s, \lambda)}{\|S_w\|}. \quad (4.2)$$

La dernière étape est le calcul de l’image de réflectance I_{ref} , définie par :

$$I_{\text{ref}}(x, y, \lambda) = \frac{I_{\text{rad}}(x, y, \lambda)}{\bar{w}_{\text{rad}}(\lambda)}. \quad (4.3)$$

Les valeurs de l’image de réflectance (voir Figures 4.1c et 4.2c) sont exprimées en pourcentage et correspondent à la proportion de la lumière qui est réfléchi.

L'image de réflectance est plus robuste aux conditions d'illumination et à la sensibilité de la caméra. Elle rend possible la comparaison de spectres acquis dans des conditions différentes.

4.3 Dérive temporelle des valeurs brutes

Une analyse des images pré-calibration permet d'apporter des informations sur des défauts potentiellement présents dans les données brutes, à corriger avant de conduire la calibration. Nous présentons en Figure 4.3 la bande 55 (1424nm) des images *DC* acquises lors de la session 5. Sur ces images, nous pouvons observer que les valeurs augmentent au cours du temps, ce qui signifie qu'il n'est pas possible d'utiliser la même image *DC* pour la calibration de toutes les images d'échantillons.

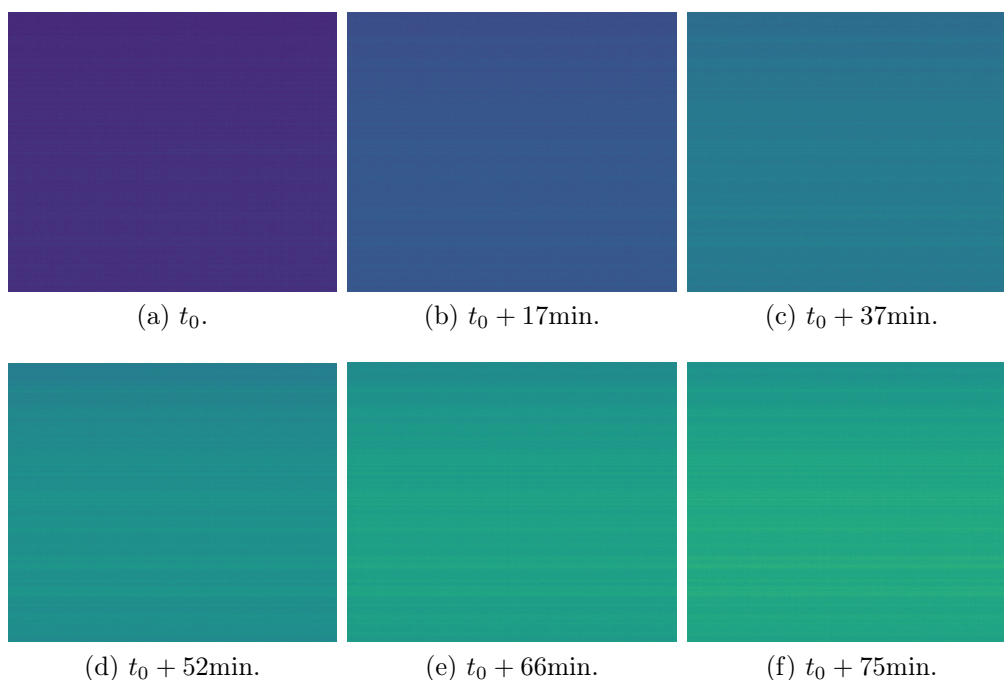


FIGURE 4.3 – Bande 55 (1424nm) de 6 images *DC* acquises à des temps différents lors de la session 5. La *colormap* utilisée est *viridis* : les pixels d'intensité faible sont présentés en violet/bleu, ceux d'intensité élevée en jaune.

Dans les histogrammes des images *DC*, présentés en Figure 4.4, les valeurs de l'image acquise en temps T_0 sont réparties de manière quasi-gaussienne. Cependant, nous pouvons constater que cette répartition s'étale au cours du temps et ne correspond plus à une unique répartition. Par simplicité de présentation, seule la bande 55 est présentée ici, mais ces constatations sont également observables sur les autres bandes spectrales des images *DC*.

4.3. Dérive temporelle des valeurs brutes

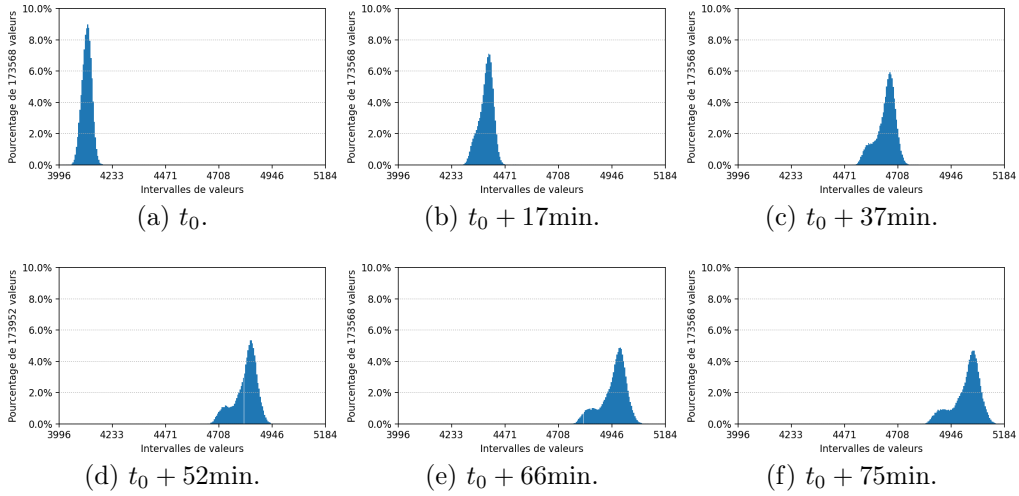


FIGURE 4.4 – Distribution des valeurs de la bande 55 (1424nm) de 6 images *DC* acquises à des temps différents lors d’une même session.

Afin de mieux comprendre la nature de l’évolution des valeurs, nous présentons en Figure 4.5 les moyennes respectives des images présentées en Fig. 4.3. La ligne bleue représente une évolution linéaire de la moyenne au cours du temps. Nous pouvons alors observer que son évolution est non-linéaire : la moyenne croît d’abord rapidement puis de plus en plus lentement et se rapproche d’un point de saturation.

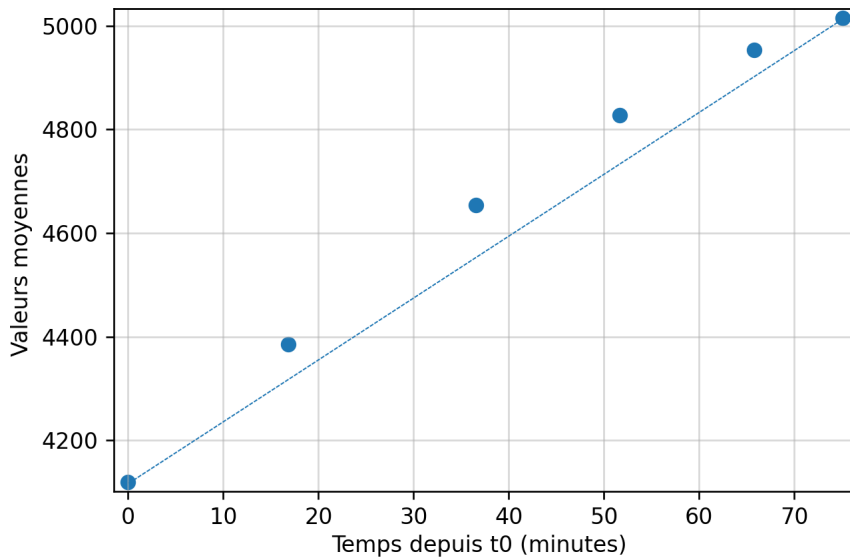


FIGURE 4.5 – Valeurs d’intensité moyenne, en fonction du temps de l’acquisition, de la bande 55 (1424nm) des 6 images *DC* de la session 5 présentées en Figure 4.3.

Afin de vérifier que cette évolution n'est pas limitée à la moyenne des images DC mais à l'ensemble de leurs valeurs, nous calculons les distances Kullback-Leibler (KL) (voir Sec. 2.5) entre les histogrammes présentés en Figure 4.4. Pour chaque paire d'images, nous comparons l'image ultérieure (Q dans l'équation) à l'image antérieure (P dans l'équation).

Le calcul est effectué en plusieurs étapes :

- Nous calculons les histogrammes (voir Sec. 2.3) des deux images à comparer.
- Ces histogrammes sont centrés en zéro à l'aide de la moyenne de ces images.
- Les histogrammes centrés sont transformés en distributions de probabilités (voir Sec. 2.3).
- Une valeur faible ϵ , ici 0.0001, est ajoutée à chaque distribution. Cette étape permet d'éviter la présence de valeurs nulles dans Q , ces valeurs nulles provoquant une erreur dans le calcul de la distance KL.
- Nous calculons ensuite la distance KL (voir équation 2.4) sur les densités de probabilité résultantes.

Les valeurs obtenues sont présentées en Fig. 4.6. Nous pouvons observer que, pour un même intervalle de temps entre les images comparées, les images acquises à un temps plus élevé ont une valeur de KL plus faible. Cela confirme que la différence entre les images DC diminue dans le temps et se rapproche d'un point de saturation.

Nous notons que la valeur ϵ a un impact important sur la magnitude des valeurs de Kullback-Leibler. Cependant, cette valeur ϵ n'a aucun impact sur les relations observées ci-dessus : parmi trois images comparées entre elles, les deux images ayant la distance KL la plus élevée seront toujours les mêmes, peu importe la valeur ϵ sélectionnée.

4.4 Origine de la dérive temporelle

Afin d'étudier le comportement de la dérive sur une période de temps plus importante, nous nous sommes intéressés à l'ensemble des acquisitions conduites au cours d'une même journée. Les moyennes des images DC de cette journée sont présentées en Figure 4.7. Pour rappel des numéros de session (voir Fig. 3.3b, p. 36), la session 1 désigne la lampe positionnée face à la caméra, 6 la lampe éteinte, 3 et 4 la lampe très proche derrière la caméra, et 2 la lampe positionnée de côté par rapport aux échantillons. La lampe était également éteinte au cours de la pause.

Nous pouvons constater sur cette Figure que la valeur moyenne augmente lors de l'allumage de la caméra et de la lampe (sessions 1, 3 et 4), et diminue à l'extinction de la lampe (session 6) ou lorsque celle-ci est éloignée de la caméra

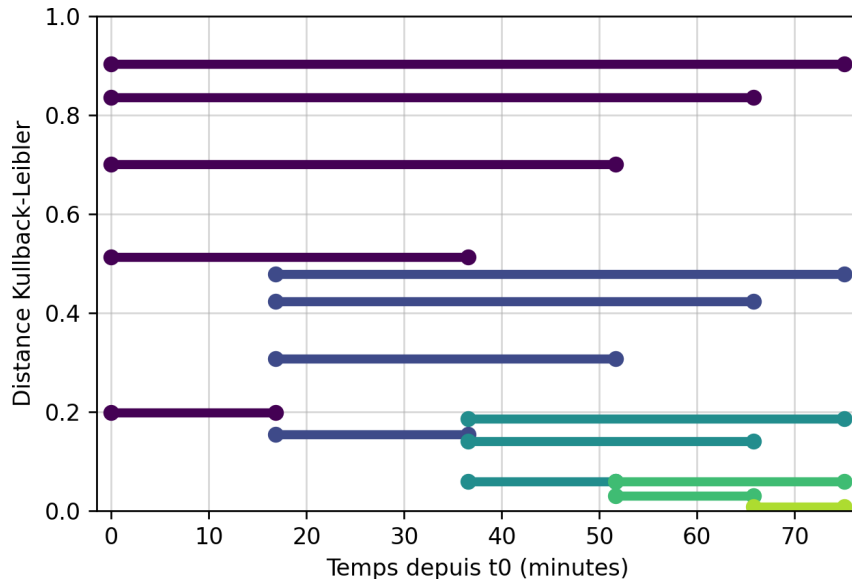


FIGURE 4.6 – Distances de Kullback-Leibler entre la bande 55 (1424nm) des 6 images *DC* de la session 5 présentées en Figure 4.3. Les images sont comparées par paires, indiquées par l’image antérieure (gauche) et l’image ultérieure (droite) de chaque segment. Les segments sont coloriés en fonction de l’heure de l’acquisition de l’image antérieure.

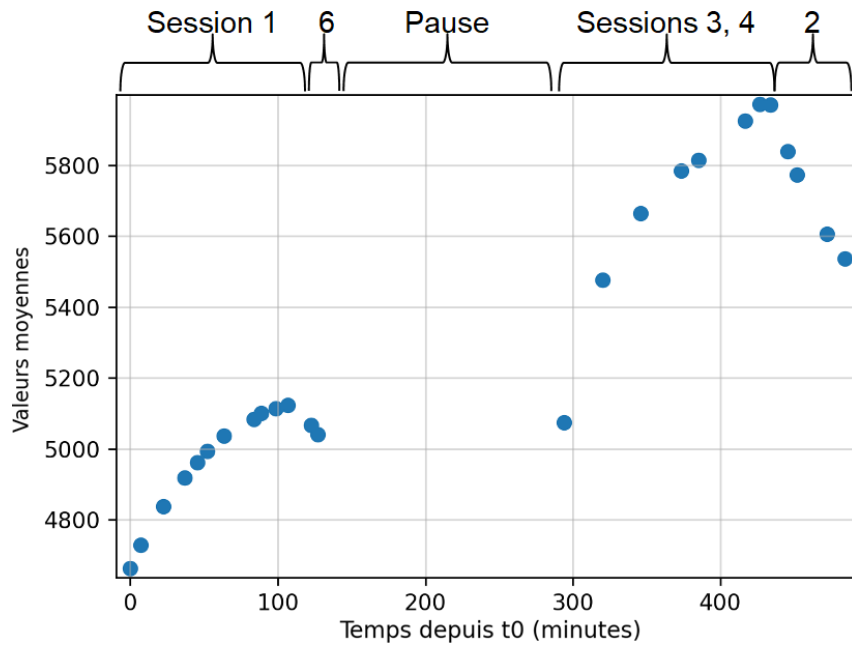


FIGURE 4.7 – Valeurs d’intensité moyenne de la bande 55 (1424nm) de 25 images *DC* acquises à des temps différents lors de plusieurs sessions consécutives.

(session 2). Après extinction de lampe et de la caméra pendant environ 1000 minutes puis redémarrage (session 5), leurs valeurs sont de nouveau faibles et augmentent rapidement (voir Fig. 4.5).

Si l'obturateur filtrait incorrectement la lumière, la chute des valeurs lorsque la caméra est éteinte devrait être instantanée et non progressive.

À la lumière de ces analyses, notre hypothèse est que ces variations progressives proviennent d'un impact de la chaleur du capteur sur son bruit. En effet, la lampe halogène produit une quantité importante de chaleur au cours de son utilisation, qui se diffuse dans l'ensemble de la pièce et augmente la température du capteur. Nous pouvons en particulier observer que l'éloignement entre la lampe et la caméra en session 2 est corrélé avec la chute progressive des valeurs DC au-delà de 440 minutes.

De plus, plusieurs publications antérieures font état d'un impact de la chaleur sur les spectres : [Shaikh et al. \[2021\]](#) décrivent le DC comme étant dépendant de la température du capteur, et [Burger et Geladi \[2005\]](#) mentionnent que la température du capteur peut impacter le signal. Dans l'analyse de température conduite par [Noviyanto et Abdulla \[2019\]](#), l'impact de la chaleur sur les spectres apparaît non-linéaire. Enfin, les fournisseurs de la caméra ont validé le fait que la chaleur de la caméra pouvait impacter directement le bruit du capteur malgré son système de refroidissement, mais n'ont pas su nous donner plus d'informations pour corriger cette dérive temporelle.

Nous pouvons cependant noter que, en conditions opérationnelles de vol de la caméra sur un drone, cette caméra serait sujette à des variations importantes de températures. Dans ces conditions, il est alors important de pouvoir modéliser ou corriger l'impact de la température sur le bruit du capteur.

Néanmoins, les données à notre disposition ne nous permettent pas de déterminer si un changement de température lié à l'émission de la lampe halogène est effectivement la seule source de la dérive que nous avons constatée. Dans le futur, il serait intéressant d'analyser l'impact de la température ambiante ou de la caméra sur le bruit du capteur. De plus, cette dérive temporelle des valeurs brutes n'est pas limitée aux images DC et elle s'observe également dans les images brutes des échantillons. Cela veut dire que cette dérive doit être prise en compte lors du calcul de la radiance.

4.5 Impact sur la radiance

Les images de radiance représentent des images isolées du bruit du capteur. Ce bruit doit donc être représenté au mieux par l'image DC utilisée pour la calibration (voir équation 4.1). Soit en Fig. 4.8 les reconstructions d'images couleur (voir Sec. 5.2.1) de 4 images hyperspectrales brutes d'échantillons et 3 images DC acquises lors de la même session.

Nous avons calculé les images de radiance en utilisant l'image $DC2$. Ces

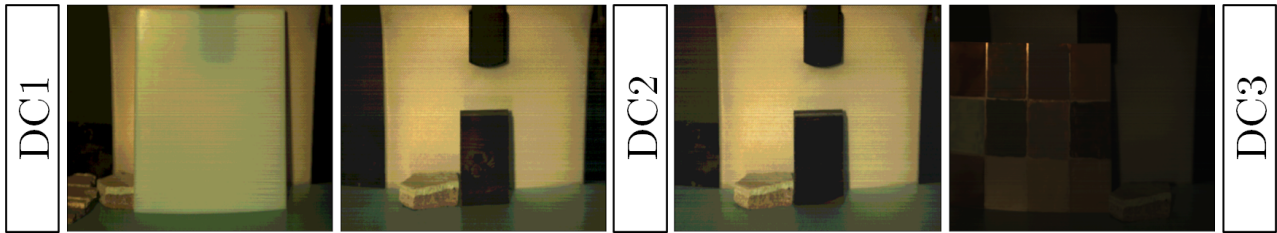


FIGURE 4.8 – Reconstructions couleur de 4 échantillons acquis à différents moments d'une même session, ainsi que 3 images *DC* acquises lors de la même session.

images de radiance sont présentées en Fig. 4.9. Les images acquises avant l'image DC2 (gauche) présentent des valeurs négatives sur l'ensemble de l'image de la bande 0 et sur une partie de la bande 55. Dans les images acquises après l'image DC2 (droite), il y a très peu de valeurs négatives, mais le bruit du capteur n'a été que partiellement éliminé.

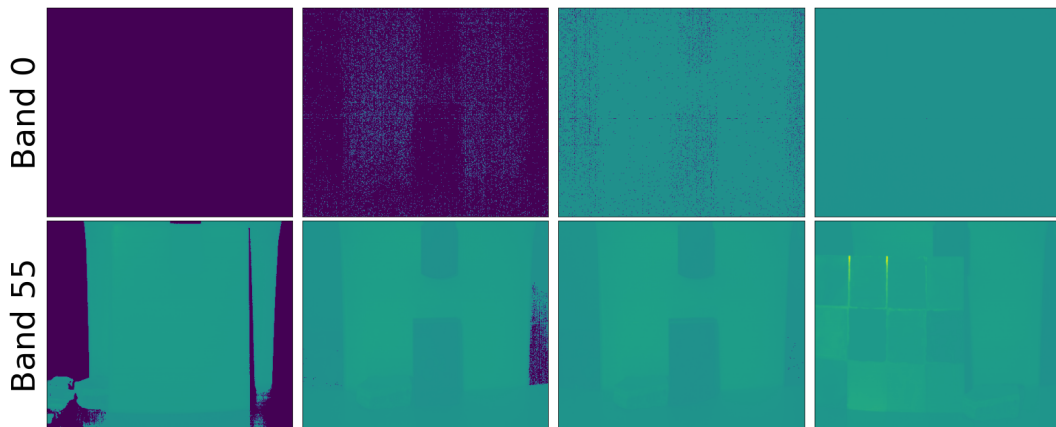


FIGURE 4.9 – Bandes spectrales 0 et 55 (899 et 1424nm) de radiance des échantillons en Figure 4.8. La radiance est calculée à partir de l'image DC2. Les pixels d'intensité négative sont affichés en violet, ceux d'intensité positive sont affichés en gradient du vert au jaune.

4.6 Méthodes de sélection d'image *DC* pour le calcul de la radiance

4.6.1 « *Single* »

Précédemment, nous avons calculé les images de radiance d'échantillons à partir d'une unique image *DC*, méthode que nous nommons « *single* » dont

le résultat est présenté en Figure 4.9. Cependant, lorsque plusieurs images *DC* ont été acquises au cours d'une même session, il est préférable d'utiliser plusieurs de ces images en fonction de leur temps d'acquisition.

4.6.2 « *Closest* »

Dans un premier temps, pour chaque image de radiance à calculer, nous considérons l'utilisation de l'image *DC* acquise la plus proche dans le temps. Nous nommons cette méthode « *closest* » et présentons en Figure 4.10 les images de radiance calculées avec cette méthode. De cette manière, le bruit du capteur est représenté de manière plus précise qu'avec la méthode « *single* », ce qui a pour effet de réduire le nombre de valeurs négatives (image de gauche) ou le bruit du capteur (image de droite).

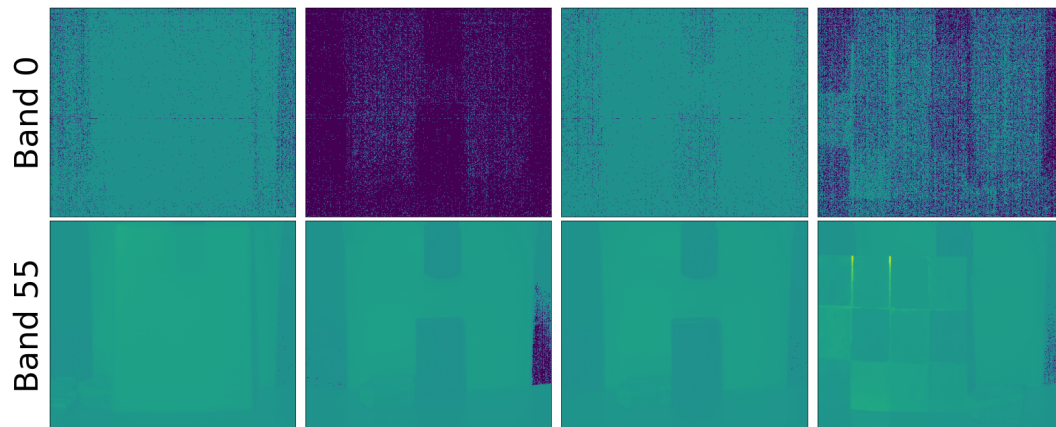


FIGURE 4.10 – Bandes spectrales 0 et 55 (899 et 1424nm) de radiance des échantillons en Figure 4.8. La radiance est calculée avec la méthode « *closest* ». Les pixels d'intensité négative sont affichés en violet, ceux d'intensité positive sont affichés en gradient du vert au jaune.

4.6.3 « *Previous* »

La méthode « *previous* », qui consiste à utiliser l'image *DC* antérieure à l'image à calibrer, ici *DC1* ou *DC2*, est proposée afin d'éviter la présence de valeurs négatives. Ses résultats sont présentés en Figure 4.11. Cette méthode élimine la majorité des valeurs négatives, mais en laisse dans les zones de faible intensité des longueurs d'onde extrêmes (par ex. l'arrière-plan la bande 0) de certaines images. De plus, cette méthode a l'inconvénient de ne pas supprimer le bruit du capteur dans les images.

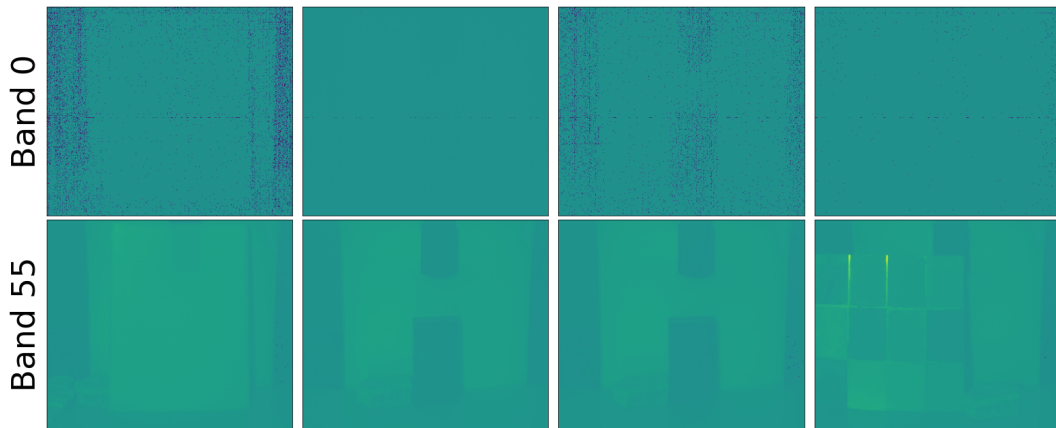


FIGURE 4.11 – Bandes spectrales 0 et 55 (899 et 1424nm) de radiance des échantillons en Figure 4.8. La radiance est calculée avec la méthode « *previous* ». Les pixels d'intensité négative sont affichés en violet, ceux d'intensité positive sont affichés en gradient du vert au jaune.

4.6.4 « *Next* »

La méthode « *next* », qui consiste à utiliser l'image DC ultérieure à l'image à calibrer, ici DC2 ou DC3, est proposée afin d'éviter la présence de bruit. Ses résultats sont présentés en Figure 4.12. Cette méthode élimine le bruit, mais produit des valeurs de radiance négatives dans l'intégralité du spectre.

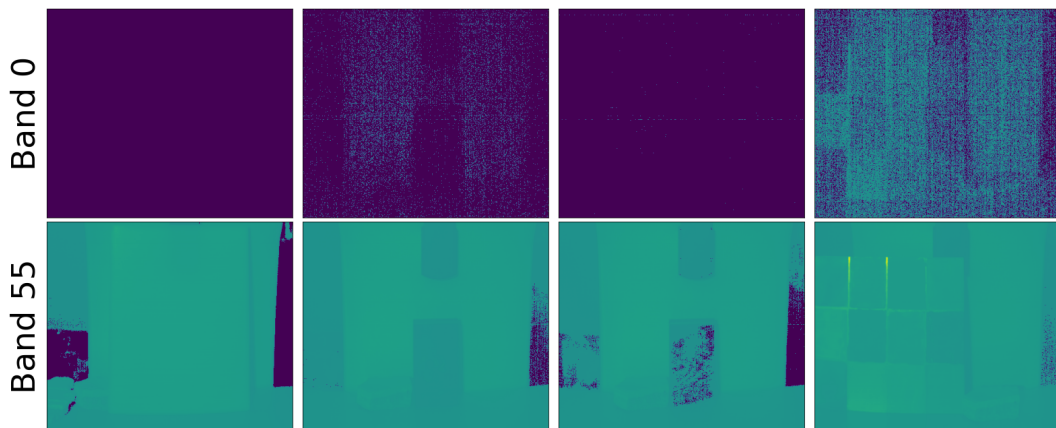


FIGURE 4.12 – Bandes spectrales 0 et 55 (899 et 1424nm) de radiance des échantillons en Figure 4.8. La radiance est calculée avec la méthode « *next* ». Les pixels d'intensité négative sont affichés en violet, ceux d'intensité positive sont affichés en gradient du vert au jaune.

4.6.5 « *Linear session* »

La méthode « *linear session* » est proposée comme compromis entre les méthodes présentées précédemment, qui créent des images de radiance avec une quantité importante de bruit du capteur ou de valeurs négatives. Cette méthode effectue une interpolation linéaire entre les deux images *DC* acquises en début et en fin de session afin de générer une image *DC* composite. Le ratio de l'interpolation linéaire est sélectionné à partir de l'heure d'acquisition de l'image à calibrer. Cela permet de générer une image composite avec des valeurs proches de celles d'une image *DC* qui aurait été acquise au même instant que l'échantillon. Ses résultats sont présentés en Figure 4.13. En raison de la non-linéarité de la dérive, cette méthode ne parvient pas à supprimer le bruit du capteur dans les images (voir les deux images du milieu, Fig. 4.13). Les résultats de cette méthode sont néanmoins meilleurs que ceux des méthodes précédentes, car la majorité des valeurs négatives a disparu, à l'exception de l'arrière-plan des bandes spectrales extrêmes.



FIGURE 4.13 – Bandes spectrales 0 et 55 (899 et 1424nm) de radiance des échantillons en Figure 4.8. La radiance est calculée avec la méthode « *linear session* ». Les pixels d'intensité négative sont affichés en violet, ceux d'intensité positive sont affichés en gradient du vert au jaune.

4.6.6 « *Linear closest* »

Afin d'approcher plus fidèlement des images de radiance "parfaites", qui seraient générées avec des images *DC* acquises simultanément aux images d'échantillon à calibrer, nous proposons une dernière méthode, « *linear closest* ». Cette méthode est similaire à la précédente dans son utilisation d'une interpolation linéaire afin de générer une image *DC* composite, mais utilise

les deux images *DC* antérieure et ultérieure pour l'interpolation. Cette méthode permet de prendre en compte une partie de la non-linéarité de la dérive temporelle. Les résultats sont présentés en Figure 4.14.

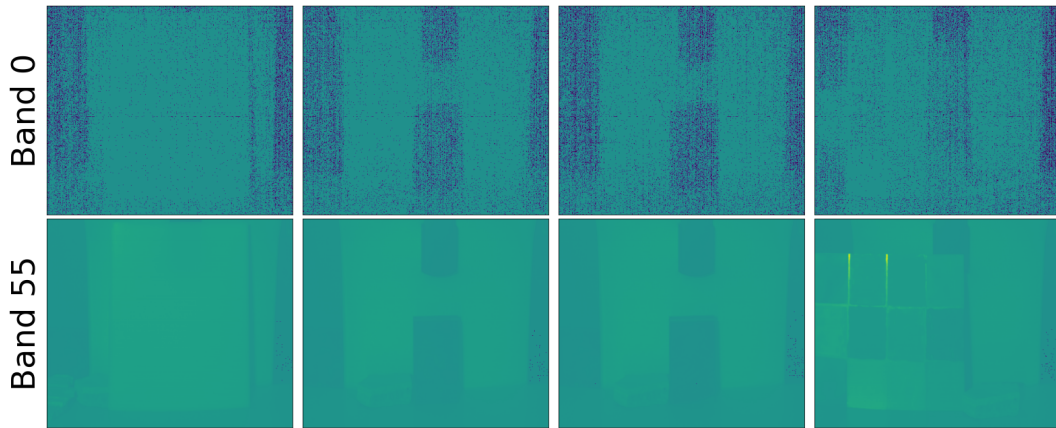


FIGURE 4.14 – Bandes spectrales 0 et 55 (899 et 1424nm) de radiance des échantillons en Figure 4.8. La radiance est calculée avec la méthode « *linear closest* ». Les pixels d'intensité négative sont affichés en violet, ceux d'intensité positive sont affichés en gradient du vert au jaune.

Sur ces images, nous pouvons constater que des valeurs négatives subsistent. Ces valeurs négatives sont majoritairement situées dans les longueurs d'onde extrêmes du spectre acquis, proches de 2500nm et surtout de 900nm, pour lesquelles la caméra présente un faible rapport signal-bruit. Spatialement, ces valeurs se situent principalement dans l'arrière plan et dans les zones très sombres de l'image.

4.6.7 Conclusion de la sélection de l'image *DC*

Nous avons présenté plusieurs méthodes d'utilisation des images *DC* pour le calcul des images de radiance. Pour chacune de ces méthodes, nous avons pu voir que la quantité de bruit du capteur retirée était soit trop importante, faisant ainsi apparaître des valeurs de radiance négatives, soit trop faible.

La dernière méthode proposée, « *linear closest* », permet de réduire l'impact de la non-linéarité de la dérive temporelle. Elle permet également d'obtenir des images de radiance avec un bruit du capteur qui reste faible tout au long de la session.

De plus, la présence de valeurs négatives est inévitable sans augmenter de manière importante le niveau de bruit du capteur dans les images. Dans les Figures 4.9 à 4.14, l'image de gauche a été acquise 29 secondes après la première image *DC*, soit un intervalle de temps durant lequel le bruit du capteur aurait

dû augmenter légèrement. Mais au contraire, les premières bandes de cette image présentent des valeurs négatives quelle que soit la méthode utilisée. Nous voyons deux raisons possibles. Il est d'abord possible que le bruit du capteur ait diminué entre l'acquisition de l'image *DC* et celle de l'échantillon. L'autre possibilité est que le bruit du capteur change très rapidement pour chaque pixel, ce qui peut être en raison de la sensibilité du capteur ou de l'obturateur ne filtrant pas suffisamment la lumière. Dans les deux cas, les fluctuations importantes du bruit du capteur sur une durée courte (inférieure à une minute) impliquent qu'il est très difficile d'effectuer une calibration supprimant tout le bruit sans générer de valeurs négatives. Une bonne méthode de calcul de la radiance implique alors soit un modèle très précis de la dérive temporelle du bruit du capteur, soit une estimation de ce bruit directement dans les images à calibrer.

4.7 Comparaison des références de blanc

La calibration nécessite la présence d'une référence de blanc. Comme mentionné en Section 3.3.2, nous avons utilisé deux objets comme références de blanc. Ces objets sont une plaque de référence de blanc en plastique PTFE (Polytétrafluoroéthylène) et une plaque en Spectralon, fixée à une plaque en bois par le fournisseur. Afin de décider de la référence à utiliser, nous avons conduit une analyse spectrale sur chacun des deux objets. Nous présentons en Figure 4.15 les spectres acquis en début et en fin d'une même session.

Nous pouvons constater que l'échantillon de Spectralon réfléchit davantage de lumière que l'échantillon de PTFE, en début comme en fin de session. Par simplicité, les spectres d'une seule session sont présentés ici. Dans toutes les sessions, pour des échantillons de Spectralon et de PTFE acquis au même moment (c'est-à-dire avec une faible dérive temporelle), l'échantillon de Spectralon réfléchit une quantité de lumière plus importante que le PTFE.

Contrairement à l'échantillon de PTFE, l'échantillon de Spectralon acquis en fin de session (courbe rouge) a une luminosité plus faible que celui du début de session (courbe orange). Cela peut être dû à des modifications de la scène ou une orientation différente du même échantillon.

De plus, sur certaines images, l'échantillon de Spectralon apparaît opaque tandis que l'échantillon de PTFE apparaît translucide, ce qui est inadéquat pour une référence de blanc. Cela s'explique par le fait que, malgré l'épaisseur similaire des deux échantillons – environ 6mm –, l'échantillon de Spectralon est beaucoup plus dense que celui de PTFE (Helmlinger et Arrecchi [2012]).

Nous décidons donc d'utiliser l'échantillon de Spectralon comme référence de blanc pour le calcul de la réflectance.

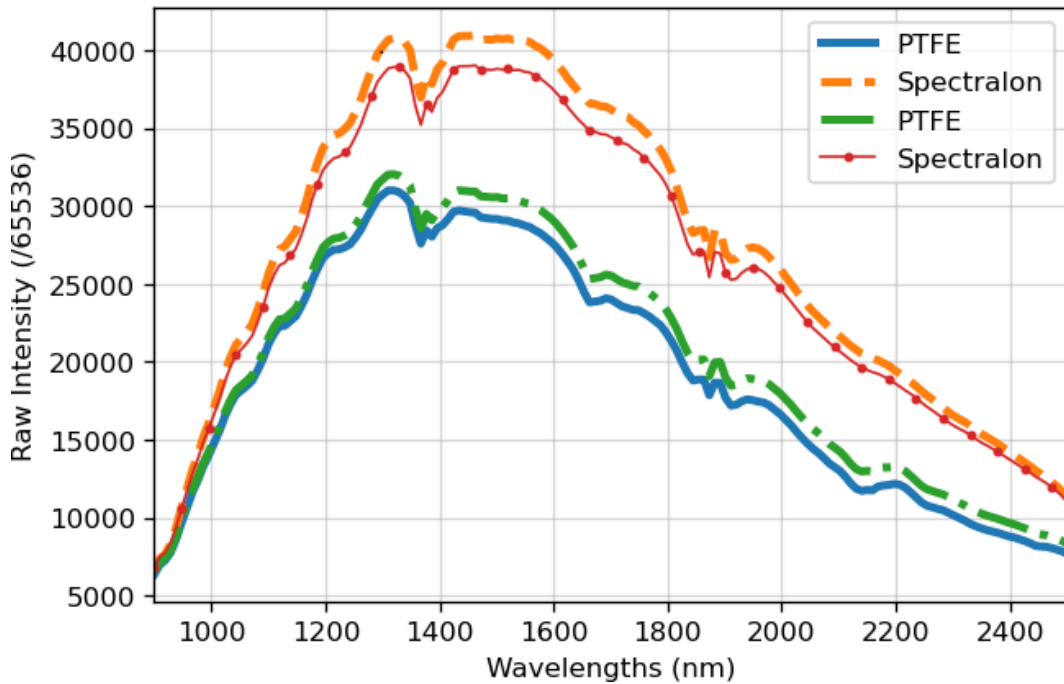


FIGURE 4.15 – Spectres bruts des cibles de références de blanc en PTFE (Polytétrafluoroéthylène) et en Spectralon en début (bleu, orange) et fin (vert, rouge) d’une même session.

4.8 Calibration par référence de gris

L’acquisition d’une image de référence de blanc dans les mêmes conditions que chaque échantillon n’est pas toujours possible. Nous souhaitons donc évaluer la possibilité d’utiliser un échantillon intermédiaire, que nous nommerons « référence de gris ». Pour cela, cette référence de gris doit être présente à la fois dans l’image à calibrer, exprimée en radiance, et dans une image déjà calibrée, exprimée en réflectance. Des spectres de gris \bar{g}_{rad} et \bar{g}_{ref} sont extraits respectivement de ces deux images à partir de l’équation 4.2. La nouvelle image de « gris-réflectance », I_{gref} , est alors calculée à partir de l’équation :

$$I_{\text{gref}}(x, y, \lambda) = I_{\text{rad}}(x, y, \lambda) * \frac{\bar{g}_{\text{ref}}(\lambda)}{\bar{g}_{\text{rad}}(\lambda)}. \quad (4.4)$$

Nous considérons ici uniquement des images de radiance calibrées à partir de la méthode « *Linear closest* » décrite en Section 4.6.

Nous présentons en Figure 4.16 les régions d’intérêt d’un échantillon de ciment (vert) et d’une référence de gris (bleu) acquis au cours des deux sessions en extérieur. La référence de gris est la plaque de PTFE posée sur une table en bois peinte en blanc. La plaque de PTFE a été sélectionnée comme référence

de gris en raison de sa réflectance importante et très proche du Spectralon, ce qui permet de tester l'utilisation d'une référence de gris quasi-idéale. Des spectres sont extraits pour chaque région d'intérêt sélectionnée, puis moyennés afin de ne garder qu'un spectre pour chaque échantillon.

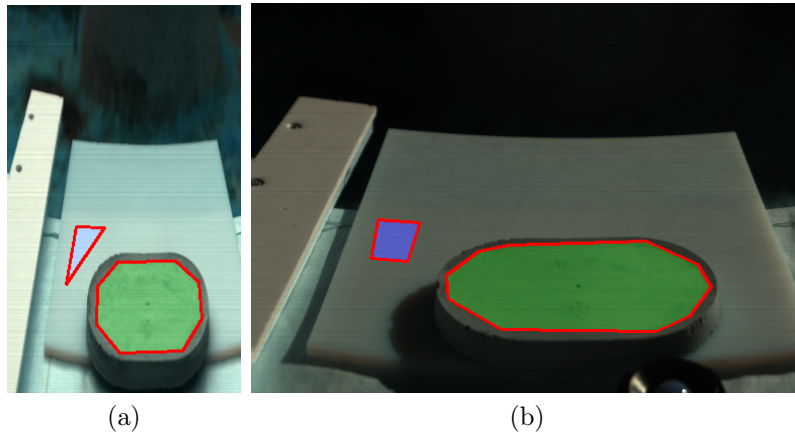


FIGURE 4.16 – Échantillon de ciment en extérieur au cours des sessions avec (a) nuages et (b) soleil. La zone bleue correspond à la région d'intérêt de la référence de gris.

Les spectres de réflectance sont présentés en Figure 4.17. Les spectres de « gris-réflectance » générés avec la session nuageuse comme référence (pour \bar{g}_{ref}) sont présentés en Figure 4.18. Le spectre de « gris-réflectance » obtenu, celui de la session ensoleillée, est proche de son spectre de réflectance. De plus, la magnitude du spectre a diminué et s'est approchée de celle du spectre de la session nuageuse.

Nous présentons également en Figure 4.19 les régions d'intérêt d'un échantillon de ciment (vert) et d'une référence de gris (bleu) acquis au cours de trois sessions en intérieur. La référence de gris est la table en bois sur laquelle sont posés les échantillons. Les spectres de réflectance de ces échantillons sont présentés en Figure 4.20 et ceux de « gris-réflectance », avec la session 5 comme référence, en Figure 4.21. Le spectre de « gris-réflectance » de la session 4 a des variations proches de celles du spectre de réflectance. En revanche, contrairement aux échantillons en extérieur, sa magnitude est très différente de celle attendue.

Le spectre de « gris-réflectance » de la session 3 présente une quantité importante de variations locales qui ne sont pas présentes dans le spectre de réflectance. Sur les images de la Figure 4.19, nous pouvons voir que la référence de gris (région en bleu) ne contient pas exactement la même information. Sur les images 4.19b et 4.19c, la table reflète l'objet contre lequel est positionné l'échantillon. De plus, la lumière n'est pas distribuée de la même manière sur l'échantillon, qui est à plat sur l'image 4.19a et face à la caméra sur les deux

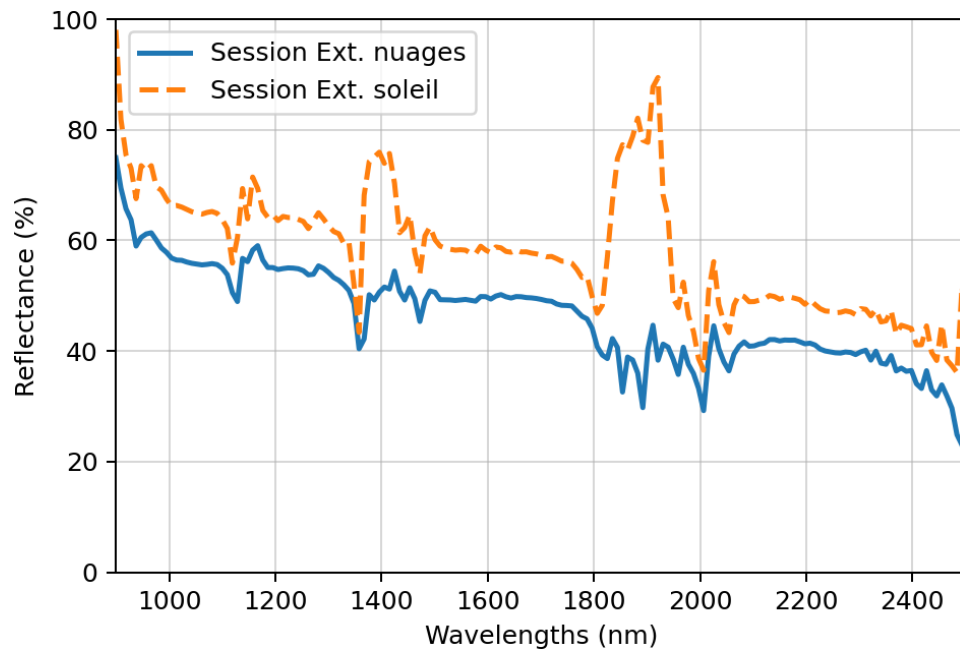


FIGURE 4.17 – Spectres de réflectance d'un échantillon de ciment au cours des deux sessions en extérieur.

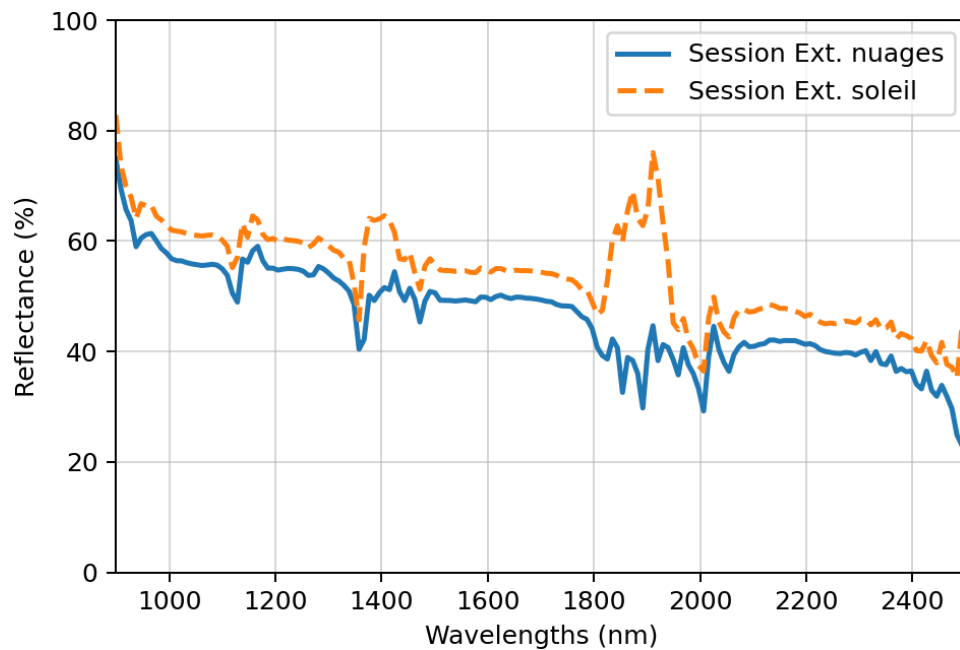


FIGURE 4.18 – Spectres de « gris-réfectance » d'un échantillon de ciment au cours des deux sessions en extérieur. La session de référence (pré-calibrée) est la session avec nuages.

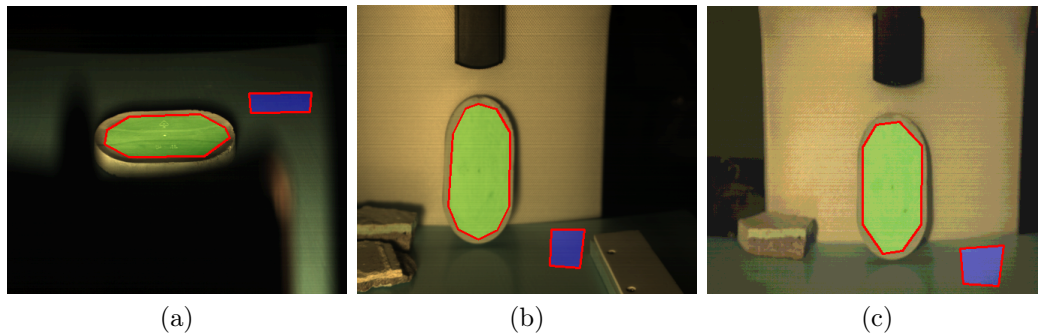


FIGURE 4.19 – Échantillon de ciment au cours des sessions (a) 3, (b) 4 et (c) 5 en intérieur. La zone bleue correspond à la région d'intérêt de la référence de gris.

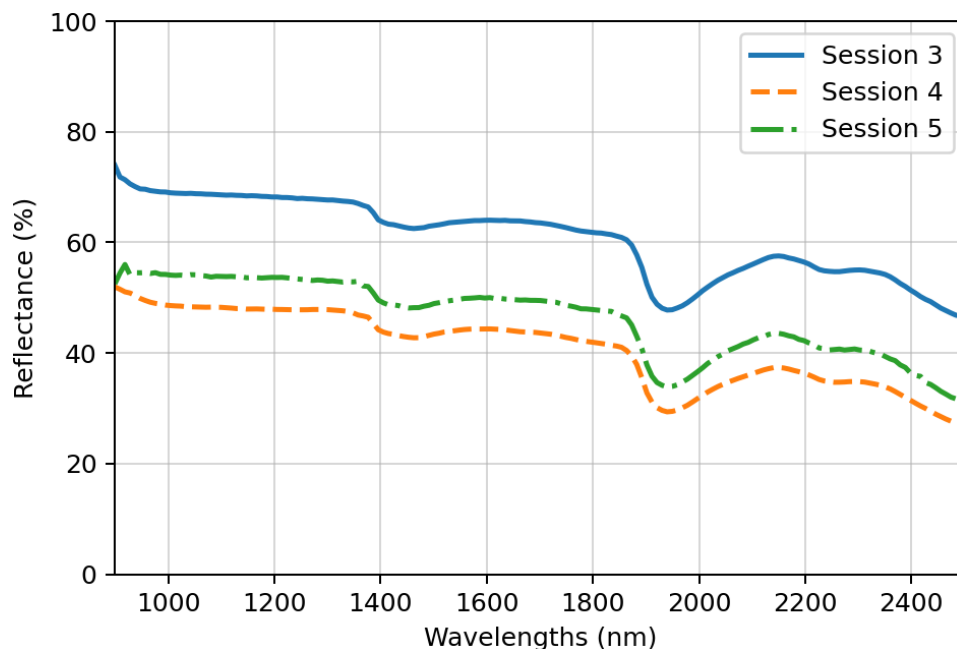


FIGURE 4.20 – Spectres de réflectance d'un échantillon de ciment au cours des sessions 3, 4 et 5.

autres images. Ces deux raisons peuvent influencer de manière importante le spectre de l'échantillon ou de la référence de gris et donc la calibration.

Pour conclure, comme nous pouvons le constater sur la Figure 4.18, l'utilisation d'une « référence de gris » peut permettre de calibrer des images pour lesquelles aucune référence de blanc n'est disponible. Pour cela, un objet ou matériau sélectionné comme référence de gris doit rester identique entre plusieurs sessions, car des variations faibles ou des reflets peuvent avoir un impact significatif sur le spectre résultant (voir Fig. 4.19a et session 3 Figs. 4.20 et

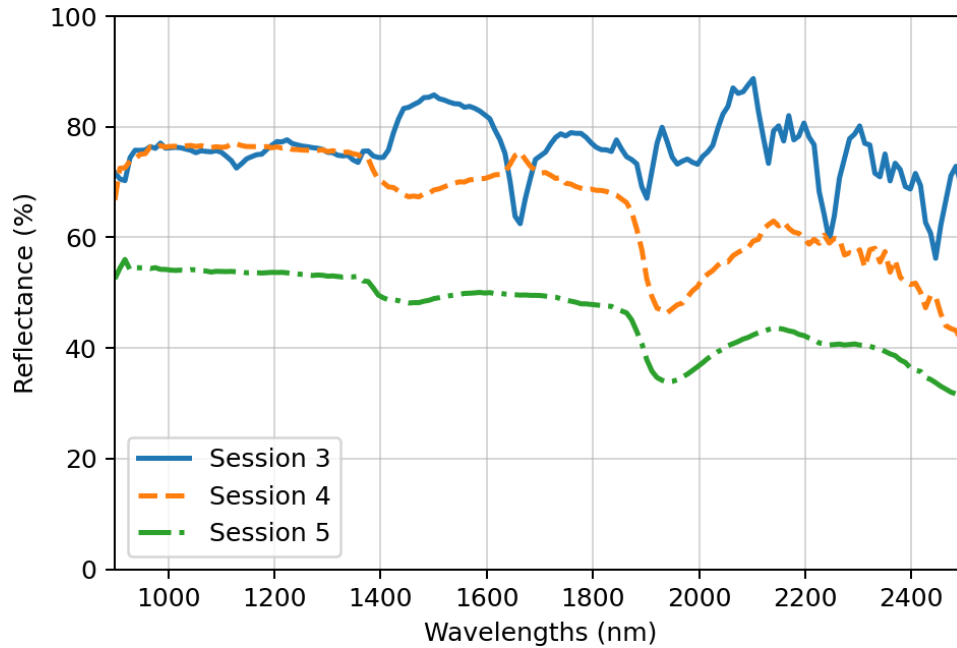


FIGURE 4.21 – Spectres de « gris-réfectance » d'un échantillon de ciment au cours des sessions 3, 4 et 5. La session de référence (pré-calibrée) est la session 5.

4.21). De plus, même dans des conditions idéales, la magnitude du spectre de « gris-réfectance » peut différer significativement de celle de la réflectance (voir session 4 Figs. 4.20 et 4.21). Cette méthode d'utilisation de cibles comme références de gris peut alors s'avérer pertinente en conditions opérationnelles, mais seulement lorsqu'une cible ou un type de matériau est identique entre plusieurs acquisitions. Enfin, il est pertinent de préciser que l'utilisation d'étapes intermédiaires pour calculer la réflectance ajoutent des incertitudes dans le spectre produit, et particulièrement lorsque la référence de gris a une magnitude faible et donc un ratio signal-bruit faible.

4.9 Limitations de la calibration

En conditions opérationnelles, la conduite de la calibration est limitée par plusieurs facteurs.

L'équation 4.3 suppose que la scène est illuminée de manière uniforme car la valeur \bar{w}_{rad} est la même pour tous les pixels (x, y) . En réalité, l'illumination est rarement uniforme, en raison de la topologie de la scène ou d'obstructions entre la source d'illumination et une partie de la scène qui génèrent des ombres (par exemple les nuages). Il est alors possible d'utiliser, à l'instar de l'équation 4.1, une image entièrement couverte par la référence de blanc afin de calibrer chaque pixel individuellement, mais l'acquisition d'une telle image est rarement

possible en conditions réelles.

De plus, la dérive temporelle (voir Sec. 4.3) montre que des facteurs extérieurs peuvent perturber la calibration de manière importante. La température ambiante varie systématiquement pendant des vols, ce qui impacte directement le bruit du capteur, et nécessite la mise en place de nouvelles méthodes de correction, telles que celles proposées en Section 4.6.

De manière similaire à la dérive temporelle, [Miura et Huete \[2009\]](#) mentionnent les variations d'illumination rapides qui peuvent se produire en extérieur et soulignent l'importance d'une nouvelle acquisition fréquente de la référence de blanc (< 30min). En revanche, l'utilisation de références de gris (voir Sec. 4.8) pourrait permettre de réduire cette contrainte, à condition que l'objet utilisé comme référence reste inchangé pendant l'ensemble des acquisitions.

4.10 Conclusion

Dans ce chapitre, nous avons présenté la calibration d'images hyperspectrales, en transformant les images brutes en images de radiance puis de réflectance. Nous avons identifié une dérive temporelle des images brutes causée par la température ambiante. L'impact de la dérive temporelle sur les images de radiance a été constaté et nous avons proposé plusieurs méthodes de sélection du courant d'obscurité (*DC*) pour atteindre un compromis entre le niveau de bruit du capteur et la quantité de valeurs négatives dans les images de radiance.

Nous avons également analysé l'utilisation de plusieurs références de blanc pour la calibration. Nous avons exploré l'utilisation de « références de gris » pour faciliter la calibration en conditions opérationnelles. Cependant, la procédure de calibration proposée n'est pas adaptée à des acquisitions en conditions opérationnelles, tel que le vol d'un drone pendant plusieurs heures, et les méthodes d'amélioration proposées ne permettent actuellement pas de s'en affranchir complètement en conditions opérationnelles.

Chapitre 5

Analyse d'images hyperspectrales

Sommaire

5.1	Introduction	58
5.2	Pré-traitements	59
5.2.1	Reconstruction d'images couleur	59
5.2.2	Segmentation et boîtes englobantes de véhicules	59
5.2.3	Segmentation d'objets	61
5.2.4	Correction atmosphérique	63
5.2.5	Normalisation des spectres	65
5.3	Analyse visuelle	65
5.3.1	Analyse d'objets de même couleur	65
5.3.2	Analyse d'objets de même matériau	69
5.3.3	Observations générales	70
5.4	Mesures de similarité	71
5.4.1	Classification de spectres	71
5.4.2	Détection de matériaux	72
5.4.3	Discrimination de pixels	74
5.4.4	Réduction de dimension	77
5.4.5	Détection de matériaux avec ACP	79
5.4.6	Discrimination de pixels avec ACP	80
5.5	Acquisitions nocturnes/sans lumière	81
5.6	Conclusion	82

5.1 Introduction

Nous avons vu au chapitre précédent le principe de la calibration et ses limitations sur nos données et pour la calibration en conditions opérationnelles de surveillance. Dans un premier temps, nous présentons les pré-traitements

conduits sur les données acquises au Chapitre 4. De manière complémentaire à celles du Chapitre 2, des analyses spectrales et par mesures de similarité sont conduites sur les images hyperspectrales acquises et sont présentées afin de confirmer l'utilité de l'hyperspectral pour la détection et la discrimination de matériaux.

5.2 Pré-traitements

5.2.1 Reconstruction d'images couleur

Les images hyperspectrales acquises avec la caméra FireflEYE S185, décrites en Section 3.4, étant acquises dans le domaine visible, il est possible de les utiliser pour reconstruire des images couleur. Ces images reconstruites sont plus proches de la vision humaine, ce qui améliore leur interprétabilité, mais également du type d'images pour lesquels sont définis une majorité des outils de traitement d'image.

La reconstruction est constituée de trois étapes :

- Les bandes hors du domaine visible (450-780nm) sont supprimées.
- Les bandes spectrales conservées sont interpolées en fonction de leur longueur d'onde en canaux RGB (*Red*, *Green*, *Blue*).
- Les canaux ne sont pas équitablement représentés, tel le canal bleu qui est sous-représenté dû à l'absence de bandes entre 380 et 450nm. Pour cette raison, nous normalisons chaque canal séparément, au risque de sur-représenter un canal qui serait peu présent dans une scène.

Nous présentons en Figure 5.1 un exemple d'image hyperspectrale (voir Fig. 5.1a) et de sa reconstruction couleur (voir Fig. 5.1b). Les images présentées en Figure 3.2 sont des images couleur reconstruites par ce procédé.

5.2.2 Segmentation et boîtes englobantes de véhicules

Afin d'analyser les résultats des méthodes de détection de véhicules, il est nécessaire d'avoir une labellisation (ou vérité terrain) de ces images sous forme de segmentation ou de boîtes englobantes. Nous séparons les véhicules en trois classes, celle des voitures, celle des camions et une pour les autres types de véhicules (motos, camionnettes, camping-cars et bus) en raison de la faible représentativité de ces autres véhicules. Cette labellisation est constituée de trois étapes.

Dans un premier temps, nous utilisons le Mask-RCNN (Abdulla [2017]), un réseau de neurones pré-entraîné à la détection d'objets et de véhicules, sur les images couleur reconstruites décrites en Section 5.2.1. Le Mask-RCNN produit un masque et un label (ou une classe) par objet détecté et pour chaque image et nous fournit donc une première segmentation des véhicules dans les images.

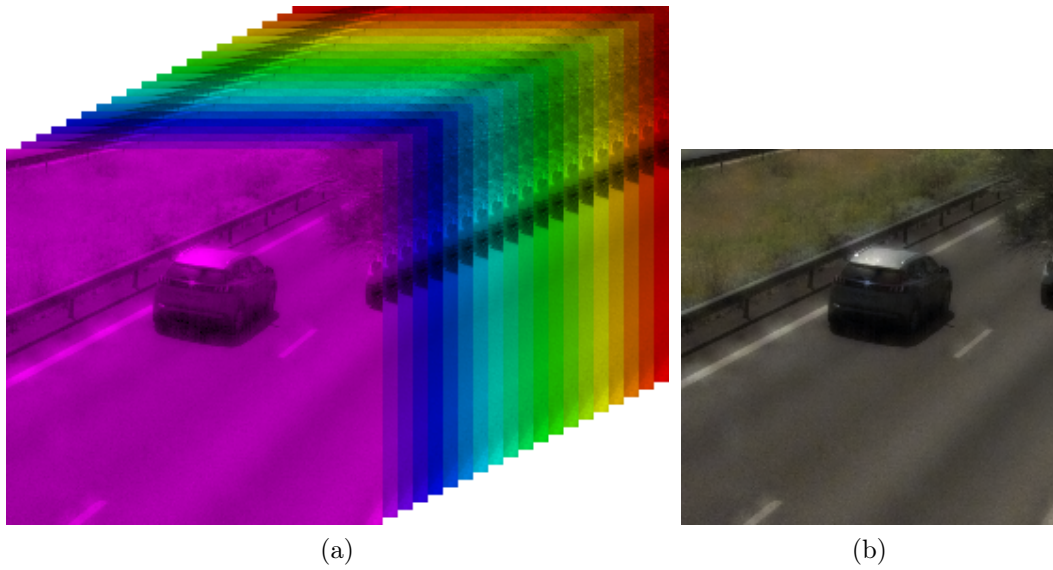


FIGURE 5.1 – (a) Image hyperspectrale d'un véhicule et (b) reconstruction couleur de cette image hyperspectrale.

Un exemple de segmentation par Mask-RCNN est présenté en Figure 5.2b. Cette méthode a l'avantage d'être automatique et ne nécessite d'intervention humaine qu'afin de vérifier les résultats. En revanche, le Mask-RCNN commet souvent des erreurs sur la forme des polygones et parfois sur les labels (voir Figure 5.2b).

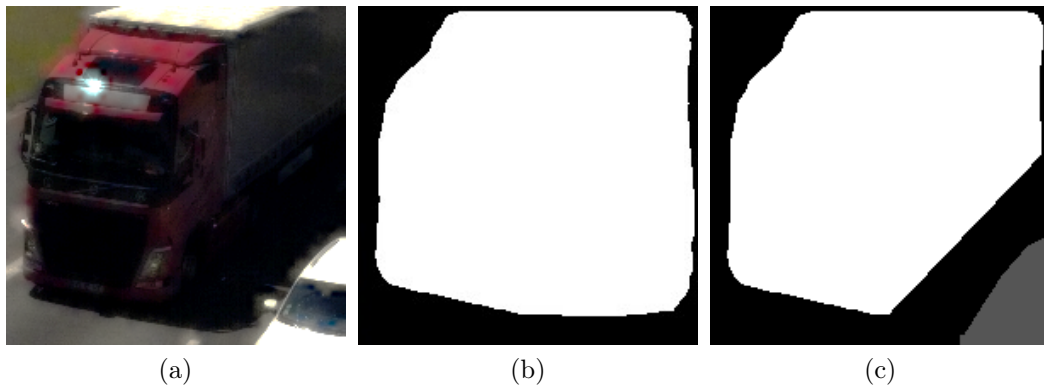


FIGURE 5.2 – Exemple de segmentation des véhicules. (a) Image reconstruite en couleur pour la génération des masques. (b) Masque généré par le Mask-RCNN et (c) masque obtenu après segmentation manuelle.

Nous conduisons ensuite une segmentation manuelle afin de corriger les erreurs du Mask-RCNN. Chaque masque est généré manuellement par-dessus le masque du Mask-RCNN, puis superposé sur le précédent afin d'éliminer les erreurs de segmentation. Un exemple de masque est présenté en Figure 5.2c,

sur laquelle les pixels en noir désignent l'arrière-plan, en blanc un camion, en gris foncé une voiture et en gris clair un autre type de véhicule.

Afin de permettre l'utilisation d'autres méthodes que celles effectuant de la segmentation, nous convertissons les masques (après segmentation manuelle) en boîtes englobantes. Pour cela, nous cherchons les groupes de pixels de même classe dans les masques, puis utilisons les coordonnées de leurs extrémités comme bords d'une boîte englobante appartenant à la même classe. Il est toutefois important de noter que, dans le cas où des objets de même classe sont adjacents dans une image, la segmentation de ces objets peut se confondre en un unique groupe de pixels. Un tel exemple est présenté par les deux camions des Figures 5.2a et 5.2c, dans quel cas notre méthode génère une unique boîte englobante pour les deux véhicules.

Le Mask-RCNN a été appliqué sur les 15000 images de véhicules. La segmentation manuelle et la génération des boîtes englobantes ont été conduites sur près de 1000 de ces masques.

5.2.3 Segmentation d'objets

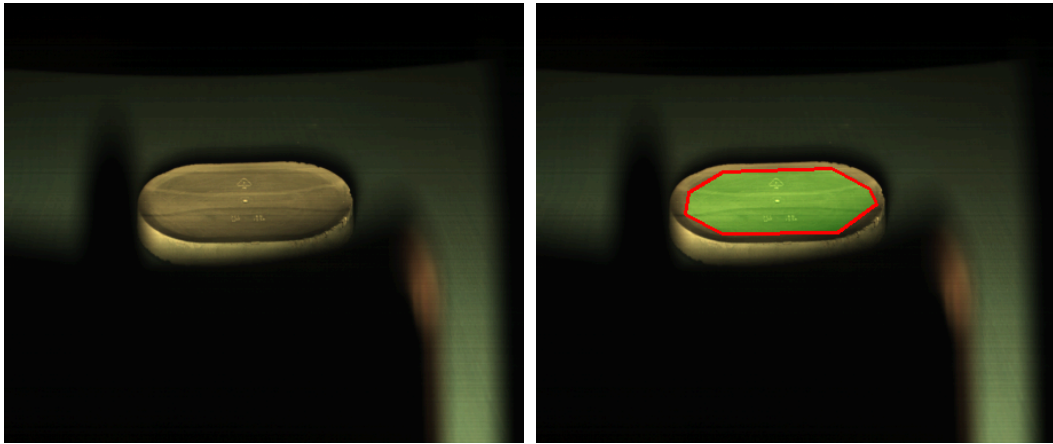
Nous conduisons ensuite la segmentation d'objets dans les images hyperspectrales du domaine 900-2500nm (voir Sec. 3.5). Dans ces images, l'objet d'intérêt (ou échantillon) est un unique objet situé au milieu d'un arrière-plan lui-même constitué de plusieurs objets ou zones différentes. Le Mask-RCNN utilisé précédemment n'étant pas adapté pour détecter ce type d'objets (bloc de ciment, planche de bois...), nous ne conduisons que l'étape de segmentation manuelle sur ces données.

Dans les images d'objets, les bandes spectrales du domaine visible ne sont pas disponibles dans les images d'objets, ne permettant ainsi pas la reconstruction d'images couleur définie en Section 5.2.1. Avant la segmentation, nous construisons une image en fausses couleurs dans laquelle les canaux rouge, vert, bleu correspondent respectivement aux bandes 73, 31 et 15 (1596nm, 1195nm et 1043nm) de l'image hyperspectrale. L'image en fausses couleurs d'un échantillon de ciment est présentée en Figure 5.3a.

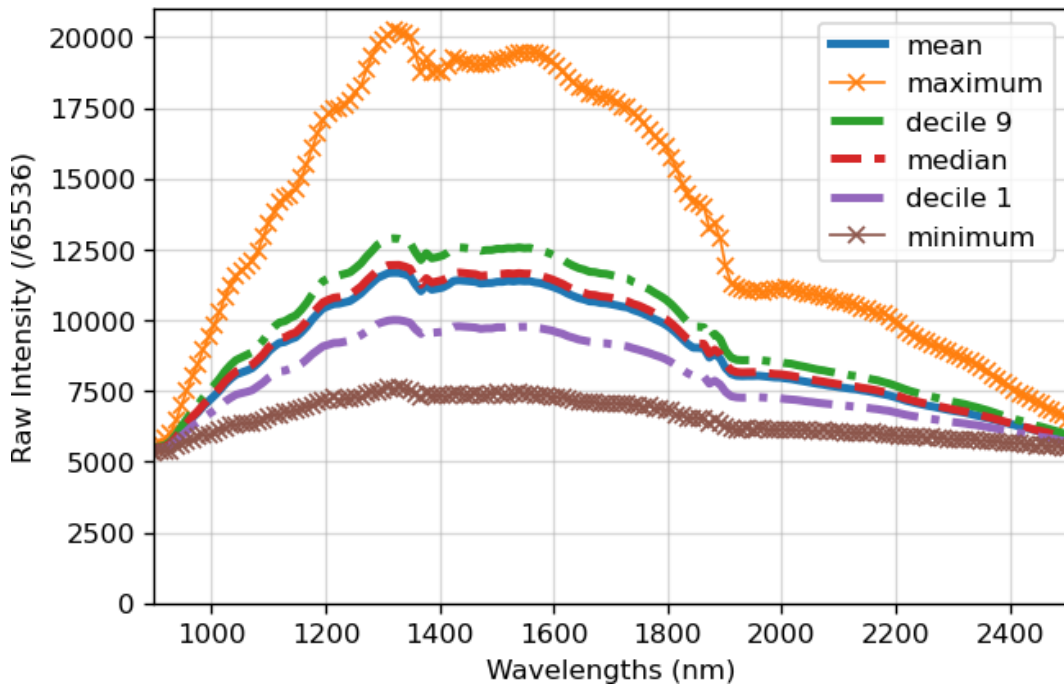
Les bords des objets sont ignorés afin de ne pas sélectionner des pixels d'autres objets que l'objet d'intérêt. Un exemple de segmentation est présenté en Figure 5.3b. Après la segmentation, il est possible d'extraire autant de profils spectraux qu'il y a de pixels dans la région sélectionnée.

La Figure 5.3c présente 6 courbes des valeurs spectrales brutes minimale à maximale de l'échantillon de ciment en Figure 5.3a. Nous pouvons observer des différences d'intensité importantes entre ces courbes, jusqu'à un facteur 3 entre les valeurs minimum et maximum des longueurs d'onde proches de 1400nm. Ces différences sont principalement dues à deux facteurs, qui sont l'illumination non-uniforme et les aspérités présentes dans les échantillons.

L'un de nos objectifs étant la discrimination de matériaux, nous souhaitons



(a) Image en fausses couleurs d'un échantillon de ciment, bandes 73, 31 et 15. (b) Polygone (en rouge) et région d'intérêt (en vert) de l'image (a).



(c) Spectres dans la région d'intérêt sélectionnée en (b) des valeurs minimum, maximum, moyenne et déciles 1, 5 (médiane) et 9 pour chaque bande spectrale.

FIGURE 5.3 – Exemple d'extraction de spectre depuis l'image hyperspectrale d'un échantillon de ciment.

pouvoir comparer entre eux les échantillons issus de différents types de matériaux ou acquis dans différentes conditions. Dans un premier temps, il sera intéressant de considérer des échantillons représentés chacun par un unique spectre représentatif afin de réduire la quantité de données à interpréter. Nous choisissons le spectre moyen comme spectre représentatif de chaque échantillon

car il représente l'ensemble des variations d'intensité de l'échantillon en Figure 5.3c.

5.2.4 Correction atmosphérique

Le soleil émet de la lumière dans l'ensemble du domaine visible et dans une partie de l'infrarouge, mais toutes ces longueurs d'onde n'atteignent pas la surface de la Terre. Certaines de ces longueurs d'onde sont absorbées ou réfléchies par l'atmosphère, et majoritairement par les molécules de H₂O. Il est alors possible de calculer la proportion de lumière transmise par l'atmosphère pour chaque longueur d'onde, aussi appelée transmittance atmosphérique. La courbe de transmittance atmosphérique est présentée en Figure 5.4. Nous pouvons particulièrement observer une transmittance proche de 0 autour de 1400nm et 1900nm.

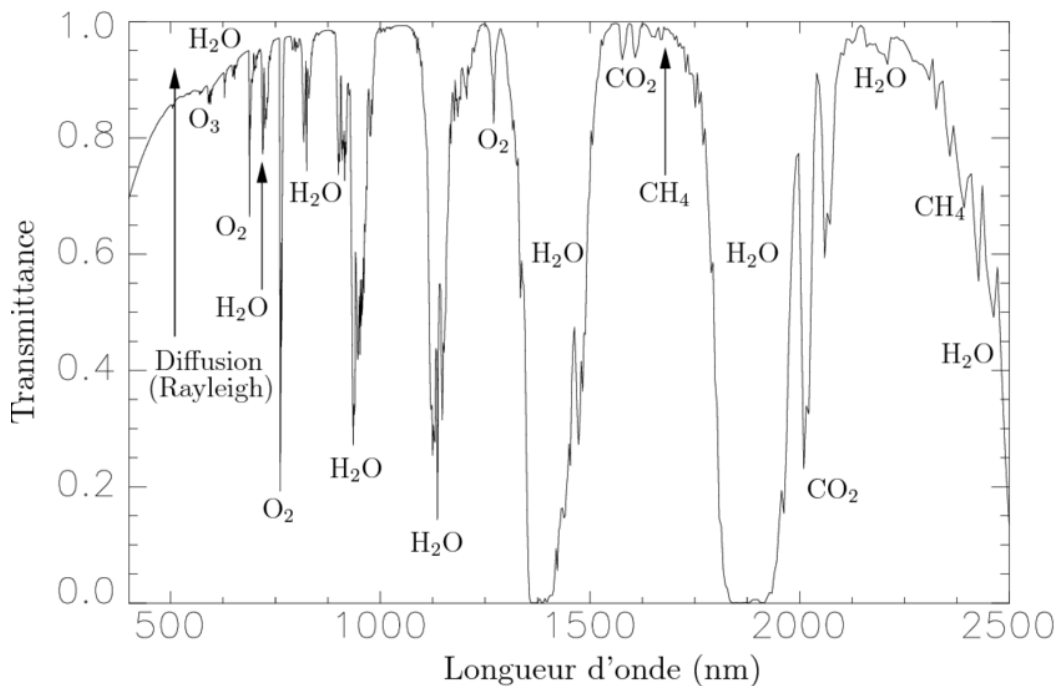


FIGURE 5.4 – Figure extraite de Marion [2003] : « Transmittance atmosphérique directe pour un trajet vertical entre le sol et l'espace (simulation MODTRAN4, modèle US 1976 Standard). »

La Figure 5.5 présente deux spectres bruts d'un même échantillon de ciment acquis en intérieur et en extérieur. Les différences d'intensité entre les deux spectres sont dues à l'illumination des deux scènes. Des creux importants sont visibles autour de 1100nm, 1400nm et 1900nm dans le spectre extérieur et

correspondent à ceux de la courbe de transmittance atmosphérique. Ces mêmes longueurs d'onde sont peu impactées dans le spectre en intérieur, en raison de la faible distance parcourue par la lumière et/ou d'une faible quantité de vapeur d'eau (H_2O) dans l'air. Ces observations sont cohérentes avec l'acquisition et l'analyse par [Schneider *et al.* \[2009\]](#) de spectres en laboratoire et en extérieur, desquels seuls les spectres acquis sous lumière naturelle sont soumis à une forte absorption atmosphérique.

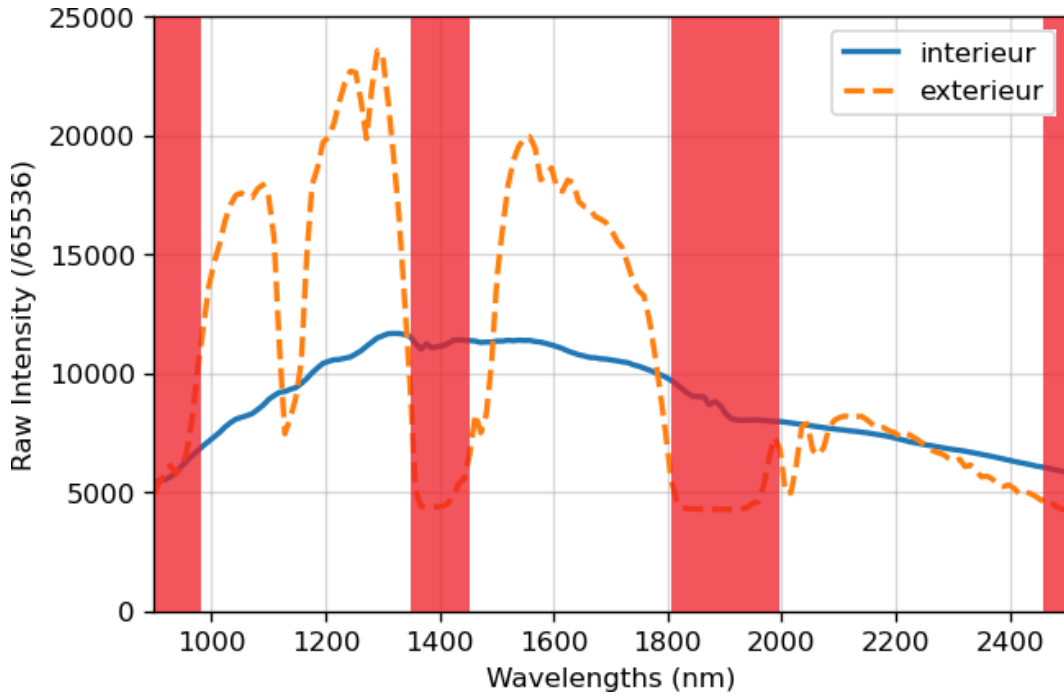


FIGURE 5.5 – Spectres bruts d'un échantillon de ciment, acquis en conditions intérieures (session 3) et extérieures (ensoleillé). Les zones rouges représentent les bandes spectrales retirées des spectres extérieurs.

En extérieur, l'absorption atmosphérique de certaines longueurs d'onde est telle que seul le bruit du capteur est acquis dans les bandes spectrales correspondantes, ce qui rend ces bandes inexploitable. Pour éviter que ces longueurs d'onde « mortes » n'affectent les analyses, nous les supprimons des spectres extérieurs. Les zones en rouge sur la Figure 5.5 marquent les bandes spectrales retirées à cet effet. Nous supprimons d'abord les bandes de 1350nm à 1450nm et de 1800nm à 2000nm, sélectionnées par l'observation des spectres affectés. Nous retirons aussi les bandes spectrales proches de 2500nm car elles sont affectées par l'absorption atmosphérique. Les bandes spectrales autour de 900nm, également affectées par l'absorption atmosphériques et par la faible sensibilité de la caméra dans ces longueurs d'onde extrêmes, sont supprimées dans l'ensemble des spectres, intérieurs et extérieurs. Nous pouvons également constater

des perturbations autour de 1100nm. Cependant, ces bandes spectrales n'ont pas été supprimées car l'absorption atmosphérique y est moins importante que dans les longueurs d'onde décrites précédemment.

Nous notons également que dans les spectres en intérieur, un traitement de l'air pour réduire la quantité de vapeur d'eau (H₂O) et de CO₂ (Kokaly *et al.* [2017]) peut permettre d'améliorer légèrement la qualité des spectres dans les longueurs d'onde affectées. Une telle correction rendrait également possible une estimation de l'humidité des échantillons par analyse spectrale. Cependant, l'impact de ces gaz sur les spectres en intérieur est faible et ce traitement de l'air n'a pas été mis en oeuvre au cours de nos expérimentations.

5.2.5 Normalisation des spectres

La normalisation permet de comparer l'évolution des spectres. Nous normalisons les spectres minimum, décile, moyen et maximum (voir Fig. 5.3c) de la même manière que dans la Section 2.3. L'objectif ici est d'observer si les différents spectres bruts d'un même matériau suivent la même évolution en fonction de la longueur d'onde : la normalisation permet de comparer les variations des spectres en ignorant leur magnitude.

Nous présentons en Figure 5.6 les spectres de ciment et de bois avant et après normalisation et acquis au cours d'une même session. Dans le cas des spectres de ciment (Fig. 5.6a et 5.6b), la normalisation permet de voir que les courbes représentatives des valeurs de l'échantillon suivent pratiquement la même évolution et que seule la magnitude change. Ce n'est en revanche pas le cas pour les spectres de l'échantillon de bois (Fig. 5.6c et 5.6d), qui présentent une variabilité plus importante au sein d'un même échantillon.

5.3 Analyse visuelle

Nous conduisons dans un premier temps des analyses spectrales sur les images de matériaux du domaine spectral 400-1000nm (Vis-NIR), présentées en Section 3.5.4.

5.3.1 Analyse d'objets de même couleur

Nous étudions des spectres de matériaux différents et de même couleur afin de constater si l'apport des caractéristiques spectrales permet de différencier ces objets, et ce pour plusieurs couleurs.

Objets noirs

Nous présentons en Figure 5.7a des objets de couleur noire, dont des échantillons sont sélectionnés et décrits en Tableau 5.1. Les spectres bruts de ces

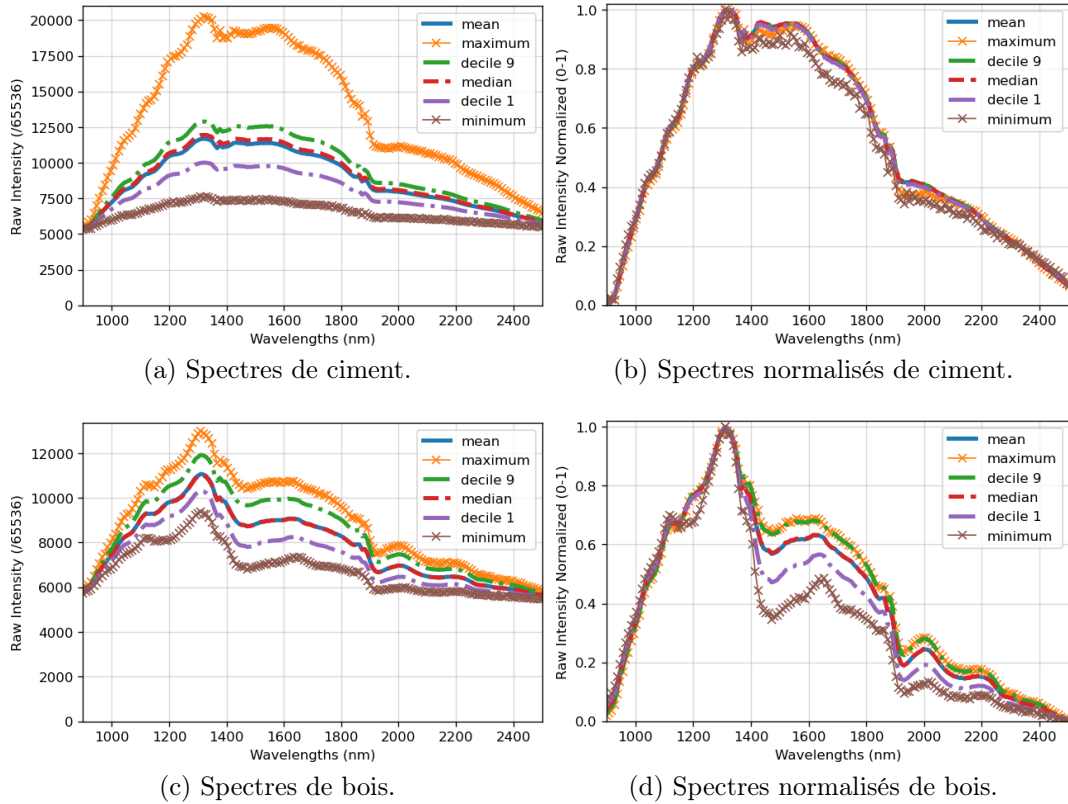


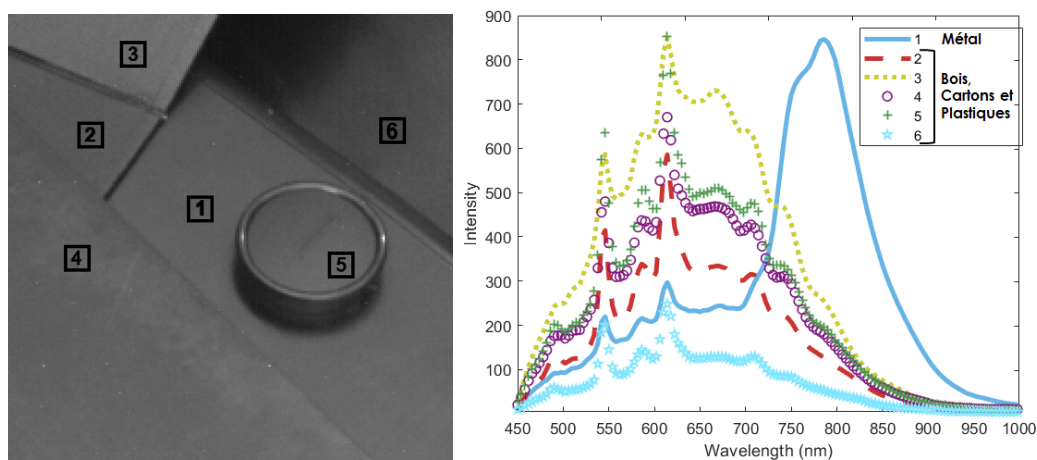
FIGURE 5.6 – Exemple de courbes minimum, maximum, moyenne et déciles d'échantillons de ciment et de bois, avant et après normalisation.

échantillons sont présentés en Figure 5.7b. Cette image n'a pas été calibrée en raison de problèmes survenus lors de l'acquisition des images *DC*, empêchant la calibration de cette image.

#	Couleur	Matériau
1	Noir	Métal
2, 3	Noir	Carton renforcé
4	Noir	Carton fin
5	Noir	Plastique
6	Noir	Bois

TABLE 5.1 – Liste des échantillons sélectionnés en Figure 5.7a.

Les échantillons 2 à 6 se distinguent visuellement très peu les uns des autres, en partie dû à l'absence de calibration. Cependant, malgré l'absence de calibration et ses valeurs faibles dans le domaine visible ($<700\text{nm}$), le spectre de l'échantillon 1 (en métal) se distingue grandement des autres à partir de ses valeurs élevées entre 700 et 1000nm, ce qui montre qu'il est possible de



(a) Image monochrome et régions d'intérêt d'objets de couleur noire. (b) Spectres moyens bruts (/65536) des régions sélectionnées dans l'image (a).

FIGURE 5.7 – (a) Image monochrome et (b) spectres d'objets de couleur noire décrits en Table 5.1.

discriminer cet échantillon des autres.

Objets rouges et jaune

Nous présentons en Figure 5.8a des objets de couleur rouge ou jaune. Des échantillons de cette image sont sélectionnés en Figure 5.8b et décrits en Tableau 5.2. Les spectres de réflectance de ces échantillons sont présentés en Figure 5.8c.

#	Couleur	Matériau
1	Rouge	Carton fin
2	Rouge	Carton épais
3	Rouge	Plastique translucide rouge sur carton épais rouge
4	Rouge	Carton épais
5	Jaune	Carton fin
6	Rouge	Plastique
7	Noir	Bois
8	Rouge	Plastique translucide rouge sur bois noir

TABLE 5.2 – Liste des échantillons sélectionnés en Figure 5.8b.

L'échantillon 3 présente une courbure similaire à celle l'échantillon 2, mais son intensité est moins élevée : l'objet translucide absorbe une partie de la lumière réfléchi par l'objet derrière lui, et réfléchit moins de lumière par lui-même.

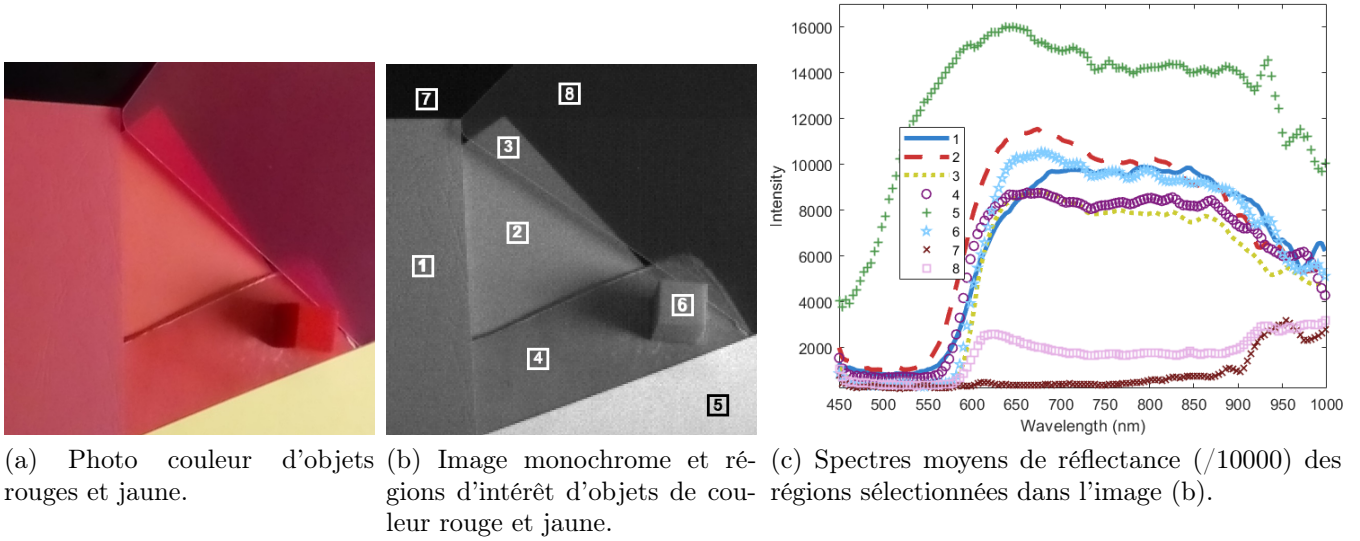


FIGURE 5.8 – (a) Photo couleur, (b) image monochrome et (c) spectres d'objets de couleur rouge et jaune décrits en Table 5.2.

Dans les spectres des échantillons 7 et 8, nous pouvons constater que l'objet translucide réfléchit effectivement une quantité faible de lumière. De plus, cet objet absorbe également une partie importante du pic d'intensité de l'objet noir autour de 950nm.

Enfin, seul un autre spectre se distingue, le spectre 5. Cependant, ce spectre correspond à l'objet jaune, et sa couleur explique cette distinction.

Les autres échantillons ont des spectres très similaires entre eux et ne sont pas visuellement discriminables dans ce domaine spectral, malgré leurs matériaux différents.

Objets bleus

Nous présentons en Figure 5.9a des objets de couleur bleue. Des échantillons de cette image sont sélectionnés en Figure 5.9b et décrits en Tableau 5.3. Les spectres de réflectance de ces échantillons sont présentés en Figure 5.9c.

#	Couleur	Matériau
1	Bleu pâle	Carton fin
2	Bleu	Papier blanc encré en bleu
3	Bleu	Carton épais plastifié
4	Bleu	Papier
5	Bleu	Carton épais
6	Bleu	Plastique

TABLE 5.3 – Liste des échantillons sélectionnés en Figure 5.9b.

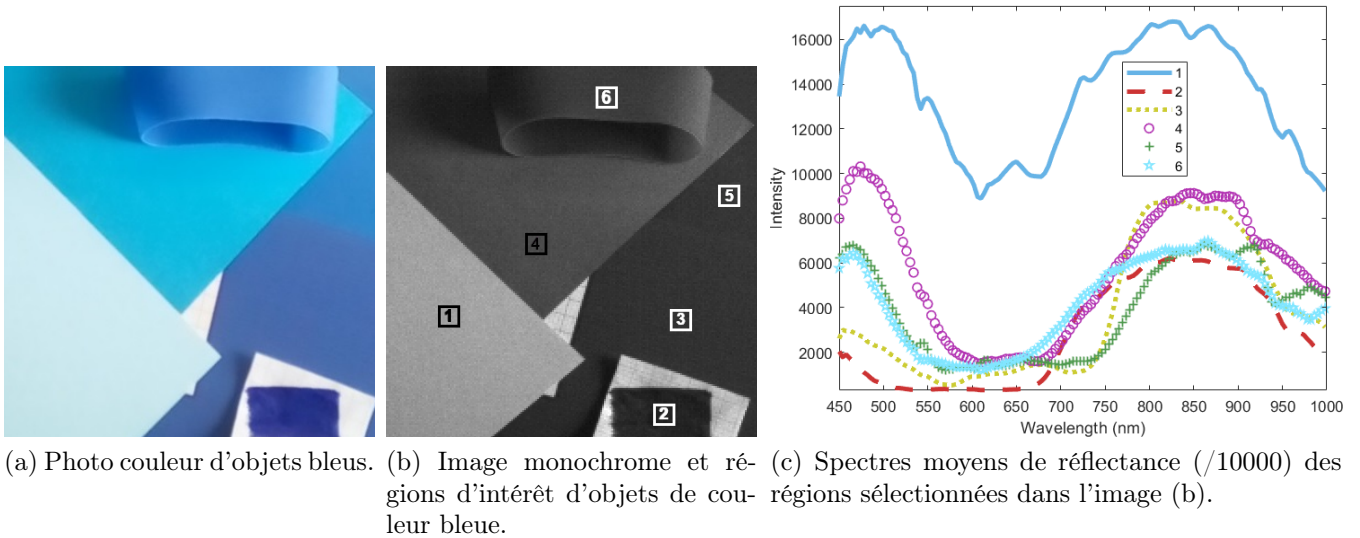


FIGURE 5.9 – (a) Photo couleur, (b) image monochrome et (c) spectres d'objets de couleur bleue décrits en Table 5.3.

Dans les spectres des échantillons, certaines caractéristiques permettant de les discriminer émergent. L'intensité des échantillons 3 et 5 commence à croître autour de 750nm, et celle des échantillons 1, 2, 4 et 6 augmente dès 680nm. Ce pic indique que ces échantillons, dont particulièrement l'échantillon 2 dont l'intensité est très faible, ont une teinte plus rougeâtre que les autres, ce qui ne se remarque pas à l'oeil nu.

Cependant, aucune caractéristique n'émerge dans le domaine spectral 450-1000nm pour permettre de discriminer ces spectres par type de matériau.

5.3.2 Analyse d'objets de même matériau

Nous analysons des spectres de même matériau et de couleur différente pour voir si des caractéristiques spectrales liées à ce matériau apparaissent.

Nous présentons en Figure 5.10a des objets en carton, de différentes couleurs. Des échantillons de cette image sont sélectionnés en Figure 5.10b et décrits en Tableau 5.4. Les spectres de réflectance de ces échantillons sont présentés en Figure 5.10c.

Les variations dans le domaine 400-750nm sont caractéristiques des couleurs des objets. Par exemple, l'échantillon 1 a une intensité élevée entre dans les longueurs d'ondes proches du bleu et dans le vert et plus faible dans les longueurs d'onde proches du rouge (600-750nm), ce qui correspond avec son apparence bleu pâle sur la Figure 5.10a. Cependant, aucune caractéristique commune à tous ces échantillons n'apparaît, y compris dans le proche infra-rouge.

#	Couleur	Matériau
1	Bleu pâle	Carton fin
2	Jaune	Carton fin
3	Rouge	Carton fin
4	Gris	Carton fin
5	Vert	Carton fin

TABLE 5.4 – Liste des échantillons sélectionnés en Figure 5.10b.

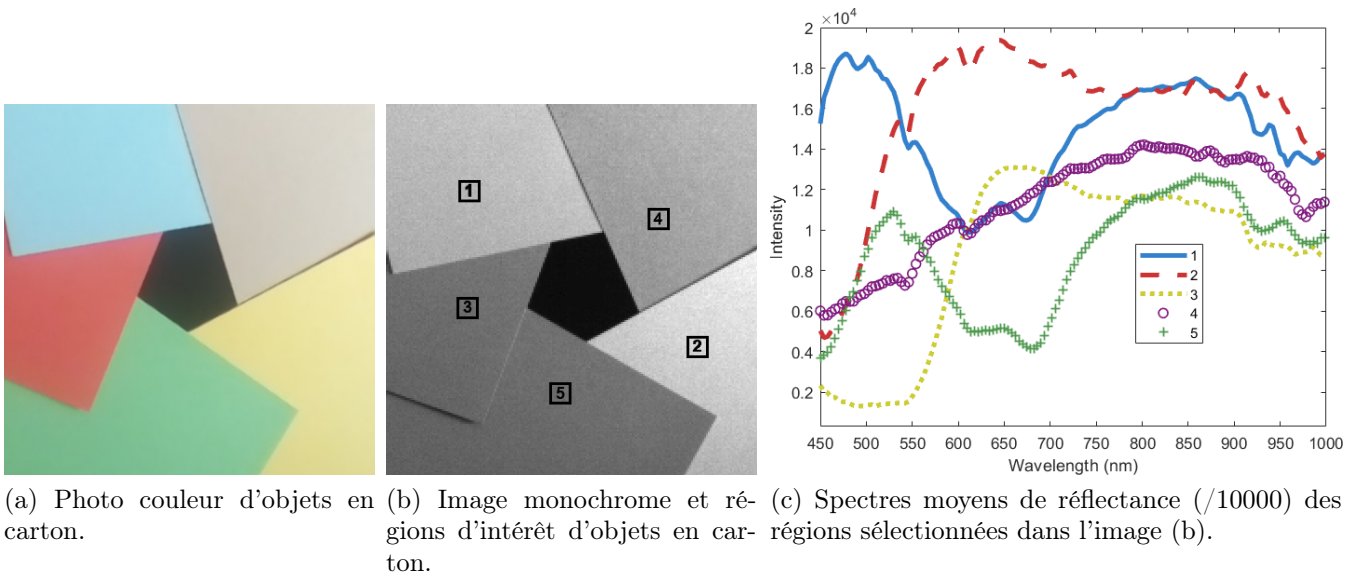


FIGURE 5.10 – (a) Photo couleur, (b) image monochrome et (c) spectres d'objets en carton décrits en Table 5.4.

5.3.3 Observations générales

De manière générale, nous pouvons également noter la présence de pics d'intensité autour de 550 et de 620nm dans les spectres bruts (Fig. 5.7b). Ces pics ne sont pas présents dans les spectres issus des acquisitions extérieures et sont vraisemblablement provoqués par l'illumination ambiante. Dans les spectres calibrés (Figs. 5.8c, 5.9c et 5.10c), des creux apparaissent aux mêmes longueurs d'onde. Ces creux sont dus à la calibration, au cours de laquelle les valeurs de ces longueurs d'onde sont sur-estimées dans la référence de blanc et sont ainsi sur-corrigées.

Le test des objets rouges montre que des objets translucides dans le visible peuvent également l'être dans le proche infrarouge.

L'analyse des objets noirs confirme qu'il est possible de discriminer certains objets visuellement similaires grâce à leurs différences dans le proche infrarouge, comme par exemple l'objet métallique dans la Fig. 5.7b. De plus,

le test des objets bleus a mis en valeur le fait que certains objets visuellement similaires peuvent être discernées par analyse spectrale.

Malgré ces analyses, les objets présentés sont en grande partie discernables dans les images couleur, sans utiliser de dimension spectrale. Des analyses conduites sur de nouveaux objets plus difficiles à discerner visuellement permettraient d'extraire de nouvelles caractéristiques.

Les objets en carton étant très discernables dans une image couleur (voir Fig. 5.10a), leurs spectres ont des différences importantes dans le domaine visible. Par conséquent, les bandes observées (400-1000nm) ne permettent pas de reconnaître des caractéristiques pour ce type de matériau. C'est également le cas dans les autres images : les caractéristiques extraites dépendent presque exclusivement de la couleur, plutôt que du matériau. Une analyse dans un nouveau domaine spectral plus éloigné du visible permettrait d'extraire des caractéristiques indépendantes de la couleur.

Enfin, des analyses quantitatives telles que présentées en Section 2.5 pourraient permettre d'extraire de nouvelles caractéristiques, différentes de celles d'une analyse quantitative.

5.4 Mesures de similarité

Nous nous intéressons alors aux données du domaine spectral infrarouge (900-2500nm) acquises en Sec. 3.5. Afin de compléter nos analyses, nous cherchons désormais à conduire des analyses quantitatives pour la comparaison de spectres et la discrimination spectrale de matériaux à l'aide de deux mesures de similarité, le SAM (*Spectral Angle Mapper*) et le TC (*Tanimoto Coefficient*).

5.4.1 Classification de spectres

Nous considérons dans un premier temps une classification de spectres de matériaux par comparaison entre les spectres représentatifs (voir Sec. 5.2.3) d'échantillons issus de plusieurs images. Cette classification a pour objectif d'évaluer la discrimination et la reconnaissance de matériaux acquis dans différentes conditions d'illumination. Les valeurs de SAM d'échantillons de ciment et de bois en intérieur et en extérieur sont présentées en Figure 5.11.

Comme attendu pour des spectres issus d'images non-calibrées, les échantillons sont discriminés en fonction de leurs conditions d'acquisition plutôt que par type de matériau. Cela signifie qu'il n'est pas possible de reconnaître des matériaux à partir de spectres de référence issus d'autres sessions.

Cependant, pour de mêmes conditions d'acquisition, les échantillons qui ont les valeurs de SAM les plus faibles sont ceux de même matériau. Dans le cadre des missions de surveillance, ce résultat indique qu'il est possible d'effectuer une discrimination des spectres présents dans plusieurs scènes acquises sous

		Intérieur		Extérieur	
		Ciment	Bois	Ciment	Bois
Intérieur	Ciment	Jaune	Vert clair	Bleu foncé	Bleu foncé
	Bois	Vert clair	Jaune	Bleu foncé	Bleu foncé
Extérieur	Ciment	Bleu foncé	Bleu foncé	Jaune	Vert clair
	Bois	Bleu foncé	Bleu foncé	Vert clair	Jaune

FIGURE 5.11 – Valeurs de SAM entre les spectres d'échantillons de ciment et de bois, acquis lors de la session 4 en intérieur et de la session nuageuse en extérieur. Les valeurs de SAM s'étendent de 0 (jaune, échantillons similaires) à 0.49 (bleu, échantillons différents).

des conditions d'illumination similaires. En revanche, la reconnaissance des matériaux d'une scène n'est pas possible à partir des informations d'une seule autre scène acquise dans des conditions d'illumination différentes.

5.4.2 Détection de matériaux

Cette méthode, binaire, a pour objectif de classifier chaque pixel comme étant similaire ou non à un spectre de référence.

Initialement, le spectre représentatif (voir Sec. 5.2.3) d'un échantillon est sélectionné comme référence. Ce « spectre de référence » est comparé avec chaque pixel d'une image à classifier, pour produire des scores de similarité. Enfin, un seuil est appliqué à ces scores afin de séparer les pixels en 2 classes, "similaire" et "différent", par rapport au spectre de référence.

Le seuil doit maximiser la séparation entre chacune des deux classes, et est calculé de la manière suivante. Les segmentations définies en Sec. 5.2.3 sont utilisées afin de générer deux histogrammes des valeurs de SAM, celui de l'échantillon (« *sample* ») et celui de l'arrière-plan (« *background* », c-à-d. le reste de l'image), voir Figure 5.12. Le seuil est alors calculé de manière à ce que le pourcentage d'éléments échantillon inférieurs à ce seuil soit égal au pourcentage d'éléments arrière-plan supérieurs à ce seuil. Dans le cas de la Figure 5.12, le seuil sélectionné est de 0.057, avec une discrimination correcte à 99% (voir Figure 5.12).

Nous présentons en Figures 5.13 et 5.14 les régions d'intérêt utilisées ainsi

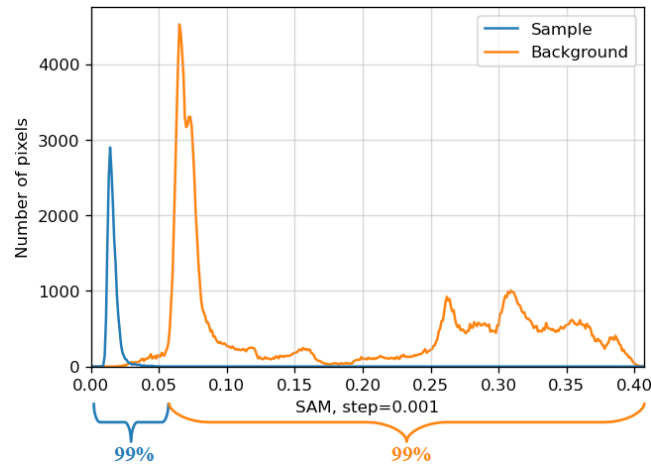


FIGURE 5.12 – Histogramme des valeurs de SAM d'un échantillon de ciment et de l'arrière-plan, dans une image hyperspectrale. Le *step* est le pas des valeurs de SAM utilisé pour tracer les histogrammes.

que les résultats de la classification binaire avec le SAM et le TC sur des images d'échantillons de ciment et d'une voiture.

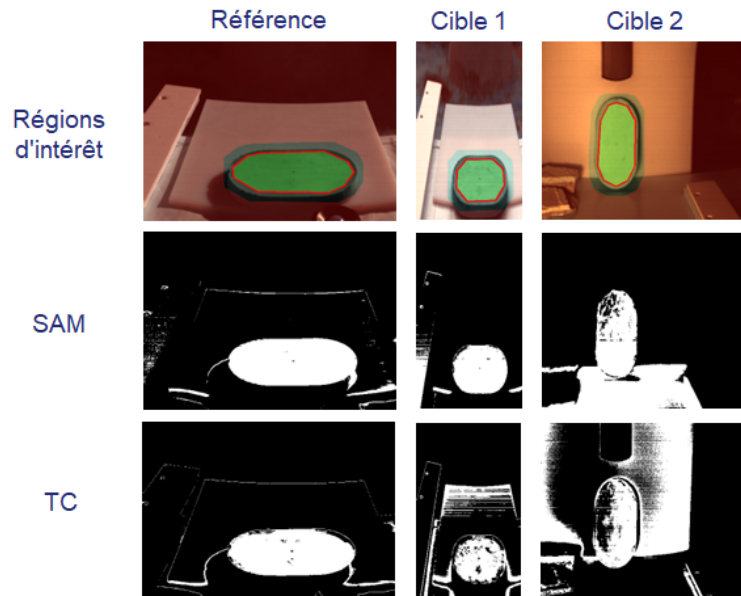


FIGURE 5.13 – Détection de ciment sur 3 images à partir du spectre de référence issu de l'une de ces images. Les pixels blancs sont les pixels qui ont été catégorisés comme "similaires" au spectre de référence (donc au type d'objet recherché), et les pixels noirs comme "différents" (donc appartenant à l'arrière-plan).

En Figure 5.13, malgré une faible portion de l'arrière-plan classifié comme

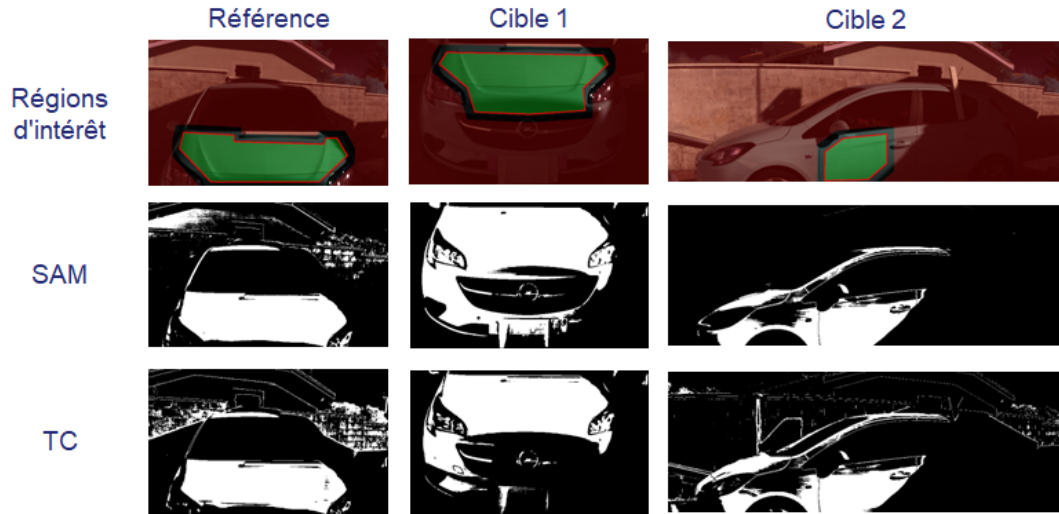


FIGURE 5.14 – Détection de voiture sur 3 images à partir du spectre de référence issu de l’une de ces images. Les pixels blancs sont les pixels qui ont été détectés comme appartenant au type d’objet recherché, et les pixels noirs comme appartenant à l’arrière-plan.

du ciment, l’échantillon de ciment est très bien détecté dans l’image de référence et dans l’image cible 1. En revanche, dans la cible 2, acquise dans des conditions d’acquisition différentes, les performances se dégradent : la détection d’une partie de l’échantillon de ciment échoue et une partie de l’arrière-plan est détectée à la place.

Dans la Figure 5.14, une partie importante de la voiture n’est pas détectée. Cela s’explique d’abord par le fait que les régions d’intérêt sélectionnées ne couvrent qu’une partie de l’objet d’intérêt, mais également par les variations importantes d’illumination dans ces images et sur certains objets. Cet effet est particulièrement visible sur la cible 2, pour lesquelles les parties ombragées du véhicule sont détectées comme appartenant à l’arrière-plan de l’image.

Nous pouvons également constater sur ces deux Figures que le TC commet davantage d’erreurs sur l’arrière plan, par exemple sur la cible 2 de la Fig. 5.14, et est moins adapté que le SAM pour effectuer la tâche de classification binaire. Cependant, lorsqu’un objet d’intérêt présente d’importantes variations d’illumination, ni le SAM ni le TC ne permettent de le détecter.

5.4.3 Discrimination de pixels

Cette autre méthode a pour objectif de regrouper les pixels d’une même image dans plusieurs classes, sans connaissance à-priori de la scène ni des classes à discriminer.

La méthode de discrimination est illustrée en Figure 5.15. Initialement, un seuil est fixé manuellement par l'utilisateur. Pour chaque itération, le premier pixel non-classifié est sélectionné, et attribué à une nouvelle classe i . Ce pixel est comparé à chaque pixel de l'image via une mesure de similarité. Les pixels non-classifiés et dont le score de similarité est inférieur au seuil prédéfini sont attribués à la classe i . Une nouvelle itération commence alors, jusqu'à ce que chaque pixel de l'image soit classifié.

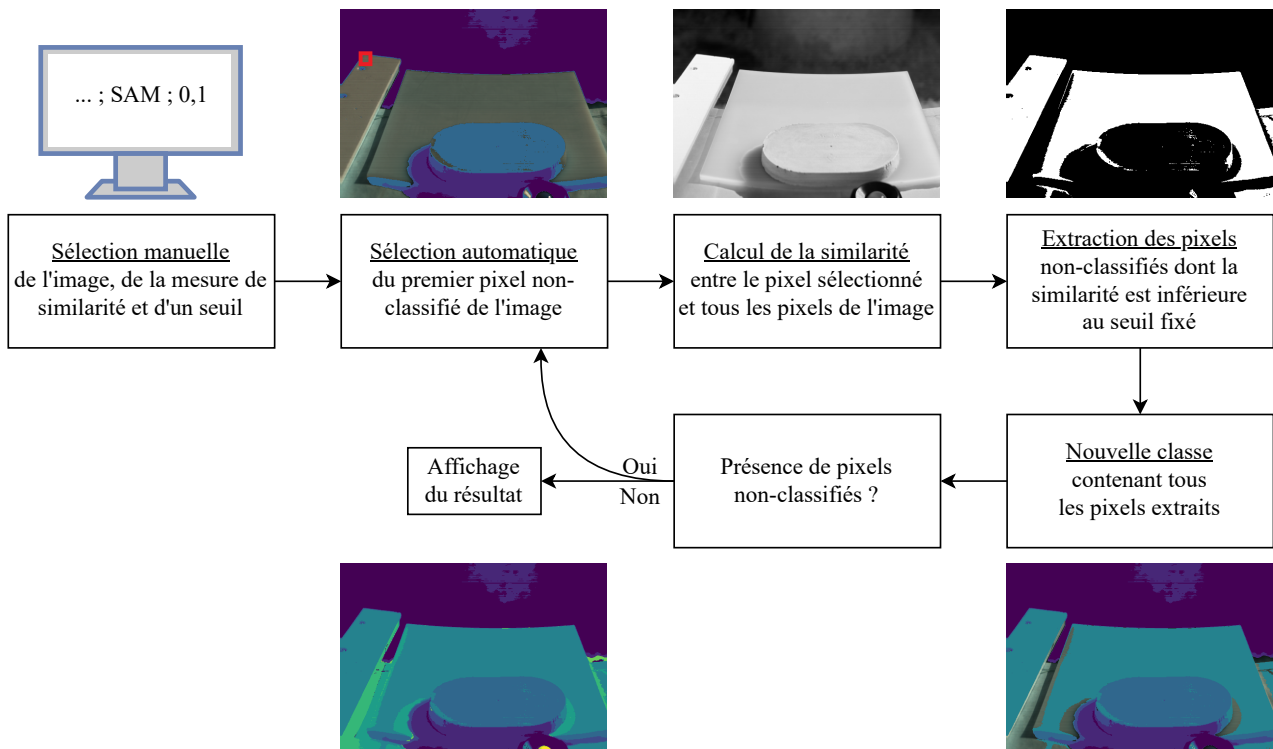


FIGURE 5.15 – Organigramme de la méthode proposée de discrimination non-supervisée de matériaux.

À la fin, les classes dont le nombre de pixels est inférieur à 100 sont regroupées en une même classe « exceptions ». Ce procédé permet d'éviter la présence d'un nombre important de classes en raison de pixels ayant une trop faible similarité avec le reste de l'image.

Nous présentons en Figures 5.16 et 5.17 les discriminations spectrales conduites avec cette méthode, ainsi que les résultats de cette discrimination sur plusieurs images de ciment et de voiture. Chaque couleur correspond à une classe différente. Les pixels colorés en blanc sont ceux de la classe « exceptions », c'est-à-dire les pixels pour lesquels il existe peu de pixels correspondant dans l'image.

Dans la Figure 5.16, la discrimination permet de différencier correctement les objets présents dans les images cibles 1 et 2. Ce n'est cependant pas le cas

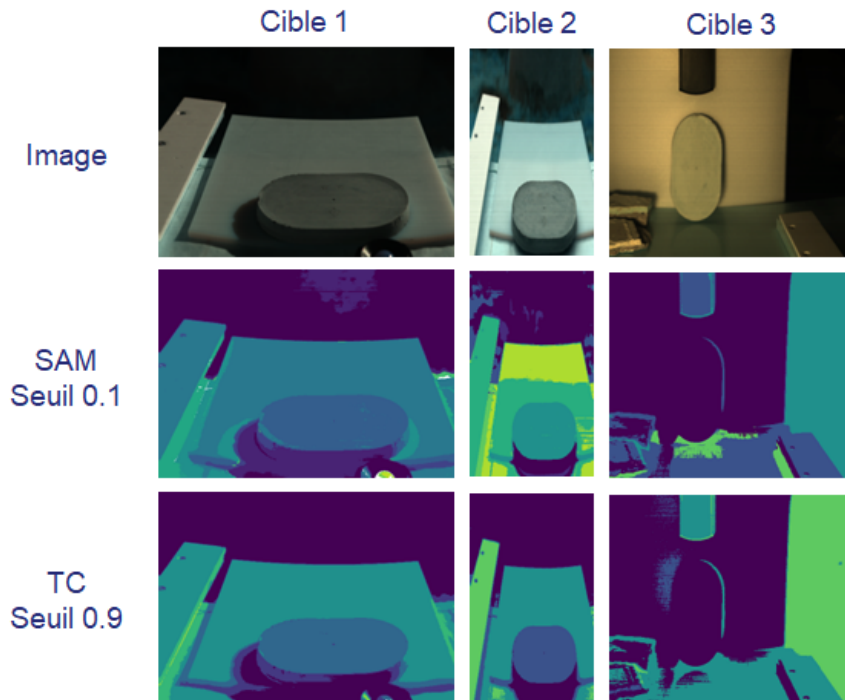


FIGURE 5.16 – Discrimination non-supervisée sur 3 images d'échantillons de ciment.

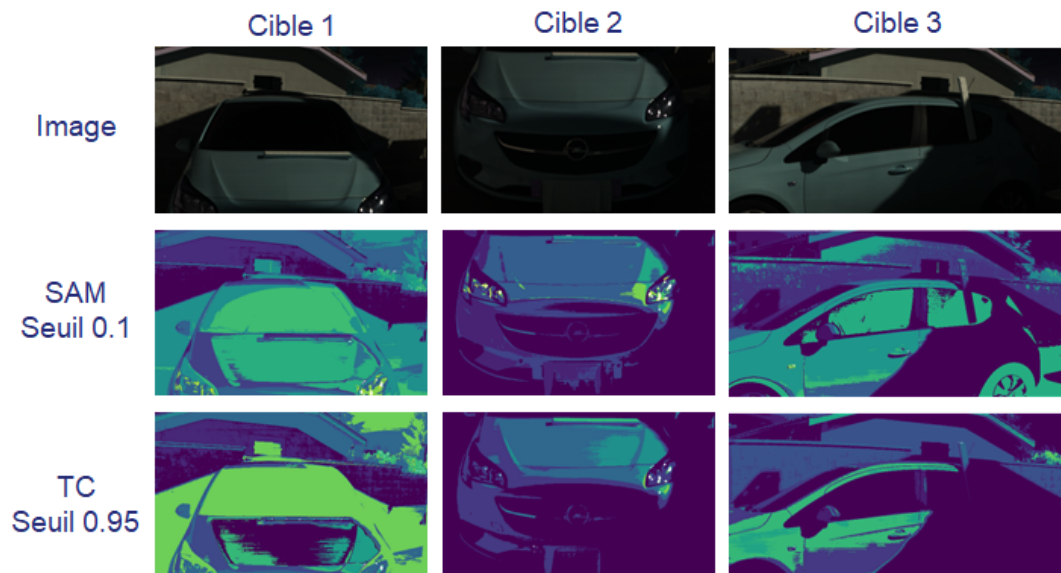


FIGURE 5.17 – Discrimination non-supervisée sur 3 images de voiture.

dans l'image cible 3 : l'échantillon de ciment est associé à la même classe que l'objet derrière lui. Cela signifie que le seuil est trop faible pour permettre de

discriminer ces objets, autant avec le SAM qu'avec le TC.

Dans la Figure 5.17, les régions de l'image sont discriminées en fonction de leur niveau de luminosité. Ainsi, deux parties d'un même objet sont classifiées différemment si leur niveau d'illumination est différent, et deux objets différents sont considérés comme une seule et même classe lorsqu'ils sont illuminés de la même manière.

5.4.4 Réduction de dimension

Les méthodes de réduction de dimension permettent de diminuer de manière importante la taille des données tout en conservant les informations pertinentes de celles-ci. Il existe deux grandes catégories des méthodes de réduction de dimension appliquées aux données spectrales.

La première catégorie est celle de la suppression de bandes spectrales afin de ne conserver que les bandes spectrales les plus pertinentes (Guo *et al.* [2006]; Nhaila *et al.* [2019]). Cette catégorie n'est pas considérée ici afin de bénéficier de toutes les informations spectrales à notre disposition.

La deuxième catégorie englobe l'ensemble des méthodes de fusion de bandes spectrales (Uzair *et al.* [2015]; Zhao et Du [2016]). Parmi ces méthodes, l'ACP (Analyse en Composantes Principales) a déjà fait ses preuves pour la réduction de dimension spectrale (Farrell et Mersereau [2005]; Liu *et al.* [2017]). L'ACP consiste à prendre des données à L dimensions, ici des spectres de taille L , et à trouver une projection vers L nouvelles dimensions qui maximisent la variance entre les échantillons. Les données projetées dans de nouvelles dimensions sont alors nommées « composantes ». Lorsque l'ACP est appliquée à la dimension spectrale d'une image hyperspectrale, elle résulte en un nouveau volume 3D avec 2 dimensions spatiales et 1 dimension de composantes. Ces composantes sont triées de manière à ce que les premières composantes contiennent la plus grande variance, d'où le nom de composantes principales. Les dernières composantes peuvent alors être supprimées pour permettre la réduction de dimension.

La quantité de variance, exprimée en « poids », cumulée contenue dans les composantes de l'ACP appliquée à une image de ciment et une image de voiture est présenté en Figure 5.18. Dans les deux cas, le poids associé à la première composante est supérieur à 95%, et celui des 5 premières composantes dépasse les 99%.

Ce poids important (> 99%) associé aux premières composantes de l'ACP signifie qu'il est possible d'extraire la majeure partie de la variance des images hyperspectrales dans seulement 5 composantes. Nous ne nous intéressons donc qu'à ces 5 premières composantes dans la suite de nos calculs.

Les Figures 5.19 et 5.20 présentent l'image couleur et les 5 premières composantes de l'ACP appliquée à des images de ciment et de voiture. Les composantes sont représentées avec la carte de couleurs *viridis* : les valeurs faibles

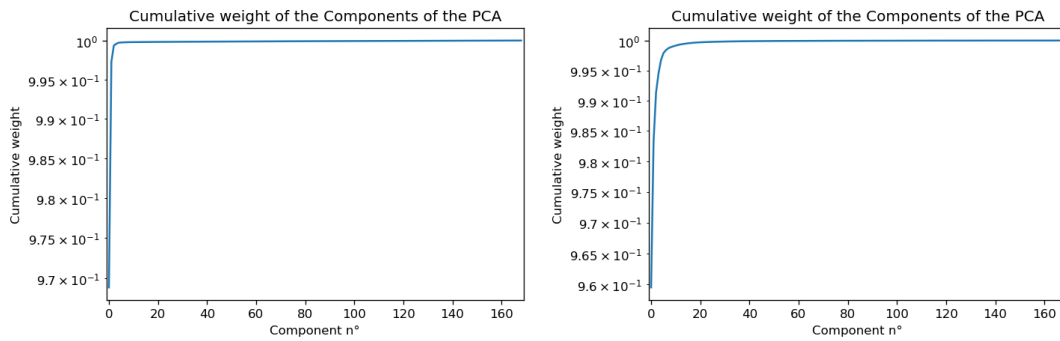


FIGURE 5.18 – Poids cumulatif des composantes de l'ACP sur (gauche) une image d'échantillon de ciment et (droite) une image de voiture.

sont affichées en violet, les valeurs moyennes en vert et les valeurs élevées en jaune.

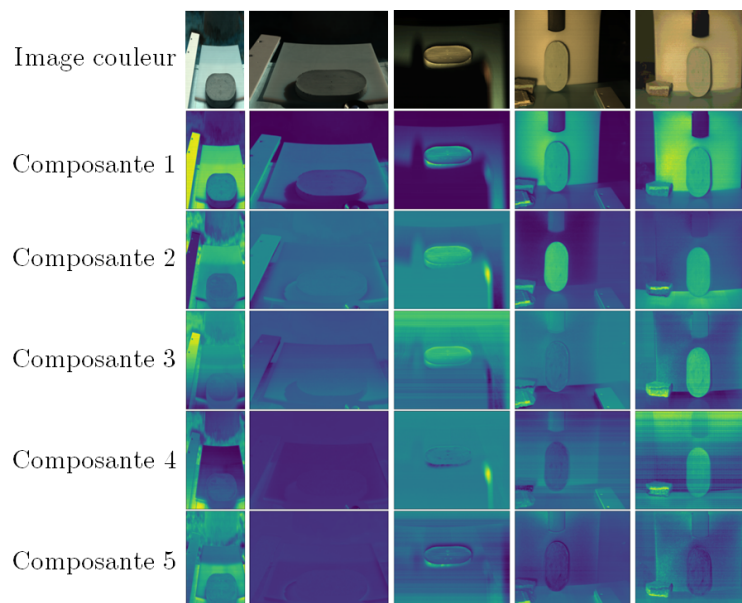


FIGURE 5.19 – Reconstructions couleur de 3 images d'échantillons de ciment et les 5 composantes principales de ces images.

Sur chaque image, la première composante correspond à la luminosité du pixel. Les autres composantes permettent d'extraire des éléments saillants du décor (végétation en arrière-plan, phares de la voiture).

Néanmoins, aucune composante ne permet de distinguer à elle seule l'objet d'intérêt de l'arrière-plan, à part dans le cas de la deuxième composante de la Fig. 5.19 (droite). Nous souhaitons donc analyser le résultat des méthodes de détection et de discrimination appliquées aux composantes extraites.



FIGURE 5.20 – Reconstructions couleur de 3 images de voiture et les 5 composantes principales de ces images.

5.4.5 Détection de matériaux avec ACP

Nous avons appliqué la méthode de détection binaire aux 5 premières composantes produites par l'ACP, et comparé ces résultats avec ceux de la détection sans ACP. Contrairement à la méthode précédente de détection binaire, un spectre de référence est sélectionné dans et pour chaque image cible, car les composantes des différentes images contiennent une information différente et ne sont donc pas comparables entre elles par les mesures de similarité.

La détection par SAM avec et sans ACP est présentée en Figure 5.21. Nous présentons l'histogramme des valeurs de SAM de l'échantillon et de l'arrière-plan, sur l'image hyperspectrale (Fig. 5.22a) et sur les 5 premières composantes extraites de cette image (Fig. 5.22b).

Les résultats du SAM sur des images hyperspectrales et sur leurs 5 premières composantes sont très proches malgré la compression importante des données par l'ACP. Cependant, une partie plus importante de l'arrière-plan est mal classifiée dans le cas de l'ACP-SAM. Cette mauvaise classification est due à l'augmentation du recouvrement entre les valeurs de SAM de l'échantillon et de l'arrière-plan après l'ACP, recouvrement visible dans les histogrammes de ces valeurs (voir Figure 5.22).

Nous notons également que le SAM performe mieux sur ces exemples que sur ceux de la Figure 5.13, en raison de la sélection d'un spectre de référence dans chaque image cible plutôt que dans une unique image de référence acquise dans des conditions d'illumination différentes.

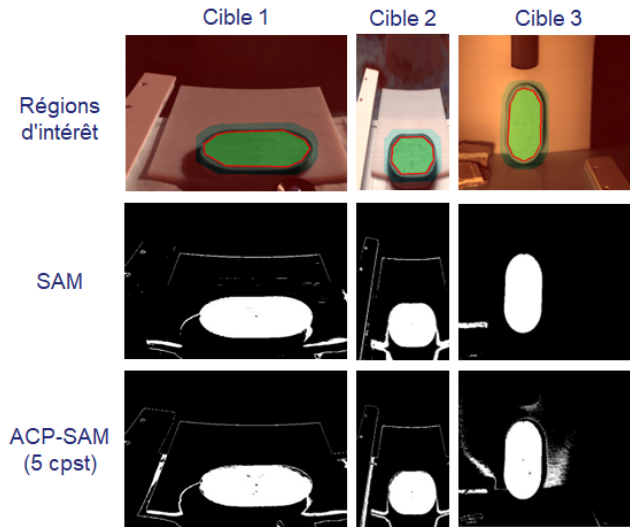


FIGURE 5.21 – Détection de ciment sur 3 images sans et avec ACP. Les pixels blancs sont les pixels qui ont été classifiés comme appartenant au type d’objet recherché, et les pixels noirs comme appartenant à l’arrière-plan.

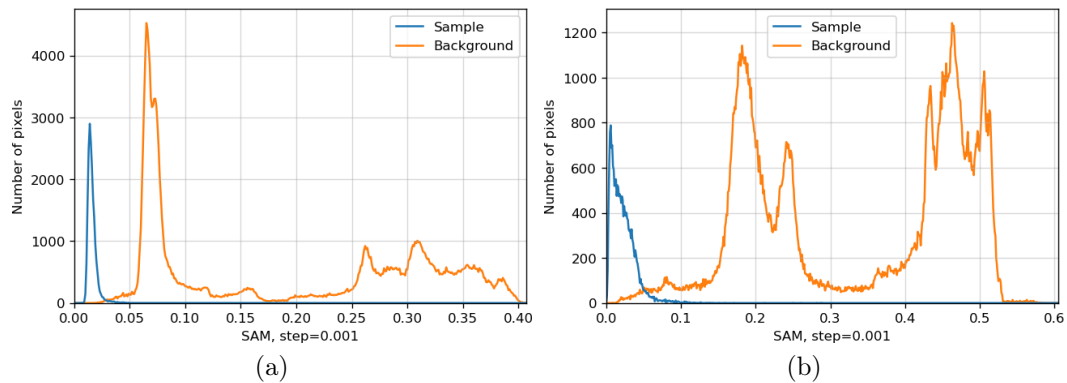


FIGURE 5.22 – Histogramme des valeurs de SAM d’un échantillon de ciment et de l’arrière-plan, dans (a) une image hyperspectrale et (b) les 5 premières composantes de cette image hyperspectrale. Le *step* est le pas des valeurs de SAM utilisé pour tracer les histogrammes.

5.4.6 Discrimination de pixels avec ACP

La discrimination par SAM a été conduite sur des images avec et sans ACP, avec différents seuils de discrimination. Les résultats de ces discriminations avec un seuil de 0.15 sont présentés en Figure 5.23. Dans le cas de l’ACP, la discrimination est appliquée sur les 5 premières composantes.

La méthode de discrimination avec SAM ne permet de discriminer correctement les objets ni sans ni avec ACP. La sélection d’un seuil unique pour la discrimination de toutes les images n’est pas pertinente, en raison des va-

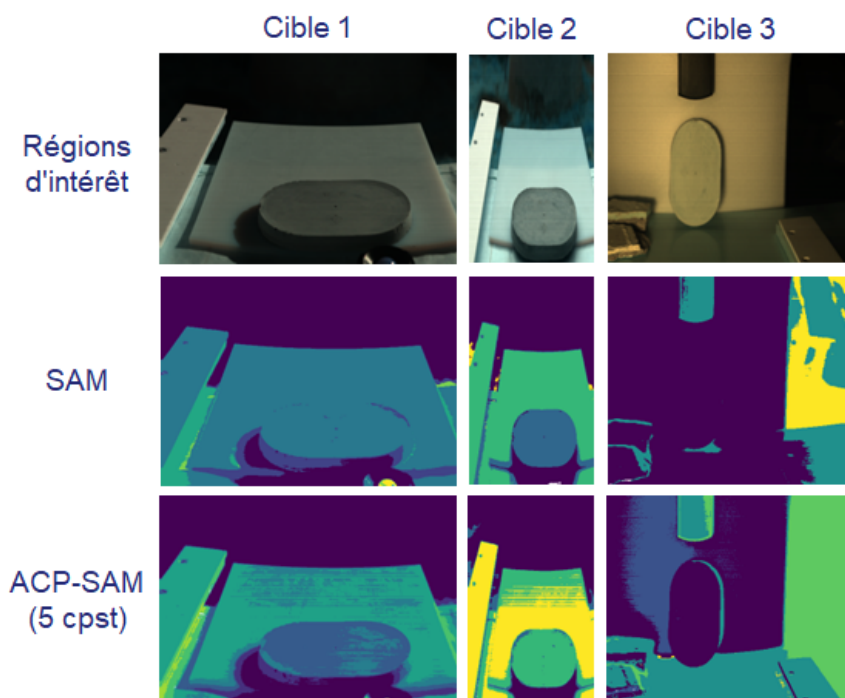


FIGURE 5.23 – Discrimination non-supervisée par SAM sur 3 images de ciment, sans et avec ACP, avec un seuil de 0.15.

riations d'illumination importantes à l'intérieur des images. Néanmoins, des expériences empiriques de sélection de seuil pour chaque image n'ont pas permis d'améliorer cette discrimination de manière significative.

De plus, la méthode naïve de sélection des pixels de référence présente des limitations. Certaines régions d'image représentant un même objet peuvent se retrouver scindées en plusieurs parties, par exemple lorsque l'un des pixels aberrants (sans classe) extérieurs à la région d'intérêt est similaire à une partie de cette région. Un pré-traitement de segmentation d'image pourrait permettre de réduire l'impact de pixels aberrants sur la mauvaise discrimination de régions d'intérêt.

5.5 Acquisitions nocturnes/sans lumière

Afin de déterminer s'il était possible d'acquérir et de traiter des images hyperspectrales acquises la nuit, nous avons également conduit l'analyse des acquisitions nocturnes (voir Sec. 3.5.3) et des images acquises sans lumière (voir Sec. 3.5.4). Contrairement aux images des autres sessions, ces images sont extrêmement sombres et ressemblent fortement aux images *DC*. De plus, les valeurs moyennes et d'écart-type des images acquises dans ces conditions sont identiques à celles des images *DC*. Enfin, un rehaussement de contraste de

ces images n'a permis de faire apparaître que le bruit du capteur, mais aucune forme ou signal distinctif des objets acquis. De plus, les valeurs moyennes et d'écart-type des images acquises dans ces conditions sont identiques à celles des images *DC*.

Nous en concluons que dans les images acquises en conditions nocturnes et sans source d'illumination dans le domaine spectral de la caméra, le signal produit est trop faible et est entièrement écrasé par le bruit du capteur.

5.6 Conclusion

Dans ce chapitre, nous avons présenté l'ensemble des méthodes de pré-traitement utilisées sur nos données. Ces pré-traitements incluent la reconstruction d'images couleur à partir d'images hyperspectrales, la segmentation et l'extraction de spectres représentatifs d'objets d'intérêt, ainsi que la correction atmosphérique et la normalisation de ces spectres.

Nous avons également présenté des analyses spectrales d'objets de même couleur et d'objets de même matériau. Cependant, le domaine spectral étudié, proche du visible, n'a pas permis d'extraire des caractéristiques pour la caractérisation de ces matériaux.

De nouvelles analyses par mesures de similarité ont été conduites sur des images hyperspectrales acquises dans un autre domaine spectral, le proche infrarouge, pour la classification, la détection et la discrimination spectrales de matériaux. Ces analyses montrent qu'il est possible de détecter et de discriminer des matériaux dans des images hyperspectrales. En revanche, les méthodes proposées souffrent actuellement de plusieurs limitations, telles que la sélection de seuils pour la détection et la discrimination, et particulièrement d'un manque de robustesse aux variations intra- et inter-images des conditions d'illumination.

Chapitre 6

Réseaux de neurones

Sommaire

6.1	Introduction	83
6.2	Détection de véhicules	84
6.2.1	Réseau spectral	84
6.2.2	Réseau colorimétrique	90
6.2.3	Réseau à convolutions 3D	91
6.2.4	Conclusion	91
6.3	Réseaux existants pour la détection d’objets	92
6.3.1	YOLOv3	92
6.3.2	Mask-RCNN	99
6.4	Réseaux pour la classification d’images hyperspectrales	102
6.4.1	Terminologie	102
6.4.2	<i>DeepHyperX</i> (Audebert, 2019)	103
6.4.3	<i>hyperspectral_deeplearning_review</i> (Paoletti, 2019)	105
6.4.4	Conclusion	105
6.5	Conclusion	106

6.1 Introduction

Dans le chapitre précédent, nous avons vu que, si les mesures de similarité permettent parfois la discrimination de matériau, elles ne sont pas robustes aux changements de conditions d’illumination. Nous cherchons donc des méthodes permettant d’aller au-delà de ces limitations. Les réseaux de neurones apparaissent comme une méthode appropriée, dû à leur habilité à représenter des espaces de caractéristiques divers. Nous présentons d’abord un réseau spectral et l’analyse des caractéristiques apprises par ce réseau pour la classification

de véhicules dans des images hyperspectrales, travaux ayant donné lieu à une publication (Caubet *et al.* [2022b]).

Deux réseaux de détection d'objets issus de l'état de l'art, YOLOv3 (Redmon et Farhadi [2018]) et Mask-RCNN (He *et al.* [2017]), sont utilisés afin de comprendre la pertinence des différentes caractéristiques des images hyperspectrales pour la détection d'objets. D'autres réseaux de neurones de l'état de l'art, permettant la classification de pixels d'images hyperspectrales, sont présentés et comparés afin d'analyser leur utilisabilité dans notre contexte ainsi que leurs performances et limitations.

6.2 Détection de véhicules

Nous souhaitons utiliser des réseaux de neurones pour détecter et discriminer des véhicules. Afin d'entraîner ces réseaux, nous utilisons les images présentées en Section 3.4 et segmentées en suivant les étapes décrites en Section 5.2.2. Au total, 400 images hyperspectrales de taille 200x200x138 sont utilisées : 350 images pour l'entraînement et 50 pour la validation. Aucune base de test n'est sélectionnée car les réseaux ne sont comparés qu'à travers leurs performances sur l'ensemble de validation.

6.2.1 Réseau spectral

Les corrélations spatiales d'images sont déjà très étudiées et de nombreuses méthodes (Harris et Stephens [1988]; Goferman *et al.* [2012]; Yan *et al.* [2013]), dont certaines à base de réseaux de neurones convolutionnels (Lecun *et al.* [1998]; Dong *et al.* [2015]; Simonyan et Zisserman [2015]), ont déjà été créées afin de bénéficier des informations spatiales. Pour cette raison, nous nous intéressons dans un premier temps seulement aux données spectrales des images, d'où le nom de « réseau spectral ».

Architecture du réseau

L'architecture de ce réseau est proche de celle proposée par Shafiee *et al.* [2016], dont le réseau atteint 95% de précision pour la classification d'échantillons de miel.

Nous présentons en Figure 6.1 le flux des données depuis l'image d'entrée jusqu'à la classification avec le « réseau spectral ». Chaque pixel ou spectre de l'image hyperspectrale est fourni indépendamment en entrée du réseau. Le « réseau spectral » est un réseau de neurones dense entièrement connecté constitué d'une couche d'entrée, une couche de neurones intermédiaire et une couche de sortie. Ce nombre de couches a été choisi afin de réduire les chances de sur-apprentissage du réseau. La sortie du réseau correspond à sa prédiction de la classe du pixel d'entrée. Ces prédictions peuvent être regroupés en

une image de classes, dont les dimensions spatiales sont identiques à celles de l'image hyperspectrale d'entrée.

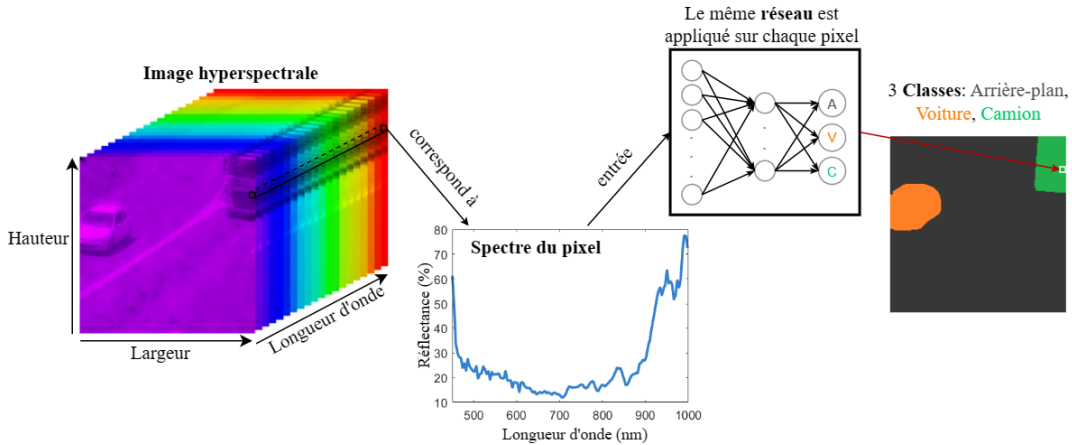


FIGURE 6.1 – Traitement pixel par pixel (ou spectre par spectre) d'une image hyperspectrale par notre « réseau spectral » pour la discrimination de véhicules.

Cette architecture est utilisée pour créer deux réseaux distincts. Le premier est un réseau de « détection » de véhicules, avec une couche intermédiaire de 15 neurones et une couche de sortie binaire (0 ou 1), pour lequel l'ensemble des véhicules (voir Section 5.2.2) sont considérés comme appartenant à une même classe et comparés à l'arrière-plan. Le deuxième réseau, pour la « discrimination » de véhicules, est représenté en Figure 6.1 et considère 3 classes : *arrière-plan*, *voiture* et *camion*. Les autres types de véhicules sont ignorés par le réseau en raison de leur très faible représentativité. Ce réseau est composé d'une couche intermédiaire de 23 neurones intermédiaires et une couche de sortie de 3 neurones (*arrière-plan*, *voiture*, *camion*).

Terminologie des analyses

« Modèles » : Pour chacun de ces deux réseaux (détection et discrimination), nous avons créé plusieurs « modèles » en faisant varier les fonctions d'activation des couches intermédiaire et de sortie. Les fonctions d'activation utilisées à cet effet sont *ReLU*, *Leaky-ReLU*, *sigmoid* et *tanh*.

« Précision » : Nous analysons les performances de chaque modèle en calculant la « précision » (*accuracy*) de la ou les classes de véhicules, définie de la manière suivante :

$$Precision = TP / (TP + FP)$$

avec TP le nombre de pixels Vrais Positifs (*True Positives*) et FP le nombre de Faux Positifs.

« Activation de neurone » : Nous présentons en Figure 6.2 un schéma du traitement des données par le réseau et une visualisation de ces données, à chaque étape de leur traitement par le réseau spectral, de l'entrée à la sortie du réseau. Les images « d'activation de neurone » correspondent aux valeurs renvoyées par un neurone du réseau pour chaque pixel de l'image, et sont colorisées avec la palette de couleur Jet : les valeurs faibles sont affichées en bleu, les valeurs moyennes en vert et les valeurs élevées en rouge. Chaque pixel d'une image étant traité individuellement par le réseau, les images présentées sont construites à partir des valeurs obtenues à une même étape du traitement pour chaque pixel. Par exemple, sur l'image d'activation du neurone intermédiaire 11, nous pouvons observer une activation faible (cyan) sur les pixels extraits de la voiture ou de la barrière de sécurité, et très faible (bleu foncé) sur les autres pixels de l'image.

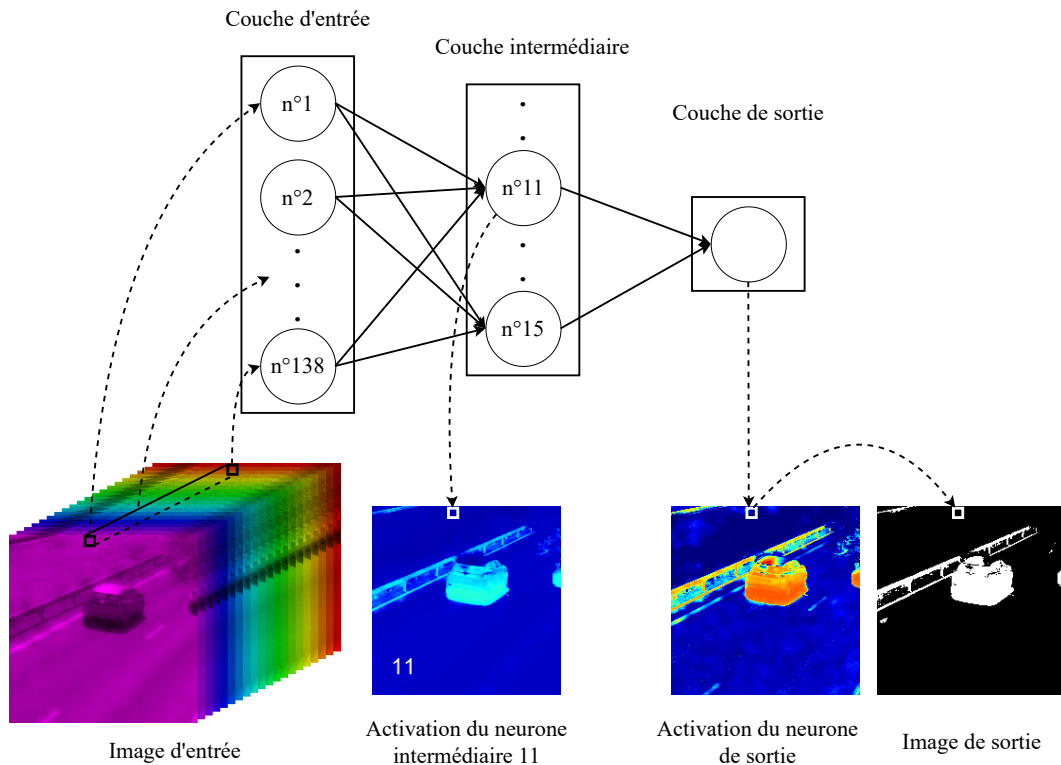


FIGURE 6.2 – Schéma du traitement des données, de l'image hyperspectrale à la carte de classification binaire, par le réseau spectral.

Résultats du réseau de détection

Nous présentons en Table 6.1 la précision obtenue, sur l'ensemble de validation, par les 16 apprentissages de réseau spectral pour la détection de véhicules.

Certains apprentissages restent à 0% de précision, car ils classifient l'ensemble des pixels comme appartenant à l'arrière-plan. Cette précision nulle peut être due au problème de disparition du gradient, empêchant toute modification de la prédiction du réseau. D'autres apprentissages atteignent une précision de 23% dès les premières centaines d'époques, mais leur précision n'évolue pas davantage par la suite. Enfin, 10 des 16 apprentissages atteignent une précision d'environ 53% dès la première centaine d'époques, malgré des variations de plusieurs pourcents au cours des itérations.

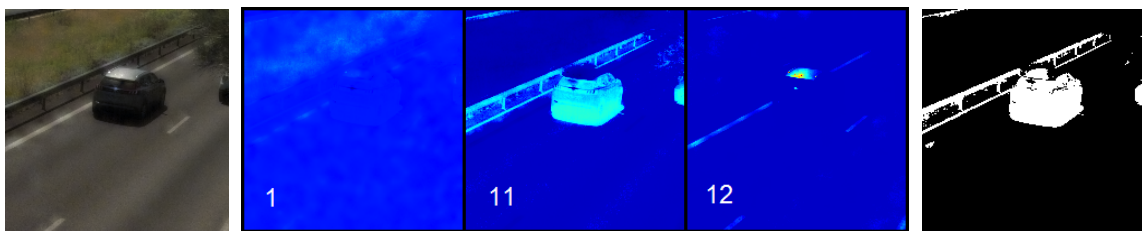
Apprentissage n°	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Précision	53%	0%	53%	56%	53%	54%	53%	55%	52%	0%	23%	22%	54%	0%	24%	53%

TABLE 6.1 – Valeurs de précision des 16 apprentissages du réseau spectral de détection avec différentes fonctions d'activation après 800 époques d'entraînement.

Nous pouvons délimiter deux seuils dans ces valeurs de précision, autour de 23% et 53%. Ces seuils sont dus à de grandes quantités de pixels dont la classe prédite est modifiée en un faible nombre d'époques, ce qui indique que ces groupes de pixels sont potentiellement très proches entre eux.

Explicabilité du réseau de détection

Nous souhaitons comprendre comment les modèles ont atteint ces valeurs de précision, et plus particulièrement les régions d'images qui ont eu un impact significatif pour la classification. Nous analysons pour cela l'un des modèles ayant obtenu une précision de 55%. Les images de l'activation des neurones de ce modèle sur une image hyperspectrale sont présentées en Figure 6.3.



(a) Image couleur. (b) Activation des neurones 1, 11 et 12 de la couche intermédiaire. (c) Image de sortie.

FIGURE 6.3 – Détection appliquée à l'image hyperspectrale d'une voiture.

Le neurone 1 réagit principalement aux bandes autour de 550nm (vert) et 750nm (rouge/NIR) tout en rejetant les bandes autour de 950nm (NIR). Sur l'image 6.3b, ce neurone détecte une partie de la végétation et ignore le reste, y compris les véhicules. Dans la détection finale, ce neurone est associé à l'arrière-plan, indiquant un rejet de la végétation pour la détection des véhicules.

Les neurones 11 et 12 mettent respectivement en valeur les zones présentant une intensité faible (ombres) et élevée (lignes blanches, reflets du soleil...). Ces deux neurones ont un poids très élevé dans la prédiction du modèle, ainsi qu'un impact important sur la précision de ce modèle, car une quantité importante des pixels de véhicule ont une intensité soit très faible, soit très élevée. Cependant, d'autres portions des images (lignes blanches, barrières de sécurité...) partagent ces caractéristiques d'intensité (voir Fig. 6.3c), et le poids important associé à ces informations provoque une détection erronée de ces zones d'arrière-plan.

À travers l'analyse de plusieurs modèles du réseau de détection, nous avons pu constater que les informations utiles à la détection de véhicules restaient principalement les trois mêmes catégories : la végétation, les ombres et les zones d'intensité élevée. De plus, nous pouvons également trouver une corrélation entre ces catégories et les seuils apparus dans les valeurs de précision présentées en Table 6.1. En effet, les réseaux pour lesquels la précision a atteint 23% sont ceux qui ont appris à détecter les zones d'intensité élevée, et les réseaux dont la précision a atteint 55% sont ceux qui ont appris à détecter à la fois les régions d'intensité élevée et faible.

Résultats du réseau de discrimination

Dans le cas de la discrimination, les trois classes recherchées sont inégalement représentées dans la base de données, ce qui peut impacter l'apprentissage. La classe *arrière-plan* représente près de 78.72% des pixels, *voiture* près de 14.10% et *camion* environ 6.41%, avec les 0.77% restants représentant les autres types de véhicules. Afin de réduire l'impact de cette représentation inégale, nous utilisons comme fonction de perte l'entropie croisée avec une pondération des classes sur l'apprentissage.

Nous présentons en Table 6.2 les valeurs de précision obtenue, sur l'ensemble de validation, par les modèles de réseau spectral pour la discrimination des classes *voiture* et *camion*. Comme pour le réseau spectral de détection, deux seuils apparaissent : un proche de 73% pour la classe *voiture* et un proche de 37% pour la classe *camion*.

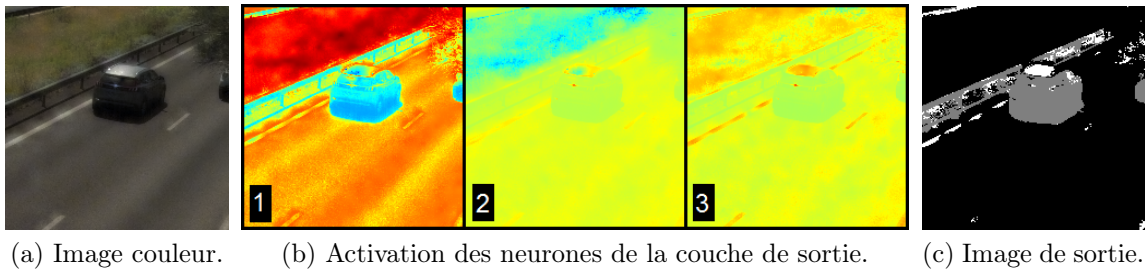
Apprentissage n°	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Précision classe <i>voiture</i>	76%	0%	0%	71%	76%	75%	0%	42%	72%	0%	0%	0%	72%	73%	0%	2%
Précision classe <i>camion</i>	36%	0%	37%	37%	36%	0%	37%	37%	37%	37%	0%	36%	44%	0%	36%	37%

TABLE 6.2 – Valeurs de précision des 16 apprentissages du réseau spectral de discrimination avec différentes fonctions d'activation après 800 epochs d'entraînement pour les classes (a) *voiture* et (b) *camion*.

Outre les 0% de précision lorsque le réseau n'a pas appris à discriminer certaines classes, deux valeurs inhabituelles apparaissent. L'apprentissage n°8 obtient une précision de 42% sur la classe *voiture*, ce qui peut s'expliquer par le fait que le réseau était en train d'apprendre les caractéristiques de cette classe. En revanche, le 44% atteint sur la classe *camion* est supérieur à toutes les autres valeurs atteintes sur cette classe, ce qui semble indiquer que ce réseau a appris de nouvelles caractéristiques permettant de discriminer les camions du reste des images.

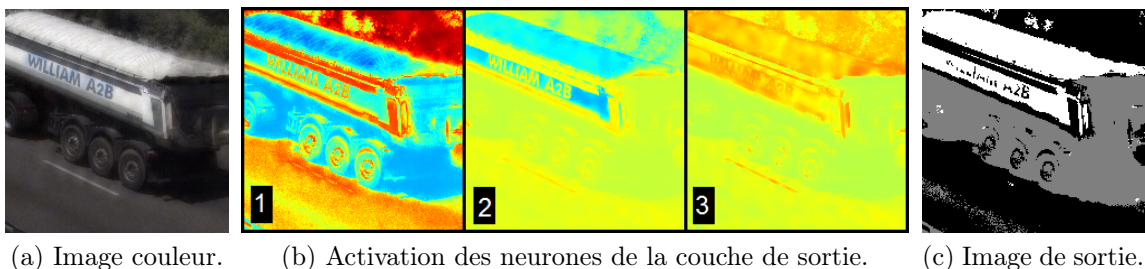
Explicabilité du réseau de discrimination

Nous souhaitons comprendre comment les modèles de discrimination ont atteint ces valeurs de précision. Nous analysons pour cela l'un des modèles ayant obtenu une précision de plus de 30% à la fois sur *voiture* et sur *camion*. Les images de l'activation des neurones de ce modèle sur une image hyperspectrale sont présentées en Figures 6.4 et 6.5. En sortie, les pixels noirs correspondent à une prédiction d'*arrière-plan*, gris pour *voiture* et blanc pour *camion*.



(a) Image couleur. (b) Activation des neurones de la couche de sortie. (c) Image de sortie.

FIGURE 6.4 – Discrimination appliquée à l'image hyperspectrale d'une voiture.



(a) Image couleur. (b) Activation des neurones de la couche de sortie. (c) Image de sortie.

FIGURE 6.5 – Discrimination appliquée à l'image hyperspectrale d'un camion.

Le neurone de sortie 1, celui de l'*arrière-plan*, a une activation importante au niveau de la végétation et de la route et une activation très faible aux régions d'intensité élevée (lignes blanches, camion) ou faible (ombres). L'activation du

neurone 2, celui de la classe *voiture*, est très faible dans les zones de végétation ou de forte intensité lumineuse. Enfin, le neurone 3, associé à la prédiction de *camion*, réagit fortement aux zones de forte intensité lumineuse.

Les caractéristiques apprises par le modèle sont les mêmes que pour les modèles de détection : la végétation, et les pixels d'intensité faible ou élevée, associés respectivement à l'*arrière-plan*, à *voiture* et à *camion*. À ces trois caractéristiques s'ajoute une quatrième, la route, qui est fortement attribuée à l'*arrière-plan* (voir neurone 1 en Figures 6.4b et 6.5b), et est caractérisée par une illumination moyenne.

Conclusion

Les modèles de réseau spectral obtiennent des résultats de précision assez faibles, et particulièrement dans le cas de la discrimination, mais indiquent la présence de caractéristiques permettant la détection et la discrimination de véhicules. En revanche, plutôt qu'apprendre des caractéristiques spectrales complexes, les modèles ont appris des caractéristiques limitées à la présence de végétation ou à la luminosité globale du pixel, c'est-à-dire des caractéristiques visibles à l'œil nu ou avec des méthodes d'imagerie couleur classique. De plus, nous rappelons que ces caractéristiques sont limitées à l'intensité d'un pixel. L'absence des dimensions spatiales empêche l'apprentissage de caractéristiques telles que les contours ou les formes, généralement nécessaires à la détection d'objets.

6.2.2 Réseau colorimétrique

Nous souhaitons entraîner un réseau sur des images couleur acquises dans les mêmes conditions que les images hyperspectrales afin de vérifier si les mêmes caractéristiques seront apprises par ce nouveau réseau. Nous utilisons pour cela les images couleur reconstruites à partir de ces images hyperspectrales (voir Section 5.2.1).

Architecture du réseau

Comme pour le réseau spectral, chaque pixel est fourni indépendamment au réseau, dans une entrée à trois neurones correspondant aux trois canaux de couleur primaires (rouge, vert, bleu) plutôt qu'au spectre. Le nombre de neurones intermédiaires et de sortie et autres paramètres du réseau colorimétrique sont identiques à ceux du réseau spectral de discrimination (voir Sec. 6.2.1).

Résultats du réseau

En 800 epochs, le réseau colorimétrique a atteint une précision comparable à celle du réseau spectral pour la détection de la classe *camion*. Néanmoins,

ce réseau n'a pas réussi à dépasser les 23% de précision sur la classe *voiture*, contre 31% pour le meilleur modèle du réseau spectral.

Nous notons également une variabilité importante de la précision dépassant parfois les 10% entre deux epochs consécutives. Cette variabilité peut être due aux données de validation étant sous-représentées par l'ensemble d'entraînement ou à une vitesse d'apprentissage (*learning rate*) trop élevée empêchant au réseau d'atteindre un minimum local au cours de son apprentissage.

6.2.3 Réseau à convolutions 3D

Afin d'évaluer l'apport des dimensions spatiales pour la détection de véhicules, nous entraînons un nouveau réseau apprenant simultanément des caractéristiques spatiales et spectrales : un réseau de neurones convolutionnel 3D (CNN-3D).

Architecture du réseau

Ce réseau prend en entrée une image hyperspectrale complète. L'entrée du réseau est constitué de deux couches de convolutions 3D de taille 3x3x5, chacune suivie d'une couche de maxpooling dans lesquelles les dimensions spatiales sont préservées. Ces convolutions sont connectées à une couche de neurones dense fournissant en sortie trois neurones, un par classe (*arrière-plan*, *voiture*, *camion*).

Résultats du réseau

En moins de 650 epochs, le réseau à convolution 3D obtient une précision comparable à celle du réseau spectral pour la détection de *voiture* (environ 31%), mais performe mieux sur *camion* avec près de 42% de précision (contre 30%). Cela nous indique que l'information spatiale, même dans un voisinage de taille faible, permet d'améliorer la détection de certains types de véhicules. Cependant, comme pour le réseau spectral et le réseau colorimétrique, la précision présente de fortes variations entre des epochs consécutives.

6.2.4 Conclusion

Nous avons pu constater que l'utilisation des caractéristiques spectrales, vis-à-vis d'une information seulement de couleur, a permis une augmentation de cette précision. Toutefois, leurs résultats proches indiquent qu'une partie importante des caractéristiques apprises par le réseau spectral sont également présentes dans des images couleur : la végétation et la luminosité. Deux pixels ayant une luminosité similaire sont alors associés à une même classe, même lorsqu'ils présentent des objets ou matériaux différents.

L'utilisation des caractéristiques spatiales a également permis d'augmenter les valeurs de précision pour la discrimination de véhicules, bien que cette augmentation soit également faible. Nous notons toutefois que l'utilisation d'un voisinage de plus grande taille permettrait au réseau d'extraire des informations locales de plus haut niveau, telles que les formes de certains objets.

Ainsi, la détection de véhicule conduite avec les architectures décrites dans cette section atteint au mieux une précision d'environ 55%. La précision de la discrimination atteint ses limites autour de 31% pour la classe *voiture*, et 42% pour la classe *camion*.

Notre détection de véhicules souffre cependant de plusieurs limitations. Tout d'abord, les données utilisées sont des images acquises dans le domaine spectral 400-1000nm (visible et très-proche infrarouge). Des images acquises dans le proche infrarouge, tel que préconisé au Chapitre 2, pourraient permettre d'obtenir des caractéristiques spectrales plus pertinentes pour la discrimination d'objets et de matériaux. Il est également possible que la labellisation de ces images soit inadaptée aux résultats recherchés : chacune des classes constitue une agrégation de nombreux matériaux de spectres différents, dont par exemple les voitures et les camions qui sont fabriqués avec les mêmes matériaux. De plus, peu de données sont utilisées au cours de l'entraînement, ce qui impacte la capacité de généralisation des réseaux de neurones.

Enfin, ces réseaux sont eux-même peu profonds, avec une unique couche intermédiaire dans le cas des réseaux spectral et colorimétrique, et trois couches intermédiaires dans le réseau à convolution 3D. Cette profondeur faible des réseaux ne permet l'apprentissage que de caractéristiques spatiales et spectrales de bas niveau, plutôt que des caractéristiques haut niveau telles que des formes d'objets ou des motifs de variations spectrales.

6.3 Réseaux existants pour la détection d'objets

Nous avons vu en Section 6.2.4 que les réseaux de neurones peu profonds ne permettaient pas d'apprendre des caractéristiques de haut niveau sur les images hyperspectrales de véhicules (400-1000nm). Nous évaluons alors les performances de réseaux de neurones profonds construits pour des tâches de détection. Nous utilisons deux réseaux issus de l'état de l'art de la détection de véhicules sur des images couleur, YOLOv3 (Redmon et Farhadi [2018]) et Mask-RCNN (He *et al.* [2017]), sur nos images hyperspectrales.

6.3.1 YOLOv3

YOLOv3 est un réseau de neurones profond de détection d'objets par prédiction de boîtes englobantes. Les boîtes englobantes de la base de données d'images de véhicules (Sec. 3.4) sont définies selon la description en Section

5.2.2.

Dans le but de proposer une analyse exhaustive des caractéristiques spectrales, nous analysons l'impact des canaux de couleur des images reconstruites et l'impact des différentes bandes spectrales sur les prédictions de YOLOv3.

YOLOv3 étant un réseau de neurones construit pour détecter divers types d'objets, nous notons que seules les prédictions de véhicules sont considérées, et les autres prédictions de YOLOv3 sont ignorées.

Évaluation

Pour comparer les résultats de YOLOv3 sur différentes données, nous présentons les caractéristiques suivantes :

- Le nombre de boîtes attendues (ou labellisées).
- Le nombre de boîtes prédites par YOLOv3.
- Le nombre de boîtes correctement prédites, et leur pourcentage parmi l'ensemble des boîtes attendues. Une prédiction correcte signifie qu'une boîte attendue et une boîte prédite ont un IoU (voir ci-dessous) supérieur à un seuil fixé. Lorsque plusieurs boîtes prédites correspondent à une même boîte attendue, seule la prédiction ayant l'IoU le plus élevé est comptée comme correctement prédite.
- L'IoU (*Intersection over Union*) est une mesure du recouvrement entre deux boîtes englobantes. Les valeurs d'IoU sont comprises entre 0 et 1 : plus cette valeur est proche de 1, et plus les boîtes sont similaires. L'IoU moyen est la moyenne des valeurs d'IoU de l'ensemble des boîtes prédites.
- L'AP (*Average Precision*) est la moyenne de la précision du réseau avec un certain seuil d'IoU utilisé pour l'acceptation des boîtes englobantes. Le mAP (*mean Average Precision*) correspond à la moyenne des valeurs d'AP obtenues avec des valeurs d'IoU de 0.5, 0.55, 0.6... et 0.95.

Comparaison de YOLOv3 sur deux bases de données

Nous souhaitons dans un premier temps comparer les performances de YOLOv3 sur notre base d'images couleur de voitures, « Reconstruction » (voir Section 5.2.1), contre une base de données publique de détection de véhicules : « NuScenes » (Caesar *et al.* [2019]). Des exemples d'images de ces deux bases de données avec les boîtes englobantes prédites par YOLOv3 sont présentées en Figure 6.6. Pour chaque objet détecté, sont indiquées la classe associée à cette boîte englobante (*voiture/car* ou *camion/truck*) ainsi que la valeur de confiance de YOLOv3 dans sa prédiction.

Nous présentons en Table 6.3 la comparaison des résultats sur ces deux bases de données. Les nombres de boîtes attendues et prédites est très différent entre les deux bases de données, en raison de la quantité d'images traitées ainsi que du nombre moyen de véhicules par image.

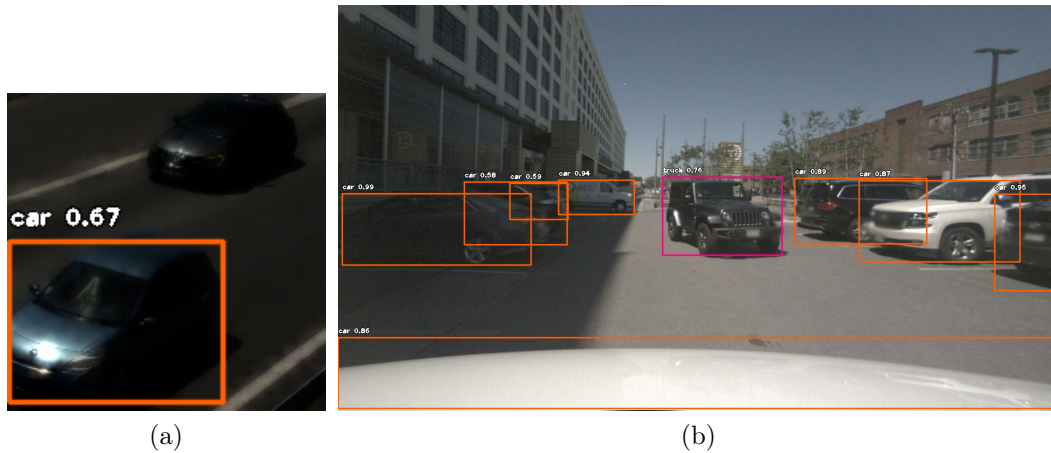


FIGURE 6.6 – Détection de véhicules par boîtes englobantes avec YOLOv3 sur des images de (a) « Reconstruction » et (b) « NuScenes ».

Données	Boîtes attendues	Boîtes prédites	Boîtes correctes	(%)	IoU moyen	mAP
Reconstruction	1446	707	553	38,2	0,721	0,605
NuScenes	29935	26927	17317	57,8	0,581	0,411

TABLE 6.3 – Comparaison des nombres de boîtes et résultats de YOLOv3 sur les bases d'images « Reconstruction » et « NuScenes ».

Le nombre de boîtes englobantes prédites est plus proche du nombre de boîtes attendues dans « NuScenes » que dans « Reconstruction ». Ainsi, le pourcentage de boîtes correctes est plus élevé dans « NuScenes » (57,8% contre 38,2%), mais le nombre de prédictions incorrectes est également très élevé. De plus, les valeurs d'IoU moyen et de mAP sont plus faible pour « NuScenes » que pour « Reconstruction ».

YOLOv3 détecte donc davantage de véhicules dans « NuScenes », mais avec un taux de précision bien inférieur à celui de « Reconstruction ». De telles différences peuvent venir de la taille des images, mais aussi de l'information contenue dans les canaux de couleur des images des deux bases de données.

Par la suite, sauf indication contraire, les données présentées sont celles de la base de données « Reconstruction ».

Pertinence des canaux de couleur

Nous cherchons ensuite à déterminer la pertinence de l'information des différents canaux de couleur reconstruits. Pour cela, chaque canal de chaque image a été sélectionné individuellement et utilisé pour générer une image en niveau de gris. Les images en niveau de gris sont ensuite fournies à YOLOv3 pour prédiction.

Les résultats de la prédiction de YOLOv3 sur les images en niveaux de gris issues des différents canaux de couleur sont présentées en Table 6.4. Nous pouvons d’abord constater que les valeurs d’IoU et de mAP sont pratiquement identiques entre les 3 canaux de couleur. Cela signifie que l’information contenue dans chacun des 3 canaux est tout aussi pertinente dans l’évaluation de l’IoU et du mAP. Le nombre de boîtes prédites et correctes est également très proches entre les canaux rouge et vert.

Données	Boîtes attendues	Boîtes prédites	Boîtes correctes	(%)	IoU moyen	mAP
Image couleur	1446	707	553	38,2	0,721	0,605
Gris (canal rouge)	1446	689	552	38,2	0,720	0,590
Gris (canal vert)	1446	698	553	38,2	0,724	0,587
Gris (canal bleu)	1446	762	601	41,6	0,717	0,592

TABLE 6.4 – Comparaison des nombres de boîtes et résultats de YOLOv3 sur les images en niveau de gris constituées à partir des canaux de couleur, sur la base de données « Reconstruction ».

En revanche, le canal bleu permet la prédiction d’un nombre plus important de boîtes prédites et correctes (+ $\sim 10\%$). L’information contenue dans ce canal est donc plus pertinente pour la tâche de détection de véhicules avec YOLOv3 sur la base de données « Reconstruction ».

De plus, chaque canal de couleur individuel permet l’obtention de résultats similaires ou meilleurs à ceux des images couleur, à l’exception du mAP. Puisqu’une image couleur est la concaténation de ces différents canaux sans perte d’information, cela signifie que cette baisse de performances est liée au traitement par YOLOv3 des images couleur.

Importance des canaux de couleur dans YOLOv3

Pour compléter l’analyse précédente, nous souhaitons déterminer l’importance qu’associe YOLOv3 à chaque canal de couleur des images qui lui sont fournies. Pour cela, nous remplaçons deux canaux de couleur par un signal vide (image entièrement noire). Si YOLOv3 accorde la même pertinence à chacun des trois canaux, les résultats des trois images devraient suivre la même relation que dans le Table 6.4 : des valeurs presque identiques entre les canaux rouge et vert, et plus élevées pour le canal bleu.

Les résultats de YOLOv3 sur ces images sont présentés en Table 6.5. Contrairement au résultat attendu, le nombre de boîtes correctement prédites dans les images à un seul canal est beaucoup plus faible que pour l’image couleur. Cela signifie que, pour un unique canal de couleur, YOLOv3 est plus adapté à le traiter comme une image en niveau de gris, qu’en vidant le signal des deux autres canaux.

Données	Boîtes attendues	Boîtes prédites	Boîtes correctes	(%)	IoU moyen	mAP
Image couleur	1446	707	553	38,2	0,721	0,605
Image rouge	1446	550	440	30,4	0,715	0,567
Image verte	1446	521	433	29,9	0,729	0,592
Image bleue	1446	443	354	24,5	0,714	0,569

TABLE 6.5 – Comparaison des nombres de boîtes et résultats de YOLOv3 sur les images couleur après suppression de certains canaux, sur la base de données « Reconstruction ».

Le nombre de boîtes prédites et correctes est plus élevé pour le canal rouge que pour le canal vert, et pour le canal vert que pour le canal bleu. Ces résultats indiquent que YOLOv3 accorde une pertinence plus élevée aux canaux rouge et vert qu'au canal bleu, malgré la pertinence de ce canal dans les images « Reconstruction » (voir Table 6.4).

Une telle différence signifie que YOLOv3 n'est pas un réseau capable de tirer parti de l'information présente dans les images de la base de données « Reconstruction », et particulièrement de l'information du canal bleu.

Impact de la compression JPEG sur YOLOv3

YOLOv3 est pré-entraîné sur ImageNet, une large base de données d'images d'objets et de véhicules, dans laquelle les images sont compressées au format JPEG. Or, nos images sont directement générées à partir des images hyperspectrales, et enregistrées dans un format sans compression afin d'éviter toute perte d'information. Nous souhaitons donc tester si la compression JPEG de nos images, c'est-à-dire dans un type de données pour lesquelles YOLOv3 a été entraîné, permet d'améliorer ses performances sur nos images.

Les résultats de YOLOv3 sur les images compressées sont présentés en Table 6.6. Nous pouvons voir une baisse significative des performances de YOLOv3 en comparaison aux données non-compressées (voir Table 6.4). De nouveau, les images couleur ont des performances similaires aux différents canaux individuels, indiquant qu'aucune information n'est gagnée à la concaténation des trois canaux en une seule image. Malgré la compression JPEG des données de l'entraînement de YOLOv3, compresser nos données vers le même format provoque une perte importante d'information qui réduit drastiquement les performances du réseau.

Contrairement aux données sans compression, le canal bleu ne permet pas la prédiction de davantage de boîtes englobantes que les autres canaux. Ce résultat signifie que la compression au format JPEG a éliminé davantage d'information pertinente sur le canal bleu que sur les deux autres canaux.

Données	Boîtes attendues	Boîtes prédites	Boîtes correctes	(%)	IoU moyen	mAP
Image couleur avec compression	1446	503	388	26,8	0,714	0,570
Gris (canal rouge) avec compression	1446	481	383	26,5	0,717	0,556
Gris (canal vert) avec compression	1446	484	390	27,0	0,718	0,564
Gris (canal bleu) avec compression	1446	485	388	26,8	0,714	0,569

TABLE 6.6 – Comparaison des nombres de boîtes et résultats de YOLOv3 après compression JPEG, sur les images en niveau de gris constituées à partir des canaux de couleur, sur la base de données « Reconstruction ».

Suppression des bandes proches du bleu

Nous souhaitons désormais analyser la pertinence de l'information contenue dans les différentes bandes spectrales.

Pour cela, nous analysons dans un premier temps l'évolution des performances de YOLOv3 en fonction des bandes spectrales utilisées pour la reconstruction d'images couleur. À toute fin utile, nous rappelons que les images couleur sont reconstruites avec les bandes spectrales de 450nm à 780nm, qui ont un pas de 4nm.

Le procédé est le suivant : Les images couleur sont reconstruites en ignorant les bandes spectrales dont la longueur d'onde est la plus faible, jusqu'à 570nm à la frontière entre le vert et le rouge. Après la reconstruction, les images sont normalisées comme si les bandes spectrales ignorées étaient quand même présentes. Ainsi, lorsque davantage de bandes spectrales du domaine bleu (puis vert) sont supprimées, l'image reconstruite devient de plus en plus jaune (puis rouge). YOLOv3 est ensuite utilisé pour la détection de véhicules sur les images, et son nombre de boîtes prédites est présenté en Figure 6.7.

Le nombre de boîtes prédites diminue de 707 à 559 (−21%) lors de la suppression des bandes inférieures à 570nm. La diminution la plus rapide du nombre de boîtes englobantes prédites se produit lors de la suppression des bandes entre 450nm et 490nm, atteignant 585 boîtes (−17%), indiquant la pertinence de ces longueurs d'onde pour la détection de véhicules dans les images reconstituées. Le nombre de boîtes prédites augmente légèrement avec la suppression des bandes spectrales entre 490nm et 530nm, puis diminue à nouveau. Cette augmentation implique que l'information contenue dans les longueurs d'onde entre 490nm et 530nm est soit redondante, soit néfaste à la détection de véhicules, par exemple en favorisant l'apparition d'autres types d'informations dans les images reconstruites.

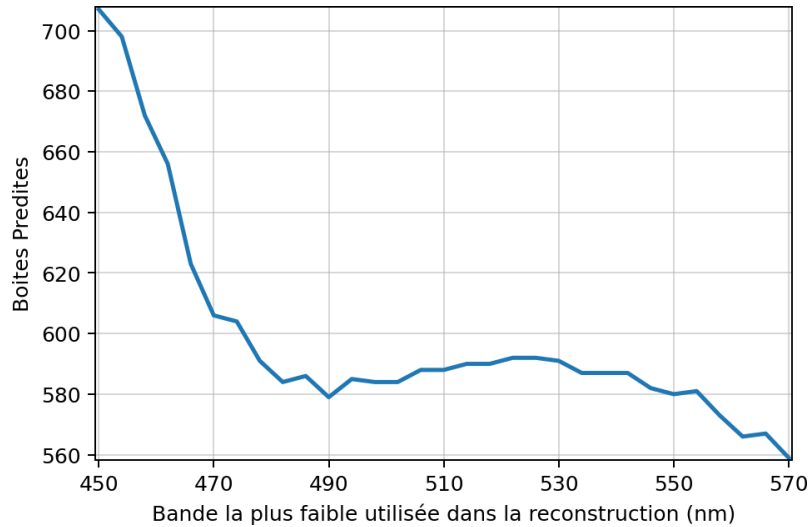


FIGURE 6.7 – Nombre de boîtes prédites par YOLOv3 sur les images couleur reconstituées, en fonction de la longueur d’onde la plus faible utilisée pour la reconstruction.

Pertinence des bandes spectrales

Dans un second temps, nous ne considérons que les bandes spectrales, individuellement. Les bandes spectrales sont des données interprétables comme des images en niveaux de gris. Chaque bande spectrale est donc fournie individuellement à YOLOv3 comme une image en niveaux de gris. Nous avons donc utilisé YOLOv3 sur chaque bande de chaque image hyperspectrale, afin d’évaluer la pertinence de ces différentes bandes spectrales pour la détection de véhicules.

Le nombre de boîtes prédites par YOLOv3 en fonction de la bande spectrale est présenté en Figure 6.8. Cette courbe indique que les bandes les plus pertinentes pour la détection de véhicules par YOLOv3 sont celles en-dessous de 700nm, suivies par les bandes au-delà de 900nm. Dans près de 40% des bandes spectrales, YOLOv3 réussit à prédire davantage de boîtes englobantes que sur les images reconstituées en couleur (707).

Impact de la taille des images

Enfin, nous souhaitons désormais analyser la pertinence de YOLOv3 à travers un dernier axe : la taille des images. Nous considérons les images couleur reconstituées à partir des images hyperspectrales de taille spatiale 200x200 et 803x803. Nous rappelons que ces images hyperspectrales sont acquises avec une dimension spatiale de 50x50, et sont agrandies à l’aide d’un canal monochromatique de taille 1000x1000 et d’une étape de pansharping (voir Sec. 3.3.1), ce qui peut entraîner des déformations dans les images produites.

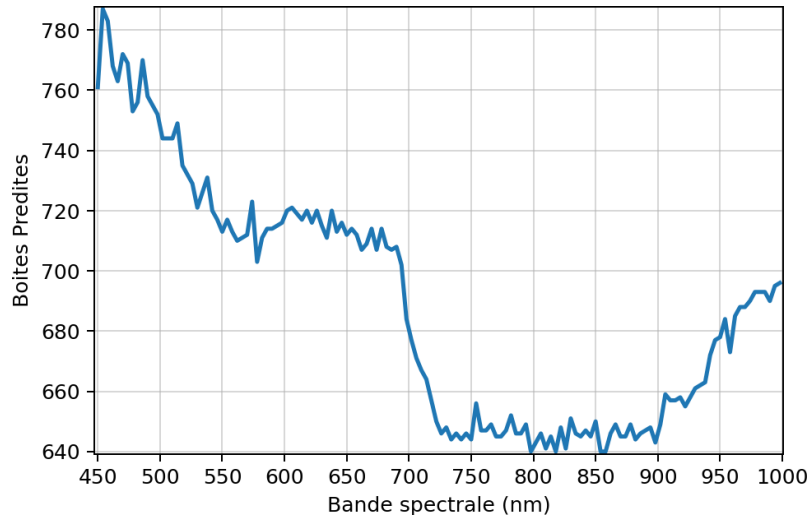


FIGURE 6.8 – Nombre de boîtes prédites par YOLOv3 sur les bandes spectrales considérées comme images en niveau de gris.

Les résultats de YOLOv3 sur ces images de tailles différentes sont présentés en Table 6.7. Le nombre de boîtes prédites et correctes est plus élevé sur les images de taille 803x803.

Données	Boîtes attendues	Boîtes prédites	Boîtes correctes	(%)	IoU moyen	mAP
Image couleur 200x200	1446	751	579	40,0	0,7214	0,605
Image couleur (803x803)	1446	1005	786	54,4	0,7207	0,602

TABLE 6.7 – Comparaison des nombres de boîtes et résultats de YOLOv3 sur les images couleur de taille 200x200 et 803x803, sur la base de données « Reconstruction ».

Cependant, YOLOv3 prédit également un nombre plus important de boîtes incorrectes. De plus, les valeurs d'IoU et de mAP sont similaires, indiquant que la précision de YOLOv3 n'est pas impactée par la taille des images pour notre base de données « Reconstruction ».

6.3.2 Mask-RCNN

Mask-RCNN (He *et al.* [2017]) est un réseau de neurones profond de détection d'objets et qui prédit simultanément les boîtes englobantes, les classes et les masques de ces objets.

Nous utilisons l'implémentation de Abdulla [2017] du Mask-RCNN pour la détection d'objets sur les images de la base de données « Reconstruction »,

de taille 200x200. Les données sont séparées en ensembles d'entraînement (350 images) et de validation (50 images). Le réseau est entièrement entraîné sur nos données, sans pré-entraînement des poids, à partir des boîtes englobantes et des masques définis en Section 5.2.2, pendant 100 epochs. Cela signifie qu'il peut exister un biais dans les résultats qui seront observés, car certaines des données d'entraînement et de validation sont celles produites par un Mask-RCNN lui-même entraîné sur une autre base de données. La labellisation (résultat attendu) et les prédictions du Mask-RCNN sur des images de l'ensemble de validation sont présentées en Figure 6.9.

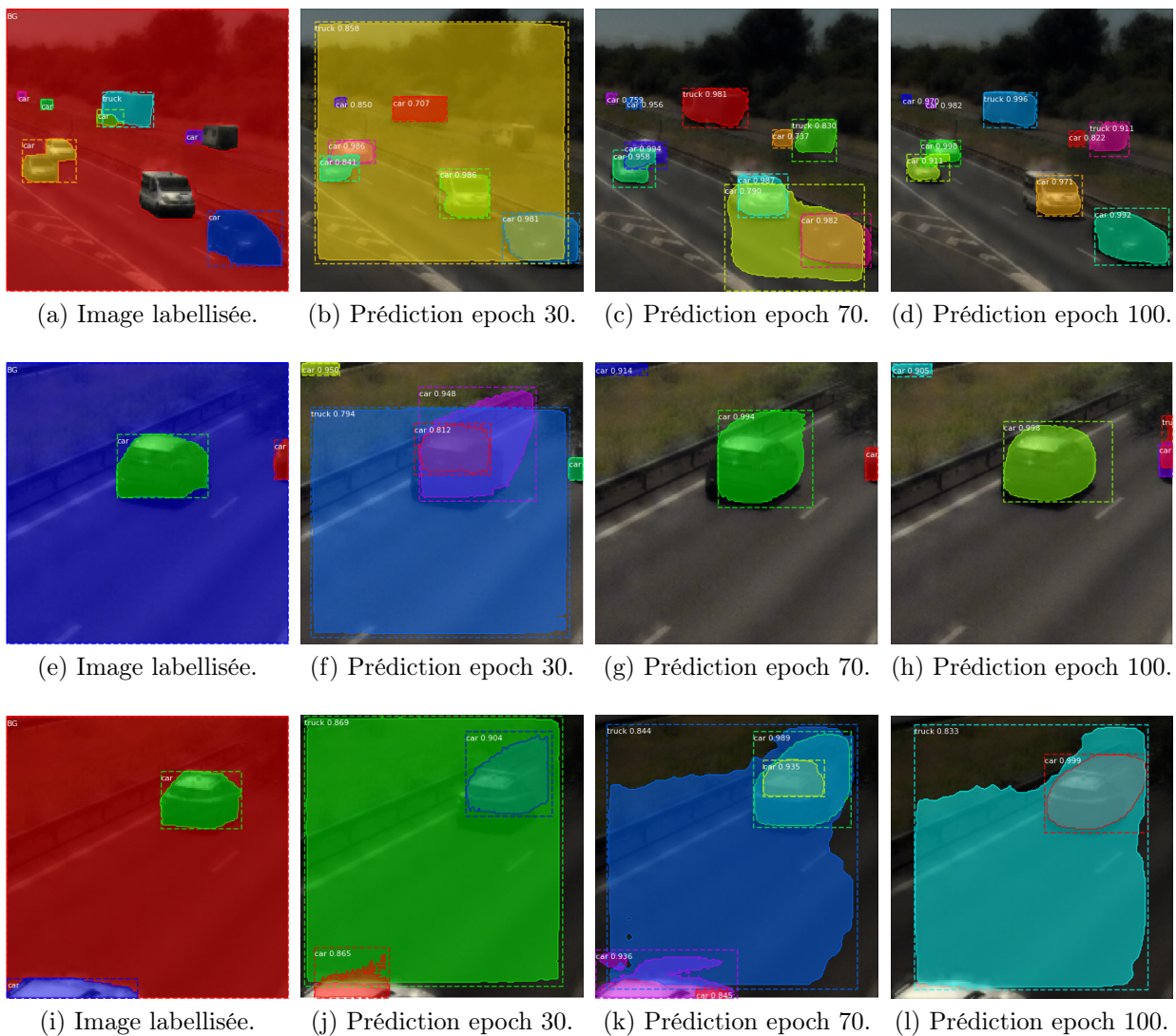


FIGURE 6.9 – Exemples de labellisation et de prédiction de véhicules sur les images de l'ensemble de validation de « Reconstruction ».

Pour la labellisation, des masques sont définis pour chaque objet comme

voiture (car) ou *camion (truck)*, et un masque est également défini pour l'ensemble de l'*arrière-plan*. Chaque masque est présenté avec une couleur différente. Les véhicules n'appartenant à aucune des deux autres catégories (camionnettes, motos...) sont ignorés pour l'apprentissage (voir Figure 6.9a). Nous rappelons que lorsque deux objets d'une même classe sont très proches, notre méthode génère une unique boîte englobante pour les deux véhicules, par exemple sur les deux voitures au milieu gauche de la Figure 6.9a.

Après 30 epochs, le Mask-RCNN parvient déjà à détecter la majeure partie des véhicules, mais commet un nombre important d'erreurs : boîtes de mauvaise taille, masque mal positionné, boîtes et masques dupliqués et surtout l'apparition d'une boîte *camion* recouvrant presque l'intégralité des images. Cette boîte est due à des images de l'ensemble d'entraînement représentant un unique camion sur toute l'image, provoquant une mauvaise généralisation du Mask-RCNN qui les prédit sur toutes les images.

Le réseau améliore ensuite la précision de ses prédictions, et atteint de bons résultats à l'epoch 100 : Sur la Figure 6.9d, tous les véhicules sont prédits séparément les uns des autres, avec des boîtes englobantes représentatives des véhicules, malgré des masques légèrement mal placés. En revanche, deux erreurs apparaissent en Figure 6.9h : sur le bord droite de l'image, deux prédictions de véhicules (1 *voiture* et 1 *camion*) se superposent, et dans le coin haut-gauche, la barrière routière est détectée comme étant une voiture. De plus, les erreurs les plus flagrantes sont dans la Figure 6.9l. Non seulement un camion est prédit sur l'ensemble de l'image comme lors des epochs précédentes, mais le réseau ne détecte plus la voiture en bas de l'image, qui était pourtant détectée à l'epoch 70. La voiture correctement prédite a également un masque allongé par rapport à sa taille réelle.

En résumé, le Mask-RCNN sans pré-entraînement obtient des résultats satisfaisants sur les images de la base de données « Reconstruction » au cours des 100 epochs et avec 350 images d'entraînement. Les erreurs qui subsistent dans la détection des boîtes englobantes et dans la taille des masques pourraient vraisemblablement être résolues avec un nombre d'epochs plus élevé et un ensemble d'entraînement plus grand. En revanche, le réseau nécessite près de 20 secondes pour la prédiction d'une image, et augmenter le nombre d'epochs et/ou le nombre d'images nécessiterait un temps d'apprentissage important. De plus, ce réseau est optimisé pour des images du domaine visible avec trois canaux de couleur, et n'est pas en l'état utilisable sur des images hyperspectrales. Il est possible de transformer sa structure pour traiter des images avec plus de 100 bandes, mais cette modification du réseau se répercuterait sur le temps de calcul, sans offrir de garantie de résultats en raison du nouveau type de données d'entrée.

6.4 Réseaux pour la classification d'images hyperspectrales

Nous souhaitons étudier l'utilisabilité de réseaux de neurones hyperspectraux issus de l'état de l'art sur nos images hyperspectrales.

À notre connaissance, seules deux revues de l'état de l'art, [Audebert *et al.* \[2019\]](#) et [Paoletti *et al.* \[2019\]](#), ont mis leur code à disposition. Ces deux revues présentent des bibliothèques pour la classification d'images hyperspectrales par réseaux de neurones.

Dans cette section, nous nous intéressons à l'analyse de ces deux bibliothèques en comparant leurs résultats à travers plusieurs implémentations des mêmes réseaux de neurones.

6.4.1 Terminologie

Dans cette section, les réseaux de neurones hyperspectraux sont analysés au travers des critères suivants :

- Le « Réseau » correspond au type d'architecture du réseau ainsi qu'à la publication d'origine de ce réseau.
- La « Source » indique la provenance des résultats (article ou implémentation). Lorsque la source indiquée est « Nous », il s'agit de notre propre réimplémentation.
- La « Base de données » utilisée, parmi *Pavia University* et *Indian Pines*¹.
- La méthode de « Sélection de l'ensemble d'entraînement », telle que présentée dans [Audebert *et al.* \[2019\]](#), utilisée pour séparer les pixels labellisés en ensembles d'entraînement, de validation et de test. La première méthode est la sélection aléatoire. Cette méthode permet de renforcer l'apprentissage grâce à des données variées des mêmes classes. Cependant, certains pixels des ensembles d'entraînement, de validation et de test sont voisins, ce qui signifie que les pixels de validation et de test sont appris pendant l'entraînement de tous les réseaux bénéficiant des informations spatiales (par ex. réseaux à convolutions 2D ou 3D). La deuxième méthode est la méthode « disjointe », qui consiste à séparer distinctement les ensembles d'entraînement, de validation et de test. Les résultats obtenus avec la méthode disjointe sont plus faibles qu'avec la méthode aléatoire sur les données de validation, mais ils ont plus représentatifs de la réalité.
- La « Précision (%) » est le pourcentage des éléments de l'ensemble de validation qui ont été correctement classifiés.

1. https://www.ehu.eus/ccwintco/index.php/Hyperspectral_Remote_Sensing_Scenes, dernière visite le 08/03/2023.

- Le « Coefficient Kappa » (ou « Kappa de Cohen ») est une mesure de la correspondance entre la labellisation et la prédiction.

6.4.2 *DeepHyperX* (Audebert, 2019)

Le constat présenté par la revue de [Musgrave et al. \[2020\]](#) est celui que les publications sur les fonctions de perte des réseaux de neurones ne se comparent pas sur une base identique et équitable, et les avancées faites sont moindres que celles décrites dans les publications. Un constat similaire est présenté par [Ghamisi et al. \[2017\]](#) sur les réseaux de neurones hyperspectraux : si ces publications prétendent à des résultats toujours plus impressionnants, le manque de base commune rend impossible la comparaison fiable des méthodes présentées. [Audebert et al. \[2019\]](#) constatent la même chose, et mentionnent que de nombreux auteurs ne comparent pas leurs résultats aux autres de manière juste ou omettent des paramètres, ce qui rend difficile l'évaluation des différentes méthodes proposées dans l'état de l'art.

[Audebert et al. \[2019\]](#) ont ré-implémenté et comparé entre eux des réseaux de neurones issus de publications antérieures. Les auteurs ont également mis à disposition la bibliothèque *DeepHyperX*² permettant d'utiliser ces réseaux de neurones et d'en implémenter de nouveaux, sur des bases de données hyperspectrales de l'état de l'art.

Afin de préserver les résultats de l'article, tous les hyperparamètres ont été définis avec leurs valeurs par défaut lors de notre utilisation de *DeepHyperX*.

Avec sélection aléatoire

Nous présentons en Table 6.8 les résultats de différentes implémentations du réseau de [Li et al. \[2017\]](#).

Réseau	Source	Base de données	Sélection de l'ensemble d'entraînement	Précision (%)	Coefficient Kappa
Li, 2017	Article (Li, 2017)	PaviaU	Aléatoire	99.39	0.9920
Li, 2017	Article (Audebert, 2019)	PaviaU	Aléatoire	96.71	/
Li, 2017	DeepHyperX	PaviaU	Aléatoire	97.12	0.962
Li, 2017	Nous	PaviaU	Aléatoire	98.37	/

TABLE 6.8 – Comparaison des résultats du réseau proposé par [Li et al. \[2017\]](#) en fonction de l'implémentation, sur la base de données *Pavia University*. La sélection de l'ensemble d'entraînement est conduite aléatoirement.

Malgré les résultats de plus de 99% de précision selon l'article original, [Audebert et al. \[2019\]](#) décrivent une précision inférieure à 97% pour le même

2. <https://github.com/nshaud/DeepHyperX>, dernière visite le 07/03/2023.

réseau : une très bonne précision pour la classification de matériaux, mais inférieure à la précision attendue. Les résultats obtenus avec *DeepHyperX* sont proches de ceux décrits dans la revue, mais légèrement plus favorables à ceux de l'article original. Enfin, la précision obtenue avec notre propre implémentation de ce réseau est un compromis entre les autres résultats. Si notre précision diffère de plus de 1% par rapport à l'article d'origine, elle est loin d'être aussi faible que celle obtenue avec *DeepHyperX* et sa publication, ce qui peut indiquer que certains paramètres de *DeepHyperX* réduisent les performances du réseau.

Avec sélection disjointe

Nous présentons en Table 6.9 les résultats de différentes implémentations du réseau de Li *et al.* [2017], avec sélection disjointe des ensembles d'entraînement, de validation et de test. Les auteurs du réseau n'ont pas présenté de résultats avec cette méthode de sélection.

Réseau	Source	Base de données	Sélection de l'ensemble d'entraînement	Précision (%)	Coefficient Kappa
Li, 2017	Article (Li, 2017)	PaviaU	Disjointe	/	/
Li, 2017	Article (Audebert, 2019)	PaviaU	Disjointe	84.32	0.799
Li, 2017	DeepHyperX	PaviaU	Disjointe	86.77	0.827
Li, 2017	Nous	PaviaU	Disjointe	90.71	/

TABLE 6.9 – Comparaison des résultats du réseau proposé par Li *et al.* [2017] en fonction de l'implémentation, sur la base de données *Pavia University*. La sélection de l'ensemble d'entraînement est conduite avec la méthode disjointe.

Comme attendu, la précision et le coefficient kappa des trois implémentations sont beaucoup plus faibles qu'avec la sélection aléatoire (voir Table 6.8), car le voisinage des pixels utilisés à l'entraînement ne contient plus les pixels des ensembles de validation et de test.

À nouveau, notre implémentation atteint une précision plus élevée que *DeepHyperX*, indiquant une meilleure sélection des hyperparamètres du réseau. De même, les résultats obtenus avec *DeepHyperX* sont supérieurs à ceux décrits par Audebert *et al.* [2019]. Ainsi, ces différences peuvent provenir de modifications du code de *DeepHyperX*, après sa publication, sur la structure du réseau ou de ses hyperparamètres.

Contrairement aux Tables 6.8 et 6.9, les résultats du réseau sur *Indian Pines* avec *DeepHyperX* présentés en Table 6.10 sont bien inférieurs à ceux décrits par Audebert *et al.* [2019]. La raison d'une différence aussi importante peut également provenir de modifications du code de *DeepHyperX* après la publication de l'article revue. De plus, les résultats obtenus sur *Indian Pines*

sont faibles en comparaison à ceux obtenus sur *Pavia University* (Table 6.9) ce qui indique que la variabilité des données est moins bien apprise au cours de l'apprentissage.

Réseau	Source	Base de données	Sélection de l'ensemble d'entraînement	Précision (%)	Coefficient Kappa
Li, 2017	Article (Li, 2017)	Indian Pines	/	/	/
Li, 2017	Article (Audebert, 2019)	Indian Pines	Disjointe	75.47	0.719
Li, 2017	DeepHyperX	Indian Pines	Disjointe	62.11	0.560
Li, 2017	Nous	Indian Pines	/	/	/

TABLE 6.10 – Comparaison des résultats du réseau proposé par Li *et al.* [2017] en fonction de l'implémentation, sur la base de données *Indian Pines*. La sélection de l'ensemble d'entraînement est conduite avec la méthode disjointe.

6.4.3 *hyperspectral_deeplearning_review* (Paoletti, 2019)

Paoletti *et al.* [2019] ont mis à disposition une autre bibliothèque, *hyperspectral_deeplearning_review*³, permettant d'utiliser différents types de réseaux de neurones et de méthodes d'apprentissage : réseaux à convolution 1D, 2D ou 3D, réseau récurrent, apprentissage par transfert... Cette publication compare également ces réseaux sur des bases de données hyperspectrales de l'état de l'art.

Nous présentons en Table 6.11 les résultats de différentes implémentations du réseau 3D-CNN de Paoletti *et al.* [2019]. Les résultats de l'article et de son code sont très proches, ayant tous deux une précision presque parfaite (> 99.9%). Cependant, notre réimplémentation de ce réseau n'atteint qu'une précision de 84.05% sur l'ensemble de validation. L'ensemble d'entraînement n'étant composé que de 15% des pixels labellisés, la raison d'un tel écart pourrait provenir d'une mauvaise sélection des données d'entraînement résultant en un manque de représentativité des données de validation. Nous n'avons à ce jour pas conduit d'expérience pour vérifier cette hypothèse.

6.4.4 Conclusion

Selon Audebert *et al.* [2019], les publications présentant des réseaux de neurones pour la classification d'images hyperspectrales comparent inégalement leurs résultats à ceux d'autres méthodes, sans présenter l'entièreté des paramètres et hyperparamètres utilisés pour l'affinement de leur réseau. Par

3. https://github.com/mhaut/hyperspectral_deeplearning_review, dernière visite le 07/03/2023.

Réseau	Source	Base de données	Sélection de l'ensemble d'entraînement	Précision (%)	Coefficient Kappa
3D-CNN (Paoletti, 2019)	Article (Paoletti, 2019)	PaviaU	Aléatoire	99.92	0.9989
3D-CNN (Paoletti, 2019)	Code (Paoletti, 2019)	PaviaU	Aléatoire	99.97	0.9996
3D-CNN (Paoletti, 2019)	Nous	PaviaU	Aléatoire	84.05	0.7910

TABLE 6.11 – Comparaison des résultats du réseau 3D-CNN proposé par [Paoletti et al. \[2019\]](#) en fonction de l'implémentation, sur la base de données *Pavia University*. La sélection de l'ensemble d'entraînement est conduite aléatoirement.

conséquent, les performances réelles de ces réseaux seraient moins élevées que celles décrites par les auteurs. Nos analyses au travers des réseaux de neurones de [Paoletti et al. \[2019\]](#) (voir Table 6.11) et de [Li et al. \[2017\]](#) (voir Table 6.8) confirment cette hypothèse. En revanche, nous montrons également que les hyperparamètres utilisés par [Audebert et al. \[2019\]](#) pour la réimplémentation de ces modèles peuvent être considérablement améliorés à partir des informations disponibles sur ces modèles.

Comme attendu, la sélection des ensembles d'entraînement par méthode disjointe plutôt qu'aléatoire permet d'obtenir des performances de classification plus réalistes, mais également beaucoup plus faibles.

De nombreuses différences tant positives que négatives entre les résultats présentés par [Audebert et al. \[2019\]](#) et leur code mis en ligne, *DeepHyperX*, laissent à penser que le code a été modifié suite à la publication de l'article revue, rendant nécessaire leur comparaison.

Les deux bibliothèques sont flexibles pour l'implémentation de nouveaux réseaux de neurones et l'utilisation de nouvelles bases de données hyperspectrales à une seule image. Cependant, elles ne permettent pas de modifier une partie importante des hyperparamètres, ce qui limite les possibilités d'amélioration des réseaux. De manière plus importante, ces bibliothèques ne permettent pas non plus l'utilisation de bases de données constituées de plusieurs images, caractéristique essentielle à notre utilisation de ces bibliothèques avec nos propres données. Or, nos bases de données d'images hyperspectrales, présentées en Sections 3.4 et 3.5, sont constituées de centaines ou de milliers d'images, pour lesquelles une classification adéquate nécessite un entraînement sur plusieurs images.

6.5 Conclusion

Dans ce chapitre, nous avons présenté et analysé des réseaux de neurones pour la détection de véhicules et la classification d'images hyperspectrales. Un

réseau intitulé « réseau spectral » et ses dérivés, un « réseau colorimétrique » et un réseau à convolutions 3D, nous ont permis d'observer une pertinence des dimensions spectrales et spatiales, mais cette pertinence reste faible sur nos données. De plus, peut-être en raison de la profondeur faible de ces réseaux ou en raison des données utilisées à leur apprentissage, les caractéristiques apprises par ces réseaux se limitent à des informations de luminosité moyenne du spectre et à la présence de végétation.

L'utilisation d'un réseau de neurones état de l'art en détection de véhicules, YOLOv3, sur nos images reconstruites en couleur, nous a permis d'analyser la pertinence de l'information contenue dans les différents canaux de couleur et dans les bandes spectrales pour la détection de véhicules. L'utilisation de YOLOv3 nous a également permis d'observer que des images avec de plus grandes résolutions spatiales étaient plus adaptées à la détection via ces réseaux. L'utilisation d'un autre réseau état de l'art, Mask-RCNN, a également offert des perspectives intéressantes en détection et classification de véhicules. Cependant, ces deux réseaux fonctionnent exclusivement sur des images couleur (3 canaux) et ne sont pas utilisables sur nos données hyperspectrales.

Enfin, nous avons exploré deux bibliothèques de réseaux de neurones pour la classification de pixel dans des images hyperspectrales. L'inconvénient principal à l'utilisation de ces deux bibliothèques est qu'elles souffrent d'un manque de flexibilité, empêchant l'ajout de bases de données constituées de plusieurs images hyperspectrales. De plus, nous avons pu constater des différences majeures entre les résultats décrits par les auteurs et les résultats obtenus avec plusieurs implémentations de ces algorithmes, montrant l'importance de certains choix de designs et de certains hyperparamètres non-décrits dans les publications scientifiques.

En conclusion, à l'heure actuelle, aucun réseau de l'état de l'art n'est adapté à l'apprentissage sur des données telles que celles que nous proposons au Chapitre 3. Nos tentatives d'utilisation de réseaux peu profonds sur des images du domaine spectral 400-1000nm (visible et proche infrarouge) ont également été infructueuses. À l'avenir, nous souhaitons développer des architectures de réseaux profondes, capable d'apprendre des caractéristiques variées sur des bases de données d'images hyperspectrales acquises dans différentes conditions d'acquisition.

Conclusion

Travaux et contributions

Le travail conduit au cours de cette thèse avait pour objectif la détection et la discrimination de matériaux, dans des images hyperspectrales acquises avec des capteurs terrestres ou à basse altitude, pour les systèmes de mission de surveillance.

Nous avons commencé par dresser une liste de bibliothèques de spectres de matériaux et d'objets existantes et avons choisi deux de ces bibliothèques pour conduire des analyses spectrales. Nous avons décrit les transformations applicables sur les spectres issus de deux de ces bibliothèques, tels que le calcul des dérivées, d'un histogramme ou le lissage du spectre. Ces transformations ont pour effet de modifier la nature des informations présentées par le spectre et ainsi permettre une plus grande variété de traitements et d'analyses. Ces spectres ainsi transformés ont été analysés à l'aide de méthodes locales, l'extraction des pics, et globales, les mesures de similarité. L'extraction des pics n'a pas permis l'extraction de caractéristiques pertinentes pour la classification de matériaux. En revanche, les méthodes basées sur des mesures de similarité nous ont permis d'analyser les problématiques liées à la variabilité spectrale et des mélanges de matériaux. Enfin, une analyse qualitative et quantitative des spectres issus de ces bibliothèques nous a permis de caractériser un domaine spectral d'intérêt pour la conduite d'acquisitions d'images hyperspectrales, entre 1200 et 2300nm.

Les caractéristiques des caméras hyperspectrales pertinentes pour notre application ont été décrites et expliquées. L'acquisition de deux bases de données hyperspectrales a ensuite été présentée. Pour la première base de données, nous avons choisi d'utiliser une caméra instantanée afin d'obtenir des images hyperspectrales de véhicules en mouvement, nous limitant ainsi au domaine spectral 450-1000nm (visible et proche infrarouge) en dépit des préconisations précédentes. Grâce à cette caméra, plus de 15000 images de véhicules ont été acquises avec plusieurs angles de vue et dans plusieurs conditions d'illumination. Pour la deuxième base de données, près de 400 images hyperspectrales d'échantillons de ciment, de bois et de métal ont été acquises, dans plusieurs conditions d'illumination avec une caméra à balayage dans le domaine spectral 900-2500nm, conformément au domaine spectral caractérisé sur les bases de données spectrales.

Les images hyperspectrales acquises peuvent être calibrées, c'est-à-dire que les images brutes doivent être converties en images de réflectance, afin de les rendre comparables entre elles. Cependant, lorsque l'utilisateur ne mesure pas

le bruit du capteur (ou courant d'obscurité ou *DC*) fréquemment, ce bruit n'est pas corrigé correctement et les valeurs des images de réflectance sont affectées par une dérive temporelle importante. Nous avons proposé plusieurs méthodes de correction de cette dérive à partir des images *DC* disponibles, dont la méthode par interpolation linéaire obtient les meilleurs résultats. Cependant, nous avons rencontré deux limitations. Premièrement, une analyse du courant d'obscurité a montré que son évolution était non-linéaire, ce qui implique que la méthode d'interpolation linéaire proposée ne permet pas de corriger la dérive temporelle avec précision. Deuxièmement, le bruit corrigé est soit trop faible, soit trop important, et la quantité de bruit corrigé est soumise à des compromis.

De plus, en conditions opérationnelles, il est contraignant ou parfois impossible de conduire la calibration. Nous avons alors proposé une méthode de calibration utilisant un échantillon intermédiaire : une « référence de gris ». Cette méthode permet alors d'obtenir des spectres de réflectance proches de ceux obtenus avec une calibration classique. En revanche, afin d'obtenir des spectres de bonne qualité, le matériau intermédiaire doit rester inchangé entre les acquisitions. De plus, l'ajout d'étapes au calcul de la calibration ajoute des imprécisions dans le calcul du spectre.

Afin de rendre exploitables les données acquises, plusieurs autres étapes de pré-traitement ont été conduites, telles que la sélection de régions d'intérêt et la correction atmosphérique. Suite à ces pré-traitements, nous avons conduit des analyses de nos images pour la discrimination de matériaux avec et sans calibration. Presque aucune des caractéristiques extraites des images acquises dans le domaine spectral 450-1000nm n'est liée au type de matériau, car le signal était principalement corrélé aux couleurs des échantillons, ce qui confirme l'intérêt faible de ce domaine spectral pour la surveillance.

Les images hyperspectrales non-calibrées du domaine 900-2500nm ont été analysées avec deux mesures de similarité, le SAM (*Spectral Angle Mapper*) et le TC (*Tanimoto Coefficient*), pour la détection et la discrimination de matériaux. Tout d'abord, l'utilisation de l'ACP (Analyse en Composantes Principales) a permis de réduire de manière importante la dimension des images contre un faible coût en performances de détection et de discrimination. Malgré l'absence de calibration, la détection d'un matériau de référence dans des images acquises dans des conditions d'illumination différentes a été possible pour certains types de matériaux. Cette détection ne fonctionne en revanche pas pour tous les types de matériaux : dans certains cas, la détection du matériau est erronée même dans l'image de référence, en raison d'importantes variations locales de l'illumination. De plus, cette détection est actuellement limitée par la méthode de sélection automatique du seuil, qui nécessite que l'image soit déjà segmentée.

La méthode de discrimination de matériaux proposée n'offre cependant que peu de résultats intéressants, et souffre de plusieurs limitations. Pour chaque image, un utilisateur doit définir un seuil de discrimination, et recommencer

jusqu'à trouver un seuil qu'il juge adapté. De plus, le parcours des pixels pour la discrimination suit une approche naïve. Une amélioration de ce parcours des pixels permettrait notamment de réduire l'impact de pixels aberrants sur le résultat de la discrimination. Enfin, la limitation principale des méthodes proposées de détection et de discrimination est le fait qu'aucune des mesures de similarité analysée n'est robuste aux variations d'illumination pour les différents matériaux analysés.

Nous avons mis en place un réseau de neurones spectral pour la détection et discrimination de véhicules dans plusieurs conditions d'illumination et analysé les caractéristiques apprises au cours de son apprentissage. Pour la classification, ce réseau de neurones se base exclusivement sur la luminosité pour classifier les pixels. Comme les méthodes décrites précédemment, ce réseau n'est donc pas robuste aux variations d'illumination. De plus, les méthodes proposées se focalisent sur l'aspect spectral des données, les dimensions spatiales étant utilisées principalement pour la labellisation et pour leur grande quantité d'informations spectrales. Plusieurs versions de ce réseau ont alors été proposées. Cela nous a permis d'observer que pour des réseaux avec une architecture similaire, le réseau ayant pour entrée des données couleur (RGB) performe moins bien que le réseau spectral, qui performe lui-même moins bien que le réseau spatio-spectral. En revanche, la précision faible de ce réseau indique qu'il n'est pas adapté à la détection de véhicules.

Deux réseaux de l'état de l'art de détection de véhicules, YOLOv3 et Mask-RCNN, ont été utilisés pour analyser l'information contenue dans ces images hyperspectrales. Nous avons pu constater que la faible dimension spatiale des images de la première base de données affecte négativement ces réseaux état de l'art. Dans l'ensemble, les résultats de détection obtenus par ces réseaux sont satisfaisants. Cependant, ces réseaux de neurones ne sont adaptés que pour des images couleur (RGB) et non pour des images hyperspectrales. De plus, ces réseaux nécessitent déjà des capacités de calcul importantes, et les adapter ces réseaux aux images hyperspectrales aurait un impact considérable sur leurs durée d'entraînement et d'inférence.

Nous nous sommes alors intéressés aux réseaux de neurones pour le traitement d'images hyperspectrales. Deux bibliothèques ont été extraites de l'état de l'art et deux limitations majeures ont été présentées. Premièrement, il n'est pas toujours possible de reproduire les résultats des réseaux de l'état de l'art car leurs paramètres sont mal décrits dans ces publications et car les codes sont rarement mis à disposition. Enfin et surtout, ces bibliothèques sont adaptées au traitement de bases de données hyperspectrales contenant une seule image, ce qui signifie qu'il n'est actuellement pas possible de les utiliser pour apprentissage sur nos bases de données contenant une multitude d'images.

Perspectives

Les méthodes proposées pour l'analyse des spectres sont soit locales, soit globales, mais aucune des méthodes proposées jusqu'ici ne permet de déterminer quels sont les domaines spectraux déterminants pour la classification de matériaux liés à la surveillance. Par exemple, une méthode consisterait à décomposer les spectres par sous-régions spectrales et d'appliquer les mesures de similarité sur ces sous-régions pour observer quelles parties du spectres sont les plus pertinentes.

Plusieurs travaux restent à effectuer sur la dérive temporelle. Cette dérive ayant été prouvée non-linéaire, il faudrait proposer une nouvelle méthode, non-linéaire, de sélection du *DC* pour le calcul de la radiance. Les algorithmes non-linéaires étant souvent plus lents que leurs équivalents linéaires, il conviendrait alors de calculer le temps de calcul nécessaire pour les différentes méthodes proposées afin de proposer un rapport entre le temps de calcul et la qualité de la radiance reconstituée. De plus, malgré de nombreuses informations qui semblent confirmer l'hypothèse de l'impact de la chaleur sur les données du capteur, il serait intéressant de conduire de nouvelles analyses pour vérifier cette hypothèse expérimentalement.

Les méthodes proposées, malgré leurs résultats encourageants, ne sont pas robustes aux variations d'illumination. Comme mentionné en Chapitre 1. [Contexte](#), ces problèmes sont encore ouverts, et le développement d'une telle méthode serait sans aucun doute une avancée majeure dans le domaine de la classification d'objets et de matériaux dans des images hyperspectrales.

La méthode de discrimination de matériaux par mesures de similarité (voir Sec. 5.4) est limitée par la sélection du seuil de discrimination par l'utilisateur. Une méthode de sélection de seuil automatique, par exemple en demandant plutôt à l'utilisateur un nombre de classes souhaité, permettrait d'améliorer la fiabilité de cette méthode en réduisant la quantité de tests empiriques nécessaires, au risque d'augmenter le temps de calcul de la discrimination.

Une nouvelle labellisation de la base de données de voitures, par exemple en fonction des types de matériaux, pourrait permettre de nouvelles analyses sur ces données et d'apprendre des caractéristiques non liées à l'illumination. Il serait également intéressant de développer des réseaux colorimétrique, spectral et 3D plus profonds afin de vérifier si les relations trouvées entre ces réseaux sont également valides pour des architectures plus complexes. De manière plus générale, le développement de réseaux de neurones profonds adaptés au traitement de multiples images hyperspectrales va devenir crucial pour le traitement de grandes quantités de données hyperspectrales, particulièrement pour la surveillance.

Enfin, une poursuite des travaux de cette thèse est envisagée sur une thématique multimodale, afin de bénéficier simultanément des avantages de capteurs couleur, infrarouges et hyperspectraux dans une même scène.

Annexes

Bibliographie

- AASEN, Helge, HONKAVAARA, Eija, LUCIEER, Arko et ZARCO-TEJADA, Pablo J., 2018. Quantitative remote sensing at ultra-high resolution with UAV spectroscopy : A review of sensor technology, measurement procedures, and data correction workflows. doi :10.3390/rs10071091.
- ABDULLA, Waleed, 2017. Mask r-cnn for object detection and instance segmentation on keras and tensorflow. https://github.com/matterport/Mask_RCNN. Dernière visite le 29/03/2023.
- ADLER-GOLDEN, Steven M., LEVINE, Robert Y., MATTHEW, Michael W., RICHTSMEIER, Steven C., BERNSTEIN, Lawrence S., GRUNINGER, John H., FELDE, Gerald W., HOKE, Michael L., ANDERSON, Gail P. et RATKOWSKI, Anthony, 2001. Shadow-insensitive material detection/classification with atmospherically corrected hyperspectral imagery. Dans Sylvia S. Shen et Michael R. Descour, rédacteurs, *SPIE Algorithms for Multispectral, Hyperspectral, and Ultraspectral Imagery VII*, pages 460–469. doi :10.1117/12.437037. URL <http://proceedings.spiedigitallibrary.org/proceeding.aspx?articleid=914125>
- ADLER-GOLDEN, Steven M., MATTHEW, Michael W., BERNSTEIN, Lawrence S., LEVINE, Robert Y., BERK, Alexander, RICHTSMEIER, Steven C., ACHARYA, Prabhat K., ANDERSON, Gail P., FELDE, Jerry W., GARDNER, J. A., HOKE, Michael L., JEONG, Laila S., PUKALL, Brian, RATKOWSKI, Anthony J. et BURKE, Hsiao-hua K., 1999. Atmospheric correction for short-wave spectral imagery based on MODTRAN4. Dans Michael R. Descour et Sylvia S. Shen, rédacteurs, *SPIE Imaging Spectrometry V*, pages 61–69. doi : 10.1117/12.366315. URL <http://proceedings.spiedigitallibrary.org/proceeding.aspx?articleid=994756>
- AMIGO, José Manuel, BABAMORADI, Hamid et ELCOROARISTIZABAL, Saioa, 2015. Hyperspectral image analysis. A tutorial. doi :10.1016/j.aca.2015.09.030.
- AUDEBERT, Nicolas, SAUX, Bertrand et LEFÈVRE, Sébastien, 2019. Deep Learning for Classification of Hyperspectral Data : A Comparative Review. *IEEE Geoscience and remote sensing magazine*. doi :10.1109/MGRS.2019.2912563. URL <http://arxiv.org/abs/1904.10674>

- BALDRIDGE, A.M., HOOK, S.J., GROVE, C.I. et RIVERA, G., 2009. The ASTER spectral library version 2.0. *Remote Sensing of Environment*, 113(4) :711–715. doi :10.1016/j.rse.2008.11.007.
URL <https://linkinghub.elsevier.com/retrieve/pii/S0034425708003441>
- BAUMGARDNER, Marion F., BIEHL, Larry L. et LANDGREBE, David A., 2015. 220 band aviris hyperspectral image data set : June 12, 1992 indian pine test site 3. doi :doi:/10.4231/R7RX991C. Dernière visite le 05/06/2023.
URL <https://purr.purdue.edu/publications/1947/1>
- BERGMAN, Steven M., 1996. *The utility of hyperspectral data to detect and discriminate actual and decoy target vehicles*. Thèse de doctorat, US Navy.
URL <https://calhoun.nps.edu/handle/10945/9130>
- BIOUCAS-DIAS, José M., PLAZA, Antonio, DOBIGEON, Nicolas, PARENTE, Mario, DU, Qian, GADER, Paul et CHANUSSOT, Jocelyn, 2012. Hyperspectral Unmixing Overview : Geometrical, Statistical, and Sparse Regression-Based Approaches. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 5(2) :354–379.
URL <http://arxiv.org/abs/1202.6294>
- BONDI, Elizabeth, SALVAGGIO, Carl, MONTANARO, Matthew et GERACE, Aaron D., 2016. Calibration of UAS imagery inside and outside of shadows for improved vegetation index computation. Dans John Valasek et J. Alex Thomasson, rédacteurs, *SPIE Autonomous Air and Ground Sensing Systems for Agricultural Optimization and Phenotyping*, page 98660J. doi : 10.1117/12.2227214.
URL <http://proceedings.spiedigitallibrary.org/proceeding.aspx?doi=10.1117/12.2227214>
- BORSOI, Ricardo Augusto, IMBIRIBA, Tales, BERMUDEZ, José Carlos Moreira, RICHARD, Cédric, CHANUSSOT, Jocelyn, DRUMETZ, Lucas, TOURNERET, Jean-Yves, ZARE, Alina et JUTTEN, Christian, 2021. Spectral Variability in Hyperspectral Data Unmixing : A Comprehensive Review. *IEEE Geoscience and Remote Sensing Magazine*, 9(4) :223–270.
URL <http://arxiv.org/abs/2001.07307>
- BURGER, James et GELADI, Paul, 2005. Hyperspectral NIR image regression part I : Calibration and correction. *Journal of Chemometrics*, 19(5-7) :355–363. doi :10.1002/cem.938.
- CAESAR, Holger, BANKITI, Varun, LANG, Alex H., VORA, Sourabh, LIONG, Venice Erin, XU, Qiang, KRISHNAN, Anush, PAN, Yu, BALDAN, Giancarlo et BEIJBOM, Oscar, 2019. nuscenes : A multimodal dataset for autonomous driving. *arXiv preprint arXiv :1903.11027*.

- CAMPS-VALLS, Gustavo et BRUZZONE, Lorenzo, 2004. Kernel-Based Methods for Hyperspectral Image Classification. *Image and Signal Processing for Remote Sensing X*, 5573(6) :226. doi :10.1117/12.601712.
- CAUBET, Christophe, GUERRINI, Gilles, DESBARATS, Pascal et DOMENGER, Jean-Philippe, 2022a. Case Study of a Calibration Problem in Acquired Hyperspectral Images. Dans *2022 IEEE International Conference on Image Processing (ICIP)*, pages 2261–2265. IEEE. ISBN 978-1-6654-9620-9. doi : 10.1109/ICIP46576.2022.9897839.
- CAUBET, Christophe, GUERRINI, Gilles, DESBARATS, Pascal et DOMENGER, Jean-Philippe, 2022b. Evaluation of the Operational Contribution of the Use of Neural Networks on Hyperspectral Images for the Benefit of Airborne Surveillance. Dans *SET-312-RSM*. Bled, Slovenia.
URL <https://hal.science/hal-03963144v1>
- CSIRO, 2020. Csiro mineral spectral libraries. <https://mineralspectrallibraries.csiro.au/home>. Dernière visite le 30/03/2023.
- DEBORAH, Hilda, RICHARD, Noel et HARDEBERG, Jon Yngve, 2015. A Comprehensive Evaluation of Spectral Distance Functions and Metrics for Hyperspectral Image Processing. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 8(6) :3224–3234. doi : 10.1109/JSTARS.2015.2403257.
- DONG, Chao, LOY, Chen Change et HE, Kaiming, 2015. Image Super-Resolution Using Deep Convolutional Networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 38(2) :295–307.
- FARAH, I R, HEMISSI, S, SAHEB ETTABAA, K et SOULEIMAN, B, 2010. Multi-temporal Hyperspectral Images Unmixing and Classification Based on 3D Signature Model and Matching. *PIERS ONLINE*, 6(5).
- FARRELL, M.D. et MERSEREAU, R.M., 2005. On the Impact of PCA Dimension Reduction for Hyperspectral Detection of Difficult Targets. *IEEE Geoscience and Remote Sensing Letters*, 2(2) :192–195. doi :10.1109/LGRS.2005.846011.
URL <http://ieeexplore.ieee.org/document/1420303/>
- FAUVEL, Mathieu, CHANUSSOT, Jocelyn et BENEDIKTSSON, Jón Atli, 2007. Classification en imagerie hyperspectrale haute résolution : utilisation conjointe de l’information spatiale et spectrale. *Gretsi*, pages 11–14.
- GELADI, Paul, BURGER, Jim et LESTANDER, Torbjörn, 2004. Hyperspectral imaging : Calibration problems and solutions. *Chemometrics and Intelligent Laboratory Systems*, 72(2) :209–217. doi :10.1016/j.chemolab.2004.01.023.

- GHAMISI, Pedram, YOKOYA, Naoto, LI, Jun, LIAO, Wenzhi, LIU, Sicong, PLAZA, Javier, RASTI, Behnood et PLAZA, Antonio, 2017. Advances in Hyperspectral Image and Signal Processing : A Comprehensive Overview of the State of the Art. doi :10.1109/MGRS.2017.2762087.
- GOFERMAN, Stas, ZELNIK-MANOR, Lihi et TAL, Ayellet, 2012. Context-Aware Saliency Detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 34(10) :1915–1926. doi :10.1109/TPAMI.2011.272.
URL <http://ieeexplore.ieee.org/document/6112774/>
- GOLDBERG, A C, STANN, B et GUPTA, N, 2003. Multispectral, Hyperspectral, and Three-Dimensional Imaging Research at the U.S. Army Research Laboratory. Dans *6th International Conference of Information Fusion*.
- GOVENDER, M, CHETTY, K et BULCOCK, H, 2007. A review of hyperspectral remote sensing and its application in vegetation and water resource studies. *Water Sa*, 33(2) :145–151.
URL <https://journals.co.za/doi/pdf/10.10520/EJC116430>
- GROUPE DE SYNTHÈSE HYPERSPECTRAL (GSH), 2008. Synthèse sur l'imagerie hyperspectrale. *CNES strategic documentation*.
- GUO, B., GUNN, S.R., DAMPER, R.I. et NELSON, J.D.B., 2006. Band Selection for Hyperspectral Image Classification Using Mutual Information. *IEEE Geoscience and Remote Sensing Letters*, 3(4) :522–526. doi :10.1109/LGRS.2006.878240.
URL <http://ieeexplore.ieee.org/document/1715309/>
- GUO, Yanhui, HAN, Siming, CAO, Han, ZHANG, Yu et WANG, Qian, 2018. Guided filter based Deep Recurrent Neural Networks for Hyperspectral Image Classification. Dans *Procedia Computer Science*, tome 129, pages 219–223. Elsevier B.V. doi :10.1016/j.procs.2018.03.048.
- HAGEN, Nathan et KUDENOV, Michael W., 2013. Review of snapshot spectral imaging technologies. *Optical Engineering*, 52(9) :090901. doi : 10.1117/1.oe.52.9.090901.
URL <http://opticalengineering.spiedigitallibrary.org/article.aspx?doi=10.1117/1.OE.52.9.090901>
- HARRIS, Chris et STEPHENS, Mike, 1988. A combined corner and edge detector. *Alvey Vision Conference*, pages 147–151.
- HE, Kaiming, GKIOXARI, Georgia, DOLLÁR, Piotr et GIRSHICK, Ross, 2017. Mask R-CNN. Dans *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, pages 2961–2969.
URL <http://arxiv.org/abs/1703.06870>

- HEALEY, Glenn, MEMBER, Senior et SLATER, David, 1999. Models and Methods for Automated Material Identification in Hyperspectral Imagery Acquired Under Unknown Illumination and Atmospheric Conditions. *IEEE Transactions on Geoscience and Remote Sensing*, 37(6).
- HEMLINGER, Mark et ARRECCHI, Angelo, 2012. Space grade spectralon. *Conference on Characterization and Radiometric Calibration for Remote Sensing (CALCON)*. Dernière visite le 17/01/2023.
URL <https://digitalcommons.usu.edu/cgi/viewcontent.cgi?article=1131&context=calcon>
- HEROLD, Martin, 2004. Understanding spectral characteristics of asphalt roads. Rapport technique, National Center for Remote Sensing in Transportation, University of California, Santa Barbara.
URL https://www.ugpti.org/smartse/research/citations/downloads/Herold-Understanding_Spectral_Characteristics_Aspphalt_Roads-2004.pdf
- HEROLD, Martin, ROBERTS, Dar A., GARDNER, Margaret E. et DENNISON, Philip E., 2004. Spectrometry for urban area remote sensing - Development and analysis of a spectral library from 350 to 2400 nm. *Remote Sensing of Environment*, 91(3-4) :304–319. doi :10.1016/j.rse.2004.02.013.
- JANICZEK, John, THAKER, Parth, DASARATHY, Gautam, EDWARDS, Christopher S, CHRISTENSEN, Philip et JAYASURIYA, Suren, 2020. Differentiable Programming for Hyperspectral Unmixing using a Physics-based Dispersion Model. Dans *European Conference on Computer Vision*.
URL <https://github.com/johnjaniczek/InfraRender>.
- JIA, Sen, JI, Zhen, QIAN, Yuntao et SHEN, Linlin, 2012. Unsupervised Band Selection for Hyperspectral Imagery Classification Without Manual Band Removal. *Ieee Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 5(2) :531–543.
- KHAN, Zohaib, SHAFAIT, Faisal et MIAN, Ajmal, 2013. Hyperspectral imaging for ink mismatch detection. Dans *Proceedings of the International Conference on Document Analysis and Recognition, ICDAR*, pages 877–881. doi : 10.1109/ICDAR.2013.179.
- KOKALY, R.F., CLARK, R.N., SWAYZE, G.A, LIVO, K.E., HOEFEN, T.M., PEARSON, N.C., WISE, R.A., BENZEL, W.M., LOWERS, H.A., DRISCOLL, R.L. et KLEIN, A.J., 2017. USGS Spectral Library Version 7 Data : U.S. Geological Survey data release. Rapport technique, U.S. Geological Survey. doi :<https://dx.doi.org/10.5066/F7RR1WDJ>.
URL <https://dx.doi.org/10.5066/F7RR1WDJ>

- KULLBACK, S. et LEIBLER, R. A., 1951. On Information and Sufficiency. *The Annals of Mathematical Statistics*, 22(1) :79–86. doi :10.1214/aoms/1177729694.
- LAGRANGE, Adrien, FAUVEL, Mathieu, MAY, Stéphane et DOBIGEON, Nicolas, 2017. Un modèle bayésien pour le démélange, la segmentation et la classification robuste d’images hyperspectrales. Dans *Colloque GRETSI*, pages 1–4.
URL <http://oatao.univ-toulouse.fr/19124>
- LANDGREBE, David, 2002. Hyperspectral Image Data Analysis as a High Dimensional Signal Processing Problem. *IEEE Signal Processing Magazine*, 19(1) :17–28.
- LECUN, Yann, BOTTOU, Le’on, BENGIO, Yoshua et HAFFNER, Parick, 1998. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11) :2278–2324.
URL <http://ieeexplore.ieee.org/document/726791/>
- LI, Qingli, HE, Xiaofu, WANG, Yiting, LIU, Hongying, XU, Dongrong et GUO, Fangmin, 2013. Review of spectral imaging technology in biomedical engineering : achievements and challenges. *Journal of Biomedical Optics*, 18(10) :100901. doi :10.1117/1.jbo.18.10.100901.
- LI, Ying, ZHANG, Haokui et SHEN, Qiang, 2017. Spectral-spatial classification of hyperspectral imagery with 3D convolutional neural network. *Remote Sensing*, 9(1). doi :10.3390/rs9010067.
- LI, Zhuanfeng, XIONG, Fengchao, ZHOU, Jun, WANG, Jing, LU, Jianfeng et QIAN, Yuntao, 2020. BAE-Net : A Band Attention Aware Ensemble Network for Hyperspectral Object Tracking. Dans *2020 IEEE International Conference on Image Processing (ICIP)*.
URL <https://ieeexplore.ieee.org/document/9191105>
- LIU, Jinchao, OSADCHY, Margarita, ASHTON, Lorna, FOSTER, Michael, SOLOMON, Christopher J. et GIBSON, Stuart J., 2017. Deep Convolutional Neural Networks for Raman Spectrum Recognition : A Unified Solution. *The Analyst*. doi :10.1039/C7AN01371J.
URL <http://arxiv.org/abs/1708.09022>
- MA, Lei, LIU, Yu, ZHANG, Xueliang, YE, Yuanxin, YIN, Gaofei et JOHNSON, Brian Alan, 2019. Deep learning in remote sensing applications : A meta-analysis and review. doi :10.1016/j.isprsjprs.2019.04.015.
- MAIRE, Florian et LEFEBVRE, Sidonie, 2015. Detecting Aircraft in Low-Resolution Multispectral Images : Specification of Relevant IR Wavelength

-
- Bands. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 8(9) :4509–4523. doi :10.1109/JSTARS.2015.2457514.
- MARION, Rodolphe, 2003. *Apport de la télédétection hyperspectrale pour l'estimation conjointe des concentrations des gaz atmosphériques et des réflectances des sols*. Thèse de doctorat, Cergy-Pontoise.
- MEERDINK, Susan K., HOOK, Simon J., ROBERTS, Dar A. et ABBOTT, Elsa A., 2019. The ECOSTRESS spectral library version 1.0. *Remote Sensing of Environment*, 230 :111196. doi :10.1016/j.rse.2019.05.015.
URL <https://linkinghub.elsevier.com/retrieve/pii/S0034425719302081>
- MEI, Shaohui, YUAN, Xin, Ji, Jingyu, ZHANG, Yifan, WAN, Shuai et DU, Qian, 2017. Hyperspectral image spatial super-resolution via 3D full convolutional neural network. *Remote Sensing*, 9(11). doi :10.3390/rs9111139.
- MELGANI, Farid et BRUZZONE, Lorenzo, 2004. Classification of Hyperspectral Remote Sensing Images With Support Vector Machines. *IEEE Trans. Geosci. Remote Sens.*, 42(8) :1778–1790.
- MENG, Ziyi, MA, Jiawei et YUAN, Xin, 2020. End-to-End Low Cost Compressive Spectral Imaging with Spatial-Spectral Self-Attention. Dans *European Conference on Computer Vision*.
URL <https://github.com/mengziyi64/TSA-Net>.
- MIAN, Ajmal et HARTLEY, Richard, 2012. Hyperspectral video restoration using optical flow and sparse coding. *Optics Express*, 20(10) :10658–10673.
- MIESCH, Christophe, POUTIER, Laurent, ACHARD, Véronique, BRIOTTET, Xavier, LENOT, Xavier et BOUCHER, Yannick, 2005. Direct and inverse radiative transfer solutions for visible and near-infrared hyperspectral imagery. *IEEE Transactions on Geoscience and Remote Sensing*, 43(7) :1552–1561. doi :10.1109/TGRS.2005.847793.
- MIURA, Tomoaki et HUETE, Alfredo R., 2009. Performance of three reflectance calibration methods for airborne hyperspectral spectrometer data. *Sensors*, 9(2) :794–813. doi :10.3390/s90200794.
- MOU, Lichao, BRUZZONE, Lorenzo et ZHU, Xiao Xiang, 2019. Learning spectral-spatial-temporal features via a recurrent convolutional neural network for change detection in multispectral imagery. *IEEE Transactions on Geoscience and Remote Sensing*, 57(2) :924–935. doi :10.1109/TGRS.2018.2863224.

BIBLIOGRAPHIE

- MOU, Lichao, GHAMISI, Pedram et ZHU, Xiao Xiang, 2017. Deep recurrent neural networks for hyperspectral image classification. *IEEE Transactions on Geoscience and Remote Sensing*, 55(7) :3639–3655. doi :10.1109/TGRS.2016.2636241.
- MUSGRAVE, Kevin, BELONGIE, Serge et LIM, Ser-Nam, 2020. A Metric Learning Reality Check. Dans *European Conference on Computer Vision*.
- NGUYEN, Hien Van, BANERJEE, Amit et CHELLAPPA, Rama, 2010. Tracking via object reflectance using a hyperspectral video camera. Dans *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition - Workshops, CVPRW 2010*, pages 44–51. ISBN 9781424470297. doi :10.1109/CVPRW.2010.5543780.
- NHAILA, Hasna, ELMAIZI, Asma, SARHROUNI, Elkebir et HAMMOUCH, Ahmed, 2019. Supervised classification methods applied to airborne hyperspectral images : Comparative study using mutual information. Dans *Procedia Computer Science*, tome 148, pages 97–106. Elsevier B.V. doi :10.1016/j.procs.2019.01.013.
- NOVIYANTO, Ary et ABDULLA, Waleed H., 2019. Segmentation and calibration of hyperspectral imaging for honey analysis. *Computers and Electronics in Agriculture*, 159 :129–139. doi :10.1016/j.compag.2019.02.006.
- OUERGHEMMI, Walid, LE BRIS, Arnaud, CHEHATA, Nesrine et MALLET, Clément, 2017. A two-step decision fusion strategy : Application to hyperspectral and multispectral images for urban classification. Dans *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives*, tome 42, pages 167–174. International Society for Photogrammetry and Remote Sensing. doi :10.5194/isprs-archives-XLII-1-W1-167-2017.
- PAOLETTI, M. E., HAUT, J. M., PLAZA, J. et PLAZA, A., 2019. Deep learning classifiers for hyperspectral imaging : A review. doi :10.1016/j.isprsjprs.2019.09.006.
- PARSONS, Mark, BRATANOV, Dmitry, GASTON, Kevin J. et GONZALEZ, Felipe, 2018. UAVs, hyperspectral remote sensing, and machine learning revolutionizing reef monitoring. *Sensors (Switzerland)*, 18(7). doi :10.3390/s18072026.
- PIETERS, Carlé M, 1983. Strength of mineral absorption features in the transmitted component of near-infrared reflected light : First results from RELAB. *Journal of Geophysical Research : Solid Earth*, 88(B11) :9534–9544. doi :10.1029/JB088iB11p09534.

- POUTIER, Laurent, MIESCH, Christophe, LENOT, Xavier, ACHARD, Véronique et BOUCHER, Yannick, 2002. COMANCHE and COCHISE : two reciprocal atmospheric codes for hyperspectral remote sensing. Dans *AVIRIS Earth Science and Applications Workshop Proceedings*.
URL <http://onera.fr/dota-en/>
- REDMON, Joseph et FARHADI, Ali, 2018. YOLOv3 : An Incremental Improvement. Rapport technique, University of Washington.
URL <http://arxiv.org/abs/1804.02767>
- ROBERTS, Dar A et HEROLD, Martin, 2004. Imaging Spectroscopy of urban materials. *Infrared Spectroscopy in Geochemistry, Exploration and Remote Sensing, Mineral Association of Canada, Short Course Series*, 33 :155–181.
URL <http://www.op.dlr.de/dais/dais-scr.htm>.
- SANDERS, Jeffrey D, OLSEN, Richard C et CLEARY, David D, 1998. *Target detection and classification at Kernel Blitz 1997 using spectral imagery*. Thèse de doctorat, US Navy.
URL <https://core.ac.uk/download/pdf/36702611.pdf>
- SCHLERF, Martin, ATZBERGER, Clement et HILL, Joachim, 2005. Remote sensing of forest biophysical variables using HyMap imaging spectrometer data. *Remote Sensing of Environment*, 95(2) :177–194. doi :10.1016/j.rse.2004.12.016.
- SCHNEIDER, Sven, MURPHY, Richard J., MONTEIRO, Sildomar T. et NETTLETON, Eric, 2009. On the development of a hyperspectral library for autonomous mining systems. Dans *Australasian Conference on Robotics and Automation (ACRA)*.
- SHAFIEE, Sahameh, POLDER, Gerrit, MINAEI, Saeid, MOGHADAM-CHARKARI, Nasrolah, VAN RUTH, Saskia et KUŚ, Piotr M., 2016. Detection of Honey Adulteration using Hyperspectral Imaging. *IFAC-PapersOnLine*, 49(16) :311–314. doi :10.1016/j.ifacol.2016.10.057.
- SHAIKH, Muhammad Saad, JAFERZADEH, Keyvan, THÖRNBERG, Benny et CASSELGREN, Johan, 2021. Calibration of a hyper-spectral imaging system using a low-cost reference. *Sensors*, 21(11). doi :10.3390/s21113738.
- SIDOROV, Oleksii et HARDEBERG, Jon, 2019. Deep Hyperspectral Prior : Single-Image Denoising, Inpainting, Super-Resolution. Dans *Proceedings of the IEEE/CVF International Conference on Computer Vision Workshops*.
URL <https://github.com/acecreamu/deep-hs-prior>
- SIMONYAN, Karen et ZISSERMAN, Andrew, 2015. Very Deep Convolutional Networks for Large-Scale Image Recognition. *arXiv preprint arXiv :1409.1556*.

BIBLIOGRAPHIE

- TANIMOTO, T, 1957. Tanimoto coefficient. Rapport technique, IBM Internal Technical Report.
- THENKABAIL, P. et ANEECE, I., 2019. Global hyperspectral imaging spectral-library of agricultural (ghisa) crops for conterminous united states v001. <https://doi.org/10.5067/Community/GHISA/GHISACONUS.001>. Dernière visite le 30/03/2023.
- THOMAS, Colin, DOZ, Stéphanie, BRIOTTET, Xavier et LACHÉRADE, Sophie, 2011. AMARTIS v2 : 3D radiative transfer code in the [0.4 ; 2.5 μm] spectral domain dedicated to urban areas. *Remote Sensing*, 3(9) :1914–1942. doi : 10.3390/rs3091914.
- UEZATO, Tatsumi, HONG, Danfeng, YOKOYA, Naoto et HE, Wei, 2020a. Guided Deep Decoder : Unsupervised Image Pair Fusion. Dans *European Conference on Computer Vision*.
- UEZATO, Tatsumi, YOKOYA, Naoto et HE, Wei, 2020b. Illumination invariant hyperspectral image unmixing based on a digital surface model. *IEEE Transactions on Image Processing*, 29 :3652–3664. doi :10.1109/TIP.2020.2963961.
URL <http://arxiv.org/abs/2007.11770><http://dx.doi.org/10.1109/TIP.2020.2963961>
- UZAIR, Muhammad, MAHMOOD, Arif et MIAN, Ajmal, 2014. Hyperspectral Face Recognition using 3D-DCT and Partial Least Squares. Dans *BMVC*, pages 1–57. British Machine Vision Association and Society for Pattern Recognition. doi :10.5244/c.27.57.
- UZAIR, Muhammad, MAHMOOD, Arif et MIAN, Ajmal, 2015. Hyperspectral face recognition with spatio-spectral information fusion and PLS regression. *IEEE Transactions on Image Processing*, 24(3) :1127–1137. doi :10.1109/TIP.2015.2393057.
- VEMURI, Anant S, WIRKERT, Sebastian et MAIER-HEIN, Lena, 2019. Hyperspectral Camera Selection for Interventional Health-care. *arXiv preprint arXiv :1904.02709*, pages 1–13.
- VILLA, A., CHANUSSOT, J., BENEDIKTSSON, J.A., JUTTEN, C. et DAMBREVILLE, R., 2013. Unsupervised methods for the classification of hyperspectral images with low spatial resolution. *Pattern Recognition*, 46(6) :1556–1568. doi :10.1016/j.patcog.2012.10.030.
URL <https://linkinghub.elsevier.com/retrieve/pii/S0031320312004967>

- WENDEL, Alexander et UNDERWOOD, James, 2017. Illumination compensation in ground based hyperspectral imaging. *ISPRS Journal of Photogrammetry and Remote Sensing*, 129 :162–178. doi :10.1016/j.isprsjprs.2017.04.010.
- WU, Hao et PRASAD, Saurabh, 2017. Convolutional recurrent neural networks for hyperspectral data classification. *Remote Sensing*, 9(3). doi :10.3390/rs9030298.
- XIONG, Fengchao, ZHOU, Jun, CHANUSSOT, Jocelyn et QIAN, Yuntao, 2019. Dynamic Material-Aware Object Tracking in Hyperspectral Videos. Dans *2019 10th Workshop on Hyperspectral Imaging and Signal Processing : Evolution in Remote Sensing (WHISPERS)*.
URL <https://ieeexplore.ieee.org/abstract/document/8921176>
- XIONG, Fengchao, ZHOU, Jun et QIAN, Yuntao, 2020. Material Based Object Tracking in Hyperspectral Videos : Benchmark and Algorithms. *IEEE Transactions on Image Processing*, 29 :3719–3733.
URL <https://ieeexplore.ieee.org/document/8960632>
- YAN, Qiong, XU, Li, SHI, Jianping et JIA, Jiaya, 2013. Hierarchical saliency detection. Dans *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*.
- YAO, Jing, HONG, Danfeng, CHANUSSOT, Jocelyn, MENG, Deyu, ZHU, Xiaoxiang et XU, Zongben, 2020. Cross-Attention in Coupled Unmixing Nets for Unsupervised Hyperspectral Super-Resolution. Dans *European Conference on Computer Vision*.
URL https://github.com/danfenghong/ECCV2020_CUCaNet.
- YASUMA, Fumihito, MITSUNAGA, Tomoo, ISO, Daisuke et NAYAR, Shree K, 2010. Generalized Assorted Pixel Camera : Postcapture Control of Resolution, Dynamic Range, and Spectrum. *IEEE Transactions on Image Processing*, 19(9) :2241–2253. doi :10.1109/TIP.2010.2046811.
URL <http://ieeexplore.ieee.org/document/5439932/>
- YUEN, P. W.T. et RICHARDSON, M., 2010. An introduction to hyperspectral imaging and its application for security, surveillance and target acquisition. *Imaging Science Journal*, 58(5) :241–253. doi :10.1179/174313110X12771950995716.
- YUHAS, Roberta H, GOETZ, Alexander F H et BOARDMAN, Joe W, 1992. Discrimination among semi-arid landscape endmembers using the Spectral Angle Mapper (SAM) algorithm. Dans *JPL, Summaries of the Third Annual JPL Airborne Geoscience Workshop. Volume 1 : AVIRIS Workshop*.
URL <https://ntrs.nasa.gov/search.jsp?R=19940012238>

- ZAMAN, Bushra, JENSEN, Austin, CLEMENS, Shannon R. et MCKEE, Mac, 2014. Retrieval of spectral reflectance of high resolution multispectral imagery acquired with an autonomous unmanned aerial vehicle : Aggieair™. *Photogrammetric Engineering and Remote Sensing*, 80(12) :1139–1150. doi : 10.14358/pers.80.12.1139.
- ZARE, Alina et HO, K. C., 2014. Endmember variability in hyperspectral analysis : Addressing spectral variability during spectral unmixing. *IEEE Signal Processing Magazine*, 31(1) :95–104. doi :10.1109/MSP.2013.2279177.
- ZHAO, Min, CHEN, Jie et HE, Zhe, 2019. A laboratory-created dataset with ground-truth for hyperspectral unmixing evaluation. *IEEE Journal of Earth Observations and Remote Sensing*.
URL <http://arxiv.org/abs/1902.08347>
- ZHAO, Wenzhi et DU, Shihong, 2016. Spectral-Spatial Feature Extraction for Hyperspectral Image Classification : A Dimension Reduction and Deep Learning Approach. *IEEE Transactions on Geoscience and Remote Sensing*, 54(8) :4544–4554. doi :10.1109/TGRS.2016.2543748.
- ZHU, Lin, CHEN, Yushi, GHAMISI, Pedram et BENEDIKTSSON, Jón Atli, 2018. Generative Adversarial Networks for Hyperspectral Image Classification. *IEEE Transactions on Geoscience and Remote Sensing*, 56(9) :5046–5063. doi :10.1109/TGRS.2018.2805286.