



HAL
open science

Estimation par stéréovision multimodale de caractéristiques géométriques d'un feu de végétation en propagation

Tom Toulouse

► To cite this version:

Tom Toulouse. Estimation par stéréovision multimodale de caractéristiques géométriques d'un feu de végétation en propagation. Modélisation et simulation. Université Pascal Paoli; Université Laval (Québec, Canada), 2015. Français. NNT : 2015CORT0009 . tel-01239762v2

HAL Id: tel-01239762

<https://theses.hal.science/tel-01239762v2>

Submitted on 29 Nov 2016

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

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



UNIVERSITÉ DE CORSE – PASQUALE PAOLI
ÉCOLE DOCTORALE ENVIRONNEMENT ET SOCIÉTÉ
UNIVERSITÉ LAVAL – (QUÉBEC)



Thèse en Convention de Cotutelle Internationale
présentée pour l'obtention du grade de

DOCTEUR EN AUTOMATIQUE, SIGNAL,
PRODUCTIQUE, ROBOTIQUE
Mention : Traitement du signal

Soutenue publiquement par

Tom TOULOUSE

le 13 novembre 2015

Estimation par stéréovision multimodale de caractéristiques
géométriques d'un feu de végétation en propagation

Directeurs :

Xavier MALDAGUE,

Professeur Université Laval

Lucile ROSSI,

Dr-HDR Université de Corse

Rapporteurs :

Guillaume-Alexandre BILODEAU,

Professeur École polytechnique de Montréal

Yvon VOISIN,

Professeur Université de Bourgogne

Jury :

Moulay AKHLOUFI,

Professeur UTFSM, Valparaso, Chili

Guillaume-Alexandre BILODEAU,

Professeur École polytechnique de Montréal

Xavier MALDAGUE,

Professeur Université Laval

Lucile ROSSI,

Dr-HDR Université de Corse

Paul-Antoine SANTONI,

Professeur Université de Corse

Yvon VOISIN,

Professeur Université de Bourgogne

Estimation par stéréovision multimodale de caractéristiques géométriques d'un feu de végétation en propagation

**Thèse en cotutelle
Doctorat en Génie Électrique**

Tom TOULOUSE

Université Laval
Québec, Canada
Philosophiæ doctor (Ph.D.)

et

Université de Corse Pasquale Paoli
Corte, France
Docteur en AUTOMATIQUE, SIGNAL,
PRODUCTIQUE, ROBOTIQUE

Résumé

Les travaux menés dans cette thèse concernent le développement d'un dispositif de vision permettant l'estimation de caractéristiques géométriques d'un feu de végétation en propagation. Ce dispositif est composé de plusieurs systèmes de stéréovision multimodaux générant des paires d'images stéréoscopiques à partir desquelles des points tridimensionnels sont calculés et les caractéristiques géométriques de feu tels que sa position, vitesse, hauteur, profondeur, inclinaison, surface et volume sont estimées.

La première contribution importante de cette thèse est la détection de pixels de feu de végétation. Tous les algorithmes de détection de pixels de feu de la littérature ainsi que ceux développés dans le cadre de cette thèse ont été évalués sur une base de 500 images de feux de végétation acquises dans le domaine du visible et caractérisées en fonction des propriétés du feu dans l'image (couleur, fumée, luminosité). Cinq algorithmes de détection de pixels de feu de végétation basés sur la fusion de données issues d'images acquises dans le domaine du visible et du proche-infrarouge ont également été développés et évalués sur une autre base de données composée de 100 images multimodales caractérisées.

La deuxième contribution importante de cette thèse concerne l'utilisation de méthodes de fusion d'images pour l'optimisation des points appariés entre les images multimodales stéréoscopiques.

La troisième contribution importante de cette thèse est l'estimation des caractéristiques géométriques de feu à partir de points tridimensionnels obtenus depuis plusieurs paires d'images stéréoscopiques et recalés à l'aide de relevés GPS et d'inclinaison de tous les dispositifs de vision.

Le dispositif d'estimation de caractéristiques géométriques à partir de systèmes de stéréovision a été évalué sur des objets rigides de dimensions connues et a permis d'obtenir les informations souhaitées avec une bonne précision. Les résultats des données obtenues pour des feux de végétation en propagation sont aussi présentés.

Mots clefs : Feux de forêt, stéréovision, traitement d'images, segmentation, multimodal

Abstract

This thesis presents the geometrical characteristics measurement of spreading vegetation fires with multimodal stereovision systems. Image processing and 3D registration are used in order to obtain a three-dimensional modeling of the fire at each instant of image acquisition and then to compute fire front characteristics like its position, its rate of spread, its height, its width, its inclination, its surface and its volume.

The first important contribution of this thesis is the fire pixel detection. A benchmark of fire pixel detection algorithms of the literature and of those that are developed in this thesis have been on a database of 500 vegetation fire images of the visible spectra which have been characterized according to the fire properties in the image (color, smoke, luminosity). Five fire pixel detection algorithms based on fusion of data from visible and near-infrared spectra images have also been developed and tested on another database of 100 multimodal images.

The second important contribution of this thesis is about the use of images fusion for the optimization of the matching point's number between the multimodal stereo images.

The second important contribution of this thesis is the registration method of 3D fire points obtained with stereovision systems. It uses information collected from a housing containing a GPS and an IMU card which is positioned on each stereovision systems. With this registration, a method have been developed to extract the geometrical characteristics when the fire is spreading.

The geometrical characteristics estimation device have been evaluated on a car of known dimensions and the results obtained confirm the good accuracy of the device. The results obtained from vegetation fires are also presented.

Key words: wildland fire, stereovision, image processing segmentation, multimodal

Stimazione

I fogu di vegetazione ripresentani un risicu maioru pa numerosi paesi ogni anu, piciani à pocu pressu un milione di ettari in Europa é dui milione é mezzu di ettari a u Canada. Pa parà issu risicu, sò rializati azzione di ghjestionu, di privenzione é di lotta. L'efficacità di issi tre tipi d'azzione hè liatu à a cunniscenza di i fenomeni chi intervenani mentre a prupagazione di i fochi é a u migliuramentu di i modeli di cumprutamenti. Pa avè una cunniscenza piu fina di i fenomeni, pudè cumparà i dati sperimentali incu i dati ottenuti da i mudeli matematiche di prupagazione é à a fina ripiglià issi dati cume valori di intrata indè i mudeli di cumpurtamenti di fochi, hè primura di fà misure di e caratteristiche geometriche di u focu in core di à so prupagazione.

Dipoi u 2007, i tacchi di ricerca sont rializati in l'UMR CNRS 6134 di l'Università di Corsica pa sviluppà i strumenti di misura dedicati à i fochi di foresta, basati nanta u tratamentu di l'immagine é a visione. Belli avanzati sò stati fati cuncernandu u sviluppu di tal stuvigli pa i fochi in prupagazione in laboratoriu o fora nantu parcelle quadrati di 10 m. Pa i scali piu maiori, a prisenzia di fumi opachi chi podani piatà u focu é i distanze maiore chi pudani esse francatu, sò problematiche da risolve.

In lu casu di issa tesa, hè statu sviluppatu un dispositivu aduprandu a visione pìrmitandu di misurà e caratteristiche geometriche di un focu di vegetazione chi si prupaga à u nuvellu di una parcella. Hè cumpostu di parechji sistemi di "stereovision multimodaux" partendu di i quali sò stimati l'informazione complementari di u focu in tre dimensione. L'adattamento di issi dati permette d'ottene una modelisazione in tre dimesione di u fronte di u focu à parte si di quale sò stimati l'infurmazione tale chi à so posizione, à so vitezza, à so altezza, à so prufundezza, à so inclinazione, à so area è u so volume.

I travagli di ricerca si sò cuncentrati à nantu à :- l'apprezzazione è a creazione di nuveli "algorithme" pa rimarcà i "pixel" cuncernendu à parte di u focu indè l'immagine presi indè u spettro di u videbule, di u "proche-infrarouge" è di u "multimodale". Pa risalizà quistu, sò stati costruiti due base d'immagine di fochi di vegetazione, una cumposta di pliu di 500 immagine indè u videbule, è quidda altra di 100 immagine presi in modu simultaneu indè u videbule è u "proche-infrarouge",- u megliuramentu di a quantità di i punti rimarcati è appariati à parte si d'immagine "stereoscopiques multimodales" ;- l'adattamento di l'informazione in tredimensione ottenuti da parte di ogni sistema di "stereovision" incù u scopu di d'ottene una forma di focu glubale.

Table des matières

Résumé	v
Abstract	vii
Stimazione	ix
Table des matières	xi
Liste des tableaux	xv
Liste des figures	xvii
Remerciements	xxiii
Introduction	1
1 État de l’art de la vision pour la mesure des feux	5
1.1 Introduction	5
1.2 Méthodes de détection de pixels de feu	6
1.2.1 Méthodes de détection de pixels de feu basées sur l’intensité de pixels dans le spectre visible	6
1.2.2 Méthodes de détection de pixels de feu basées sur l’intensité de pixels dans le spectre infrarouge	15
1.2.3 Méthodes de détection de pixels de feu basées sur la fusion d’informations d’intensité de pixels obtenues dans différents spectres	24
1.2.4 Méthodes de détection de pixels de feu utilisant des critères complémentaires à l’intensité de pixels	24
1.2.5 Métriques d’évaluation des performances des méthodes de détection de pixels de feu	29
1.2.6 Évaluations comparatives des performances de certains algorithmes de détection de pixels de feu	31
1.3 Techniques de mesure pour les feux	32
1.3.1 Caractéristiques géométriques d’un front de feu	32
1.3.2 Mesures estimées à partir d’images obtenues depuis un même point de vue	33
1.3.3 Mesures estimées à partir d’images obtenues depuis plusieurs points de vue	34
1.3.4 Mesures par stéréovision	36

1.4	Conclusion	38
2	Architecture de la solution proposée	39
2.1	Introduction	39
2.2	Bandes spectrales sélectionnées pour l'observation des émissions électromagnétiques des feux de végétation	40
2.2.1	Rayonnements électromagnétiques des feux de végétation	40
2.2.2	Comparaison d'images de feux de végétation acquises dans différentes bandes spectrales	44
2.3	Choix d'un dispositif de vision et de ses caractéristiques	50
2.3.1	Choix de la focale de l'objectif	51
2.3.2	Choix de la distance entre les deux caméras du système de stéréovision	51
2.4	Solution proposée	52
2.4.1	Matériels composant le dispositif	52
2.4.2	Description du processus de traitement	53
2.5	Conclusion	55
3	Détection des pixels de couleur feu dans les images du spectre visible	57
3.1	Introduction	57
3.2	Présentation de la base de données d'images intitulée « Corsican Fire Database »	58
3.2.1	Images de la base	59
3.2.2	Caractérisation des images	59
3.2.3	Catégorisation des images de la base	67
3.2.4	Création d'un ensemble de pixels d'apprentissage	69
3.2.5	Publication de la base d'images sur Internet	71
3.3	Adaptation des méthodes de la littérature	72
3.3.1	Règles de détection	72
3.3.2	Méthodes basées sur un apprentissage	79
3.4	Nouvelles méthodes proposées	82
3.4.1	Nouvelles règles	83
3.4.2	Nouvelles méthodes	83
3.5	Évaluation des performances des méthodes	88
3.5.1	Évaluation des performances des règles	88
3.5.2	Évaluation des performances des différentes versions de la méthode de détection basée sur un apprentissage par histogramme gaussien lissé	92
3.5.3	Évaluation des performances de la nouvelle méthode basée sur un apprentissage par machines à vecteurs de support	93
3.5.4	Comparaison des performances des méthodes	95
3.6	Conclusion	99
4	Détection multimodale de pixels	101
4.1	Introduction	101
4.2	Base d'images multimodales	102
4.2.1	Triplet d'images	103
4.2.2	Recalage des images	103
4.2.3	Caractérisation des images	106
4.2.4	Création d'un ensemble d'images d'apprentissage	106

4.2.5	Synthèse d'images avec différents types de fumée	107
4.3	Fusion de données : principes théoriques	109
4.3.1	Méthode statistique	110
4.3.2	Méthode du Modèle des Croyances Transférables (MCT)	113
4.4	Méthodes de fusion appliquées à la détection de pixels « feu »	115
4.4.1	Probabilités conditionnelles	116
4.4.2	Détection multimodale par fusion de probabilités	123
4.4.3	Détection multimodale par régression logistique	130
4.5	Évaluation des performances des algorithmes de détection de pixels « feu » et comparaison de leurs résultats	134
4.5.1	Évaluation des performances des algorithmes de détection de pixels « feu » sur la base d'images multimodales et comparaison de leurs résultats	135
4.5.2	Évaluation de la robustesse aux fumées des algorithmes de détection travaillant sur des images du visible et multimodales	138
4.6	Conclusion	141
5	Optimisation par fusion d'images multimodales du nombre de points appariés	143
5.1	Introduction	143
5.2	État de l'art de la fusion d'images	144
5.2.1	Approches classiques	145
5.2.2	Approches multirésolutions	146
5.3	Appariement d'images multimodales	155
5.3.1	Détection et appariement de points d'intérêt	156
5.3.2	Fusion de données d'images pour l'appariement	161
5.3.3	Concaténation de correspondances	163
5.4	Évaluation de l'efficacité d'appariement de points d'intérêt en fonction du type d'image	163
5.5	Conclusion	172
6	Estimation de caractéristiques géométriques de fronts de feu à partir d'images stéréoscopiques	173
6.1	Introduction	173
6.2	Principes théoriques de la stéréovision	174
6.2.1	Modèle géométrique de la caméra	174
6.2.2	Estimation des paramètres intrinsèques d'une caméra	176
6.2.3	Système stéréoscopique	179
6.2.4	Estimation de la matrice essentielle	182
6.2.5	Reconstruction 3D des points appariés	184
6.3	Estimation des paramètres intrinsèques et extrinsèques d'un système de vision	185
6.4	Reconstruction tridimensionnelle de points de feu	188
6.5	Estimation des caractéristiques géométriques d'un feu	188
6.5.1	Estimation du plan de base	190
6.5.2	Calcul des caractéristiques géométriques	191
6.6	Conclusion	198

7 Estimation de caractéristiques géométriques de fronts de feux à partir de plusieurs systèmes de stéréovision	199
7.1 Introduction	199
7.2 Recalage de points tridimensionnels obtenus depuis différents systèmes de vision	200
7.2.1 Matériel utilisé	200
7.2.2 Détermination de la matrice de passage entre les repères des différents systèmes de stéréovision	202
7.3 Reconstruction tridimensionnelle de la forme de feu	208
7.4 Calcul de caractéristiques géométriques	210
7.4.1 Périmètre du feu	211
7.4.2 Surface du végétal participant à la combustion	211
7.4.3 Direction principale de propagation	212
7.4.4 Lignes de front	212
7.4.5 Epaisseur	212
7.4.6 Volume	214
7.5 Incertitude de mesure	214
7.6 Conclusion	216
8 Suivi d'une propagation de feu avec le dispositif de vision	219
8.1 Introduction	219
8.2 Suivi d'une propagation de feu réalisée sur un petit terrain plat	219
8.2.1 Description de l'expérience	219
8.2.2 Résultats obtenus et commentaires	220
8.3 Suivi de brûlages dirigés	220
8.4 Conclusion	226
Conclusion et perspectives	227
Publications et distinctions	229
Bibliographie	231

Liste des tableaux

1.1	Espaces de couleur des méthodes présentées.	7
1.2	Caractéristiques de couleur des flammes dans le <i>TSI</i>	11
1.3	Méthodes de seuillage utilisées pour la détection de pixels « feu » dans le spectre infrarouge.	15
1.4	Indice de qualité de Hafiane obtenu par cinq méthodes de détection de pixels de feu évaluées sur des bases d'images différentes.	31
2.1	Rapport entre des radiations émises par un feu, à différentes températures et dans différentes bandes spectrales, et celles émises par un objet à une température de 30°C.	44
2.2	Calcul par modèle statistique de la transmittance atmosphérique d'un corps noir à 1500 K à une distance de 100 m	44
2.3	Valeur de l'erreur théorique en z (Δz) en fonction des deux focales choisies et de la distance du feu par rapport à la caméra	52
3.1	Répartition par catégorie des images de la base CFDB.	68
3.2	Distribution des pixels de la base CFDB par catégorie.	69
3.3	Valeurs minimale et maximale des canaux des différents espaces de couleurs utilisés.	80
3.4	Scores des vrais positifs et vrais négatifs pour les 34 règles de détection	90
3.5	Performance des différentes versions de la méthode de détection par histogramme gaussien lissé	93
3.6	Statistiques des caractéristiques d'images des neuf séquences utilisées	94
3.7	Comparaison de la méthode SVM sur deux espaces de couleurs sur neuf séquences d'images de propagation	95
3.8	Valeurs de F-Score des méthodes de détection testées sur la base d'images	97
4.1	Répartition par catégorie des images de la base multimodale	107
4.2	Valeurs de F-score obtenues par les différentes méthodes de seuillage	118
4.3	Valeurs moyenne de F-score obtenues par la méthode de détection par apprentissage pour différentes tailles d'image.	122
4.4	Noms associés aux classes d'information et de source.	126
4.5	Valeurs des $p(w_j d_{s,i})$ choisies pour la détection de pixels de classe « feu ».	126
4.6	Formule de calcul de $p(w_j \mathbf{x}_s)$ en fonction des sources et des classes.	127
4.7	Valeurs de F-Score des méthodes de détection testées sur la base d'images multimodales	137
4.8	Ratio des scores en fonction de la transparence de fumée pour les algorithmes de détection travaillant sur des images du visible et multimodales	140

4.9	Ratio des scores des algorithmes de détection multimodaux avec évaluation préalable du taux de superposition de fumée en fonction de la transparence de fumée.	141
5.1	Pourcentage de pixels de feux appariés et temps d'exécution moyens pour les paires d'images du spectre visible et proche-infrarouge	164
5.2	Pourcentages d'appariement pour les images fusionnées	166
5.3	Pourcentages de pixels de feux appariés par catégorie d'image obtenus pour des images fusionnées par la méthode F_1 et pour des images couleur du spectre du visible (V).	167
5.4	Pourcentage d'appariements pour les concaténations d'appariement	168
5.5	Pourcentage de pixels de feu appariés par catégorie pour des concaténations particulières	170
5.6	Pourcentage de pixels appariés pour la concaténation des points obtenus par fusion	171
6.1	Focale et distance inter-caméra mesurées et estimées	187
6.2	Erreurs de reconstruction à différentes distances	188
7.1	Comparaison des mesures réelles et estimées	216
8.1	Caractéristiques estimées du feu présenté dans la figure 8.5	224
8.2	Caractéristiques estimées du feu présenté dans la figure 8.9	224

Liste des figures

1.1	Exemple de probabilité d'un pixel d'appartenir au feu en fonction de ces valeurs I_{L^*} et I_{a^*}	12
1.2	Superficie de décision du système flou de Martínez.	21
1.3	Schéma explicatif de la procédure permettant l'obtention de la fonction f_1	21
1.4	Schéma explicatif de la méthode de seuillage multi-résolution de Martínez.	21
1.5	Comparaison par Martínez de Dios <i>et al.</i> de l'efficacité des algorithmes de seuillage de la littérature sur une image infrarouge de feu	22
1.6	Classification des zones de feu par Meléndez <i>et al.</i>	25
1.7	Détection multimodale de zones de feu par ACP	25
1.8	Extraction d'arrière-plan grâce aux caractéristiques dynamiques	26
1.9	Distribution de fréquences de vacillement dans les spectres visible et infrarouge	27
1.10	Segmentation dynamique de flamme	28
1.11	Caractéristiques géométriques principales d'un front de feu	32
1.12	Calculs de paramètres sur une image multispectrale	34
1.13	Caractéristiques géométriques de fronts de feux obtenues par Martínez-De Dios <i>et al.</i>	35
1.14	Caractéristiques obtenues par Martínez-De Dios <i>et al.</i>	36
1.15	Estimation de la vitesse d'un front de flamme	37
1.16	Surface d'un front calculé avec deux systèmes de stéréovision	37
2.1	Sensibilité spectrale de la caméra JAI AD080 GE dans le spectre proche-infrarouge	40
2.2	Mesure des intensités des émissions pour des feux réalisés en laboratoire	41
2.3	Spectres de flamme mesurés lors d'essais en tunnel à feux pour des flammes d'épaisseur de cinquante centimètres à quatre mètres	42
2.4	Transmittance atmosphérique dans le spectre infrarouge	43
2.5	Schéma du montage de la caméra JAI AD-080GE	45
2.6	Caméra DFK 72BUC02 (à gauche) et caméra IR-TCM 384 (à droite) fixées côte à côte sur un axe	46
2.7	Images de feu simultanément acquises dans les domaines du visible et de l'infrarouge à grandes longueurs d'onde	47
2.8	Images de feu simultanément acquises dans les domaines du visible et de l'infrarouge à longueurs d'onde moyennes	47
2.9	Caméra AD-080GE (à gauche) et caméra Jade3MW fixées côte à côte sur un axe	48
2.10	Images de feu simultanément acquises dans les domaines du visible, du proche-infrarouge et de l'infrarouge à longueurs d'onde moyennes avec filtre à $4,2 \mu m$	49
2.11	Images de feu avec fumée simultanément acquises dans les domaines du visible et du proche-infrarouge	49

2.12	Fumée masquant le feu dans une image le spectre proche-infrarouge	50
2.13	Système de stéréovision proposé	52
2.14	Exemple de positionnements de deux systèmes de stéréovision	53
2.15	Schéma fonctionnel du système proposé.	54
3.1	Sélection d'images de feux contenues dans la base CFDB.	60
3.2	Exemple de vérité terrain	61
3.3	Couleurs dans l'espace de couleurs <i>TSI</i>	62
3.4	Classement par couleur des pixels d'une image de feu	63
3.5	Exemple d'images d'un feu présentant un haut niveau de texture et un faible niveau de texture	64
3.6	Exemple d'image de feu sans fumée et avec fumée	65
3.7	Caractérisation automatique de la présence de fumées	66
3.8	Images de feu avec différents niveaux de luminosité de l'environnement	68
3.9	Ensemble des pixels sélectionnés pour la base d'apprentissage	71
3.10	Nombre de pixels « feu » de la base d'apprentissage distribués en fonction de l'intensité dans les canaux <i>T</i> , <i>S</i> , et <i>I</i> des pixels /classefeu d'apprentissage : (a) histogramme dans le canal <i>T</i> , (b) histogramme dans le canal <i>S</i> et (c) histogramme dans le canal <i>I</i>	76
3.11	Valeurs de C_b et C_r des pixels « feu » de la base d'apprentissage.	77
3.12	Exemple de régions d'intérêt sélectionnés dans une image de feu	82
3.13	Exemple de détection de la méthode SVM	82
3.14	Exemples de résultat de détection de pixels obtenus par les règles proposées	84
3.15	Exemples de résultat de détection de pixels obtenus par la méthode par vote	85
3.16	Schéma fonctionnel de la régression logistique.	86
3.17	Exemples de résultat de détection de pixels obtenus par les deux versions de la méthode par régression logistique	88
4.1	Exemple d'images multimodales	103
4.2	Transformation directe d'image par étirement	105
4.3	Transformation inverse d'image par étirement	105
4.4	Exemple de feu masqué par de la fumée dans des conditions réelles	107
4.5	Images de fumées utilisées pour construire les images de feu avec fumée	108
4.6	Exemple de feu masqué synthétiquement par de la fumée	109
4.7	Principe général de la fusion de résultats pour un pixel \mathbf{x} ayant des données obtenues à partir de n sources.	110
4.8	Schéma fonctionnel de la fusion de résultats par méthode statistique.	111
4.9	Probabilité conditionnelle de la classe « feu » sachant la valeur de pixel dans le spectre proche-infrarouge pour un seuil $t = 230$	117
4.10	Résultats de détection de pixels « feu » dans des images du proche-infrarouge	118
4.11	Schéma fonctionnel de la méthode pour calculer les probabilités conditionnelles associées au spectre visible.	120
4.12	Probabilité conditionnelle $p(\mathcal{F} \mathbf{x}_2)$ obtenue à partir des pixels contenus dans la boîte englobante présentée dans la figure figure 4.11	121
4.13	Classement de pixels pour l'image couleur présentée dans la figure 4.11	122
4.14	Détection multimodale par moyenne pour les images de la figure 4.1	125
4.15	Détection de pixels « feu » par méthode statistique	128
4.16	Détection multimodale par MCT.	130

4.17	Schéma fonctionnel de l'architecture 1 de régression logistique.	132
4.18	Détection multimodale par Régression Logistique (Architecture 1).	133
4.19	Schéma fonctionnel de l'architecture 2.	134
4.20	Détection multimodale par Régression Logistique (Architecture 1).	135
5.1	Schéma des opérations réalisées pour la fusion multirésolution de deux images.	147
5.2	Représentation de l'algorithme de la pyramide laplacienne	149
5.3	Exemple de détails laplaciens sur 4 niveaux pour l'image de Lenna	150
5.4	Exemple de détails de ratios sur 4 niveaux pour l'image de Lenna	151
5.5	Exemple de détails de contraste sur 4 niveaux pour l'image de Lenna	151
5.6	Schéma de la décomposition en ondelettes discrète sur deux dimensions.	152
5.7	Exemple d'une image décomposée par ondelettes	153
5.8	Exemple de détails horizontaux du premier niveau de décomposition en onde- lettes de l'image de Lenna	154
5.9	Schéma de la transformation inverse en ondelettes discrètes sur deux dimensions.	155
5.10	Détection de points d'intérêt Harris sur une image de feu en niveau de gris acquise dans le domaine du visible	158
5.11	Exemple de descripteur pour un point d'intérêt particulier.	159
5.12	Recherche de point autour d'une ligne épipolaire	160
5.13	Exemple de points appariés dans une image de feu du domaine du visible	160
5.14	Exemples de fusion d'images de feu	162
5.15	Concaténation des correspondances de points pour une image de feu	171
6.1	Modèle géométrique d'une caméra	174
6.2	Système de stéréovision	180
6.3	Géométrie épipolaire	181
6.4	Mires de calibrage utilisées dans la littérature	185
6.5	Projection canonique d'une image de calibrage	187
6.6	Mur de la salle de métrologie	187
6.7	Reconstruction tridimensionnelle des points du feu	189
6.8	Vue avant et arrière d'une propagation	190
6.9	Points appariés d'un feu détectés comme étant au sol	191
6.10	Estimation d'un plan de propagation à partir des points 3D bas d'un feu	192
6.11	Position du front de feu sur le plan de base	192
6.12	Estimation des vitesses de trois points au sol.	193
6.13	Estimation de la direction de propagation	194
6.14	Estimation automatique de la hauteur d'un front de feu	195
6.15	Estimation automatique de la largeur d'un front de feu	195
6.16	Estimation automatique de la longueur et de l'inclinaison d'un front de feu	196
6.17	Interface graphique permettant la saisie de points nécessaires au calcul de la hauteur, longueur et inclinaison d'un front de feu	196
6.18	Triangulation de Delaunay appliquée aux points du feu permettant la détermi- nation du profil de hauteur du front	197
6.19	Estimation de la surface d'un front de feu.	198
7.1	Positionnement de la carte de navigation de l'AR Parrot 2.0	201
7.2	Angles de roulis, de tangage et de lacet d'une caméra	201
7.3	Exemple de données acquises par la carte de navigation sur 120 secondes	202

7.4	Repères de référence	204
7.5	Reprojection de points 3D	207
7.6	Reprojection des points après optimisation	207
7.7	Interface proposée pour la correction manuelle d'angles.	208
7.8	Schéma des étapes de la reconstruction des ponts du feu	209
7.9	Forme tridimensionnelle d'un feu obtenue par triangulation de Delaunay	210
7.10	Reconstruction complète du feu	210
7.11	Périmètre du feu	211
7.12	Direction de propagation du feu	213
7.13	Position du front de feu au sol	213
7.14	Tétraèdre quelconque.	214
7.15	Véhicule utilisé pour l'évaluation de l'incertitude des mesures	215
7.16	Configurations de prise d'images	215
7.17	Forme tridimensionnelle obtenue depuis les images de quatre systèmes de stéréovision positionnés autour d'une voiture	216
8.1	Configuration des l'expérience	220
8.2	Estimation des paramètres géométriques d'un feu au cours de sa propagation	221
8.3	Évolution temporelle du périmètre du feu.	222
8.4	Brûlage dirigé réalisé sur la commune de Corte	223
8.5	Image multimodale obtenue lors du premier brûlage dirigé	224
8.6	Brûlage dirigé réalisé sur la commune de Piedicorte	225
8.7	Images des spectres visible et proche-infrarouge du second brûlage	225
8.8	Brûlage dirigé réalisé sur la commune de Solaro	226
8.9	Image multimodale obtenue lors du troisième brûlage dirigé	226

Qui veut élever en un instant une
flamme puissante commence par
l'allumer avec de faibles brins de
paille.

William Shakespeare
Jules César (1599)

Remerciements

Les travaux présentés dans ce document sont le fruit de trois ans de recherche entre la Corse et le Québec. J'ai eu la chance d'être entouré de personnes qui m'ont beaucoup apporté et aidé durant ces années et je tiens ici à les remercier.

En premier lieu, je remercie chaleureusement ma directrice de thèse Lucile Rossi, sans qui cette thèse n'aurait pas vu le jour. Après m'avoir accueilli dans son équipe, elle a toujours su m'orienter efficacement dans mes recherches. Merci pour son investissement, sa bienveillance, son soutien et pour sa relecture minutieuse de ce manuscrit.

Je remercie mon directeur de thèse de l'Université Laval Xavier Maldague pour ses encouragements et son aide au sein de l'Université Laval. Un grand merci aussi à mon co-directeur Moulay Akhloufi pour son accueil au CRVI de Lévis, où j'ai pu travailler dans d'excellentes conditions lors de mes séjours à Québec. Je le remercie aussi pour ses conseils pertinents tout au long de cette thèse.

Merci à Guillaume-Alexandre Bilodeau et Yvon Voisin d'avoir accepté d'être rapporteurs de cette thèse. Je remercie aussi Paul-Antoine Santoni de faire partie du jury ainsi que pour son accueil au sein de l'équipe feu de l'Université de Corse.

Je souhaite aussi remercier Turgay Celik de l'Université de Witwatersrand à Johannesburg pour ses conseils sur les travaux de cette thèse.

Je remercie les personnes de Corse qui ont contribué au bon déroulement de cette thèse en commençant par Antoine Pieri pour sa grande aide lors des expériences et pour la construction du dispositif ainsi qu'à Yves Tison pour la conception des parties électroniques du dispositif. Je remercie également Fred Morandini pour son aide à l'utilisation de la caméra infrarouge. Merci aussi à Augustin et Iratxe qui ont travaillé dans notre équipe temporairement pour leur aide lors des expériences et pour leur bonne humeur. Je remercie Antoine Campana pour son travail sur la mise en ligne de la base d'image. Merci également aux stagiaires Mouhamadou, Yacine et Youssef que j'ai eu la chance d'encadrer ainsi qu'à Alexandre pour son aide dans nos expériences. Un grand merci à Joseph Chatelon et Thierry Marcelli pour leur aide dans la préparation de la soutenance orale. Je remercie aussi tous les doctorants de l'Université de Corse (dont certains sont aujourd'hui docteurs) avec lesquels j'ai sympathisé durant ces

années : Andreï, Christelle, Gauthier, Jean-Baptiste, Laetitia, Louis-Félix, Marie, Mathilde, Mohamed, Melissa, Raphaël ainsi que Wani avec qui j'ai pu échanger de précieux conseils sur L^AT_EX.

Je remercie également l'Unité d'Instruction et d'Intervention de la Sécurité Civile N°5 de Corte et les forestiers sapeurs de Corse pour avoir rendu possible les expériences de feu réalisées au cours de cette thèse et pour avoir assuré notre sécurité durant ces brûlages. Merci aussi à tous nos partenaires nous ayant transmis des images de feu de végétation pour construire la base.

Je tiens aussi à remercier les personnes de l'Université Laval. Merci à Jean-François Lalonde pour ses cours captivants qui m'ont aidé pour certains aspects de cette thèse. Je remercie Julien Fleuret et Denis Ouellet pour leurs conseils avisés. Et merci aussi à Farnoosh, François, Grzegorz, Mana, Maxime et Yohan.

Je remercie également les personnes du CRVI de Lévis qui m'ont aidé pendant mes séjours dans leurs locaux et avec qui j'ai sympathisé : Benjamin, Gao, Louis-André, Michael, Nassim, et Rivonala ainsi qu'aux stagiaires que j'ai co-encadrés : Jérôme, Maxime, Pierre et Thibaud.

Enfin, merci à mon père pour sa relecture et ses encouragements, à ma mère de m'avoir toujours soutenu dans ce que j'entreprenais et ma grand-mère pour m'avoir transmis son admiration pour les merveilleux paysages de Corse. Merci à tous mes proches pour les beaux moments que j'ai pu passer en leur compagnie durant ces trois années et merci à ma compagne Géraldine pour son grand soutien particulièrement dans les derniers mois de cette thèse.

Introduction

Les feux de forêt représentent un risque majeur pour de nombreux pays dans le monde. Ils causent des dégâts environnementaux (déforestation, désertification, pollution de l'air par émissions de CO_2 , recirculation des métaux lourds et radionucléides) ainsi que des pertes humaines et financières (destruction de logements et autres bâtiments, destruction de matériel de lutte). Chaque année, plus de 350 millions d'hectares de végétation sont détruits sur notre planète par ce phénomène [1]. Les événements de l'année 2013 donnent un exemple de l'étendue de ce fléau. Ainsi, en juin aux États-Unis, dix-neuf pompiers sont morts et cent maisons ont été détruites en Arizona et en juillet 104 200 hectares de végétation ont été détruits en Californie avec un coût de lutte estimé à 127,35 millions de dollars [2, 3]. Cette même année, en Australie, 248 logements et autres bâtiments ont été endommagés par des incendies et plus de 118 000 km² de végétation ont été brûlés dans le pays [4, 5]. En Europe, l'année 2013 a également été dramatique. Ainsi au Portugal, ils ont occasionné la mort de six pompiers et la destruction de 31 000 hectares de végétation. En Espagne, plus de 1 200 hectares ont été brûlés en moins de vingt-quatre heures et en Grèce dix maisons près d'Athènes ont été détruites par les flammes.

Le Canada et la France sont également chaque année très touchés par ce phénomène. Ainsi, au Canada, 8600 incendies détruisent annuellement 2,5 millions d'hectares dont 735 000 hectares de forêts aménagées, entraînant ainsi une perte d'environ un milliard de dollars de bois [6]. En France, 4 000 feux ont lieu en moyenne chaque année et détruisent 24 000 hectares. La Corse est particulièrement touchée par ce phénomène puisque en 2014, 2 353 hectares ont été brûlés par 541 incendies de végétation [7]. Si des moyens de lutte efficace ne sont pas mis en place dans les pays concernés, ces chiffres vont augmenter avec le réchauffement climatique [8, 9]

Pour lutter contre ce risque majeur, des actions de gestion, de prévention et de lutte sont réalisées. L'efficacité de ces trois types d'action est liée à la connaissance des phénomènes qui interviennent lors de la propagation des feux.

Depuis plus de vingt ans, les chercheurs de l'Université de Corse travaillent sur la problématique des feux de forêt pour comprendre les phénomènes intervenant lors de la propagation d'un incendie de végétation, les modéliser afin de développer un modèle comportemental et l'intégrer dans un simulateur.

Un modèle de comportement simplifié a été développé au sein de l'équipe « Feux de forêt » de l'Unité Mixte de Recherche CNRS 6134 Sciences pour l'Environnement de l'Université de Corse [10, 11, 12]. Il existe encore d'autres modèles tels que ceux présentés dans les publications [13, 14, 15, 16, 17, 18].

Pour tous les modèles en général, il est nécessaire d'obtenir des données expérimentales afin de mieux comprendre les phénomènes étudiés, de comparer les données numériques et mesurées et ainsi améliorer et/ou valider les formulations mathématiques. Les caractéristiques géométriques d'un feu décrivent l'évolution d'un incendie ; mais la dangerosité du phénomène, ses dimensions et les distances parcourues rendent leur estimation difficile.

Depuis une dizaine d'années, des travaux de recherche sont menés pour développer des instruments de mesure basés sur la vision par ordinateur dédiés aux feux de forêt. Des méthodes très intéressantes ont été proposées pour des feux réalisés à l'échelle du laboratoire [19, 20, 21, 22]. Toutefois ces algorithmes ne sont pas applicables à des feux de terrain principalement en raison de leurs dimensions, de l'étendue des distances parcourues, de la diversité de leur couleur et des textures présentées et enfin de la présence de fumée.

Cette thèse présente les travaux qui ont été menés pour l'estimation par stéréovision multimodale (visible et proche-infrarouge) de caractéristiques géométriques de feux de végétation.

Ce rapport est organisé en huit chapitres dont voici le détail :

Le premier chapitre est dédié à la présentation de l'état de l'art de la vision pour la mesure des feux. Il décrit, dans une première partie, toutes les méthodes de détection de pixels de feu de la littérature, puis présente, dans une deuxième partie, les techniques d'estimation par vision des caractéristiques géométriques des feux de végétation. Il se termine par l'énoncé des difficultés scientifiques à résoudre pour atteindre l'objectif de cette thèse.

Le second chapitre décrit l'analyse qui a été menée afin de déterminer le matériel de vision et le fonctionnement les mieux adaptés à la mesure de caractéristiques géométriques des feux de végétation.

Le troisième chapitre est ciblé sur la première étape du processus proposé de mesure par vision, à savoir la détection automatique des pixels de feu dans des images de modalité visible. Les travaux concernant l'adaptation de certaines méthodes existantes de détection de pixels et le développement de nouvelles méthodes sont présentés ainsi que leurs performances sur une nouvelle base d'images de feux de végétation créée dans la cadre de cette thèse.

Le quatrième chapitre traite la détection de pixels de feu à partir d'images du visible et du proche-infrarouge. Plusieurs techniques de fusion de données sont présentées ainsi que leur utilisation dans de nouvelles méthodes de détection de pixels. Ce chapitre se termine par une analyse des performances des différentes méthodes proposées sur une nouvelle base d'images

multimodale de feux de végétation.

Le cinquième chapitre décrit l'utilisation des modalités proche-infrarouge et visible permettant l'optimisation de la recherche et l'appariement de points de correspondance dans des images stéréoscopiques.

Le sixième chapitre est dédié à l'estimation des caractéristiques géométriques de feu à partir d'images stéréoscopiques. Les principes théoriques de la stéréovision, la méthode d'obtention de points tridimensionnels de feu et de reconstruction tridimensionnelle de forme 3D sont exposés ainsi que le calcul des caractéristiques géométriques.

Le septième chapitre expose les moyens et méthodes utilisés pour l'obtention des points tridimensionnels de feu à partir de plusieurs systèmes de stéréovision et l'estimation de caractéristiques tel que le périmètre et le volume. Les performances du dispositif développé sont montrées et analysées dans une deuxième partie à partir d'exemples de reconstruction d'objets rigides de dimensions connues.

Le huitième chapitre présente un exemple de reconstruction tridimensionnelle et d'estimation de caractéristiques géométriques estimées d'un feu de végétation expérimental.

Chapitre 1

État de l'art de la vision pour la mesure des feux

1.1	Introduction	5
1.2	Méthodes de détection de pixels de feu	6
1.2.1	Méthodes de détection de pixels de feu basées sur l'intensité de pixels dans le spectre visible	6
1.2.2	Méthodes de détection de pixels de feu basées sur l'intensité de pixels dans le spectre infrarouge	15
1.2.3	Méthodes de détection de pixels de feu basées sur la fusion d'informations d'intensité de pixels obtenues dans différents spectres	24
1.2.4	Méthodes de détection de pixels de feu utilisant des critères complémentaires à l'intensité de pixels	24
1.2.5	Métriques d'évaluation des performances des méthodes de détection de pixels de feu	29
1.2.6	Évaluations comparatives des performances de certains algorithmes de détection de pixels de feu	31
1.3	Techniques de mesure pour les feux	32
1.3.1	Caractéristiques géométriques d'un front de feu	32
1.3.2	Mesures estimées à partir d'images obtenues depuis un même point de vue	33
1.3.3	Mesures estimées à partir d'images obtenues depuis plusieurs points de vue	34
1.3.4	Mesures par stéréovision	36
1.4	Conclusion	38

1.1 Introduction

Depuis plus d'une vingtaine d'années, des images acquises dans les domaines spectraux visible et infrarouge sont traitées pour détecter la présence d'incendie ou estimer certaines caractéristiques géométriques d'un feu telles que sa position ou la longueur du front. Ce premier chapitre présente un état de l'art des travaux menés en vision par ordinateur pour la mesure des feux.

La section 1.2 décrit les méthodes de la littérature de détection de pixels de feu dans des images du spectre visible et du spectre infrarouge dans un but de mesure.

La section 1.3 présente les différentes techniques de mesure de caractéristiques de feu de végétation en extérieur par des techniques de vision utilisant une ou plusieurs caméras.

1.2 Méthodes de détection de pixels de feu

La première étape des méthodes de mesure de feu par vision est la détection des pixels appartenant au feu dans les images. Les difficultés rencontrées pour réaliser cette phase sont dues à la couleur des flammes (pouvant être rouge, orange, jaune ou blanche), à la texture variable, à la possibilité de présence de fumée blanche ou noire plus ou moins dense et à la forme chaotique et non prévisible du contour des zones de feu. Les méthodes de détection de pixels reposent sur l'utilisation de l'intensité de pixels (valeur unique comprise entre 0 et 255 pour un pixel appartenant aux images acquises dans l'infrarouge et trois valeurs correspondant aux canaux rouge, vert et bleu dont les valeurs sont comprises entre 0 et 255 pour un pixel appartenant à une image acquise dans le domaine visible) et/ou sur l'exploitation de variations entre images successives (formes, texture, oscillations).

Cette section est organisée de la façon suivante : les trois premières parties décrivent des méthodes de détection exploitant l'intensité de pixels d'images autonomes (non séquencées) : la première partie expose les méthodes utilisant des images du spectre visible, la deuxième partie présente les méthodes travaillant avec des images du spectre infrarouge et la troisième partie est dédiée aux méthodes de détection de pixels de feu utilisant des informations multimodales c'est-à-dire obtenues dans des images simultanément acquises dans différents domaines spectraux. La quatrième partie est dédiée aux techniques qui utilisent, en plus de la valeur d'intensité des pixels, d'autres critères tels que le mouvement ou la texture. La cinquième partie est consacrée aux critères d'évaluation qui sont employés pour évaluer la performance de méthodes de détection de pixels. Enfin, la dernière partie présente les résultats de la littérature concernant la comparaison de performances de certaines méthodes de détection de pixels de feu.

Nous utiliserons l'expression pixels « feu » pour désigner les pixels appartenant au feu et pixels « environnement » pour les autres pixels.

1.2.1 Méthodes de détection de pixels de feu basées sur l'intensité de pixels dans le spectre visible

La problématique de détection de départ de feux a été à l'origine de nombreux travaux de recherche en vision [23, 24, 25, 26]. Dans les méthodes par vision proposées dans la littérature, des zones sont sous surveillance et le premier traitement est souvent déclenché après la détection de modifications entre une image de référence et l'image traitée. L'expansion de

zones de pixels identifiées comme du feu ou de la fumée, la variation temporelle d'intensité des pixels dans ces zones sont très fréquemment utilisées et nécessitent le traitement de séquences d'images. La problématique de la mesure de feux est différente. D'abord, parce qu'un postulat de présence de feu est fait, ensuite, parce qu'il y a une exigence de grande précision dans la segmentation de zones de flammes car c'est à partir de ces zones que les mesures sont estimées. Aussi, nous présenterons dans cette partie uniquement les méthodes de détection de pixels de feu développées dans le cadre de la problématique de la détection de départ de feux qui peuvent être utilisées pour la mesure de caractéristiques géométriques de feux.

Différents systèmes de couleurs sont utilisés pour la détection de pixels de feu. Le système de couleur le plus employé est le système *RVB* (Rouge, Vert, Bleu) [27, 28, 29, 30, 31, 32]. D'autres systèmes sont aussi exploités. Parmi ces systèmes, il y a ceux de luminance-chrominance, tels que YC_bC_r (luminance, chrominance bleue et chrominance rouge) [33, 34, 35], *YUV* (luminance, différence de la luminance avec le canal bleu, différence de la luminance avec le canal rouge) [36] et $L^*a^*b^*$ (clarté, chrominance rouge-verte, chrominance jaune-bleue) [37]. On trouve aussi des travaux utilisant des systèmes dits cylindriques tels que *TSI* (Teinte, Saturation, Intensité) [38] et *TSV* (Teinte, Saturation, Valeur) [39]. Enfin, certains algorithmes utilisent des combinaisons de différents espaces de couleurs [40, 41, 42].

Les méthodes présentées dans cette section sont classées selon le système de couleur utilisé pour la détection de pixels de feu. Un récapitulatif des méthodes qui seront décrites dans ce document avec leurs espaces de couleur associés est présenté dans le [tableau 1.1](#).

TABLEAU 1.1 : Espaces de couleur des méthodes présentées.

	<i>RVB</i>	YC_bC_r	<i>TSI</i>	<i>YUV</i>	$L^*a^*b^*$
Phillips <i>et al.</i> 2002 [27]	✓				
T.H. Chen <i>et al.</i> 2004 [41]	✓		✓		
Martínez <i>et al.</i> 2005 [23]	✓				
Horng <i>et al.</i> 2005 [38]			✓		
Celik <i>et al.</i> 2007 [28]	✓				
Celik et Demirel 2009 [33]		✓			
Ko <i>et al.</i> 2009 [31]	✓				
J. Chen <i>et al.</i> 2010 [40]	✓		✓		
Rossi et Akhloufi 2010 [43]	✓			✓	
Celik 2010 [37]					✓
Collumeau <i>et al.</i> 2011 [30]	✓		✓		
Rudz <i>et al.</i> 2013 [35]	✓	✓			

Voici les notations qui seront utilisées dans la suite de ce document : On notera $\mathbf{I}_R(\mathbf{x})$, $\mathbf{I}_V(\mathbf{x})$

et $\mathbf{I}_B(\mathbf{x})$ les intensités du pixel \mathbf{x} de l'image \mathbf{I} dans les canaux respectifs rouge, vert et bleu du système de couleur RVB . Ce principe de notation sera également utilisé pour les autres systèmes de couleur ; ainsi par exemple, $\mathbf{I}_T(\mathbf{x})$, $\mathbf{I}_S(\mathbf{x})$ et $\mathbf{I}_I(\mathbf{x})$ sont les intensités du pixel \mathbf{x} de l'image \mathbf{I} dans les canaux T , S et I du système de couleur TSI .

1.2.1.a Méthode utilisant le système de couleur RVB

Une des premières méthodes de détection de feu proposées dans la littérature est celle de Phillips *et al.* [27]. C'est une méthode qui comprend une phase d'apprentissage de couleur de pixels de feu réalisée à partir d'images dont les zones de flammes ont été segmentées à la main. Au cours de cette phase, un histogramme gaussien lissé de dimension $256 \times 256 \times 256$ est construit. Les indices des cellules de cet histogramme correspondent aux valeurs d'intensité possibles dans les canaux rouge, vert et bleu de l'espace RVB . L'histogramme est construit de la façon suivante : les intensités $\mathbf{I}_R(\mathbf{x})$, $\mathbf{I}_V(\mathbf{x})$ et $\mathbf{I}_B(\mathbf{x})$ de chaque pixel \mathbf{x} identifié comme appartenant à la zone de flamme dans les images de la base d'apprentissage sont utilisés comme indices de position dans l'histogramme. Une gaussienne tridimensionnelle d'écart type $\sigma = 2$ centrée à la position considérée est additionnée à l'histogramme. De façon analogue, des gaussiennes tridimensionnelles d'écart type $\sigma = 1$ centrées aux positions données par les intensités $\mathbf{I}_R(\mathbf{x})$, $\mathbf{I}_V(\mathbf{x})$ et $\mathbf{I}_B(\mathbf{x})$ des pixels identifiés comme appartenant à l'environnement sont soustraites à l'histogramme. Un seuil est ensuite appliqué à l'histogramme gaussien lissé pour générer une table de consultation. Ce seuil, dont la valeur n'a pas été spécifiée par les auteurs, permet d'associer à chaque couleur une valeur booléenne indiquant si un triplet $(\mathbf{I}_R(\mathbf{x}), \mathbf{I}_V(\mathbf{x}), \mathbf{I}_B(\mathbf{x}))$ représente une couleur « feu » ou « environnement ».

Un pixel \mathbf{x} est donc labellisé « feu » si la cellule de la table de consultation en position données par les intensités $\mathbf{I}_R(\mathbf{x})$, $\mathbf{I}_V(\mathbf{x})$ et $\mathbf{I}_B(\mathbf{x})$ est de valeur « VRAI » ; il est classé « environnement » si la valeur de cette case est « FAUX ».

Dans [23], Martínez-de Dios *et al.* proposent une autre méthode de détection de pixels « feu ». Celle-ci se base sur le fait que les pixels « feu » ont une grande valeur d'intensité sur le canal rouge de l'espace RVB et que cette dernière est plus grande que les valeurs d'intensité sur les canaux vert et bleu. Les règles de détection sont les suivantes :

$$\begin{aligned} \mathbf{I}_R(\mathbf{x}) - k_v \mathbf{I}_V(\mathbf{x}) &\geq 0 \\ \mathbf{I}_R(\mathbf{x}) - k_b \mathbf{I}_B(\mathbf{x}) &\geq 0 \\ \mathbf{I}_R(\mathbf{x}) &\geq \tau_R \end{aligned} \tag{1.1}$$

Où k_v et k_b sont des coefficients dont la valeur n'est pas précisée dans l'article [23] et τ_R est obtenu en utilisant la méthode de détection de seuil de Ridler et Calvard [44] (détaillée dans la section 1.2.2) appliquée à l'image \mathbf{I}_R

Celik *et al.* [28] proposent un ensemble de règles sur l'espace de couleur RVB à partir des-

quelles la détection de pixels « feu » peut être réalisée. Ces règles ont été établies en utilisant cent-cinquante images de feu collectées sur Internet. La première règle permet une sélection des pixels en se basant sur leur intensité et sur la moyenne des intensités dans l'image \mathbf{I}_R . La deuxième règle traduit le fait que les auteurs ont observé que les couleurs de feu ont une composante rouge supérieure à la composante verte et une composante verte supérieure à la bleue.

Les trois règles suivantes traduisent le fait que si l'illumination d'une scène change, la luminance et la chrominance de l'image de cette scène changent aussi. Les auteurs proposent donc des règles utilisant des ratios RVB qui permettent de représenter simultanément la luminance et la chrominance. Ainsi, un pixel \mathbf{x} est labellisé « feu » si les conditions suivantes sont satisfaites :

$$\begin{aligned}
 \mathbf{I}_R(\mathbf{x}) &> \mu_R \\
 \mathbf{I}_R(\mathbf{x}) &> \mathbf{I}_V(\mathbf{x}) > \mathbf{I}_B(\mathbf{x}) \\
 0,25 &\leq \frac{\mathbf{I}_V(\mathbf{x})}{\mathbf{I}_R(\mathbf{x}) + 1} \leq 0,65 \\
 0,05 &\leq \frac{\mathbf{I}_B(\mathbf{x})}{\mathbf{I}_R(\mathbf{x}) + 1} \leq 0,45 \\
 0,20 &\leq \frac{\mathbf{I}_B(\mathbf{x})}{\mathbf{I}_V(\mathbf{x}) + 1} \leq 0,60
 \end{aligned} \tag{1.2}$$

Où μ_R est la moyenne d'intensité dans l'image \mathbf{I}_R : $\mu_R = \frac{1}{N} \sum_{\mathbf{x} \in \mathbf{I}} \mathbf{I}_R(\mathbf{x})$, avec N le nombre de pixels de l'image.

Une autre méthode de détection de feu est celle de Ko *et al.* [31]. Les auteurs proposent d'utiliser un modèle probabiliste gaussien généré à partir des intensités de pixel dans les canaux du système RVB . Les paramètres de la gaussienne sont estimés à partir de cent images de feu sélectionnées dans douze vidéos et segmentées à la main. La probabilité associée au canal $C \in \{R, V, B\}$ qu'un pixel soit un pixel « feu » est donnée par :

$$p_C = \frac{1}{\sqrt{2\pi}\sigma_C} \exp\left(-\frac{(\mathbf{I}_C(\mathbf{x}) - \mu_C)^2}{2\sigma_C^2}\right) \tag{1.3}$$

Avec $\mathbf{I}_c(\mathbf{x})$ la valeur d'intensité, $\mathbf{I}(\mathbf{x})$ du pixel \mathbf{x} dans le canal C de l'image, μ_C et σ_C la moyenne et l'écart type des intensités dans le canal C .

La probabilité totale p_f qu'un pixel soit classé « feu » est calculée comme le produit des probabilités p_C associée à chaque canal C :

$$p_f(\mathbf{x}) = \prod_{c \in \{R, V, B\}} p_c(\mathbf{I}_c(\mathbf{x})) \tag{1.4}$$

Si $p_f(\mathbf{x}) \geq \tau$, avec τ un seuil, alors le pixel est classé « feu ».

1.2.1.b Méthode utilisant le système de couleur YC_bC_r

Une méthode de détection de pixels de feu utilisant le système YC_bC_r est proposée par Celik et Demirel [33]. Neuf règles sont utilisées pour le classement d'un pixel en pixel « feu » :

$$\begin{aligned}
 \mathbf{I}_Y(\mathbf{x}) &> \mathbf{I}_{C_b}(\mathbf{x}) \\
 \mathbf{I}_{C_r}(\mathbf{x}) &> \mathbf{I}_{C_b}(\mathbf{x}) \\
 \mathbf{I}_Y(\mathbf{x}) &> \mu_Y \\
 \mathbf{I}_{C_b}(\mathbf{x}) &< \mu_{C_b} \\
 \mathbf{I}_{C_r}(\mathbf{x}) &> \mu_{C_r} \\
 |\mathbf{I}_{C_b}(\mathbf{x}) - \mathbf{I}_{C_r}(\mathbf{x})| &\geq \tau; \\
 \mathbf{I}_{C_b}(\mathbf{x}) &\geq f_u(\mathbf{I}_{C_r}(\mathbf{x})); \\
 \mathbf{I}_{C_b}(\mathbf{x}) &\leq f_l(\mathbf{I}_{C_r}(\mathbf{x})); \\
 \mathbf{I}_{C_b}(\mathbf{x}) &\leq f_d(\mathbf{I}_{C_r}(\mathbf{x})),
 \end{aligned} \tag{1.5}$$

Avec μ_Y, μ_{C_b} et μ_{C_r} les moyennes des valeurs d'intensité de pixels dans les canaux respectifs Y, C_b et C_r ; τ un seuil déterminé en utilisant une fonction d'efficacité du récepteur (ou ROC de l'anglais *Receiver Operating Characteristic*) [45] sur une base de mille images de feux provenant d'Internet, dont certaines avec des objets qui ont des couleurs proches de celles du feu. La courbe calculée est composée de cent points correspondants au taux vrais positifs/faux positifs des mille images pour τ compris entre un et cent. Selon le contexte, à savoir si on cherche à éviter les faux-positifs ou à avoir un maximum de vrai-positifs, plusieurs valeurs de τ peuvent être choisies. Les auteurs choisissent $\tau = 40$. $f_u(\cdot), f_l(\cdot)$ et $f_d(\cdot)$ sont des fonctions polynomiales calculées à partir d'une analyse de mille images segmentés manuellement. La distribution C_b-C_r de tous les pixels « feu » forme une zone dont les limites sont caractérisées par les trois polynômes.

Chino *et al.* [46] proposent une autre méthode de détection sur l'espace YC_bC_r . Ce papier n'étant pas encore publié lors du rendu de cette thèse nous ne le décrirons que très brièvement et n'en tiendrons pas compte dans les autres chapitres. Les auteurs proposent une méthode de détection de pixels couplant de l'apprentissage de couleur et de l'apprentissage de texture. L'apprentissage de couleur se fait sur l'espace YC_bC_r grâce à une classification naïve bayésienne [47]. L'apprentissage de texture est faite par classification de caractéristiques sur des zones définies par une méthode de « superpixel ». Pour l'apprentissage, quatre-vingt images de feu et cent-soixante images d'environnement de taille 50×50 pixels ont été utilisées. L'algorithme consiste à détecter les pixels ayant une couleur de feu ainsi que les superpixels ayant une texture de feu à partir de l'apprentissage. Les pixels ayant été détectés par les deux méthodes sont catégorisés « feu ».

1.2.1.c Méthode utilisant le système de couleur *TSI*

Le système *TSI* (Teinte, Saturation et Intensité) est un système approprié pour la description de couleurs telles que perçues par l'œil humain car la teinte et la saturation sont très liées à la perception humaine des couleurs [48]. Pour cette raison, Horng *et al.* [38] proposent une détection des pixels « feu » dans cet espace de couleur. Ils considèrent un pixel comme étant un pixel « feu » si les valeurs d'intensité dans les canaux H , S et I sont dans des intervalles estimés empiriquement à partir de soixante-dix images de flammes. Les auteurs considèrent deux cas d'images qui diffèrent par la taille de la zone de feu dans l'image (car ils estiment que la zone de feu a une influence sur l'intensité lumineuse de l'image). Les valeurs des intervalles pour ces deux cas sont données dans le tableau 1.2.

TABLEAU 1.2 : Caractéristiques de couleur des flammes dans le système de couleur *TSI* estimées empiriquement par Horng *et al.* [38]

Aire du feu dans l'image	T (Teinte)	S (Saturation)	I (Intensité)
Grande	$0^\circ - 60^\circ$	40 – 100	127 – 255
Petite	$0^\circ - 60^\circ$	20 – 100	100 – 255

Aucune précision n'est donnée dans [38] quant aux critères permettant de définir une zone de feu dans l'image comme étant grande ou petite.

1.2.1.d Méthode utilisant le système de couleur $L^*a^*b^*$

Celik [37] propose un ensemble de règles dans l'espace de couleur CIE $L^*a^*b^*$ afin de détecter les pixels de feu d'une image. Un pixel doit satisfaire les règles suivantes pour être labellisé « feu » :

$$\begin{aligned}
 \mathbf{I}_{L^*}(\mathbf{x}) &\geq \mu_{L^*} \\
 \mathbf{I}_{a^*}(\mathbf{x}) &\geq \mu_{a^*} \\
 \mathbf{I}_{b^*}(\mathbf{x}) &\geq \mu_{b^*} \\
 \mathbf{I}_{b^*}(\mathbf{x}) &\geq \mathbf{I}_{a^*}(\mathbf{x}) \\
 P(\mathbf{I}_{L^*}(\mathbf{x}), \mathbf{I}_{a^*}(\mathbf{x}), \mathbf{I}_{b^*}(\mathbf{x})) &\geq \tau
 \end{aligned} \tag{1.6}$$

Où μ_{L^*} , μ_{a^*} et μ_{b^*} sont les valeurs moyennes dans les canaux L^* , a^* et b^* respectivement, et $P(\mathbf{I}_{L^*}(\mathbf{x}), \mathbf{I}_{a^*}(\mathbf{x}), \mathbf{I}_{b^*}(\mathbf{x}))$ est la probabilité jointe que \mathbf{x} soit un pixel « feu » en considérant sa valeur dans les canaux L^* , a^* et b^* .

Pour calculer $P(\mathbf{I}_{L^*}(\mathbf{x}), \mathbf{I}_{a^*}(\mathbf{x}), \mathbf{I}_{b^*}(\mathbf{x}))$, la corrélation entre les valeurs \mathbf{I}_{L^*} , \mathbf{I}_{a^*} et \mathbf{I}_{b^*} des pixels « feu » doit être étudiée en considérant un grand nombre de pixels (Celik [37] a travaillé sur 500 images segmentées manuellement). Un histogramme est construit pour chacun des plans (L^*, a^*) , (L^*, b^*) et (a^*, b^*) de l'espace $L^*a^*b^*$. Pour cela, une quantification des valeurs

I_{L^*} , I_{a^*} et I_{b^*} est faite sur 24 niveaux, ce qui implique que les histogrammes sont des tableaux de dimension (24×24) .

Le procédé d'actualisation des histogrammes à partir d'un nouveau pixel « feu » d'intensité I_{L^*} , I_{a^*} et I_{b^*} est le suivant : la valeur 1 est additionnée à celle contenue dans l'histogramme à la position (I_{a^*}, I_{b^*}) pour l'histogramme relatif au plan (a^*, b^*) , à la position (I_{L^*}, I_{b^*}) pour l'histogramme relatif au plan (L^*, b^*) et à la position (I_{L^*}, I_{a^*}) pour l'histogramme relatif au plan (L^*, a^*) . La dernière étape du processus de construction des histogrammes est une normalisation : dans chaque histogramme, toutes les valeurs sont divisées par la plus haute valeur qu'il contient de sorte d'obtenir des données dans l'intervalle $[0; 1]$. La figure 1.1 présente un exemple de probabilité $P(I_{L^*}(\mathbf{x}), I_{a^*}(\mathbf{x}))$ d'un pixel d'appartenir au feu en fonction de ces valeurs I_{L^*} et I_{a^*} .

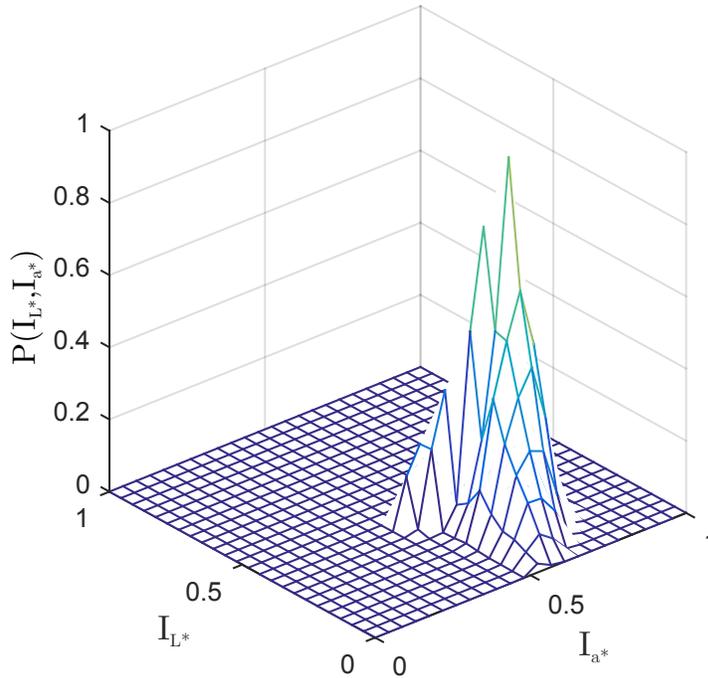


FIGURE 1.1 : Exemple de probabilité d'un pixel d'appartenir au feu en fonction de ces valeurs I_{L^*} et I_{a^*} .

La probabilité qu'une couleur soit une couleur de feu est estimée à partir des histogrammes en considérant les valeurs d'intensité dans les canaux correspondants. La probabilité finale est donnée par :

$$P(I_{L^*}(\mathbf{x}), I_{a^*}(\mathbf{x}), I_{b^*}(\mathbf{x})) = P(I_{L^*}(\mathbf{x}), I_{a^*}(\mathbf{x})) \times P(I_{L^*}(\mathbf{x}), I_{b^*}(\mathbf{x})) \times P(I_{a^*}(\mathbf{x}), I_{b^*}(\mathbf{x})) \quad (1.7)$$

Dans la dernière règle de (1.6), τ est un seuil que les auteurs fixent à 0,005. Cette valeur est calculée à l'aide d'une analyse ROC sur l'ensemble des images considérées pour l'obtention

des histogrammes.

1.2.1.e Méthodes utilisant plusieurs systèmes de couleur

T. H. Chen *et al.* [41] proposent un algorithme de détection de pixels de feu qui se base sur les critères présentés dans [49] : un feu à basse température émet de la lumière avec des couleurs très saturées alors qu'un feu à haute température émet de la lumière peu saturée. Dans l'article [50], T. H. Chen *et al.* supposent que, compte tenu des propriétés colorimétriques du feu, un pixel de feu aura une teinte entre 0 et 60 degrés, ce qui se traduit par $I_R(\mathbf{x}) \geq I_V(\mathbf{x}) > I_B(\mathbf{x})$ dans le système RVB . De plus, comme la couleur rouge est généralement très présente dans les flammes, les auteurs proposent de considérer un pixel comme faisant partie du feu si son intensité I_R est supérieure à un certain seuil R_T , ce dernier devant être compris entre 115 et 135. La troisième condition est proposée pour éviter que la détection soit affectée par une illumination trop forte de l'environnement extérieur. Pour cela, il faut que la saturation des pixels « feu » soit en dessous d'un certain seuil S_T choisi entre 55 et 65. Un pixel \mathbf{x} est donc labellisé « feu » s'il vérifie les conditions suivantes :

$$\begin{aligned} I_R(\mathbf{x}) &\geq I_V(\mathbf{x}) > I_B(\mathbf{x}); \\ I_R(\mathbf{x}) &> R_T; \\ I_S(\mathbf{x}) &\geq (255 - I_R(\mathbf{x})) \frac{S_T}{R_T}, \end{aligned} \tag{1.8}$$

Une autre méthode de détection hybride de pixels de feu est proposée par Rossi et Akhloufi [43] en combinant les informations des espaces de couleurs RVB et YUV [51]. Une étude qu'ils ont menée sur des images de feux de végétation en extérieur a conduit les auteurs à considérer le canal V (noté V_2) du système YUV pour réaliser une pré-sélection des pixels de feu. Ils utilisent l'algorithme des K-moyennes sur ce canal (avec un nombre de classe égal à deux), suivi d'un seuillage par valeur d'intensité pour détecter un premier ensemble de pixels. À partir de ces pixels présélectionnés, un apprentissage de couleur dans le système RVB est ensuite réalisé en considérant un modèle de couleur tridimensionnel gaussien dont les paramètres sont μ_R, μ_V, μ_B et $\sigma_R, \sigma_V, \sigma_B$ qui symbolisent respectivement les moyennes et écart-types des valeurs d'intensité dans les canaux R , V et B . Ainsi, un pixel \mathbf{x} d'intensité $I_R(\mathbf{x})$, $I_V(\mathbf{x})$ et $I_B(\mathbf{x})$ est classé « feu » s'il vérifie la condition suivante :

$$\sqrt{(I_R(\mathbf{x}) - \mu_R)^2 + (I_V(\mathbf{x}) - \mu_V)^2 + (I_B(\mathbf{x}) - \mu_B)^2} \leq k \sigma \tag{1.9}$$

Avec k une constante et σ la valeur maximale parmi les valeurs d'écart-types obtenues pour les trois canaux.

J. Chen *et al.* [40] proposent une méthode de détection de pixels « feu » en utilisant les valeurs RVB et de saturation des pixels. Les auteurs utilisent les deux premières règles de T.H. Chen

et al. [41] auxquelles ils rajoutent une condition sur la saturation. Cette condition a été obtenue à partir d'une étude portant sur le lien entre la saturation et le canal bleu. Il a été montré que les pixels « feu » étaient dans une région compacte délimitée par deux courbes f_1 et f_2 définies ainsi :

$$f_1(x) = \begin{cases} 100 - 0,48x & \text{si } x \geq 117 \\ 79,27 - 0,311x & \text{sinon} \end{cases} \quad (1.10)$$

$$f_2(x) = -2,0147 + 90,59435 e^{-x/77;6027} \quad (1.11)$$

Au final, un pixel est labellisé « feu » si les trois règles suivantes sont vérifiées :

$$\begin{aligned} \mathbf{I}_R(\mathbf{x}) &\geq \mathbf{I}_V(\mathbf{x}) > \mathbf{I}_B(\mathbf{x}); \\ \mathbf{I}_R(\mathbf{x}) &> R_T; \\ f_2(\mathbf{I}_B(\mathbf{x})) &\geq \mathbf{I}_S(\mathbf{x}) \geq f_1(\mathbf{I}_B(\mathbf{x})) \end{aligned} \quad (1.12)$$

Collumeau *et al.* [30] proposent une méthode d'apprentissage supervisée sur l'espace RVB par machine à vecteurs de support (ou SVM de l'anglais *Support Vector Machine*) [52] avec une fonction de base radiale. Le principe des SVM est détaillé dans la partie 3.3.2.b. L'apprentissage des pixels « feu » est réalisé à partir de douze mille pixels composés à part égale de pixels « feu » et de pixels « non feu ».

La première étape de la détection s'effectue donc par machine à vecteurs de support. Dans une deuxième étape, un opérateur morphologique d'ouverture est appliqué aux pixels retenus afin d'enlever les faux positifs qui ne forment pas de régions homogènes. Un filtrage est finalement réalisé sur les régions de pixels détectés afin de ne sélectionner que celles dont l'intensité moyenne par région est inférieure à 0,8; l'intensité d'un pixel étant donnée par la troisième composante du système *TSI*.

Rudz *et al.* [35] proposent d'appliquer l'algorithme des K-moyennes sur le canal C_b de l'espace de couleur YC_bC_r . Après ce classement, une amélioration est faite sur l'espace *RVB* pour éliminer les faux positifs en fonction de la taille des régions segmentées. Si une région est large, les règles suivantes sont appliquées :

$$\begin{aligned} \|\text{hist}_R^{ref} - \text{hist}_R\| &< \tau_R; \\ \|\text{hist}_V^{ref} - \text{hist}_V\| &< \tau_V; \\ \|\text{hist}_B^{ref} - \text{hist}_B\| &< \tau_B, \end{aligned} \quad (1.13)$$

Avec hist_c^{ref} , hist_c et τ_c respectivement l'histogramme de référence, l'histogramme de la région considérée et le seuil du canal c . Si une région est petite, les règles suivantes sont utilisées :

$$\begin{aligned}
\|\mu_R^{ref} - \mu_R\| &< \rho_R \sigma_R^{ref}; \\
\|\mu_V^{ref} - \mu_V\| &< \rho_V \sigma_V^{ref}; \\
\|\mu_B^{ref} - \mu_B\| &< \rho_B \sigma_B^{ref},
\end{aligned}
\tag{1.14}$$

Avec μ_c^{ref} et σ_c^{ref} respectivement la moyenne et l'écart type du canal c pour les pixels feu de référence, μ_c la moyenne des pixels de la région considérée pour le canal c et ρ_c un coefficient pour le canal c . $hist_c^{ref}$, μ_c^{ref} et σ_c^{ref} sont calculés sur des images de feu segmentées manuellement. Les seuils τ_c et ρ_c sont optimisés pour obtenir les meilleures segmentations sur les images de feu en se basant sur le F-score des résultats obtenus. Les notions d'histogramme de référence et d'histogramme de région n'étant pas clairement définie dans l'article, nous avons choisi de ne pas considérer cette méthode dans les chapitres suivants afin d'éviter toute erreur d'interprétation.

1.2.2 Méthodes de détection de pixels de feu basées sur l'intensité de pixels dans le spectre infrarouge

Si la détection de pixels de feu dans le domaine visible est une tâche complexe, elle reste assez simple dans le domaine infrarouge du fait que le feu a une forte intensité comparée aux autres éléments de la scène. Ainsi, la première approche pour détecter les zones de feu dans une image de l'infrarouge est de trouver le seuil qui différencie les pixels appartenant au feu à ceux de l'arrière-plan. Plusieurs algorithmes de recherche de seuil pouvant être appliqués à la détection des pixels de feu sont proposés dans la littérature. Ces méthodes sont connues comme des algorithmes de seuillage d'images monochromatiques et ont la particularité de séparer les zones lumineuses de l'image des zones sombres. Ces méthodes peuvent donc être utilisées sur des images du spectre infrarouge pour déterminer les pixels appartenant au feu. Le [tableau 1.3](#) présente les méthodes de seuillage décrites dans cette partie ainsi que des articles qui les ont utilisées pour la détection des pixels de feu dans le spectre infrarouge.

TABLEAU 1.3 : Méthodes de seuillage utilisées pour la détection de pixels « feu » dans le spectre infrarouge.

Méthode de seuillage	A été utilisée pour la détection de pixels de feu dans
Otsu (1975) [53]	Gouverneur <i>et al.</i> [54], Verstockt <i>et al.</i> [55], Martínez <i>et al.</i> [56], Torabnezhad <i>et al.</i> [57]
Ridler et Calvard (1978) [44]	Martínez <i>et al.</i> [56]
Kittler et Illingworth (1986) [58]	Martínez <i>et al.</i> [56]
Martínez (2001) [59]	Martínez <i>et al.</i> [56, 23, 60]

1.2.2.a Seuillage d'Otsu

La méthode d'Otsu [53] fait partie des méthodes de recherche de seuil les plus utilisées en traitement d'image. Cette technique permet de choisir un seuil qui minimise la variance des intensités des pixels entre deux classes de pixels et qui maximise la différence des moyennes d'intensité de ces deux classes. Dans notre cas, les deux classes sont le feu et l'environnement. On considère que les images ont 256 niveaux de gris. Soit h , l'histogramme d'une image. h a donc 256 classes. On pose p_k la probabilité de distribution du niveau k , définie par :

$$p_k = \frac{h(k)}{N} \quad (1.15)$$

Avec N le nombre total de pixels de l'image.

La variance inter-classe pondérée est définie par :

$$\sigma_w^2(s) = \omega_1(s)\sigma_1^2(s) + \omega_2(s)\sigma_2^2(s) \quad (1.16)$$

Avec s le seuil de séparation des deux classes et :

$$\omega_1(s) = \sum_{k=1}^s p_k ; \omega_2(s) = \sum_{k=s+1}^{256} p_k = 1 - \omega_1(s) \quad (1.17)$$

et

$$\sigma_1^2(s) = \sum_{k=1}^s (k - \mu_1(s))^2 \frac{p_k}{\omega_1(s)} \quad (1.18)$$

$$\sigma_2^2(s) = \sum_{k=s+1}^{256} (k - \mu_2(s))^2 \frac{p_k}{\omega_2(s)} \quad (1.19)$$

Avec les moyennes de classe définies ainsi :

$$\mu_1(s) = \sum_{k=1}^s \frac{k p_k}{\omega_1(s)} ; \mu_2(s) = \sum_{k=s+1}^{256} \frac{k p_k}{\omega_2(s)} \quad (1.20)$$

L'algorithme d'Otsu consiste à trouver le seuil s qui minimise $\sigma_w^2(s)$.

Cet algorithme a été utilisé dans les méthodes de détection de pixels « feu » développées par Gouverneur *et al.* [54], Verstockt *et al.* [55] et Martínez-de-Dios *et al.* [56]. Toutefois, Martínez-De-Dios et Ollero [61] considèrent que cette méthode n'est pas toujours efficace dans la recherche d'un seuil pour détecter les pixels de feu dans les images du spectre infrarouge. En

effet, cet algorithme détecte des pixels de l'arrière-plan si l'illumination de la scène n'est pas uniforme et si certains pixels de l'environnement ont des valeurs d'intensité nettement plus hautes que les autres pixels de l'arrière-plan.

1.2.2.b Seuillage de Ridler et Calvard

Ridler et Calvard [44] proposent une méthode de sélection de seuil itérative qui peut s'appliquer à la détection de pixels de feu. Le principe est le suivant : un seuil initial τ_0 égal à la moyenne des intensités des pixels de l'image est estimé. Puis, un nouveau seuil est calculé par itération jusqu'à convergence en suivant la formule suivante :

$$\tau_{k+1} = \frac{\mu_1(\tau_k) + \mu_2(\tau_k)}{2} \quad (1.21)$$

Où μ_1 et μ_2 sont respectivement la moyenne des intensités de pixels supérieurs à τ_k et celle des intensités inférieures à τ_k dans l'image considérée.

D'après les tests de Martínez-de Dios et Ollero [61], cette méthode n'est pas robuste au bruit et produit un grand nombre de pixels faux positifs.

1.2.2.c Seuillage de Kittler et Illingworth

La méthode de seuillage de Kittler et Illingworth [58] suppose que les niveaux de gris d'un objet sont distribués sur $[0 ; 255]$ selon une loi normale. Son principe est de minimiser l'erreur moyenne de classement des pixels en se servant du fait que les niveaux de gris de chaque classe doivent être distribués normalement. Comme pour la méthode d'Otsu, l'histogramme h de l'image est utilisé. Dans la suite, i désigne la classe étudiée et peut donc prendre la valeur 0 ou 1 respectivement pour l'environnement et le feu. Les auteurs définissent une probabilité *a priori* $P_i(s)$ du seuil s ainsi que la moyenne $\mu_i(s)$ et l'écart type $\sigma_i(s)$ des parties de l'histogramme associées à chaque classe i :

$$P_i(s) = \sum_{k=a}^b h(k) \quad (1.22)$$

$$\mu_i(s) = \frac{\sum_{k=a}^b h(k)k}{P_i(s)} \quad (1.23)$$

$$\sigma_i(s) = \frac{\sum_{k=a}^b (s - \mu_i(s))^2 h(k)}{P_i(s)} \quad (1.24)$$

Où

$$a = \begin{cases} 0 & \text{si } i = 0 \\ s + 1 & \text{si } i = 1 \end{cases} \quad (1.25)$$

et

$$b = \begin{cases} s & \text{si } i = 0 \\ N & \text{si } i = 1 \end{cases} \quad (1.26)$$

Un indice de performance ϵ est ensuite défini ainsi :

$$\epsilon(k, s) = \left(\frac{s - \mu_i(s)}{\sigma_i(s)} \right)^2 + 2 \log \sigma_i(s) - 2 \log P_i(s) ; \quad \text{avec } i = \begin{cases} 0 & \text{si } k \leq s \\ 1 & \text{si } k > s \end{cases} \quad (1.27)$$

La performance moyenne sur toute l'image est ensuite caractérisée par la fonction de critère suivante :

$$J(s) = \sum_k h(k) \epsilon(k, s) \quad (1.28)$$

L'algorithme de Kittler et Illingworth consiste à minimiser J en fonction de s . De même que pour la méthode de Ridler et Calvard, cette méthode n'est pas robuste au bruit et produit un grand nombre de faux positifs dans les tests de Martínez-de Dios et Ollero [61].

1.2.2.d Seuillage multirésolution de Martínez-de Dios

Martínez-de Dios *et al.* [60] proposent un système de vision réalisant une compensation lumineuse et un nouvel algorithme de recherche de seuil pour la détection de feux de forêt.

Les auteurs utilisent des caméras de l'infrarouge moyen ($[3 ; 5] \mu m$) pour leurs expériences. Les paramètres des caméras tels que le gain et le niveau de luminosité sont modifiés automatiquement en fonction de la luminosité et du niveau de contraste de l'image. Ce principe permet de garder des paramètres d'intensité lumineuse identiques dans les images acquises au cours du temps et de ne pas les saturer si des changements d'intensité lumineuse dans la scène interviennent. Des valeurs d'intensité lumineuse et de contraste de référence sont choisies et utilisées pour mesurer l'écart avec les paramètres de l'image courante. Les paramètres de la caméra sont ensuite modifiés en fonction de cet écart. Par exemple, le niveau de luminosité de la caméra est réduit si l'intensité lumineuse de l'image est plus élevée que celle de référence.

Pour détecter les pixels « feu » dans les images du spectre infrarouge, les auteurs proposent un algorithme de seuillage utilisant l'histogramme h de l'image à traiter [59]. Une technique de décomposition multirésolution est appliquée à l'histogramme afin de le simplifier et de rechercher le seuil plus efficacement. En effet, si la résolution d'un histogramme est réduite,

celui-ci aura moins de détails et il sera plus simple de le découper en régions. L'algorithme recherche donc sur plusieurs itérations le meilleur seuil pour détecter le feu d'une image avec un niveau de résolution croissant $l \in \{N, N-1, \dots, 0\}$ où N est le niveau minimal de résolution. La décomposition en ondelettes est utilisée pour réduire la résolution de l'histogramme grâce à l'ondelette de *Haar* [62]. Le principe de cette décomposition est détaillée dans la partie 5.2.2.d. La recherche de seuil consiste à trouver la région de l'histogramme dont les intensités correspondent à l'intensité du feu (dite « région d'intérêt »). Voici l'algorithme de cette recherche :

À chaque niveau de résolution, la région d'intérêt est choisie à l'aide d'une fonction f_1 définie par :

$$f_1 : \{\Phi_l\} \rightarrow \{\Phi_{l-1}\} \quad (1.29)$$

Où $\{\Phi_l\}$ est un ensemble de valeurs entières d'intensité qui représente la région d'intérêt de l'histogramme qui a été estimée à la résolution antérieure et $\{\Phi_{l-1}\}$ est un ensemble de valeurs d'intensités qui représente la région d'intérêt de l'histogramme au niveau de décomposition supérieur. L'algorithme est initialisé avec $\{\Phi_N\} = \left[0; \frac{L}{2^N}\right]$ avec L le nombre de valeurs d'intensité de l'image (256 pour des images codées sur 8 bits).

Le processus d'obtention de la fonction f_1 nécessite plusieurs étapes. La première étape consiste à chercher des minimums locaux p_l^i dans $\{\Phi_l\}$. Une valeur d'intensité $k \in \{\Phi_l\}$ est considérée comme un minimum local de l'histogramme h si les conditions suivantes sont vérifiées :

$$h(k+1) \geq 1,05 \times h(k) \quad (1.30)$$

$$h(k-1) \geq 1,05 \times h(k) \quad (1.31)$$

$$h(k) > 0,001 \times n_C \times n_L \quad (1.32)$$

Avec n_C le nombre de colonnes de l'image et n_L le nombre de lignes de l'image.

Les valeurs d'intensité contenues dans $\{\Phi_l\}$ respectant les conditions ci-dessus sont considérées comme des minimas locaux pour le niveau l et sont notées p_l^i . L'ensemble $\{\Phi_l\}$ est alors subdivisé en sous-ensembles $\{\omega_l^i\} = [p_l^{i-1}; p_l^i] \subseteq \{\Phi_l\}$. Les histogrammes de plus faible résolution auront un faible niveau de bruit, permettant d'obtenir un nombre restreint de minimas locaux dans les premières étapes de la recherche de seuil.

Une fois que les minimas locaux de $\{\Phi_l\}$ ont été trouvés, le calcul de $\{\Phi_{l-1}\}$ se fait par la sélection d'une ou plusieurs régions $\{\omega_l^i\}$ estimés dans l'étape précédente. $\{\Phi_{l-1}\}$ est donc l'union de certaines régions $\{\omega_l^i\}$. Pour décider quelles sont les régions $\{\omega_l^i\}$ qui doivent être

retenues pour former l'ensemble $\{\Phi_{l-1}\}$, un système de logique floue [63] est utilisé. Les entrées de ce système flou sont la largeur $\mathcal{L}_{\omega_l^i}$ et l'intensité maximale $\mathcal{I}_{\omega_l^i}$ de la région $\{\omega_l^i\}$ considérée. Ces valeurs sont données par :

$$\mathcal{L}_{\omega_l^i} = p_l^i - p_l^{i-1} \quad (1.33)$$

$$\mathcal{I}_{\omega_l^i} = \arg \max_{k \in \{\omega_l^i\}} (h_l(k)) \quad (1.34)$$

La sortie $y \in [0, 1]$ du système flou représente la possibilité d'unir la région $\{\omega_l^i\}$ à la région $\{\omega_l^{i-1}\}$. Les régions sont unies si $y > \alpha$, avec $\alpha = 0,6$ un seuil choisi empiriquement par les auteurs. Les valeurs $\mathcal{L}_{\omega_l^i}$ et $\mathcal{I}_{\omega_l^i}$ sont normalisées entre 0 et 1 en les divisant par 2^{8-l} . Les règles permettant de construire le système flou pour déterminer y sont les suivantes :

$(\mathcal{I}_{\omega_l^i} \text{ est bas})$	et	$(\mathcal{L}_{\omega_l^i} \text{ est bas})$	\Rightarrow	$y \text{ est haut}$
$(\mathcal{I}_{\omega_l^i} \text{ est bas})$	et	$(\mathcal{L}_{\omega_l^i} \text{ est moyen})$	\Rightarrow	$y \text{ est haut}$
$(\mathcal{I}_{\omega_l^i} \text{ est bas})$	et	$(\mathcal{L}_{\omega_l^i} \text{ est haut})$	\Rightarrow	$y \text{ est moyen}$
$(\mathcal{I}_{\omega_l^i} \text{ est moyen})$	et	$(\mathcal{L}_{\omega_l^i} \text{ est bas})$	\Rightarrow	$y \text{ est haut}$
$(\mathcal{I}_{\omega_l^i} \text{ est moyen})$	et	$(\mathcal{L}_{\omega_l^i} \text{ est moyen})$	\Rightarrow	$y \text{ est moyen}$
$(\mathcal{I}_{\omega_l^i} \text{ est moyen})$	et	$(\mathcal{L}_{\omega_l^i} \text{ est haut})$	\Rightarrow	$y \text{ est moyen}$
$(\mathcal{I}_{\omega_l^i} \text{ est haut})$	et	$(\mathcal{L}_{\omega_l^i} \text{ est bas})$	\Rightarrow	$y \text{ est moyen}$
$(\mathcal{I}_{\omega_l^i} \text{ est haut})$	et	$(\mathcal{L}_{\omega_l^i} \text{ est moyen})$	\Rightarrow	$y \text{ est bas}$
$(\mathcal{I}_{\omega_l^i} \text{ est haut})$	et	$(\mathcal{L}_{\omega_l^i} \text{ est haut})$	\Rightarrow	$y \text{ est bas}$

La figure 1.2 présente la surface de décision du système de logique flou résultant des neuf règles.

La région obtenue est ensuite reconstruite au niveau supérieur pour obtenir $\{\Phi_{l-1}\}$. Cette région est l'initialisation du niveau supérieur et ainsi de suite jusqu'au niveau 0. Le schéma de la figure 1.3 récapitule les étapes de la fonction f_1 pour obtenir $\{\Phi_{l-1}\}$ à partir de $\{\Phi_l\}$.

Une fois que le processus a été accompli du niveau N au niveau 0, la valeur du seuil est ensuite calculée de la façon suivante :

Soit g le nombre de régions assemblées dans la région $\{\omega_0^i\}$; On a : $\{\Phi_0\} = [p_0^g; p_0^1]$.

Le seuil s est donné par :

$$s = \begin{cases} p_0^1 & \text{si } g = 0 \\ p_0^g + (1 - \beta)p_0^1 & \text{sinon} \end{cases} \quad (1.35)$$

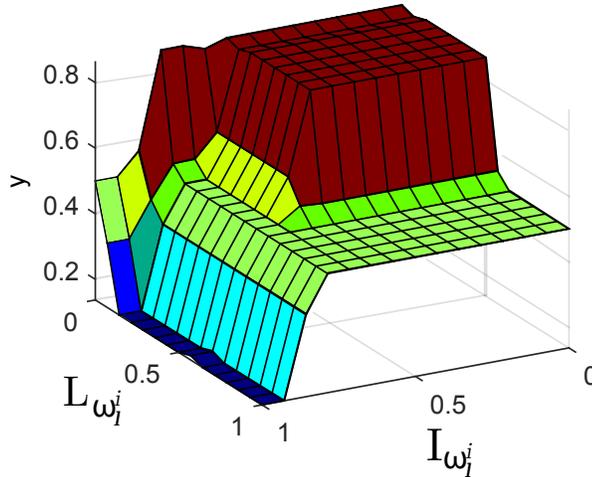


FIGURE 1.2 : Superficie de décision du système flou de Martínez.

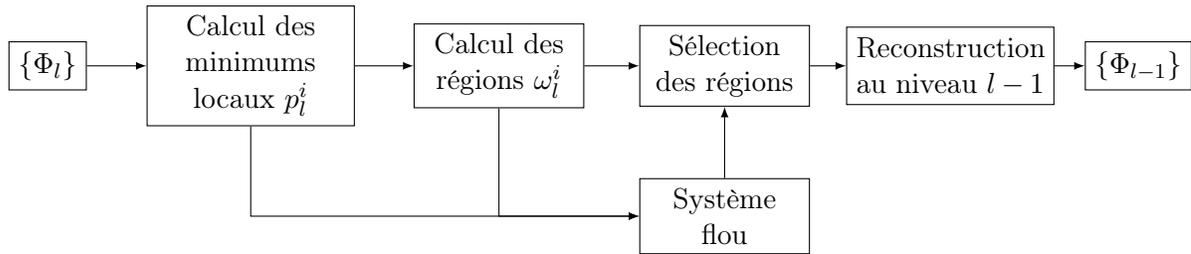


FIGURE 1.3 : Schéma explicatif de la procédure permettant l'obtention de la fonction f_1 .

Avec $\beta = 0,7$ une variable fixé par les auteurs.

Le schéma de la figure 1.4 résume les étapes de la recherche du seuil.



FIGURE 1.4 : Schéma explicatif de la méthode de seuillage multi-résolution de Martínez.

Dans [56] Martínez teste les algorithmes précédemment cités sur une image de feu en infrarouge (voir figure 1.5). Dans cet exemple la zone de pixels de feu de l'image de l'infrarouge est très mal détectée par les algorithmes d'Otsu et de Ridler et Calvard. Une des raisons qui peut expliquer la faible efficacité de ces algorithmes est le fait que l'illumination de la scène n'est pas uniforme, alors que l'uniformité d'illumination dans une scène est une condition nécessaire pour l'utilisation du seuillage d'Otsu. Cela montre les inconvénients de ces deux derniers

algorithmes car ils ne sont pas adaptés à tout type d'image de l'infrarouge.

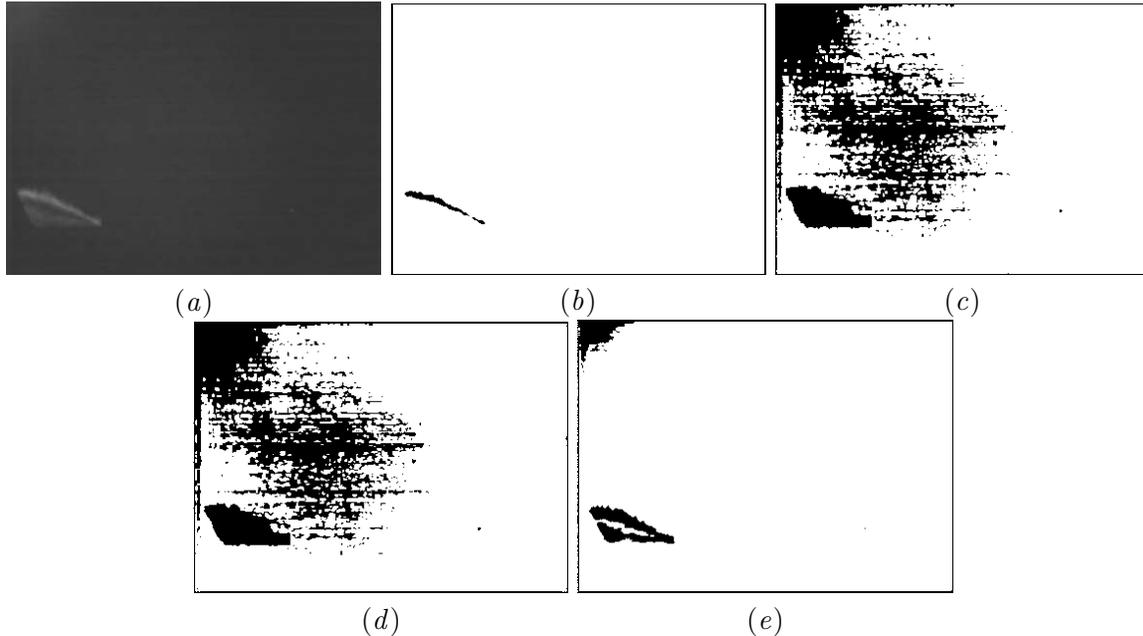


FIGURE 1.5 : Comparaison par Martínez de Dios *et al.* [56] de l'efficacité des algorithmes de seuillage de la littérature sur une image infrarouge de feu. (a) image de l'infrarouge, (b) seuillage de Martínez *et al.* , (c) seuillage d'Otsu, (d) seuillage de Ridler et Calvard et (e) seuillage de Kittler et Illingworth.¹

1.2.2.e Algorithme des contours actifs

La détection par contours actifs [64] est une technique fréquemment utilisée pour la segmentation d'image. Cette technique permet d'obtenir une détection qui se base sur les contours des objets. Le feu ayant des contours très marqués dans les images du spectre infrarouge, il nous a semblé intéressant d'évaluer son efficacité de détection de pixels de feu.

L'algorithme de contours actifs détecte les contours d'un objet dans une image en s'adaptant itérativement à la forme de celui-ci. Une description détaillée de cette méthode est faite par Rousselle dans [65]; en voici une description succincte.

Le but de l'algorithme des contours actifs est de trouver un contour c minimisant la fonction E^* définie par :

$$E^*(c) = E_i(c) + E_e(c) + E_c(c) \quad (1.36)$$

1. Dans la figure 1.5, les légendes (d) et (e) sont l'inverse de celles présentées dans [56] mais d'après nos tests sur l'image (a), l'image (e) correspond au seuillage de Kittler et Illingworth et l'image (d) à celui de Ridler et Calvard.

Où $E_i(c)$, $E_e(c)$ et $E_c(c)$ sont respectivement des fonctions d'énergie interne, externe et de contexte.

La fonction d'énergie interne se calcule à partir de la norme des dérivées première et seconde du contour. Le contour c contient N valeurs tel que $c_i = (u_i, v_i)$ soit la coordonnée du $i^{\text{ème}}$ point de c . La norme de la dérivée première discrète du contour au point s est donnée par la formule suivante :

$$\left\| \frac{dv}{ds}(s) \right\|^2 = \|c_s - c_{s-1}\|^2 = (u_s - u_{s-1})^2 + (v_s - v_{s-1})^2 \quad (1.37)$$

Et la norme de la dérivée seconde discrète au point s est donnée par la formule :

$$\left\| \frac{d^2v}{ds^2}(s) \right\|^2 = \|c_{s-1} - 2c_s + c_{s+1}\|^2 = (u_{s-1} - 2u_s + u_{s+1})^2 + (v_{s-1} - 2v_s + v_{s+1})^2 \quad (1.38)$$

La formule permettant de calculer la fonction d'énergie interne est :

$$E_i(c) = \frac{1}{2} \sum_{s=0}^N \alpha(s) \left\| \frac{dv}{ds}(s) \right\|^2 + \beta(s) \left\| \frac{d^2v}{ds^2}(s) \right\|^2 \quad (1.39)$$

Avec $\alpha(s)$ et $\beta(s)$ des poids définis par l'utilisateur. α permet de pénaliser les discontinuités le long de la courbe, tandis que β favorise le caractère lisse de la courbe.

L'énergie externe de la courbe se calcule avec les caractéristiques de l'image. Ceci peut se faire par des calculs de gradient ou d'intensité de l'image ou encore en utilisant un flux de vecteurs gradient généralisé (ou GGVF de l'anglais *Generalized Gradient Vector Flow*) [66].

L'énergie de contexte sert à introduire des connaissances *a priori* sur le contour. En effet, lors de l'utilisation de l'algorithme des contours actifs, le contour a tendance à se rétracter [65]. Or, il est généralement recherché d'étendre le contour au fil des itérations ou bien d'accélérer la rétractation. L'énergie de contexte peut être définie par l'utilisateur. Dans certaines implémentations, l'énergie ballon [67] est utilisée comme énergie de contexte.

Pour initialiser cet algorithme il faut spécifier une zone d'intérêt. Le contour de cette zone sera l'initialisation de l'algorithme. Dans le cas de la détection de pixels « feu », comme les pixels de plus grande intensité appartiennent au feu, l'algorithme peut être initialisé avec la plus grande région contenant des pixels dont l'intensité est supérieure à un seuil.

Pour trouver un contour minimisant la fonction E^* , il faut appliquer une méthode de minimisation. Les trois méthodes les plus utilisées pour la minimisation sont l'algorithme de descente en gradient [64], la programmation dynamique [68] et l'algorithme glouton [69].

1.2.3 Méthodes de détection de pixels de feu basées sur la fusion d'informations d'intensité de pixels obtenues dans différents spectres

La détection multimodale de pixels de feu consiste à utiliser l'intensité de pixel d'images acquises simultanément dans plusieurs modalités spectrales.

Meléndez *et al.* [70] proposent de travailler avec des images multimodales acquises avec une caméra munie d'un filtre rotatif dans des sous-bandes de l'infrarouge moyen (MIR) dont les centres sont $F1 = 4,8 \mu m$, $F2 = 4,2 \mu m$, et $F3 = 3,6 \mu m$ et de largeur $0,8 \mu m$. Les auteurs proposent de classer les pixels des images en cinq zones : flamme (N°1), front de feu (N°2), braises chaudes (N°3), braises froides (ou suies) (N°4) et arrière-plan (N°5). Pour classer les pixels dans ces zones, ils ont considéré dans un premier temps la différence d'intensité de pixel entre les images obtenues dans les bandes $F2$ et $F3$. Cette méthode obtient de faibles scores de détection pour des images de feux lointains du fait des atténuations atmosphériques (voir partie 2.2.1.a). Pour pallier ce problème, les auteurs ont développé une autre méthode qui utilise les images acquises dans les trois sous-bandes $F1$, $F2$ et $F3$ en leur appliquant une fusion par Analyse en Composantes Principales (ACP) [71]. Le principe de l'analyse en composantes principales est donné dans la partie 5.2 de cette thèse. Dans le cadre de cette explication, nous retiendrons que cette analyse permet de calculer les poids (ACP1, ACP2, ACP3, ACP4, ACP5 et ACP6) des combinaisons linéaires d'images générant deux nouvelles images décorréliées (PC1 et PC2) selon les formules :

$$PC1 = ACP1 \times F1 + ACP2 \times F2 + ACP3 \times F3 \quad (1.40)$$

$$PC2 = ACP4 \times F1 + ACP5 \times F2 + ACP6 \times F3 \quad (1.41)$$

L'image de la première composante PC1 contient les données avec la plus grande variance d'intensité de pixel et l'image de la seconde composante PC2 contient les données avec la deuxième plus grande variance.

Un nuage de points $P_{\mathbf{x}}$ est alors construit tel que $P_{\mathbf{x}} = (PC1(\mathbf{x}), PC2(\mathbf{x}))$. Une étude a été menée par les auteurs pour classer ces points dans les cinq zones évoquées précédemment et en fonction de la distance feu-caméra. La figure 1.6 présente le résultat de cette étude. A titre d'exemple, la figure 1.7 montre l'exploitation des données d'un nuage de points pour la détection des différentes zones d'une image de feu.

1.2.4 Méthodes de détection de pixels de feu utilisant des critères complémentaires à l'intensité de pixels

Les paragraphes précédents ont montré que la couleur (ou l'intensité dans les images de l'infrarouge) de pixel permet de détecter des pixels « feu » avec une bonne efficacité. Il existe un nombre important de travaux traitant de la détection d'incendie qui utilisent des caractéristiques dynamiques supplémentaires telles que la détection de mouvement, la variation de taille,

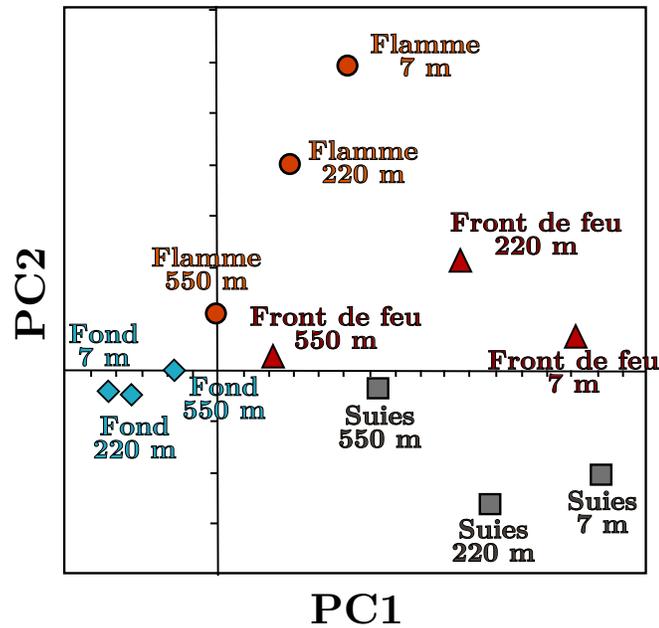


FIGURE 1.6 : Classification des zones de feu par Meléndez *et al.* [70]

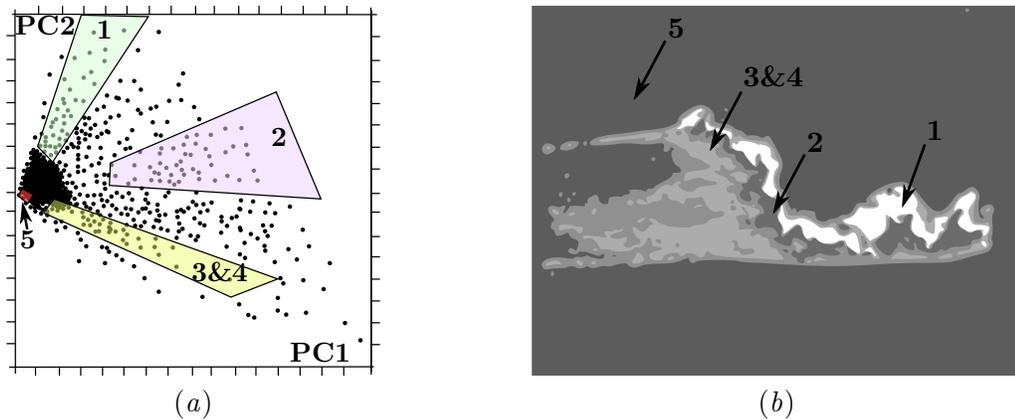


FIGURE 1.7 : Détection multimodale de zones de feu par ACP : (a) Nuage de points PC1-PC2 et (b) détection du feu en zones par Meléndez *et al.* [70]

de texture et l'oscillation de régions de pixels [72, 25] afin de limiter les fausses détections de départ de feu.

La détection de région de pixels en mouvement dans l'image peut être réalisée par différentiation d'images successives [73] acquises avec une fréquence d'au moins 25 Hz. Cette technique a pour inconvénient d'être sensible aux changements de luminosité, aux petits mouvements (comme ceux produits par la végétation ou de petits mouvements de caméra) et de considérer souvent le centre de feu comme faisant partie de l'environnement car il bouge peu. Pour extraire l'arrière-plan avec plus de robustesse, une méthode de densité de mélange gaussien (*Gaussian Mixture Models* : GMM) [74] peut être envisagée. La figure 1.8 présente les résultats

obtenus par Truong et Kim [75] de l'extraction de l'arrière-plan d'une image de flamme par les méthodes de différenciation et de GMM. Dans cet exemple, la détection par GMM donne de bons résultats. Il est à noter toutefois qu'en présence de fumée et d'objets en mouvement, cette méthode peut générer des pixels faux positifs. Son efficacité peut être améliorée en la couplant à une détection couleur de pixels pour les images acquises dans le spectre visible.

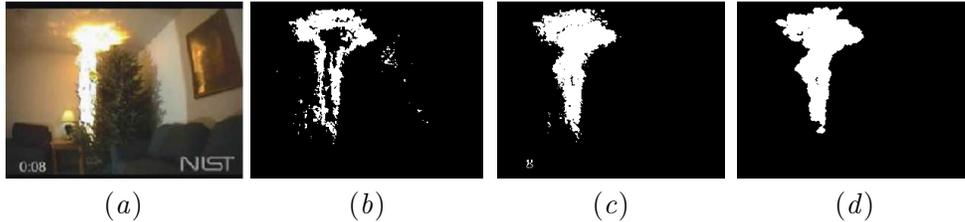


FIGURE 1.8 : Extraction d'arrière-plan grâce aux caractéristiques dynamiques : (a) Image originale, (b) image obtenue par différenciation d'images, (c) résultat de la GMM et (d) résultat d'opérations morphologiques sur le résultat (c) [75].

Une flamme a un mouvement de vacillement rapide autour de sa source avec une fréquence de 10 Hz quels que soient les matériaux et combustibles impliqués dans le processus [76]. Töreyn *et al.* [29] proposent d'identifier les contours de régions de pixels de feu en détectant, grâce à une décomposition en ondelettes, des variations rapides de l'intensité de pixel dans le canal rouge du système RVB (ou le canal Y du système YUV). Dans [77], ces mêmes auteurs utilisent également ce procédé pour détecter les pixels de contour de flammes dans l'infrarouge (8-12 μm). La figure 1.9 montre les spectres des signaux d'intensité d'un pixel d'images acquises avec une fréquence de 10 Hz et présentant des amplitudes hautes à des fréquences très basses dues à un repliement spectral de contributions autour de 8 Hz.

Une autre caractéristique de flamme peut être considérée pour éliminer les zones de pixels faux positifs : il s'agit de la forme des régions de feu. En effet, les contours des zones de feu ont la particularité d'avoir des formes complexes qui peuvent être identifiées grâce à leur décomposition en série de Fourier. Ceci a été proposé par Liu et Ahuja [39] qui utilisent les coordonnées des pixels du contour (u_i, v_i) pour obtenir les coordonnées complexes de la forme $\{z_i | z_i = u_i + jv_i\}_{(i=1)}^N$ avec N le nombre de points de contour et calculer les coefficients de Fourier de la forme du contour. Un apprentissage par machines à supports vectoriels sur 1319 contours de région de l'environnement et 1089 contours de zones de feu a permis de déterminer les coefficients spécifiques des contours de feux.

On peut aussi distinguer le contour de flamme du contour d'autres objets et enlever des fausses détections en étudiant la signature des objets [77]. La signature d'un objet est la représentation en une dimension de son contour. Elle est calculée à partir des distances entre le contour et le centre de gravité de la forme. Dans [75], les moments invariants de Hu [78] sont appliqués pour détecter les régions de feu. Dans [79], le périmètre P et l'aire A du feu sont utilisés pour calculer la métrique T avec :

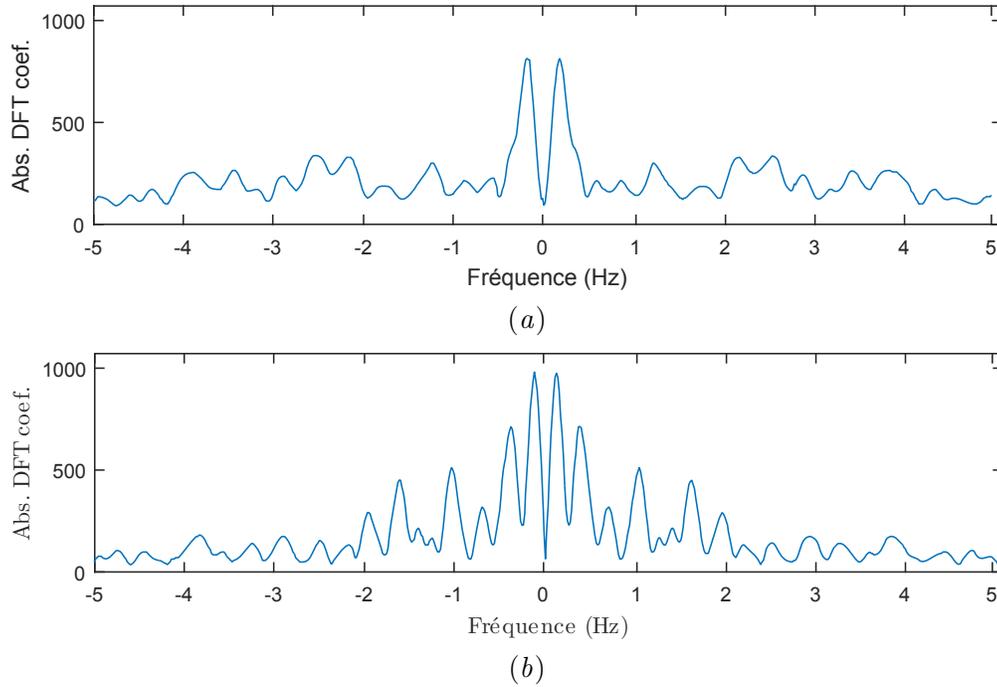


FIGURE 1.9 : Distribution de fréquences de vacillement pour une fréquence d'acquisition de dix Hertz (a) dans le spectre visible, et (b) dans le spectre infrarouge [77].

$$T = \frac{P}{2\sqrt{\pi A}} \quad (1.42)$$

La valeur de cette métrique est proche de 1 pour les objets circulaires, ce qui permet d'éliminer les zones brillantes générant de potentielles fausses détections comme celles correspondant au soleil ou à ses réflexions dans les vitres de camion.

La variation temporelle de la forme de contour de zones de feu peut également être étudiée car des changements aléatoires et rapides sont aussi caractéristiques. Dans [79], les variations d'aire ΔA sont calculées à partir de la formule suivante :

$$\Delta A = \frac{|A_i - A_{i-1}|}{A_i} \quad (1.43)$$

Avec A_i l'aire à l'instant i et A_{i-1} l'aire à l'instant $i - 1$

Si ΔA est grand, la zone est considérée comme étant une zone de feu.

Un autre moyen d'éliminer les zones de pixels ayant les mêmes caractéristiques de couleur et de mouvement que les zones de pixels de feu est d'étudier l'uniformité des couleurs. En effet, les variations temporelles et spatiales des couleurs des zones de flamme sont bien plus

importantes que celles d'objets. Un calcul de la variance appliqué sur les couleurs permet de discerner la flamme d'objets qui ont des couleurs rouges-orangé [75].

L'approche de segmentation de texture dynamique peut également être utilisée [80] pour détecter des régions de pixels de feu. Cette approche suppose que les pixels d'une même zone suivent un processus de Gauss-Markov et consiste à segmenter une image en étudiant les statistiques spatio-temporelles des régions texturées.

Les caractéristiques spatio-temporelles des textures des feux dépendent particulièrement du temps d'ouverture des caméras utilisées et de la fréquence d'acquisition des images. De plus, le vent, le combustible, l'environnement et les fumées jouent aussi un rôle dans la valeur de ces caractéristiques. Une approche pour segmenter les textures de feu est de procéder à un apprentissage des statistiques spatio-temporelles des zones de feu à partir des premières images de la vidéo. La figure 1.10 présente la segmentation de flamme réalisée avec cette méthode par Doretto *et al.* et présentée dans [80] pour un cas de segmentation complexe d'une texture dynamique de feu sur une texture dynamique de mer. Dans cet article, il apparaît que les zones de feu ne sont pas toujours correctement segmentées dans les images car la méthode suppose que les zones de textures restent aux mêmes emplacements ce qui n'est pas le cas pour les flammes dont la forme varie constamment.

Pour envisager l'application de cette méthode sur des images de l'infrarouge il faut considérer des images avec une bonne résolution thermique afin de distinguer les variations de température au sein des zones de feu.

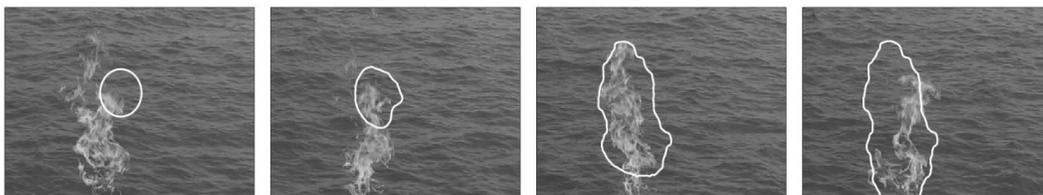


FIGURE 1.10 : Segmentation dynamique de flamme. Le cercle de la première image représente la zone d'apprentissage dans un contexte de segmentation sans *a priori* [80].

Cette sous-section a présenté des caractéristiques de pixels de feu exploitables pour leur détection. Les travaux sur l'exploitation de ces dernières ont principalement été effectués sur des images du spectre visible. L'extraction d'arrière-plan et la segmentation de texture permettent directement de générer une carte binaire de pixels détectés pouvant être considérée comme une première étape d'un processus plus complexe de détection. Les propriétés de contour et de forme peuvent être utilisées sans tenir compte du caractère dynamique du feu et permettent d'éliminer de fausses détections. Les caractéristiques de vacillement et de changement de taille de feu peuvent également être utilisées pour éliminer de fausses détections mais nécessitent des séquences d'images. Enfin, bien que ces techniques aient été testées majoritairement sur des images du spectre visible, la plupart d'entre elles sont envisageables sur des images de

l'infrarouge.

1.2.5 Métriques d'évaluation des performances des méthodes de détection de pixels de feu

Cette sous-section présente les métriques permettant d'évaluer les performances des méthodes présentées ci-avant dans le cadre de la détection de pixels de feu. L'évaluation de ces techniques peut se rapporter à l'évaluation de leur capacité de segmenter des zones de feu dans les images. Aussi, on considèrera que le résultat de performance de détection de pixel « feu » est équivalent au résultat de performance de segmentation de zones de feu dans les images.

Il existe deux familles de méthodes d'évaluation de techniques de segmentation : la première utilise une approche non supervisée qui permet de faire l'évaluation sans avoir recours à une évaluation par un expert [81], [82] ; la deuxième est basée sur une approche supervisée qui repose sur une segmentation faite à la main par un expert [83], [84]. Dans le cadre de cette thèse, la deuxième famille a été utilisée.

Une image segmentée manuellement est appelée « Vérité Terrain » (en anglais *Ground Truth*) et notée I_{VT} ; $I_{VT}(\mathbf{x}) = 1$ si \mathbf{x} est un pixel « feu » et $I_{VT}(\mathbf{x}) = 0$ sinon. Plusieurs métriques sont utilisées dans la littérature pour comparer la performance d'une segmentation manuelle par rapport à celle obtenue automatiquement par traitement d'image. Le choix d'une métrique dépend du contexte de segmentation. Des comparaisons de performances de méthodes de détection de pixels de feu ont été faites dans [85, 35, 30, 86] et trois métriques d'évaluation utilisées dans ces articles sont décrites dans cette sous-section. Ces travaux utilisent aussi les critères de Martin [87] que nous ne détaillerons pas dans cette partie. En effet, ces critères ne prenant pas en compte la sur ou sous-segmentation [88] et ne nous semblent pas adaptés à notre étude.

1.2.5.a Coefficient de corrélation de Matthews

Le coefficient de corrélation de Matthews (MCC) est utilisé pour l'évaluation de méthodes de segmentation d'images [89] et de manière spécifique pour l'évaluation de méthodes de détection de pixels de feu dans [30]. Il se définit par :

$$MCC = \frac{(VP * VN) - (FP * FN)}{\sqrt{(VN + FN)(VN + FP)(VP + FN)(VP + FP)}} \quad (1.44)$$

Où VP est le nombre de vrais positifs, c'est-à-dire le nombre de pixels « feu » correctement détectés ; VN est le nombre de vrais négatifs, c'est-à-dire le nombre de pixels « environnement » correctement détectés ; FP est le nombre de faux positifs, c'est-à-dire le nombre de pixels détectés comme « feux » alors que ce sont des pixels « environnement » et FN est le nombre de pixels détectés comme « environnement » alors que ce sont des pixels « feux ».

1.2.5.b F-Score

La métrique F-Score (F), aussi appelée F1-Score, est principalement utilisée dans le domaine de la recherche d'information [90] et dans le domaine de la segmentation d'images [91, 30]. Le F-Score fait intervenir deux mesures appelées « précision » (Pr) et « rappel » (Ra) définies ainsi :

$$Pr = \frac{TP}{TP + FP}, Ra = \frac{TP}{TP + FN} \quad (1.45)$$

Le F-Score est la moyenne harmonique de la précision et du rappel :

$$F = 2 * \frac{Pr * Ra}{Pr + Ra} \quad (1.46)$$

1.2.5.c Indice de qualité de Hafiane

L'indice de qualité de Hafiane (HAF) a été développé pour l'évaluation de méthodes de détection de pixels de feu [83]. Il prend en compte la position, la forme et la taille des régions segmentées ainsi que la sur et sous-segmentation. Il nécessite d'une part, le calcul d'un indice d'appariement défini par :

$$M = \frac{1}{\text{Card}(\mathbf{I}_S)} \sum_{j=1}^{NR_s} \frac{\text{Card}(R_{i^*}^{VT} \cap R_j^S) \times \text{Card}(R_j^S)}{\text{Card}(R_{i^*}^{VT} \cup R_j^S)} \quad (1.47)$$

Où NR^S est le nombre de régions segmentées dans l'image \mathbf{I}_S , R_j^S est la j^{ime} région segmentée et $R_{i^*}^{VT}$ est la région dans la vérité terrain \mathbf{I}_{VT} qui a la plus grande surface en commun avec la région R_j^S : $i^* = \max_i(\text{Card}(R_i^{VT} \cap R_j^S))$

D'autre part, le calcul d'un indice intégrant la sur ou sous-segmentation dont la formule est :

$$\eta = \begin{cases} NR^{VT}/NR^S & \text{si } NR^S \geq NR^{VT} \\ \log(1 + NR^S/NR^{VT}) & \text{autrement} \end{cases} \quad (1.48)$$

Où NR^{VT} et NR^S sont le nombre de régions segmentées respectivement dans les images \mathbf{I}_{VT} et \mathbf{I}_S .

Enfin, sa formule est donnée par l'équation suivante :

$$HAF = \frac{M + m \times \eta}{1 + m} \quad (1.49)$$

Où m est un facteur de pondération fixé à 0,5.

Cet indice est compris entre 0 et 1 et plus sa valeur est haute, meilleure est la méthode de détection évaluée.

1.2.6 Évaluations comparatives des performances de certains algorithmes de détection de pixels de feu

Trois articles de la littérature présentent des évaluations comparatives de performance des algorithmes de détection de pixels de feu de Chen [41], Celik [33], Ko [31], Rossi [43] et Collumeau [30] [85, 30, 35]. Les trois études présentées ont été faites sur des ensembles d'images différents : la base d'images utilisée dans [85] est constituée de dix images capturées lors d'un brûlage expérimental sur une surface de végétal de cinquante mètres de long. Celle utilisée dans [30] contient soixante-seize images de feux de végétation (échelle terrain) collectées sur Internet. Enfin, la base d'images utilisée dans [35] est constituée d'images de feu (nombre non précisé) récupérées sur Internet et d'expériences de feux de végétation.

Le tableau 1.4 présente les indices de qualité de Hafiane recueillis par ces méthodes dans les articles cités. Les résultats de ces évaluations sont hétérogènes et le peu d'informations fournies sur les images utilisées pour les évaluations ne permettent pas d'établir des conclusions quant aux performances des méthodes de détection.

TABLEAU 1.4 : Indice de qualité de Hafiane obtenu par cinq méthodes de détection de pixels de feu évaluées sur des bases d'images différentes.

	Chen [41]	Celik [33]	Ko [31]	Rossi [43]	Collumeau [30]	Rudz [35]
Rudz 2009 [85]	0,793	0,429	0,369	0,320	-	-
Collumeau 2011 [30]	0,767	0,742	0,607	0,685	0,785	-
Rudz 2013 [35]	0,597	0,572	-	0,528	-	0,659

On citera encore une fois brièvement le travail Chino *et al.* [46] qui comparent leur méthode de détection avec celles de Celik *et al.* [33], T.H. Chen *et al.* [41] Rossi *et al.* [43] et Rudz *et al.* [35] en utilisant un ensemble d'images composé de cent dix-neuf images de feu et cent sept images sans feu et le F-Score comme métrique d'évaluation. Il apparait dans [46] que les méthodes de Celik et de Rudz obtiennent de meilleurs scores que celle de Chen et de Rossi (les données chiffrées des résultats ne sont pas fournies). Le classement des méthodes est donc encore différent de celui des autres évaluations, confirmant ainsi que les évaluations présentées dans la littérature ne peuvent pas servir de référence car elles ne sont pas faites sur des images caractérisées.

L'évaluation de détection de pixels de feu dans le domaine infrarouge n'a quant à elle, jamais été faite.

1.3 Techniques de mesure pour les feux

Plusieurs techniques ont été utilisées dans la littérature pour extraire les caractéristiques géométriques d'un feu à partir d'outils de vision. Certaines n'utilisent qu'une seule caméra, d'autres un réseau de caméra et d'autres encore de la stéréovision. Des travaux de mesure du feu ont été faits en laboratoire [20, 42, 21]. Les conditions en laboratoire sont spécifiques car l'environnement est maîtrisé, les distances feu - caméra sont petites et il n'y a généralement pratiquement pas de fumée. Aussi, dans cette partie du rapport, nous choisissons de ne présenter que les travaux ayant été conduits en extérieur à une échelle d'une parcelle de terrain ou d'une vallée.

Cette section présente dans une première partie les différentes caractéristiques du feu qui intéressent les physiciens travaillant sur des modèles de comportement de feu. La seconde partie détaille les méthodes de la littérature estimant certaines de ces caractéristiques à partir d'une seule caméra. La troisième sous-section est consacrée aux techniques utilisant plusieurs caméras et la dernière partie à celles qui utilisent la stéréovision. Chacune des techniques présentées fait suite à une première étape de traitement qui a permis de segmenter des zones de feux dans les images.

1.3.1 Caractéristiques géométriques d'un front de feu

Dans cette partie nous présentons les différentes caractéristiques utilisées dans les modèles de comportement des feux de surface [13, 10, 11, 14, 15, 16, 17, 18].

Un feu en propagation est caractérisé par plusieurs données géométriques telles que la hauteur, la longueur et l'angle d'inclinaison de la flamme, la position et l'épaisseur du front de feu. Ces caractéristiques sont présentées dans la figure 1.11.

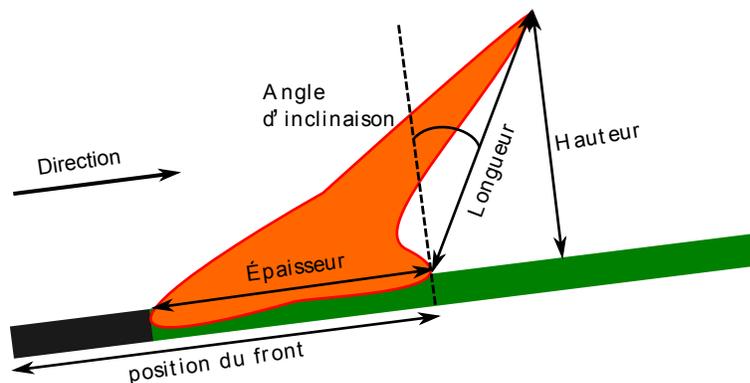


FIGURE 1.11 : Caractéristiques géométriques principales d'un front de feu.

La hauteur de la flamme correspond à la longueur du segment compris entre le sommet de la flamme et la projection orthogonale de ce dernier sur le haut de la couche de combustible [92]. La longueur de la flamme peut être définie de plusieurs manières [93] ; la plus communément

utilisée est celle qui la présente comme la distance entre le sommet de la flamme et la base du front de feu délimitant la zone brûlée et imbrulée de la végétation. L'angle d'inclinaison représente l'angle compris entre la perpendiculaire au sol et la droite passant par le sommet de la flamme et la base de cette dernière. Enfin, l'épaisseur de la flamme est la distance entre l'arrière et l'avant du front de feu.

D'autres grandeurs permettent de caractériser un incendie comme par exemple sa vitesse de propagation (ou ROS de l'anglais *Rate Of Spread*) son périmètre, sa surface et son volume. La vitesse définit la rapidité de propagation du feu et permet d'évaluer la distance parcourue par l'incendie pendant une durée donnée. Cette grandeur représente une moyenne du fait des changements continus des conditions météorologiques et topographiques intervenant au cours de la propagation d'un feu de végétation. Le périmètre du feu correspond à la longueur du contour du feu au sol. La surface du front de flamme se calcule avec l'aire de la surface frontale du feu. Cette surface est parfois simplifiée pour correspondre à un parallélogramme [11], facilitant ainsi le calcul de l'aire. Enfin, le volume du feu est la caractéristique la plus difficile à estimer et correspond au volume en mètre cube occupé par le feu considéré.

Les modèles de comportement de feux de végétation nécessitent d'estimer certaines caractéristiques géométriques afin d'obtenir une meilleure connaissance des phénomènes intervenant lors de la propagation d'un incendie et de les améliorer et/ou les valider.

1.3.2 Mesures estimées à partir d'images obtenues depuis un même point de vue

A notre connaissance, il n'existe pas de travaux de recherche relatifs à l'utilisation de caméras monomodes ou multimodes positionnées au sol au même endroit pour l'estimation de caractéristiques d'un feu en propagation en extérieur. Aussi, cette section va présenter les travaux qui se rapprochent le plus de la problématique de cette thèse et qui concernent la détermination de la vitesse de propagation d'un front et de sa direction à partir d'images aériennes obtenues par satellite ou par avion.

Ononye *et al.* présentent dans [94] une méthode d'estimation du périmètre, de la ligne active et de la direction de front de feux de grandes dimensions à partir d'images satellitaires multimodales (figure 1.12). Pour estimer la ligne du front de feu active, les auteurs utilisent une opération de gradient sur chaque bande spectrale et calculent la courbe du front à partir des contours obtenus. Le calcul de la direction de propagation se fait en deux étapes : d'abord le calcul de l'orientation du vecteur de direction puis du sens de propagation. La direction est calculée pour plusieurs points du front de feu. Pour estimer l'orientation en certains points de la ligne de front, la normale de la courbe du front de feu en ces points est calculée. La ligne de front obtenue sépare la zone de végétation en feu de la zone qui n'a pas encore été brûlée. Cette information permet de déterminer le sens de propagation.

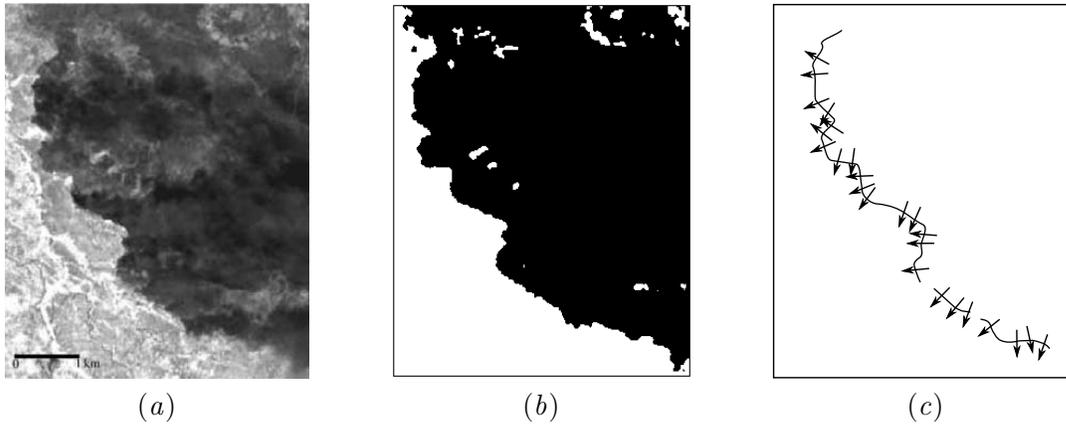


FIGURE 1.12 : Calculs de paramètres sur une image multispectrale. (a) l'image multispectrale, (b) la segmentation de la zone brûlée (en noir) et (c) la direction du front de feu en certains point de la ligne de front [94].

Paugam *et al.* décrivent dans [95] une méthode pour estimer la ligne de front, la vitesse et la direction d'un feu en propagation à partir d'images d'une caméra travaillant dans l'infrarouge moyen et portée par un avion. Les résultats sont présentés pour une expérience d'un feu se propageant sur une parcelle de neuf cent quarante-cinq mètres carrés. Cette technique utilise des points de contrôle statiques au sol et géo-référencés qui permettent d'avoir des informations entre le terrain et la caméra alors que la caméra bouge. La ligne de front et la direction sont trouvées par des méthodes similaires à celles utilisées par Ononye *et al.* [94]. Grâce à ces données, la distance entre deux fronts pour certains points est obtenue et permet d'estimer la vitesse de propagation de ces points en divisant cette distance par la fréquence d'acquisition des images.

1.3.3 Mesures estimées à partir d'images obtenues depuis plusieurs points de vue

Cette partie présente des méthodes permettant la mesure de caractéristiques géométriques de feux en propagation en extérieur à partir d'images obtenues depuis plusieurs points de vue. L'augmentation du nombre de points de vue permet d'accroître le nombre de caractéristiques mesurées [96]. Ainsi une vue de face pourra donner des indications sur la vitesse et la position du front alors qu'une vue latérale sera utile pour mesurer la hauteur de flamme et son inclinaison ainsi que l'épaisseur du front.

Martínez-De Dios *et al.* [97] utilisent deux caméras travaillant dans le spectre visible positionnées frontalement et latéralement par rapport à la direction de propagation supposée du feu et une caméra travaillant dans l'infrarouge moyen placée en face du feu. Quand la forme du terrain peut être approximée par un plan (terrain plat), les auteurs proposent de calculer la relation entre le plan image d'une caméra et le plan terrain par une homographie à l'aide de balises géo-référencées sur le sol. Ces balises permettent d'estimer une matrice de projection

de chaque caméra vers le plan du terrain où se propage le feu. Notons que seulement quatre balises suffisent à estimer ces matrices, à condition qu'elles soient toutes vues par toutes les caméras. Grâce à ces matrices, les pixels des images de chaque caméra peuvent être projetés sur le plan du terrain.

Dans le cas d'un terrain qui ne peut pas être approximé par un plan, un modèle numérique des élévations du terrain est créé. La projection des points des images sur ce modèle nécessite de calibrer les caméras et de disposer de leurs positions et orientations relatives au repère global. Pour réaliser ce calibrage, les coordonnées GPS des caméras et de balises sur le terrain sont utilisées pour estimer les paramètres d'orientation (roulis tangage et lacet). Grâce à ces informations, un calcul de la forme tridimensionnelle du feu pour chaque vue est obtenue. Cette forme tridimensionnelle permet l'estimation des paramètres tels que la hauteur, la largeur et l'inclinaison du feu ainsi que la position du front. La figure 1.13 présente les caractéristiques estimées par Martínez *et al.* et présentées dans [97].

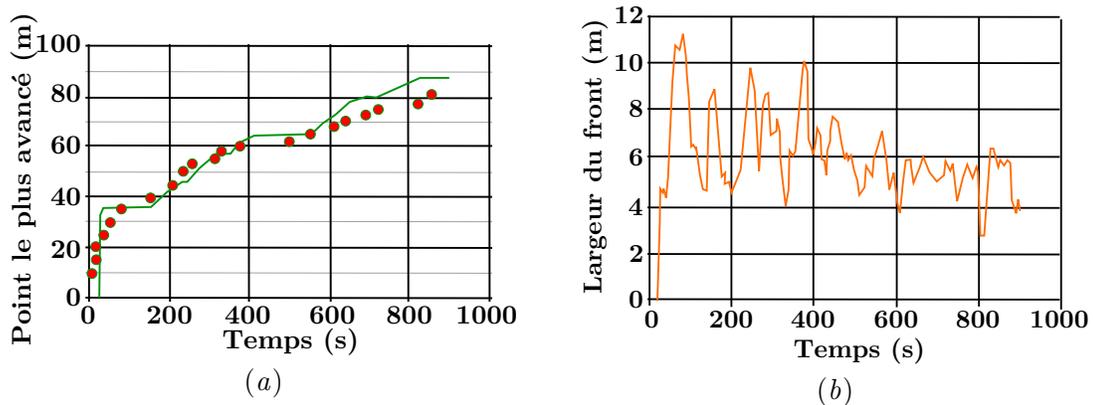


FIGURE 1.13 : Caractéristiques géométriques de front de feu obtenues par Martínez-De Dios *et al.* [97] : (a) position du point le plus avancé du feu, (b) largeur du front de feu.

Dans [98], Merino *et al.* présente une amélioration de la méthode décrite précédemment qui consiste à fusionner les formes tridimensionnelles du front de feu. Pour cela, des correspondances entre les nuages de points donnés par les caméras sont recherchées avec un algorithme du plus proche voisin. Pour les points de correspondances trouvés dans les différentes caméras, la position moyenne est calculée. Ceci permet de corriger en partie les erreurs dues aux imprécisions des capteurs.

Dans [60], Martínez-De Dios *et al.* prennent en compte le fait que la base d'un feu émet beaucoup plus de radiation que la flamme [99] et est généralement moins masquée par les fumées que la flamme. Ils proposent d'utiliser les images acquises dans l'infrarouge pour faire des mesures relatives à la base du feu et les images acquises dans le visible pour estimer les caractéristiques géométriques de la flamme. Les procédures d'estimation de caractéristiques géométriques décrites dans [97] sont utilisées avec une amélioration qui consiste à prédire les

nouvelles données en appliquant un filtre Bayésien récursif sur les caractéristiques calculées. Cette estimation permet de calculer l’incertitude des mesures. D’après les résultats présentés dans [60], ce système fournit de bonnes estimations pour la position du front et la hauteur de flamme (voir figure 1.14).

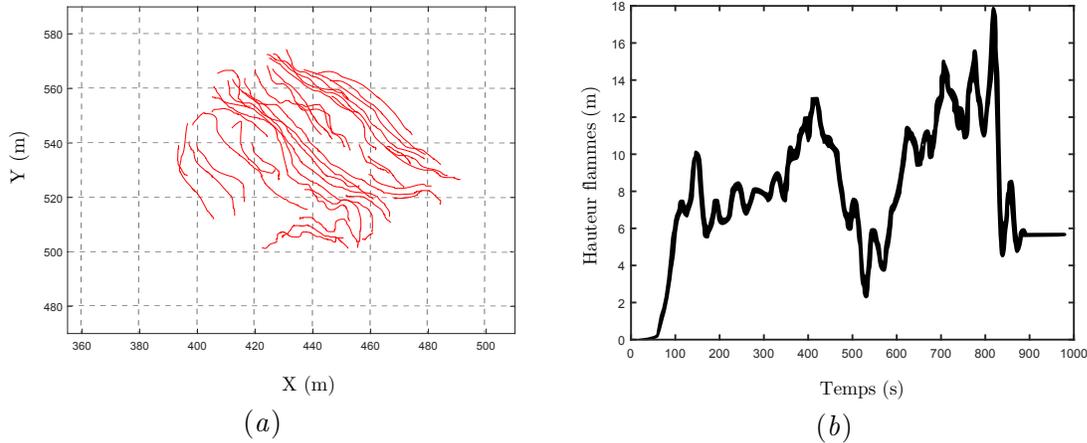


FIGURE 1.14 : Caractéristiques obtenues par Martínez-De Dios *et al.* [60] : (a) position du feu, (b) hauteur de flamme.

1.3.4 Mesures par stéréovision

La stéréovision est un procédé qui permet d’obtenir une information de profondeur à partir de deux images simultanées et décalées d’une même scène. Le principe de la reconstruction d’un feu par stéréovision est présenté en détails dans le chapitre 6.

Rossi *et al.* dans [100] proposent d’utiliser un système de stéréovision composé de deux caméras travaillant dans le domaine du visible pour estimer les caractéristiques géométriques d’un feu de végétation. Le système Bumblebee XB3 de chez *Point Grey* formé de deux caméras pré-calibrées espacées de vingt-quatre centimètres est utilisé. Les points d’intérêt des images stéréoscopiques sont cherchés uniquement sur les contours des zones de feu obtenues après segmentation. Une procédure de triangulation [101] est réalisée après l’appariement des points d’intérêt afin de calculer les coordonnées tridimensionnelles des points de l’espace correspondants et les caractéristiques géométriques du feu. La ligne de front, la hauteur, la profondeur et l’inclinaison du feu sont ainsi estimées sans aucune mesure préalable et positionnement de repère sur le terrain. Les auteurs proposent aussi de calculer la vitesse de propagation du feu en considérant certains points sur la ligne de front de feu. Pour cela, ils estiment la distance entre les points équivalents de deux lignes de front consécutives et divise cette longueur par la période d’acquisition. La figure 1.15 présente l’estimation de la vitesse de points du front de flamme à partir de deux front consécutifs.

Dans [102, 103], Rossi *et al.* ont proposé l’estimation de caractéristiques géométriques d’un front de feu en propagation sur de grandes distances à partir de deux systèmes de stéréovi-

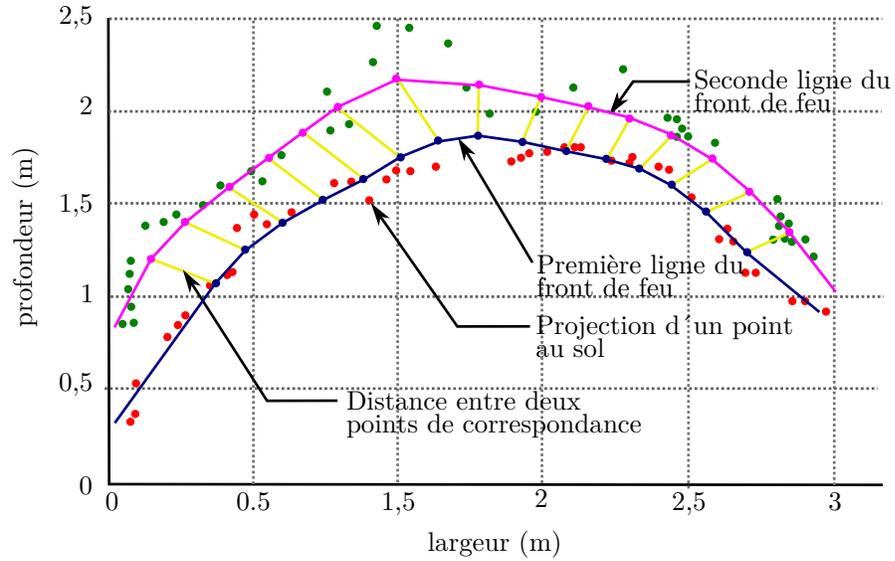


FIGURE 1.15 : Estimation de la vitesse d'un front de flamme [100].

sion déclenché simultanément et travaillant dans le domaine visible ; ceci afin d'améliorer la précision des mesures et d'obtenir simultanément la position, la vitesse, la largeur, la hauteur, l'inclinaison et la surface du front.

Le calcul de la transformation permettant de passer d'un repère caméra à l'autre s'effectue à partir du traitement d'images contenant un ballon porté à bout de bras par un homme et situé dans la zone de propagation du feu. La matrice correspondant à cette transformation est estimée à l'aide d'un algorithme itératif du point le plus proche (ou ICP de l'anglais *Iterative Closest Points*) [104] appliqué aux deux ensembles de points tridimensionnels correspondant aux positions du ballon estimées par stéréovision.

Grâce à cette matrice, les points tridimensionnels du feu obtenus depuis chaque système de vi-

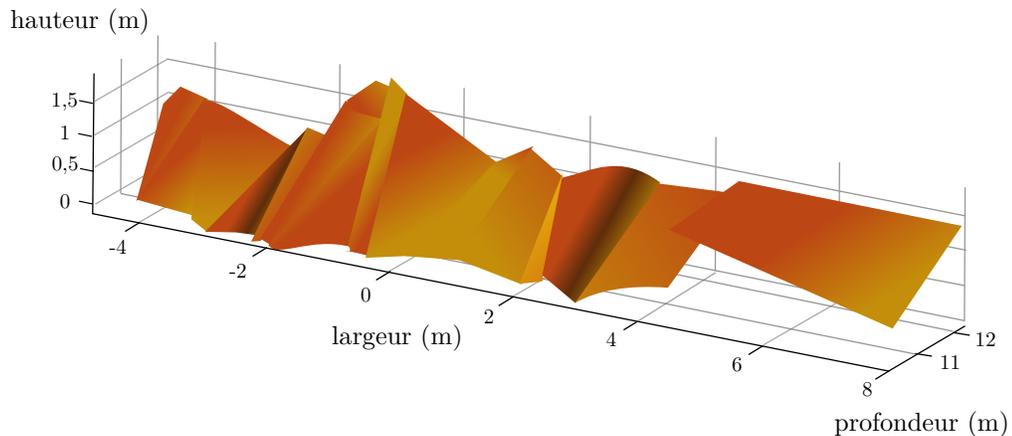


FIGURE 1.16 : Surface d'un front calculé avec deux systèmes de stéréovision [102].

sion sont ensuite projetés dans le même repère, ce qui permet l'obtention d'un nuage de points à partir duquel les caractéristiques géométriques du front de feu sont estimées. La figure 1.16 représente la surface d'un front de feu obtenue à l'aide de cette méthode.

1.4 Conclusion

Dans ce chapitre nous avons décrit les méthodes existantes dans la littérature portant sur la mesure des feux par des techniques de vision par ordinateur.

Dans la première partie de ce chapitre, les méthodes de détection de pixels « feu » ont été détaillées. Il a été montré qu'il est très difficile de comparer les performances de ces algorithmes car les articles présentant ces méthodes et leurs résultats donnent très peu d'informations sur les images utilisées pour les évaluations d'efficacité. Il est donc apparu nécessaire de construire une base d'images de feux de forêt caractérisées afin de permettre un travail d'analyse des performances des méthodes existantes et d'en proposer des améliorations ou de nouvelles.

Dans la seconde partie de ce chapitre, les différentes méthodes utilisées dans la littérature pour estimer les caractéristiques du feu ont été présentées. Ces méthodes ne permettent pas d'obtenir toutes les caractéristiques d'un front de flamme nécessaires aux physiciens pour comprendre les phénomènes intervenant lors de la propagation d'un feu, améliorer et/ou valider les modèles de comportement de feu qu'ils ont développés. Les méthodes qui ont été présentées nécessitent pour la plupart de positionner des repères géo-référencés dans les champs de vision des caméras lors des prises d'image. La technique de mesure par stéréovision présente l'avantage de ne nécessiter aucun positionnement de repère ou prises de mesures sur le terrain et permet d'obtenir des informations telles que la surface du front de feu qui n'est pas estimée par les autres méthodes. Toutefois, comme la précision des estimations par stéréovision est directement liée à la distance inter-caméra, il est nécessaire de travailler avec des dispositifs pour lesquelles une distance importante sépare les deux caméras.

Ce chapitre a donc présenté les bases des travaux de la littérature en relation avec le sujet de recherche développé dans cette thèse. Il a permis de montrer les principales idées qui ont été publiées pour résoudre le problème de la mesure des feux par vision ainsi que les limites des méthodes existantes. Ces travaux serviront de base pour cette thèse et d'éléments justifiant les solutions choisies pour le développement de nouvelles méthodes d'estimation de caractéristiques géométriques de feux de végétation.

Chapitre 2

Architecture de la solution proposée

2.1	Introduction	39
2.2	Bandes spectrales sélectionnées pour l'observation des émissions électromagnétiques des feux de végétation	40
2.2.1	Rayonnements électromagnétiques des feux de végétation	40
2.2.2	Comparaison d'images de feux de végétation acquises dans différentes bandes spectrales	44
2.3	Choix d'un dispositif de vision et de ses caractéristiques	50
2.3.1	Choix de la focale de l'objectif	51
2.3.2	Choix de la distance entre les deux caméras du système de stéréovision	51
2.4	Solution proposée	52
2.4.1	Matériels composant le dispositif	52
2.4.2	Description du processus de traitement	53
2.5	Conclusion	55

2.1 Introduction

Ce chapitre présente l'étude qui a été menée afin de déterminer l'architecture de la solution proposée pour l'estimation de caractéristiques géométriques de feux de végétation telles que sa position, vitesse, hauteur et surface. Cette solution intègre d'une part une partie matériel, avec le choix des attributs des systèmes de vision (focale et bandes spectrales des caméras, distance inter-caméras et position relative des dispositifs de vision) et d'autre part un aspect fonctionnel avec l'exploitation des données issues de plusieurs systèmes de vision.

Ce chapitre est construit de la manière suivante :

La section 2.2 est dédié, dans une première partie, à une étude bibliographique relative aux émissions de longueur d'onde des feux et aux bandes spectrales propices à leur observation. Dans une deuxième partie, elle présente une argumentation relative au choix de bandes spectrales que nous avons fait intégrant les éléments théoriques donnés dans la première partie et

à partir d'images de petits feux de végétation que nous avons menés avec des caméras que nous possédions travaillant dans différents domaines spectraux.

La section 2.3 explique le choix de la stéréovision pour mesurer les caractéristiques géométriques des feux de végétation et détaille les calculs qui ont permis de déterminer la focale des caméras et la distance inter-caméras.

La section 2.4 est dédiée à la description détaillée du matériel choisi et de l'algorithme qui sera développé afin d'exploiter les différentes données issues de chaque système de vision.

2.2 Bandes spectrales sélectionnées pour l'observation des émissions électromagnétiques des feux de végétation

Chaque objet, dont la température est supérieure à 0° K (zéro absolu), émet des ondes électromagnétiques qui lui sont caractéristiques et qui constituent sa signature spectrale.

Les capteurs de caméra sont sensibles à des bandes de rayonnement. Par exemple, la figure 2.1 montre la sensibilité spectrale de la caméra JAI AD080 GE dans la bande du proche-infrarouge.

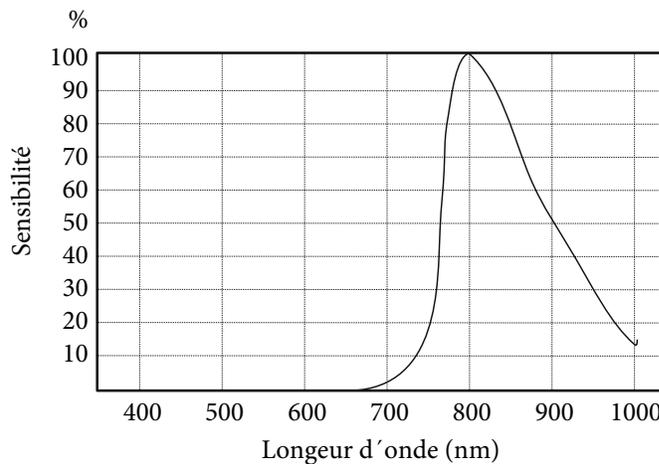


FIGURE 2.1 : Sensibilité spectrale de la caméra JAI AD080 GE dans le spectre proche-infrarouge.

2.2.1 Rayonnements électromagnétiques des feux de végétation

Un feu de forêt est un processus de combustion qui génère des températures comprises entre 400°C et 1500°C . La combustion rougeoyante du charbon se fait entre 500°C et 600°C et les flammes atteignent des températures de 1000°C [105]. Dans le cas des feux de forêt, plus de 90% de la masse de gaz émise par la combustion est constituée de vapeurs d'eau (H_2O) et de dioxyde de carbone (CO_2). La signature spectrale du feu dépend essentiellement de celle des gaz émis lors du processus de combustion ainsi que des solides chauds présents tels que le bois en combustion, les cendres chaudes, les braises et les particules comme les suies (C_2)

(dont le rayonnement dépendent de la formule moléculaire du combustible [106]). Ces solides émettent dans toutes les longueurs d'onde et ont des températures comprises entre 230°C et 730°C [107]. D'une manière générale, la température et l'émissivité du feu peuvent varier selon le vent, la topographie ou encore le type de végétal [54].

Les résultats de Boulet *et al.* dans [108] montrent que les émissions électromagnétiques des feux de petite échelle dépendent plus de la taille des flammes que des combustibles végétaux. La figure 2.2 présente les signatures spectrales de feux de 45 centimètres de diamètre sur deux combustibles différents enregistrées par un spectromètre infrarouge à transformée de Fourier (en anglais, *Fourier Transform InfraRed* (FTIR)).

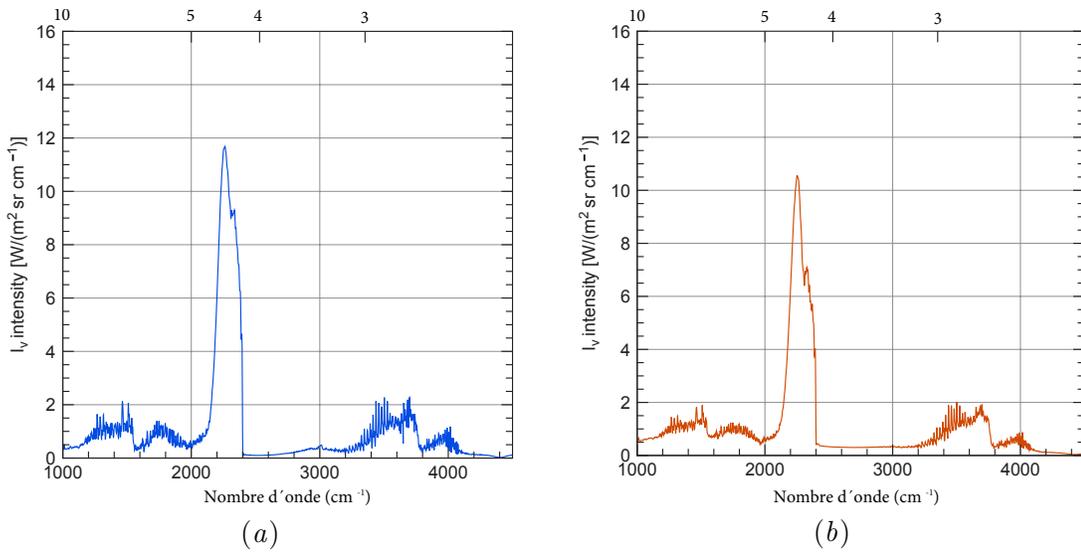


FIGURE 2.2 : Mesure des intensités des émissions pour des feux réalisés en laboratoire par Boulet *et al.* [108]. (a) Combustibles composés de branches de vignes et (b) de branches de hêtres secs.

Des études de signatures spectrales ont aussi été menées sur différentes épaisseurs de flammes [109]. La figure 2.3 présente les résultats de cette étude. Des différences de signature ont été notées en fonction de ces épaisseurs et sont principalement dues aux émissions de suies qui augmentent en fonction de la superficie du combustible.

D'une manière générale, les spectres des feux de végétation sont toujours composés de raies d'émission liées au dioxyde de carbone ($4,3 \mu m$) et à la vapeur d'eau ($2,7 \mu m$ et entre $5 \mu m$ et $10 \mu m$) plus ou moins distinguables en fonction des émissions de suies. La raie d'émission liée au dioxyde de carbone est particulièrement visible dans le cas de feux de végétation sèche (figure 2.2) alors que les raies d'émission de la vapeur d'eau sont plus importantes pour de la végétation humide.

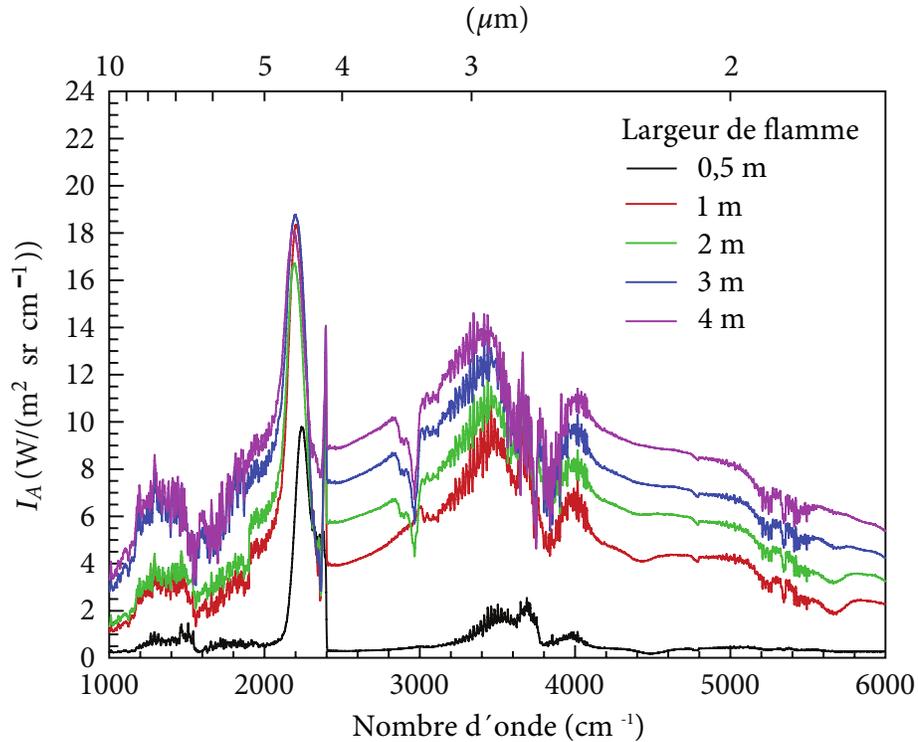


FIGURE 2.3 : Spectres de flamme (luminance en fonction du nombre d'ondes) mesurés lors d'essais en tunnel à feu pour des flammes d'épaisseur de cinquante centimètres à quatre mètres. [109].

2.2.1.a Bandes spectrales propices à l'observation des feux

Le domaine du visible est le domaine de référence de la communauté travaillant sur la problématique des feux de forêt compte tenu de l'utilisation plus simple de caméras travaillant dans ce domaine et de leur prix très abordable.

L'infrarouge est un domaine spectral invisible pour l'œil humain. Il y a dans la littérature plusieurs découpages du spectre infrarouge en bandes spectrales. Nous adopterons dans ce document les notations proposées par le MIVIM [110] :

- Proche-infrarouge (en anglais, *Near Infra-Red* : NIR) : $[0,74 ; 1] \mu m$
- Infrarouge à courtes longueurs d'onde (en anglais, *Short Wave Infra-Red* : SWIR) : $[1 ; 3] \mu m$
- Infrarouge à longueurs d'onde moyennes (en anglais, *Medium Wavelength Infra-Red* : MWIR) : $[3 ; 5] \mu m$
- Infrarouge à grandes longueurs d'onde (en anglais, *Long Wavelength Infra-Red* : LWIR) : $[8 ; 14] \mu m$

La fenêtre $[5 ; 8] \mu m$ n'est pas considérée dans ces notations compte tenu de la faible intensité des ondes de cette bande qui traversent l'atmosphère. La transmittance atmosphérique d'une onde est le rapport entre son intensité incidente et son intensité transmise. De manière générale, l'atmosphère absorbe toutes les ondes dont les longueurs d'onde sont celles des gaz qui la composent. Celle-ci est constituée de plusieurs gaz tels que l'azote (N_2), l'oxygène (O_2), l'argon (Ar), la vapeur d'eau (H_2O), le dioxyde de carbone (CO_2), le méthane (CH_4), le monoxyde de carbone (CO), le protoxyde d'azote (N_2O), les chlorofluorocarbones (CFC) ou l'ozone (O_3). Chacun de ces gaz absorbe le rayonnement qui le traverse correspondant à celui de son émission, ce qui produit de nombreuses bandes d'absorption. Les bandes d'absorption les plus larges se situent dans l'infrarouge et sont dues aux molécules de dioxyde de carbone et d'eau. La figure 2.4 présente la transmittance atmosphérique dans les différentes bandes spectrales de l'infrarouge.

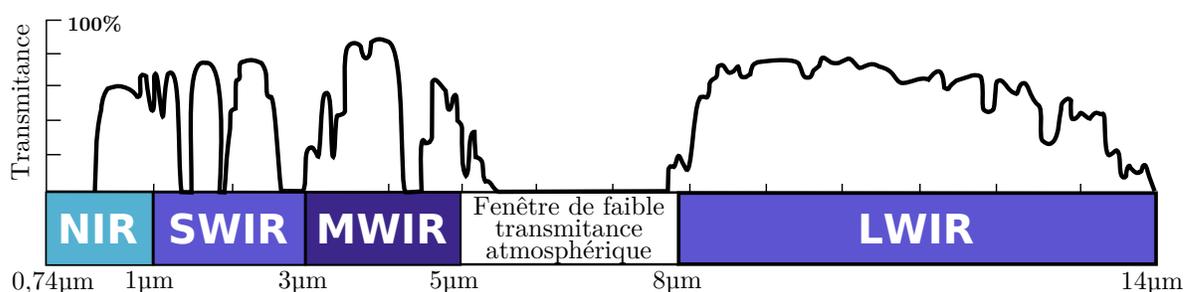


FIGURE 2.4 : Transmittance atmosphérique dans le spectre infrarouge [110]

Les gaz qui constituent les flammes sont également dans l'atmosphère et cette dernière absorbe donc les rayonnements leur correspondant. Comme le feu est un phénomène très chaud il reste toutefois facilement observable à courte distance dans toutes les bandes spectrales qui le caractérisent. Par contre, comme la transmittance de l'atmosphère diminue avec la distance parcourue par l'onde, le rayonnement associé au feu peut être fortement réduit sur de grandes distances de propagation d'onde [111].

Gouverneur *et al.* proposent dans [54] une étude qui présente, pour différentes températures de feu et bandes spectrales, le rapport théorique entre les radiations émises par un feu et celles émises par un objet de température $30^\circ C$ ou le soleil. Leurs résultats sont présentés dans le tableau 2.1.

Les données de ce tableau montrent que les radiations émises par un feu dans les bandes LWIR, MWIR et SWIR (à partir de $730^\circ C$) génèrent un ratio supérieur à 5 qui permet de différencier les radiations du feu de celles des autres éléments considérés dans cette étude. Dans la bande du visible, les radiations émises par un feu ne permettent pas de les distinguer avec celles émises par le soleil. Cette étude n'intègre pas la bande proche-infrarouge mais celle-ci étant proche du visible, on peut supposer que les ratios sont proches de ceux obtenus dans le

TABLEAU 2.1 : Rapport entre des radiations émises par un feu, à différentes températures et dans différentes bandes spectrales, et celles émises par un objet à une température de 30°C. [54].

Spectre	Type de ratio	430°C	730°C	930°C	1230°C
Visible [390 ; 700] nm	Feu/Soleil	0	0,0001	0,004	0,16
SWIR [0,9 ; 1,7] μm	Feu/Soleil	0,1	5	28	160
MWIR [3 ; 5] μm	Feu/Arrière-plan	136	682	1310	2585
LWIR [8 ; 12] μm	Feu/Arrière-plan	19	41	58	85

domaine du visible. Verstockt *et al.* dans [25] justifie l'utilisation de caméras travaillant dans le spectre LWIR par le fait que les images obtenues pour des feux à grande distance et en présence de fumée permettent une très bonne segmentation des zones de flamme. Le Maout *et al.* dans [106], quant à eux, en faisant référence aux données présentées dans le tableau 2.2, suggèrent que le domaine du proche-infrarouge est une bonne fenêtre spectrale pour la vision des feux à grande distance.

TABLEAU 2.2 : Calcul par modèle statistique de la transmittance atmosphérique d'un corps noir à 1500 K à une distance de 100 m pour une atmosphère standard (330 ppm de CO₂ et 50% d'humidité relative) [106]

Fenêtre spectrale	NIR [0,75 ; 1.1] μm	SWIR [1 ; 3] μm	MWIR [3 ; 5] μm	LWIR [8 ; 12] μm
τ_{atm}	0,95	0,75	0,78	0,93

Enfin, des études sur la signature spectrale de la fumée générée par des feux de végétation ont montré que celle-ci est peu visible dans le NIR et le SWIR [112, 113] ce qui est une information très intéressante à considérer dans le cadre de nos travaux relatifs aux feux de forêt qui génèrent beaucoup de fumée.

En résumé, les bandes spectrales visible, NIR, MWIR et LWIR semblent être adaptées à la mesure par vision de feux avec toutefois dans le visible un problème de masquage de feu par les fumées.

2.2.2 Comparaison d'images de feux de végétation acquises dans différentes bandes spectrales

Les images d'un même feu acquises dans différentes bandes spectrales ne sont pas identiques. La référence de la communauté scientifique travaillant sur la problématique des feux de forêt est

faite à partir d'images du visible. Or, en présence de fumée, générée très fréquemment par les feux de végétation, les zones de flamme sont masquées et donc les images sont inexploitable pour de la mesure. De plus, les images thermiques (NIR, MWIR et LWIR) peuvent faire apparaître des zones semblables aux zones de feu et générées par les températures élevées des gaz chauds qui ne sont pas présentes dans les images du visible.

Cette partie décrit une petite étude qui a été réalisée pour déterminer, à partir de caméras disponibles à l'Université de Corse et à l'Université Laval, la bande spectrale la plus propice aux mesures par vision de feu. Plusieurs expériences de prise d'images simultanées dans différentes bandes spectrales ont été réalisées afin d'obtenir des images à partir desquelles le choix de bande spectrale a été fait.

2.2.2.a Informations concernant les caméras utilisées

Voici les informations concernant les caméras utilisées lors des différences prises d'images de feux de végétation :

- Caméra DFK 72BUC02 d'Imaging Source [114] : ce système travaille dans le domaine du visible et génère des images de taille 1280×960 pixels.
- Caméra AD-080GE de JAI [115] : ce dispositif travaille dans le domaine du visible et du proche-infrarouge (700 nm - 900 nm). Les images des différents spectres sont acquises simultanément selon le procédé illustré dans la figure 2.5 et ont une taille de 1024×768 pixels.

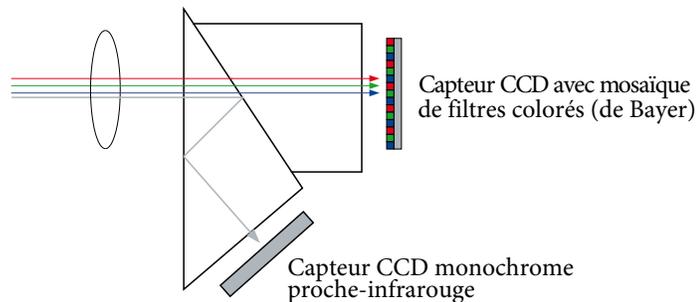


FIGURE 2.5 : Schéma du montage de la caméra JAI AD-080GE [115]. Le faisceau de lumière entrant est décomposé en deux rayons impactant un capteur de couleur et un capteur proche-infrarouge

- Caméra Jade3MW de CEDIP [116] : ce système travaille dans le spectre infrarouge à longueurs d'onde moyennes. Un filtre centré sur $4,2 \mu m$ qui permet de sélectionner uniquement les émissions du CO_2 et de favoriser ainsi la vision des flammes plutôt que des braises a été utilisé cette caméra pour une des expériences décrites dans la partie suivante. Ses images ont une taille de 256×256 pixels.
- Caméra IR-TCM 384 de Jenoptik [117] : Ce dispositif travaille dans le spectre infrarouge à grandes longueurs d'onde ($7,5 \mu m - 14 \mu m$). La plage de températures mesurées se situe entre -40 et 300 degrés Celsius. Ses images ont une taille de 384×576 pixels.

Des difficultés techniques ne nous ont pas permis de synchroniser ces différents équipements et les acquisitions d'images ont été déclenchées manuellement avec un maximum d'écart égal à 0,04 secondes.

2.2.2.b Comparaison d'images

Une première expérience a été menée pour comparer les images acquises dans le domaine du visible et celles acquises dans le spectre infrarouge à grandes longueurs d'onde. Cette expérience a été menée sur la commune de Sainte Claire dans la région de Québec sur un tas de bois de dix mètres carrés embrasé. Les dispositifs de vision utilisés ont été la caméra du spectre visible DFK 72BUC02 et celle du spectre infrarouge à grandes longueurs d'onde IR-TCM 384 fixées l'une contre l'autre sur un axe. La figure 2.6 présente ce montage situé à sept mètres du tas de bois.



FIGURE 2.6 : Caméra DFK 72BUC02 (à gauche) et caméra IR-TCM 384 (à droite) fixées côte à côte sur un axe.

La figure 2.7 présente les deux images acquises « simultanément » par les deux caméras. Ces images font apparaître de grandes différences. L'image acquise dans le domaine du visible montre de toutes petites zones de feu de couleur rouge-orange, une zone blanche de cendres, de la fumée. L'image acquise dans le spectre infrarouge ne permet pas de faire la distinction entre les cendres, le feu, et les gaz chauds. Compte tenu de la valeur maximum de la plage de températures mesurée par la caméra thermique utilisée, tous les objets de température supérieure à 300°C apparaissent avec une couleur blanche et ne peuvent pas être distingués. Cette caméra n'est donc pas adaptée à la vision du feu dans un contexte de mesure.

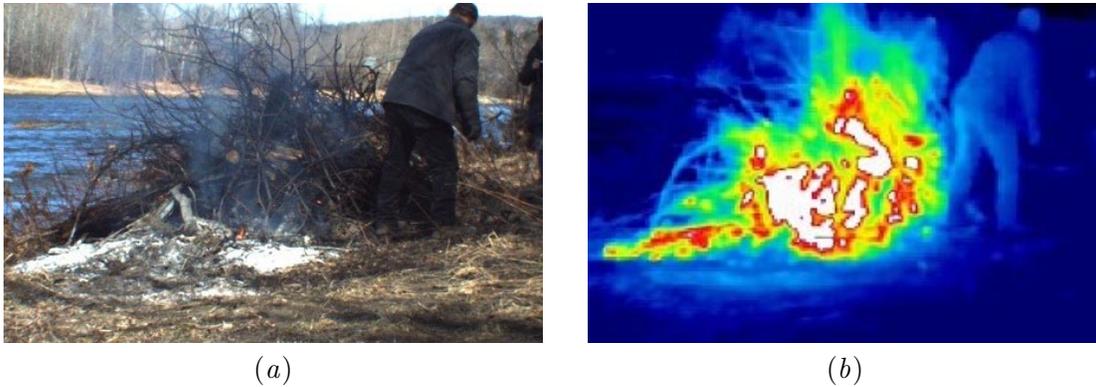


FIGURE 2.7 : Images de feu simultanément acquises dans les domaines du visible et de l'infrarouge à grandes longueurs d'onde. (a) Image du spectre visible et (b) image du spectre infrarouge à grandes longueurs d'onde.

Une seconde expérience a été conduite afin de comparer les images acquises dans le domaine du visible avec celles acquises dans les domaines du proche infrarouge et de l'infrarouge à longueurs d'onde moyennes. Cette expérience a été menée sur le plateau technique de l'Unité d'Instruction et d'Intervention de la sécurité civile N°5 située sur la commune de Corte en Corse. Le combustible employé a été de la frisure de bois. Les caméras utilisées ont été la caméra Jade3MW (sans filtre) et la caméra DFK 72BUC02 fixées côte à côte sur une tige métallique et situées à une distance de dix mètres du feu. La figure 2.8 présente les deux images acquises « simultanément » par les deux caméras. La zone de feu apparaît clairement dans les deux images. Deux points importants sont à noter : d'une part, la zone de feu est plus facilement distinguable de l'environnement (qui apparaît en noir) dans l'image de l'infrarouge que dans l'image du visible, et d'autre part, la région de feu qui apparaît dans l'image de l'infrarouge



FIGURE 2.8 : Images de feu simultanément acquises dans les domaines du visible et de infrarouge à longueurs d'onde moyennes. (a) Image du spectre visible et (b) image du spectre infrarouge à longueurs d'onde moyennes.

est plus étendue que celle correspondant au feu dans l'image du visible du fait de la présence de gaz chauds.

Une troisième expérience a été menée afin de comparer les images acquises dans les domaines du visible, du proche-infrarouge et de l'infrarouge à longueurs d'onde moyennes. Cette expérience a également été menée sur le plateau technique de l'Unité d'Instruction et d'Intervention de la sécurité civile N°5. Le combustible employé a été de la frisure de bois. Les caméras utilisées ont été la caméra Jade3MW avec son filtre centré sur $4,2 \mu\text{m}$ et la caméra multimodale AD-080GE fixées côte à côte sur une tige métallique et situées à une distance de dix mètres du feu. La figure 2.9 présente le montage utilisé pour cette expérience.



FIGURE 2.9 : Caméra AD-080GE (à gauche) et caméra Jade3MW (à droite) fixées côte à côte sur un axe.

Les images obtenues dans les différents spectres sont présentées dans la figure 2.10. On peut remarquer que les images du visible et du proche-infrarouge présentent des zones de feu très similaires. Par contre, La région de feu qui apparaît dans l'image de l'infrarouge à longueurs d'onde moyennes avec filtre diffère de celle qui est identifiable dans l'image du visible. Cette bande spectrale n'est donc pas propice à la mesure des feux par vision.

Une dernière expérience a été faite afin de déterminer si les zones de feu étaient clairement identifiables dans des images du proche-infrarouge en présence de fumée. Cette expérience a été menée sur le plateau technique de l'Unité d'Instruction et d'Intervention de la sécurité civile N°5 et la caméra utilisée a été la caméra AD-080GE. Le combustible employé a été du genêt vert. Ce combustible génère des fumées et dans le cadre de cette expérience elles ont masqué la zone de feu dans l'image du visible (comme ceci apparaît dans l'image (a) de la figure 2.11) et n'apparaissent pas ou très peu dans les images du proche-infrarouge. La figure 2.11 présente un exemple de feu masqué par de la fumée et la zone de flamme est très

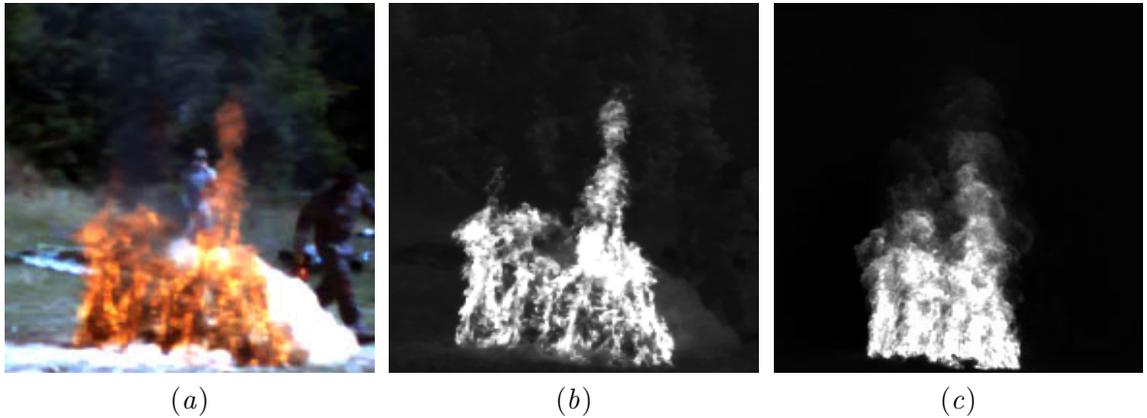


FIGURE 2.10 : Images de feu simultanément acquises dans les domaines du visible, du proche-infrarouge et de l'infrarouge à longueurs d'onde moyennes avec filtre à $4,2 \mu m$.

difficilement identifiable dans l'image du visible (image (a)) alors qu'elle apparaît clairement dans l'image du proche-infrarouge (image (b)).



FIGURE 2.11 : Images de feu avec fumée simultanément acquises dans les domaines du visible et du proche-infrarouge. (a) Image du spectre visible et (b) image du spectre proche-infrarouge.

La figure 2.12 montre deux autres images acquises dans les domaines du visible et du proche-infrarouge au cours de cette même expérience. L'image (a) montre une zone blanche qui recouvre totalement le feu. L'image (b) présente une zone de feu partiellement recouverte de fumée. Ce phénomène a été observé dans douze images parmi les deux cent cinquante-huit acquises au cours de cette expérience de brûlage. On peut supposer que cette fumée a été générée par des suies dont la quantité dépend de la combustion. Nous n'avons pas pu obtenir des images semblables dans d'autres feux expérimentaux et donc la robustesse du proche-infrarouge à ce type de fumée n'a pas pu être évaluée.

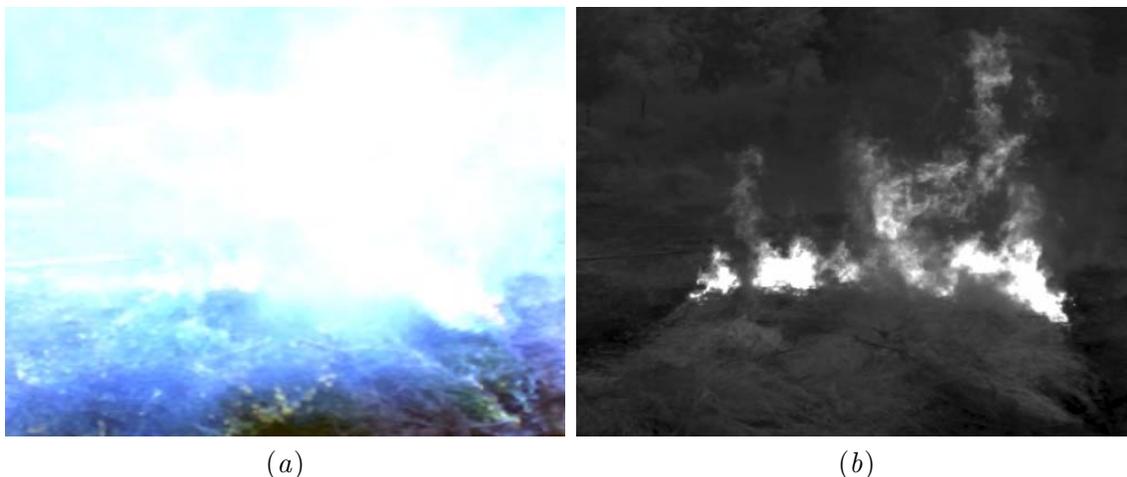


FIGURE 2.12 : Fumée masquant le feu dans le spectre proche-infrarouge. (a) image du spectre visible et (b) image du spectre proche-infrarouge.

En conclusion, les images de feux obtenues dans le domaine du proche-infrarouge font apparaître des zones de flammes présentant un fort contraste d'intensité avec l'environnement et dont les formes sont proches de celles obtenues dans les images du visible. Le soleil, certaines surfaces réfléchissantes et nuages peuvent avoir un niveau d'intensité dans ces images comparables aux zones de feu. La fusion d'images du visible et du proche-infrarouge devrait permettre d'extraire uniquement les zones de feu et de conserver l'aspect texturé des flammes.

2.3 Choix d'un dispositif de vision et de ses caractéristiques

La mesure d'objets en extérieur à partir d'une image peut s'avérer très fastidieuse car cela nécessite de positionner dans le champ de vision des repères dont les positions relatives sont connues. Ceci est encore plus difficile sur des terrains non plans, difficilement accessibles ou contenant une végétation dense. La stéréovision est une méthode de mesure qui consiste à se servir de la prise d'images sous différents angles de vue pour déterminer les dimensions, les formes ou les positions d'objets sans nécessiter le placement de repères sur le terrain étudié. Cette méthode a été choisie pour la mesure de caractéristiques géométriques des feux de végétation. Afin d'avoir une vision plus globale, le principe que nous utilisons est de positionner plusieurs systèmes stéréovision autour du feu.

Cette section présente dans une première partie les méthodes qui ont permis de déterminer les focales à utiliser pour les caméras de nos systèmes et dans une seconde partie le choix de la distance entre les deux caméras des systèmes.

2.3.1 Choix de la focale de l'objectif

La caméra JAI AD-080GE dispose de capteurs 1/4 pouces avec une taille de pixel de $4,9 \mu$, ce qui donne une image de longueur $T_u = 0,032m$ et de largeur $T_v = 0,024m$. Deux valeurs de focale ont été calculées à partir de l'équation (2.1) en considérant deux situations de brûlage : des expériences réalisées sur un terrain de $100 m^2$ et des expériences réalisées sur un terrain de $1500 m^2$.

$$f = \frac{zT_u}{h + T_u} = \frac{zT_v}{l + T_v} \quad (2.1)$$

avec :

- f : la focale du système (en mètres)
- z : la position en profondeur de l'objet (en mètres)
- l : la largeur maximum du feu que l'on souhaite observer (en mètres)
- h : la hauteur maximum du feu que l'on souhaite observer (en mètres)

Les deux focales choisies sont $f_1 = 6mm$ et $f_2 = 8mm$ pour des distances de démarrage de feu à $10m$.

2.3.2 Choix de la distance entre les deux caméras du système de stéréovision

Le système de stéréovision considéré dans cette thèse est composé de deux caméras de même focale fixées rigidement sur une barre métallique et espacées d'une distance B . : L'incertitude théorique en stéréo sur la distance estimée entre la caméra et le feu z est inversement proportionnelle à B comme le montre l'équation suivante :

$$\Delta z = \left| \frac{z^2}{Bf} \right| \Delta d \quad (2.2)$$

Avec :

- Δz : l'incertitude en z (en mètres)
- Δd : l'incertitude de disparité (en pixels)
- z : la position en profondeur de l'objet (en mètres)
- B : la distance entre les deux caméras (en mètres)
- f : la focale du système (en pixels)

La distance B a été choisie égale à $1 m$ de sorte d'obtenir une précision de mesure acceptable à des distances compatibles avec nos expérimentations tout en étant facilement transportable

et utilisable sur le terrain. Le [tableau 2.3](#) présente la valeur de l'erreur en z en fonction des deux focales choisies et de la distance du feu par rapport à la caméra.

TABLEAU 2.3 : Valeur de l'erreur théorique en z (Δz) en fonction des deux focales choisies et de la distance du feu par rapport à la caméra.

Focale	10 m	15 m	20 m	25 m	30 m
6 mm	0,08 m	0,18 m	0,33 m	0,51 m	0,74 m
8 mm	0,06 m	0,14 m	0,25 m	0,38 m	0,55 m

2.4 Solution proposée

Cette section présente la solution proposée pour la mesure par vision de caractéristiques géométriques de feux de végétation. Sa première partie est dédiée aux aspects matériels des entités de stéréovision composant le dispositif global. Ces éléments ont été déterminés à partir de l'analyse présentée dans la section précédente. Sa deuxième partie décrit le fonctionnement de la solution proposée.

2.4.1 Matériels composant le dispositif

La solution proposée est un dispositif composée de plusieurs systèmes de stéréovision identiques et synchronisés. Chaque système est composé de deux caméras JAI AD-080GE distantes d'un mètre et fixées rigidement sur un axe métallique comme cela apparaît dans la [figure 2.13](#).



FIGURE 2.13 : Système de stéréovision proposé. L'appareil au centre est l'antenne réceptrice permettant de déclencher les caméras.

Notre choix de caméra s'est porté sur la caméra JAI AD-080GE car c'est un système de vision multimodale qui permet l'acquisition simultanée d'images dans le visible et le proche-infrarouge et dont le prix est abordable (2500 € HT) ; ce qui est un point important dans une perspective de transfert technologique.

Pour des raisons économiques, notre dispositif global est composé de deux systèmes de stéréovision répartis sur le terrain de sorte d'acquérir des images complémentaires du feu. Le schéma de la figure 2.14 présente deux exemples de positionnement pour les systèmes de stéréovision : dans la configuration présentée dans l'image (a) les systèmes de stéréovision sont placés à l'arrière et à l'avant du front de feu, dans la configuration présentée dans l'image (b) les systèmes sont placés à l'avant et sur le côté droit du front de flamme.

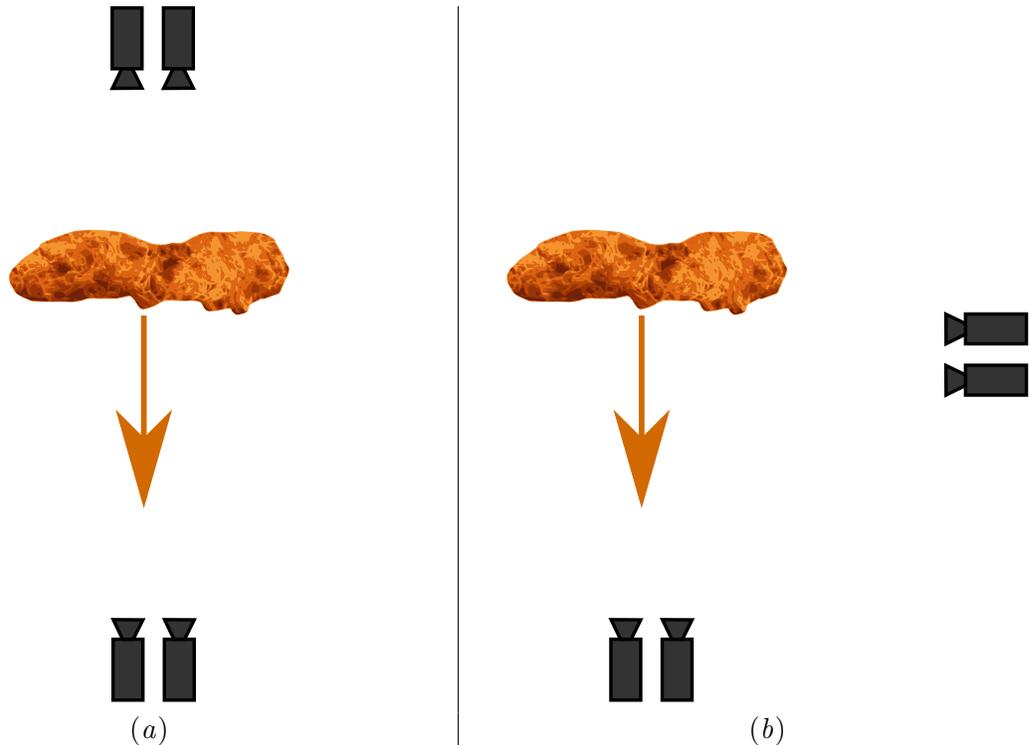


FIGURE 2.14 : Exemple de positionnements de deux systèmes de stéréovision pour le suivi d'une propagation. (a) Positionnement à l'avant et à l'arrière du front et (b) positionnement à l'avant et sur le côté droit du feu.

La procédure d'acquisition d'image de toutes les caméras du dispositif global est déclenchée simultanément à partir d'une onde hertzienne à 400 Mhz. La photo de la figure 2.13 montre le récepteur haute fréquence positionné sur la barre portant les caméras. La fréquence d'acquisition des images est réglable et peut atteindre 25 Hz. Compte tenu de la vitesse de propagation des feux, nous avons travaillé à 1 Hz.

2.4.2 Description du processus de traitement

Cette sous-section présente la description du processus de traitement proposé pour estimer les caractéristiques géométriques de feux à partir d'images stéréoscopiques. La figure 2.15 présente le schéma de fonctionnel de ce processus.

La première étape est un traitement de détection de pixels « feu » dans les images acquises

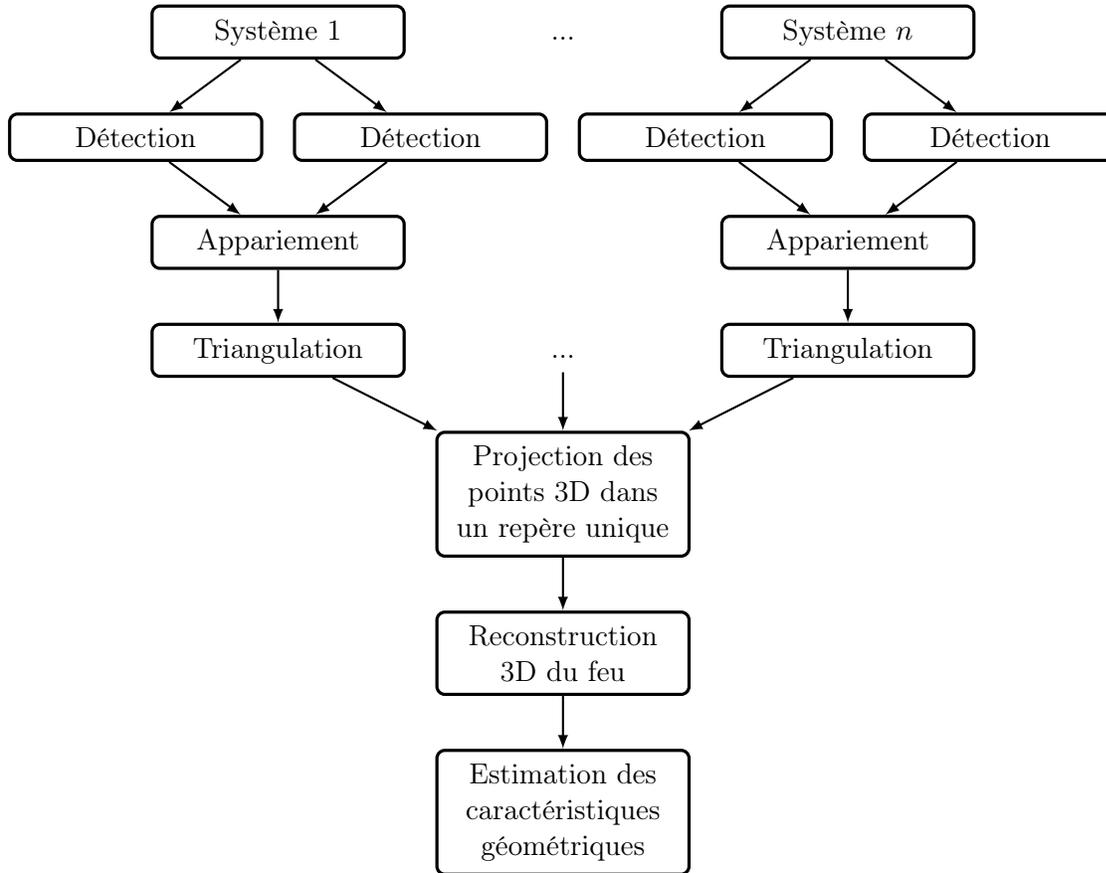


FIGURE 2.15 : Schéma fonctionnel du système proposé.

par les dispositifs de vision. Cette étape est primordiale pour l'estimation des caractéristiques car elle conditionne la précision des données qui seront traitées par les autres procédures du processus. Les chapitres 3 et 4 présentent de nouvelles méthodes que nous proposons pour la détection de pixels « feu » dans les images acquises dans le spectre visible mais également dans les images acquises simultanément dans les différents domaines et fusionnées.

La seconde étape est une procédure d'appariement de points d'intérêt détectés dans les images stéréoscopiques du visible, du proche-infrarouge et fusionnées. Dans le chapitre 5 nous proposons une nouvelle méthode permettant l'optimisation des appariements stéréoscopiques.

La troisième étape permet le calcul des coordonnées tridimensionnelles de points à partir des pixels de feu appariés. Cette étape est détaillée dans le chapitre 6 ainsi que les méthodes permettant d'obtenir certaines caractéristiques de feu à partir de points 3D obtenus depuis un seul système de stéréovision positionné à l'avant du front.

La quatrième étape correspond à un processus de projection des points 3D dans un repère commun. La cinquième étape est la reconstruction de la forme tridimensionnelle de feu à partir

des points 3D. Les méthodes employées dans ces deux étapes sont décrites dans le chapitre 7 de cette thèse.

Enfin, la dernière étape du processus est celle correspondant à l'estimation de toutes les caractéristiques géométriques de feu. Les résultats d'estimation de caractéristiques géométriques obtenus pour des feux expérimentaux sont présentés dans le chapitre 8.

2.5 Conclusion

Ce chapitre a présenté la solution proposée dans cette thèse pour l'estimation des caractéristiques géométrique d'un feu en propagation. Le dispositif présenté se compose de plusieurs systèmes de stéréovision positionnés sur le terrain de façon à obtenir des vues complémentaires du feu. La caméra de ces systèmes a été choisie après une étude bibliographique et expérimentale sur la vision du feu dans les différentes bandes spectrales. Il s'agit d'une caméra multimodale réalisant des acquisitions simultanée d'images dans les domaines du visible et du proche-infrarouge, optimisant ainsi l'obtention d'informations relatives au feu même en présence de fumée.

Ce chapitre a aussi présenté les calculs qui nous ont permis de choisir les paramètres des systèmes de stéréovision tels que la focale des caméras et la distance entre les caméras.

Le processus global de traitement des images stéréoscopiques permettant l'estimation des caractéristiques géométriques de feu a aussi été présenté. Nous proposons dans la suite de ce rapport de développer chaque étape de ce dernier.

Chapitre 3

Détection des pixels de couleur feu dans les images du spectre visible

3.1	Introduction	57
3.2	Présentation de la base de données d'images intitulée « Corsican Fire DataBase »	58
3.2.1	Images de la base	59
3.2.2	Caractérisation des images	59
3.2.3	Catégorisation des images de la base	67
3.2.4	Création d'un ensemble de pixels d'apprentissage	69
3.2.5	Publication de la base d'images sur Internet	71
3.3	Adaptation des méthodes de la littérature	72
3.3.1	Règles de détection	72
3.3.2	Méthodes basées sur un apprentissage	79
3.4	Nouvelles méthodes proposées	82
3.4.1	Nouvelles règles	83
3.4.2	Nouvelles méthodes	83
3.5	Évaluation des performances des méthodes	88
3.5.1	Évaluation des performances des règles	88
3.5.2	Évaluation des performances des différentes versions de la méthode de détection basée sur un apprentissage par histogramme gaussien lissé	92
3.5.3	Évaluation des performances de la nouvelle méthode basée sur un apprentissage par machines à vecteurs de support	93
3.5.4	Comparaison des performances des méthodes	95
3.6	Conclusion	99

3.1 Introduction

L'étape de détection de pixels de feu dans une image est primordiale dans le processus de mesure d'incendie par vision car elle conditionne la précision avec laquelle les caractéristiques géométriques du processus peuvent être estimées. Pour cette raison, nous avons fait le choix de focaliser nos travaux sur cet aspect. Comme nous l'avons évoqué dans les chapitres précédents,

toutes les mesures de référence concernant les feux de végétation obtenues par vision et utilisées par la communauté scientifique ont été estimées à partir d'images acquises dans le domaine du visible. Il est donc important d'étudier les performances et les limitations des méthodes existantes de détection de pixels de feu utilisant des images du visible afin de les améliorer ou de proposer de nouvelles méthodes. Ce chapitre présente les travaux qui ont été menés afin de réaliser ceci.

Dans le chapitre 1, les méthodes de la littérature de détection de pixels appartenant au feu ont été présentées. Il a été mis en évidence que ces méthodes n'ont jamais été évaluées sur un même grand nombre d'images ce qui rend très difficile l'étude et la comparaison de leurs performances ainsi que leurs améliorations ou la proposition de nouvelles procédures. Cette observation a servi de fondement pour les recherches présentées dans ce chapitre.

La section 3.2 présente la base de 500 images de feux de végétation qui a été construite dans le cadre de cette thèse. Le travail de classification d'images en fonction des caractéristiques visuelles du feu permettant notamment l'évaluation des algorithmes en fonction des types de feu y est présenté en détails.

La section 3.3 se focalise sur les adaptations des méthodes de détection existantes qui ont été proposées à partir de l'analyse des résultats des méthodes sur la base d'images.

La section 3.4 présente les nouvelles méthodes de détection qui ont été développées dans cette thèse.

Dans la section 3.5, une évaluation des algorithmes de détection de pixels de couleur feu est faite. Cette évaluation permet de comparer les algorithmes proposés entre eux ainsi qu'avec ceux de la littérature afin de déterminer les meilleurs algorithmes.

3.2 Présentation de la base de données d'images intitulée « Corsican Fire DataBase »

Après recherche, il n'a été trouvé aucune base de données d'images de feux de végétation libre d'accès et d'exploitation et possédant un grand nombre d'images caractérisées permettant une analyse comparative des performances des méthodes de détection de pixels de feu. Très récemment, Bedo *et al.* ont proposé la base Flickr-Fire [118] qui regroupe deux mille images dont la moitié contient des flammes de feux (de végétation, urbains, de voiture) et l'autre non. Les images ont été extraites de Flickr (site de dépôt et partage de photos et vidéos) par sélection à partir de mots clefs. Les éléments choisis sont sous la licence *Creative Commons* qui garantit leur libre usage. Chacune des images a été annotée « feu » ou « non feu » manuellement par sept experts mais aucune information spécifique aux zones de pixels de feu n'est donnée. La base Dyntax [119] est un très bon exemple de base dont les éléments sont convenablement caractérisés. Elle propose une collection de six cent cinquante vidéos de haute qualité de texture

dynamique. Les séquences contiennent des images de taille 720×576 , ont une fréquence d'acquisition de vingt-cinq images par seconde et durent au moins dix secondes. A chaque séquence, sont associées des informations générales (nom, date, lieu, ...), des informations sur les conditions d'acquisition (paramètres de la caméra, intérieur ou extérieur) ainsi que des informations sur les propriétés de la texture contenue dans la séquence (propriétés dynamique et spatiale). Cette base ne contient malheureusement pas d'images de feux de végétation.

Ces observations nous ont incités à construire une base de données contenant un grand nombre d'images caractérisées de feux de végétation que nous avons appelée « Corsican Fire DataBase » (CFDB) car elle possède un grand nombre d'images de feux de la région Corse (ainsi que des images de feux du monde entier).

Cette section présente dans un premier temps les images contenues dans la base, puis les techniques utilisées pour caractériser les images de la base. Enfin, les caractéristiques de la base sont détaillées et la construction d'un ensemble d'images d'apprentissage est présentée.

3.2.1 Images de la base

La base de données CFDB est composée de cinq cents images de feux de végétation. Ces images sont de différentes tailles (de 183×242 pixels pour la plus petite image à 4000×3000 pixels pour la plus grande), et formats (jpeg, png et bmp). Elles proviennent de sources variées. La majorité de ces images a été acquise par les chercheurs de l'UMR CNRS 6134 SPE - Université de Corse ainsi que par leurs partenaires (chercheurs, forestiers, membres de la sécurité civile et pompiers) lors d'expérimentations de propagation de feux, de brûlages dirigés ou de feux de forêt. Elles ont été prises dans différents lieux, à différents instants de la journée, depuis différents points de vue et avec des environnements et luminosités variés. Les autres images ont été récupérées sur le web lorsqu'elles avaient la licence *Creative Commons* avec l'autorisation d'utilisation, de modification, et de diffusion. Dans toutes les images de cette base, les contours des zones de feu sont distinguables et les couleurs de feu identifiables à « l'œil nu ». Une sélection de dix-sept images de cette base est présentée en figure 3.1.

3.2.2 Caractérisation des images

Dans la base CFDB, chaque image est caractérisée par les informations suivantes :

1. Domaine spectral de l'image : visible, proche-infrarouge, infrarouge à courtes longueurs d'onde, infrarouge à longueurs d'onde moyennes, infrarouge à grandes longueurs d'onde
2. Modèle de la caméra ou de l'appareil photo
3. Sensibilité
4. Temps d'exposition
5. Heure de prise d'image



FIGURE 3.1 : Sélection d'images de feux contenues dans la base CFDB.

6. Moment de la journée : jour / nuit
7. Distance caméra-feu : Proche / lointain
8. Lieu
9. Région
10. Position GPS
11. Direction de propagation du feu : vers droite, vers gauche, se rapproche, s'éloigne
12. Type de végétation : frisure de bois, maquis bas, maquis haut, arbres, inconnu
13. Couleur des fumées
14. Présence de nuages
15. Présence de camions, hommes
16. Vérité terrain
17. Dimensions d'image
18. Taux d'occupation des zones de feu dans les images
19. Couleur dominante du feu
20. Indice de texture du feu
21. Taux de superposition fumée-feu
22. Luminosité de l'environnement

Les données numérotées de 1 à 15 sont saisies à la main lorsque l'information est disponible. La vérité terrain (information N°16) est une image en noir et blanc dont la partie blanche segmentée manuellement à partir de l'image originale correspond aux pixels de feu. Elle est notée I^{VT} où $I^{VT}(\mathbf{x}) = 1$ si \mathbf{x} est un pixel « feu » et $I^{VT}(\mathbf{x}) = 0$ sinon. La figure 3.2 présente une image de feu et sa vérité terrain.



FIGURE 3.2 : Exemple de vérité terrain. (a) Image d'un feu et (b) sa vérité terrain.

Les caractéristiques numérotées de 17 à 22 sont calculées automatiquement. Les procédures d'obtention de chacune d'elles sont décrites ci-après.

3.2.2.a Dimensions d'image

Les dimensions des images sont obtenues très simplement lors du chargement des images dans le logiciel Matlab et font parties des métadonnées attachées aux images telles que son auteur, la date d'acquisition et la position GPS si la caméra permet de les avoir.

3.2.2.b Taux d'occupation des zones de feu dans les images

Le taux d'occupation des zones de feu dans les images est le rapport du nombre de pixels de feu (identifiés à partir de la vérité terrain de cette image) sur le nombre total de pixels de l'image. Le taux moyen sur la base CFDB est de 20 %.

3.2.2.c Couleur dominante du feu

Les feux de végétation ont des caractéristiques colorimétriques particulières. En fonction de la végétation, de l'environnement mais aussi des caractéristiques de la caméra, la couleur des feux peut être rouge, orange, jaune ou blanche. En nous basant sur les travaux de Zhao *et al.* [120], les pixels appartenant au feu (et identifiés grâce à la vérité terrain) sont classifiés en trois catégories : « rouge », « orange » et « jaune-blanc ». Chaque pixel « feu » est labellisé dans une de ces trois couleurs en utilisant l'espace de couleurs TSL qui permet de distinguer aisément les couleurs grâce à leur teinte (canal T). La couleur dominante de feu associée à une image est celle des pixels qui sont en plus grand nombre dans l'image.

Pour déterminer la valeur des seuils de teinte et saturation à considérer pour classifier les pixels de feu, une étude a été réalisée en considérant les données présentées sur le site *WorkWithColor* [121]. Ce site répertorie les différentes appellations de couleur en les classant par nuance de couleur et saturation. Ainsi, nous avons étudié toutes les couleurs de nuance rouge, orange et jaune et choisi les frontières de teinte et saturation entre chaque classe de couleur. La figure 3.3 est une synthèse de cette étude.

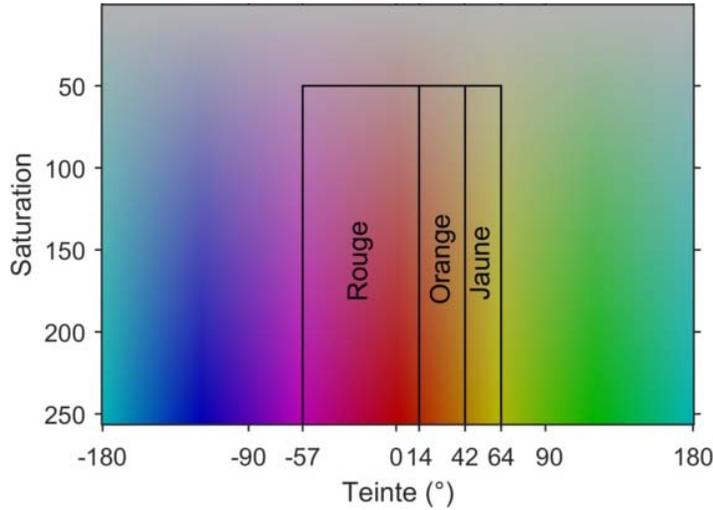


FIGURE 3.3 : Couleurs dans l'espace de couleurs *TSI*.

Il apparaît qu'aucune des couleurs répertoriées n'a une saturation inférieure à 50. Par conséquent, la première condition à vérifier par tout pixel \mathbf{x} classé « feu » est la suivante :

$$I_S(\mathbf{x}) \geq 50 \quad (3.1)$$

Si l'équation (3.1) n'est pas satisfaite, alors le pixel \mathbf{x} est catégorisé « autre couleur ». Si elle est vérifiée, le pixel est classé dans un des trois groupes de couleur « rouge », « orange » ou « jaune-blanc » selon la procédure suivante :

- Un pixel est catégorisé « jaune-blanc » s'il vérifie au moins une des deux conditions suivantes :

$$I_L(\mathbf{x}) \geq 200 \quad (3.2)$$

$$42^\circ < I_T(x) \leq 64^\circ \quad (3.3)$$

L'équation (3.2) est une condition portant sur la nuance de blanc et l'équation (3.3) est relative à la nuance de jaune.

- Un pixel est classé « orange » si l'équation (3.2) est fautive et si la couleur du pixel est dans les teintes orangées :

$$14^\circ < I_T(\mathbf{x}) \leq 42^\circ \quad (3.4)$$

- Un pixel est catégorisé « rouge » si l'équation (3.2) est fautive et si la couleur du pixel est dans les teintes rouges :

$$- 57^\circ < I_T(\mathbf{x}) \leq 14^\circ \quad (3.5)$$

Si aucune de ces conditions n'est vérifiée alors le pixel est catégorisé « autre couleur ».

La figure 3.4 présente un exemple de classement par couleur des pixels d'une image de la base CFDB (figure 3.4 *a*). Les images (*b*), (*c*) et (*d*) de cette figure sont composées respectivement uniquement des pixels labellisés « rouge », uniquement des pixels labellisés « orange » et uniquement des pixels labellisés « jaune-blanc ». Dans cet exemple, le plus grand nombre de pixels « feu » est orange, et donc la couleur « orange » est la couleur dominante du feu estimée.

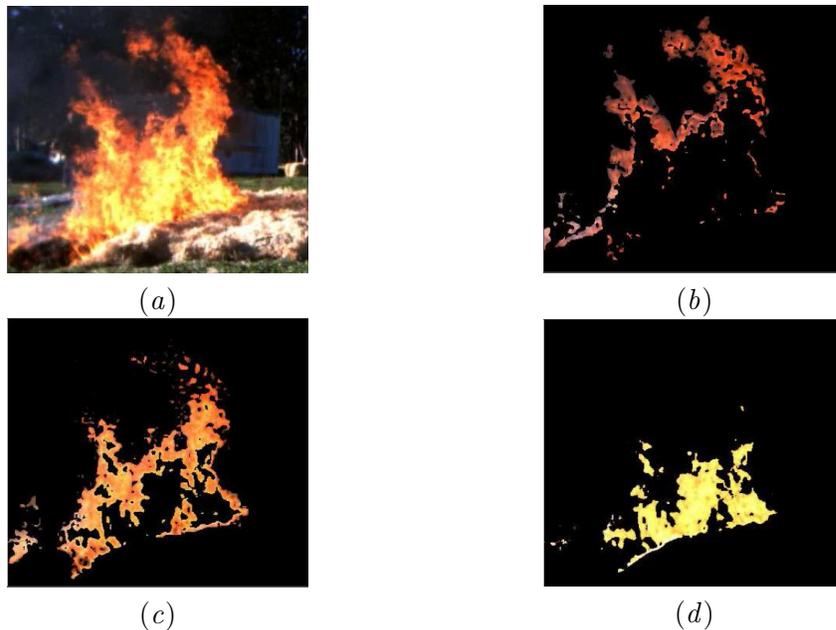


FIGURE 3.4 : Classement par couleur des pixels d'une image de feu : (*a*) image du feu, (*b*) pixels rouges, (*c*) pixels orange et (*d*) pixels jaune-blanc. Le feu de cette image est composé à 31% de pixels rouges, à 39% de pixels orange à 25% de pixels jaunes-blancs et à 5% de pixels d'autres couleurs

3.2.2.d Indice de texture du feu

Un traitement basé sur une méthode statistique d'histogramme est utilisé afin de déterminer si les zones de feu présentes dans les images ont un haut niveau de texture ou un faible niveau

de texture.

Un calcul d'entropie sur le voisinage de chaque pixel « feu » est réalisé. Ce calcul est effectué à partir de l'intensité des pixels de l'image convertie en niveau de gris et quantifiée de sorte d'obtenir uniquement huit niveaux selon la méthode suivante : la valeur 1 est affectée aux valeurs d'intensité contenues dans $[0; 31]$, la valeur 2 est affectée aux valeurs d'intensité contenues dans $[32; 64]$ et ainsi de suite.

Soit ν l'ensemble des pixels voisins à \mathbf{x} contenus dans une fenêtre de taille $i \times j$. La matrice de co-occurrence $G_{\mathbf{x}}$ associée à \mathbf{x} est construite à partir de la valeur d'intensité des éléments de ν en tenant compte de la position relative des pixels. Les éléments de cette matrice de taille 8×8 sont dans un premier temps initialisés à zéro puis modifiés de la manière suivante : pour tout pixel de l'ensemble ν d'intensité a , l'intensité b du pixel situé à sa droite est utilisée pour localiser la case $G_{\mathbf{x}}(a,b)$ de la matrice de co-occurrence où sera ajouté un 1 au contenu. A partir de $G_{\mathbf{x}}$, est ensuite calculée $p_{\mathbf{x}}(a,b)$ qui est la probabilité qu'une valeur b soit à droite d'une valeur a dans le voisinage de \mathbf{x} :

$$p_{\mathbf{x}}(a,b) = \frac{G_{\mathbf{x}}(a,b)}{N_{G_{\mathbf{x}}}} \quad (3.6)$$

Où $N_{G_{\mathbf{x}}}$ est la somme des éléments de $G_{\mathbf{x}}$.

A partir de $p_{\mathbf{x}}(a,b)$, l'entropie locale $E(\mathbf{x})$ est calculée selon la formule suivante :

$$E(\mathbf{x}) = - \sum_{a,b \in [1;8]} p_{\mathbf{x}}(a,b) \log_2(p_{\mathbf{x}}(a,b)) \quad (3.7)$$

L'attribut « haut niveau de texture » est associé une image de feu si la moyenne des entropies locales de tous ses pixels « feu » est supérieure à un certain seuil s . La figure 3.5 présente des images d'un feu présentant un haut niveau de texture et un faible niveau de texture.



(a)



(b)

FIGURE 3.5 : Exemple d'images d'un feu présentant un haut niveau de texture et un faible niveau de texture. (a) feu classé avec un haut niveau de texture et (b) feu classé avec un faible niveau de texture.

Une petite étude a été réalisée concernant l'influence de la taille de la fenêtre de voisinage sur la valeur moyenne des entropies locales. Dix images de taille différente ont été catégorisées à l'œil nu et le résultat de classement a été comparé à celui trouvé par traitement informatique. Il a été estimé que la taille de fenêtre permettant d'avoir un classement optimum était celle correspond au 200^{ième} de la taille de l'image. Donc par exemple, la taille de fenêtre de voisinage optimum pour une image de taille 1024×768 pixels est de $1024/200 \times 768/200$, c'est-à-dire 5×4 .

En considérant la taille de fenêtre de voisinage optimum, la valeur du seuil s a été estimée également empiriquement à 4,95 à partir de dix images.

3.2.2.e Taux de superposition fumée-feu

La présence de fumée entre les zones de flamme et la caméra modifie la couleur apparente du feu (comme cela apparaît dans la figure 3.6) et rend difficile la détection de ses pixels.



FIGURE 3.6 : Exemple d'image de feu sans fumée et avec fumée. (a) Image de feu sans fumée ; (b) image de feu avec fumée.

Lors de l'évaluation de méthodes de détection ou le développement de nouvelles techniques, il est intéressant de considérer l'information relative au taux de superposition fumée-feu, c'est à dire au pourcentage de pixels détectés avec de la fumée et faisant partie de zones de feu. La méthode d'apprentissage de Machines à Vecteurs de Support (en anglais, *Support Vector Machine* : SVM) a été utilisée dans cette thèse afin de discriminer parmi les pixels d'une zone de feu identifiée par la vérité terrain ceux qui sont superposés par de la fumée. Il s'agit donc d'un problème de discrimination binaire avec deux classes considérées : celle des pixels « feu » avec fumée et celle des pixels « feu » sans fumée.

Le principe des SVM découle directement des travaux de Vapnik [52] et repose sur l'utilisation de fonctions noyaux qui séparent des points d'un plan en deux classes (dans le cas d'une discrimination binaire). Un classificateur linéaire optimal qui sépare avec la plus grande efficacité les points est cherché par apprentissage à l'aide des données des deux classes. Dans le cadre de cette thèse, cet apprentissage a été réalisé à partir de cinq régions d'intérêt de taille 5×5 (ou ROI de l'anglais *Region Of Interest*) sélectionnées manuellement dans cinq images de feu avec

fumée et cinq images de feu sans fumée présentant des couleurs dominantes et des textures de feu variées. Afin que l'utilisateur puisse choisir facilement ces régions, une interface graphique a été développée en C++ avec la librairie `OpenCV` [122] de sorte que chaque clic de souris dans l'image détermine la position d'une région d'intérêt à laquelle est associée l'information « feu avec fumée » ou « feu sans fumée ». Le classement des pixels « feu » d'une nouvelle image associée à sa vérité terrain est rendu possible une fois la phase d'apprentissage terminée. Pour chaque pixel « feu » de cette image (identifié grâce à la vérité terrain de l'image), la région 5×5 qui l'entoure est considérée pour le classer à l'aide des SVM parmi les catégories « feu avec fumée » ou « feu sans fumée ». La figure 3.7, dans laquelle les pixels « feu avec fumée » sont représentés en bleu et les pixels « feu sans fumée » sont représentés en rouge, donne un exemple de ce classement.



FIGURE 3.7 : Caractérisation automatique de la présence de fumées : (a) image originale ; (b) image avec classement des pixels. Les pixels bleus sont ceux qui ont été catégorisés avec fumée et les pixels rouges sans fumée.

Le taux de superposition fumée-feu est ensuite obtenu en calculant le quotient du nombre de pixels « feu avec fumée » détectés par le nombre total de pixels « feu ».

Au final, l'attribut « avec fumée » est associée à une image de la base si plus de 50 % de ses pixels « feu » ont été classés « feu avec fumée ».

Une évaluation de performance de cette méthode a été réalisée en considérant quatre-vingt images composées de quarante images classées visuellement « avec fumée » et quarante images classées visuellement « sans fumée ». Sur ces quatre-vingt images seules quatre images (soit 5 %) ont été mal catégorisées.

3.2.2.f Luminosité de l'environnement

La luminosité de l'environnement d'un feu dans une image est un des facteurs pouvant faire varier les performances d'une méthode de détection de pixels de feu. Par exemple, si l'environnement a une haute intensité lumineuse, certaines parties de cet environnement peuvent être détectées comme du feu. S'il est sombre, le feu est une grande source de lumière et l'environne-

ment est éclairé par celui-ci ce qui génère des couleurs aux teintes rouges-orangé pour certaines parties de l'environnement (particulièrement lors de la présence de fumée). Aussi, il nous a semblé intéressant de catégoriser les images en fonction de la luminosité de l'environnement pour intégrer ce point dans l'analyse des performances d'une méthode de détection de pixels de feu. L'environnement est défini comme étant composé des pixels de l'image situés en dehors des zones de feu délimitées dans la vérité terrain et donc pour lesquels $\mathbf{I}^{VT}(\mathbf{x}) = 0$. La luminosité de l'environnement est défini comme étant la somme des intensités dans le canal L du système TSL des pixels de l'environnement divisée par le nombre de pixels de l'environnement :

$$Lum_{env} = \frac{\sum_{\mathbf{x} \in \mathbf{I} : \mathbf{I}^{VT}(\mathbf{x})=0} \mathbf{I}_L(\mathbf{x})}{Card(\mathbf{x} \in \mathbf{I} : \mathbf{I}^{VT}(\mathbf{x}) = 0)} \quad (3.8)$$

L'environnement est catégorisé « faible luminosité » si la condition suivante est vérifiée :

$$Lum_{env} \leq \tau_l \quad (3.9)$$

L'environnement est catégorisé « luminosité moyenne » si :

$$\tau_l \leq Lum_{env} \leq \tau_h \quad (3.10)$$

Enfin, l'environnement est catégorisé « grande luminosité » si :

$$\tau_h \leq Lum_{env} \quad (3.11)$$

Avec $\tau_l = 20$ et $\tau_h = 45$ des seuils choisis empiriquement à partir de cinq images de chaque catégorie.

Des exemples d'images des trois catégories sont présentées dans la figure 3.8. Dans l'image (c) le ciel et la neige sont d'une haute intensité et peuvent donc être confondus avec de la couleur feu par certains algorithmes de détections.

3.2.3 Catégorisation des images de la base

Les cinq cents images de la base CFDB ont été caractérisées et catégorisées automatiquement en fonction de leurs couleurs principales, du niveau de texture, de la présence de fumée et de l'intensité de l'environnement par les méthodes décrites précédemment. Le tableau 3.1 présente la répartition des images de la base par catégorie. On peut remarquer que, conformément à l'observation faite Zhao *et al.* dans [120] concernant la couleur des feux de végétation en extérieur, seule une petite proportion d'images est « jaune-blanc »

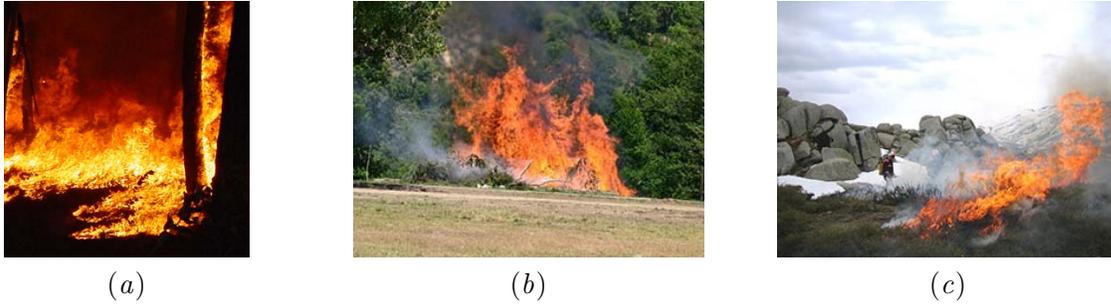


FIGURE 3.8 : Images de feu avec différents niveaux de luminosité pour l’environnement. Environnement catégorisé avec (a) une faible luminosité, (b) une luminosité moyenne et (c) une haute intensité

TABLEAU 3.1 : Répartition par catégorie des images de la base CFDB.

		Nombre d’images	Pourcentage
Catégories		500	100%
Couleur	Rouge	131	26%
	Orange	349	70%
	Jaune blanc	20	4%
	Autre	0	0%
Fumée	Avec	263	53%
	Sans	237	47%
Texture	Haute	209	42%
	Faible	291	58%
Luminosité de l’environnement	Faible	36	7%
	Moyenne	227	45%
	Haute	237	47%

Les répartitions par type de pixels ont aussi été étudiées en détails. Les cinq cents images contiennent en tout plus d’un milliard de pixels qui ont été catégorisés automatiquement. La distribution de ces pixels est présentée dans le [tableau 3.2](#) qui donne une bonne idée des pixels sur lesquels les évaluations des méthodes de détection peuvent être réalisées. L’information relative à la texture n’est pas présente dans ce tableau car elle ne peut pas être associée à un pixel seul.

On peut souligner à la lecture des données de ce tableau les points suivants concernant les pixels « feu » de la base CFDB :

- Ils sont majoritairement de couleur orange ($13,8\% + 48,4\% = 62,2\%$),
- Ceux de couleur « Autre » sont très peu nombreux ($0,2\% + 2,2\% = 2,4\%$),

TABLEAU 3.2 : Distribution des pixels de la base CFDB par catégorie.

Catégories		Nombre de pixels	Pourcentage	
Pixels « feu »	Rouge	Avec fumée	19 334 066	8.1 %
		Sans fumée	43 373 644	18.2 %
	Orange	Avec fumée	32 812 298	13,8 %
		Sans fumée	115 227 509	48,4 %
	Jaune-	Avec fumée	18 519 389	7,8 %
	Blanc	Sans fumée	2 426 830	1,0 %
	Autre	Avec fumée	578 832	0,2 %
		Sans fumée	5 332 683	2,2 %
Tous		237 950 619	100 %	
Pixels « non feu »	Faible luminosité		170 088 686	21,9 %
	Moyenne luminosité		259 874 877	33,5 %
	Haute luminosité		345 107 529	44,5 %
	Tous		775 071 092	100%

- 29,9 % des pixels (8,1 % + 13,8 % + 7,8 % + 0,2 %) ont été labellisés « avec fumée » alors que le tableau 3.1 indique que 53 % des images ont été classées « avec fumée ». Ceci peut être expliqué par le fait que les images catégorisées « sans fumée » contiennent généralement très peu de pixels « avec fumée » alors que les images catégorisées « avec fumée » contiennent des pixels « sans fumée »,

- Les pixels d'environnement labellisés « moyenne intensité » (45 %) et « haute intensité » (47 %) sont les plus nombreux. Ceci est dû au fait d'une part que les images de la base ont majoritairement été prises de jour et d'autre part que les pixels associés aux zones de ciel dans les images sont souvent de haute intensité,

- Les pixels « non-feu » sont 1,45 fois plus nombreux que les pixels « feu ». Ceci est en conformité avec le taux moyen d'occupation des zones de feu dans les images qui est de 20 %.

3.2.4 Création d'un ensemble de pixels d'apprentissage

Certaines méthodes de détection de pixels de feu nécessitent un ensemble de pixels semblables à ceux qui seront traités pour réaliser des apprentissages, calculer des seuils ou déterminer des constantes. Aussi, une base d'apprentissage formée de 1 000 000 pixels contenant un nombre équivalent de pixels « feu » et « non-feu » a été créée à partir des images de la base CFDB. Les observations faites à partir des données du tableau 3.2 et présentées précédemment ont

été considérées pour sélectionner les pixels. Ainsi :

- La base contient plus de pixels « feu orange » que de pixels « feu rouge » et plus de pixels « feu rouge » que de pixels « feu jaune-blanc ». Les proportions suivantes ont été appliquées : 50% de pixels « orange », 33 % de pixels « rouge » et 17 % de pixels « jaune-blanc »,
- L'ensemble des pixels « feu » correspondant à chaque couleur contient autant de pixels « avec fumée » que de pixels « sans fumée »,
- La base contient plus de pixels « non-feu » de « moyenne » ou de « haute intensité » que de pixels « non-feu » de « faible intensité ». Les proportions suivantes ont été appliquées : 20 % de pixels « faible luminosité », 40 % de pixels « luminosité moyenne » et 40 % de pixels « haute luminosité ».

L'ensemble des pixels d'apprentissage a été obtenu de la manière suivante :

1. Cinquante images ont été sélectionnées aléatoirement parmi les cinq cents images de la base CFDB,
2. Les pixels « feu » et « non-feu » de ces cinquante images ont tous été classés dans une des catégories présentées dans le tableau 3.2,
3. Des pixels de chaque catégorie ont été sélectionnés selon le protocole décrit ci-après.

Les pixels sont représentés par un vecteur de onze composantes noté v qui contient les valeurs de tous les canaux des systèmes de couleur RVB , $L^*a^*b^*$, YC_bC_r ainsi que les canaux S et L de l'espace TSL .

Le vecteur moyen m_v de tous les vecteurs v associés aux pixels d'une même catégorie est ensuite calculé ; ce vecteur est utilisé pour ordonner dans une liste les pixels d'une même catégorie en fonction de leur distance d_v à m_v . Cette distance est définie par :

$$d_v = |v - m_v| \quad (3.12)$$

où $|\cdot|$ représente la norme Euclidienne.

Les pixels sont finalement sélectionnés à partir de cette liste en choisissant des éléments régulièrement espacés avec un pas d'espacement égal au nombre total de pixels de cette liste divisé par le nombre de pixels à sélectionner dans cette catégorie. Ce tirage garantit l'homogénéité des couleurs de pixels dans chaque catégorie.

Enfin, pour faciliter leurs traitements, les pixels d'apprentissage sont organisés afin de constituer une image de taille 1000×1000 dont la moitié supérieure correspond aux pixels « feu » et la moitié inférieure correspond aux pixels « non-feu ». Cette image est présentée en figure 3.9 complétée par des numéros indiquant les différentes régions de pixels. Ainsi, la région N° 1 correspond aux pixels « rouge avec fumée », la région N° 2 contient les pixels « rouge sans fumée », la région N° 3 comprend les pixels « orange avec fumée », la région N° 4 est relative aux pixels « orange sans fumée », la région N° 5 correspond aux pixels « jaune-blanc avec fumée » et la région N° 6 correspond aux pixels « jaune-blanc sans fumée ». On peut remarquer dans les pixels « non-feu » de « faible » et « moyenne luminosité » sont composés de nombreux

pixels de couleur verte (dont on peut supposer qu'ils proviennent des zones de végétation des images). Les pixels « non-feu » avec une « haute luminosité » représentent le panel des couleurs du ciel et des fumées.

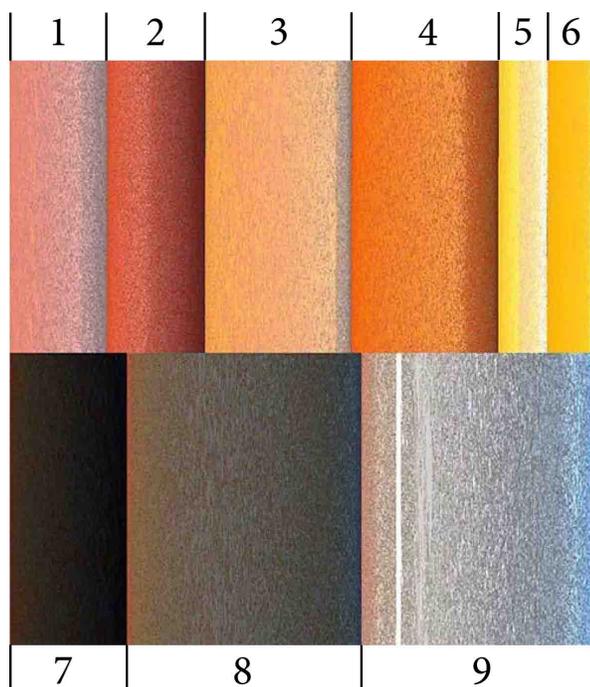


FIGURE 3.9 : Ensemble des pixels sélectionnés pour la base d'apprentissage. On distingue nettement les neuf catégories ; en haut de gauche à droite : Rouge avec fumée, rouge sans fumée, orange avec fumée, orange sans fumée, jaune-blanc avec fumée et jaune-blanc sans fumée. En bas de gauche à droite : luminosité faible, moyenne et haute.

3.2.5 Publication de la base d'images sur Internet

Une partie de la base d'images CFDB (100 images) est actuellement présente sur Internet mais de façon indirecte. En effet, une collaboration avec le Pr. Turgay Celik a permis de développer un site web dont l'adresse est <http://firetest.cs.wits.ac.za/benchmark/> qui donne la possibilité à tout utilisateur de déposer son programme de détection de pixels de feu (écrit soit en langage Matlab soit en langage Octave) et d'obtenir son évaluation sur les images catégorisées. Ce premier site a été développé alors que le nombre d'images était encore faible et avant d'avoir leurs documents d'autorisation de diffusion signés mais il a donné de nouvelles possibilités à la communauté scientifique travaillant sur la problématique de la détection des feux.

Un deuxième travail a été mené au sein de l'UMR CNRS 6134 SPE - Université de Corse par M. Antoine Paul Campana, ingénieur CNRS, afin de développer un site web dans lequel la base d'images CFDB est mise à la disposition de la communauté travaillant sur les feux

à l'adresse <http://cfdb.univ-corse.fr/>. Un ensemble de mots-clefs étant associé à chaque image, il est possible de réaliser une sélection d'images à partir de plusieurs attributs et de ne télécharger que cette sélection ou bien de télécharger l'ensemble de la base d'images CFDB.

Cette base CFDB devrait s'enrichir chaque année de nouvelles images et vidéos acquises dans le domaine du visible et de l'infrarouge grâce à une interface de dépôt présent sur le site web de l'Université de Corse.

3.3 Adaptation des méthodes de la littérature

La base d'images CFDB a été utilisée pour analyser les performances des règles et des méthodes de détection de pixels de feu présentes dans la littérature dans le but de déterminer celles qui ont les meilleures performances, de calculer leurs paramètres sur cette base spécifique aux feux de végétation (si nécessaire), et éventuellement d'en proposer de nouvelles.

Dans une première partie cette section présente une reformulation des règles de détection qui en permet un usage aisé et qui nécessite le calcul de seuils sur la base d'images étudiée. Elle développe dans une deuxième partie les adaptations des méthodes d'apprentissage proposées dans la littérature.

3.3.1 Règles de détection

Il est important de distinguer la notion de « règle de détection » (ou critère) de celle de « méthode » : une règle est composée d'opérations mathématiques basiques et de comparaisons logiques tandis qu'une méthode est la combinaison d'une ou plusieurs règles avec l'opérateur booléen ET.

Douze méthodes de détection de feux ont été présentées dans la section 1.2.1 et parmi celles-ci dix utilisent des règles que nous allons détailler dans cette partie et pour lesquelles une nouvelle formulation est proposée.

Soit $r_k(\mathbf{x})$ la $k^{\text{ième}}$ règle de détection, un pixel \mathbf{x} est labellisé « feu » selon la règle r_k s'il vérifie l'équation :

$$r_k(\mathbf{x}) \geq 0 \tag{3.13}$$

Certaines règles utilisent des constantes dont il est expliqué dans la littérature qu'elles ont été déterminées à partir d'un ensemble d'images « échantillons ». En utilisant les pixels d'apprentissage représentatifs de la base d'images CFDB, ces constantes ont été calculées en utilisant la méthode d'optimisation par recherche de motifs (en anglais, *pattern search*) décrite dans [123] et programmée avec la fonction `patternsearch` de `Matlab`. Le principe de cette méthode consiste à déterminer les paramètres des fonctions étudiées afin de les optimiser. Dans le cadre de ce travail, la méthode d'optimisation a permis de déterminer les constantes de chaque règle de sorte à maximiser la valeur de la métrique d'évaluation F-score qui traduit

l'efficacité de segmentation d'une méthode en comparant les pixels détectés à ceux des zones de feu délimitées dans la vérité terrain de chaque image.

Cette partie présente les règles de la littérature reformulées selon l'équation (3.13).

La méthode de Phillips *et al.* [27] se résume à une règle qui peut être formulée de la façon suivante :

$$r_1(\mathbf{x}) = HGL(\mathbf{x}) \quad (3.14)$$

où $HGL(\mathbf{x})$ est la valeur de l'histogramme gaussien lissé pour les coordonnées RVB du pixel \mathbf{x} .

Il a été montré dans le chapitre 1 que $HGL(\mathbf{x})$ est positive si la couleur du pixel \mathbf{x} est celle du feu et négative dans le cas contraire. L'équation (3.14) est donc bien de la forme (3.13).

Le premier critère de la méthode de T.H. Chen *et al.* [41] décrite en section 1.2.1 permet une détection rapide des pixels de couleur rouge-jaune. Ce critère est présenté dans cette partie sous la forme de deux règles afin de respecter les notations de l'équation (3.13) :

$$r_2(\mathbf{x}) = I_R(\mathbf{x}) - I_V(\mathbf{x}) \quad (3.15)$$

$$r_3(\mathbf{x}) = I_V(\mathbf{x}) - I_B(\mathbf{x}) \quad (3.16)$$

Le deuxième critère de cette méthode est utilisé pour sélectionner les pixels possédant une grande intensité dans le canal rouge du système de couleur RVB . Il peut s'exprimer sous la forme suivante :

$$r_4(\mathbf{x}) = I_R(\mathbf{x}) - \tau_R \quad (3.17)$$

Où τ_R est un seuil.

La dernière règle de la méthode de T.H. Chen *et al.* utilise les valeurs de saturation des pixels pour éliminer les régions de l'environnement dont la luminosité est semblable au feu :

$$r_5(\mathbf{x}) = I_S(\mathbf{x}) - (255 - I_R(\mathbf{x})) \frac{\tau_S}{\tau_R} \quad (3.18)$$

Où τ_R est le seuil déjà utilisé dans l'équation (3.17) et τ_S est une constante à déterminer. Les valeurs de τ_R et τ_S qui maximisent l'efficacité de la méthode de T.H. Chen *et al.* ont été estimés en utilisant la méthode d'optimisation et les pixels d'apprentissage de la base CFDB

évoquée précédemment. Leur valeur est $\tau_R = 135$ et $\tau_S = 53$. Les règles de la méthode de Martínez-de Dios [23] s'expriment ainsi :

$$r_6(\mathbf{x}) = I_R(\mathbf{x}) - \tau_V I_V(\mathbf{x}) \quad (3.19)$$

$$r_7(\mathbf{x}) = I_V(\mathbf{x}) - \tau_B I_B(\mathbf{x}) \quad (3.20)$$

$$r_8(\mathbf{x}) = I_R(\mathbf{x}) - \tau_{R2} \quad (3.21)$$

Où τ_V , τ_B sont des constantes et dont les valeurs déterminées par optimisation sont $\tau_V = 1$ et $\tau_B = 1,515$. τ_{R2} est un seuil calculé par l'algorithme de Ridler et Calvard [44] (détaillé dans la section 1.2.2) appliqué au canal rouge du système de couleur *RVB* en ne considérant que les pixels respectant les règles (3.19) et (3.20).

On peut remarquer que la règle (3.19) est la même que la règle (3.15) de T.H. Chen *et al.*

Horng *et al.* [38] proposent des règles de détection utilisant les canaux du système *TSI*. Les auteurs définissent des intervalles sur chaque canal de l'espace de couleur *TSI* dans lesquels sont contenus les valeurs d'intensité des pixels « feu ». Ainsi pour le canal *T*, on a :

$$\tau_{T1} \leq \mathbf{I}_T(\mathbf{x}_f) \leq \tau_{T2} \quad (3.22)$$

avec τ_{T1} et τ_{T2} les limites d'intervalle.

Ces règles ont été reformulées afin qu'elles correspondent à l'équation (3.13). Voici un exemple de reformulation réalisée sur l'équation (3.22) :

$$\begin{aligned} & \tau_{T1} \leq \mathbf{I}_T(\mathbf{x}_f) \leq \tau_{T2} \\ \Leftrightarrow & 0 \leq \mathbf{I}_T(\mathbf{x}_f) - \tau_{T1} \leq \tau_{T2} - \tau_{T1} \\ \Leftrightarrow & -\frac{\tau_{T2} - \tau_{T1}}{2} \leq \mathbf{I}_T(\mathbf{x}_f) - \tau_{T1} - \frac{\tau_{T2} - \tau_{T1}}{2} \leq \tau_{T2} - \tau_{T1} - \frac{\tau_{T2} - \tau_{T1}}{2} \\ \Leftrightarrow & -\frac{\tau_{T2} - \tau_{T1}}{2} \leq \mathbf{I}_T(\mathbf{x}_f) - \frac{\tau_{T2} + \tau_{T1}}{2} \leq \frac{\tau_{T2} - \tau_{T1}}{2} \\ \Leftrightarrow & \left| \mathbf{I}_T(\mathbf{x}_f) - \frac{\tau_{T2} + \tau_{T1}}{2} \right| \leq \frac{\tau_{T2} - \tau_{T1}}{2} \\ \Leftrightarrow & f_T(\mathbf{x}_f) \leq c_T \\ \Leftrightarrow & f_T(\mathbf{x}_f) - c_T \leq 0 \\ \Leftrightarrow & -f_T(\mathbf{x}_f) + c_T \geq 0 \\ \Leftrightarrow & r_T(\mathbf{x}_f) \geq 0 \end{aligned}$$

Avec $c_T = \frac{\tau_{T2} - \tau_{T1}}{2}$, $f_T(\mathbf{x}_f) = |\mathbf{I}_T(\mathbf{x}_f) - \frac{\tau_{T2} + \tau_{T1}}{2}|$, $r_T(\mathbf{x}_f) = -f_T(\mathbf{x}_f) + c_T$.

Les règles de la méthode de Horng *et al.* peuvent donc s'exprimer ainsi :

$$r_9(\mathbf{x}) = -|I_H(\mathbf{x}) - \tau'_{T1}| + \tau'_{T2} \quad (3.23)$$

$$r_{10}(\mathbf{x}) = -|I_S(\mathbf{x}) - \tau'_{S1}| + \tau'_{S2} \quad (3.24)$$

$$r_{11}(\mathbf{x}) = -|I_I(\mathbf{x}) - \tau'_{I1}| + \tau'_{I2} \quad (3.25)$$

Avec $\tau'_{C1} = \frac{1}{2}(\tau_{C2} + \tau_{C1})$ et $\tau'_{C2} = \frac{1}{2}(\tau_{C2} - \tau_{C1})$ pour $C \in \{T, S, I\}$.

Ces constantes sont calculées à partir des valeurs des bornes d'intervalle (par exemple τ_{C1} et τ_{C2}) estimées par optimisation sur l'ensemble des pixels d'apprentissage.

Les histogrammes présentant le nombre de pixels « feu » de la base d'apprentissage distribués en fonction de l'intensité dans les canaux T , S , et I sont présentés en figure 3.10. A partir de ces courbes, les intervalles ont été estimés visuellement. Ils sont égaux à $[0^\circ ; 62,5^\circ] \cup [330^\circ ; 360^\circ]$ soit $[-30^\circ ; 62,5^\circ]$ (en utilisant les égalités $180^\circ = -180^\circ$ et $360^\circ = 0^\circ$) pour le canal H , $[22,6 ; 255]$ pour le canal S et $[81 ; 247]$ pour le canal I . Leurs bornes ont ensuite été utilisées pour calculer les valeurs suivantes : $\tau'_{T1} = 16,25$, $\tau'_{T2} = 46,25$, $\tau'_{S1} = 138,8$, $\tau'_{S2} = 116,8$, $\tau'_{I1} = 164$ et $\tau'_{I2} = 83$.

Çelik *et al.* [28] proposent une méthode de détection de pixels composée de six règles qui s'appliquent sur les canaux du système de couleur RVB et dont deux sont identiques aux équations (3.15) et (3.16). Une règle porte sur le canal R et peut se reformuler ainsi :

$$r_{12}(\mathbf{x}) = I_R(\mathbf{x}) - \mu_R \quad (3.26)$$

Où μ_R est la moyenne des valeurs d'intensité du canal R calculée en considérant tous les pixels de l'image traitée.

Trois règles utilisent des divisions inter-canaux et peuvent être présentées sous la forme :

$$r_{13}(\mathbf{x}) = -\left| \frac{I_V(\mathbf{x})}{I_R(\mathbf{x}) + 1} - \tau_{RV1} \right| + \tau_{RV2} \quad (3.27)$$

$$r_{14}(\mathbf{x}) = -\left| \frac{I_B(\mathbf{x})}{I_R(\mathbf{x}) + 1} - \tau_{RB1} \right| + \tau_{RB2} \quad (3.28)$$

$$r_{15}(\mathbf{x}) = -\left| \frac{I_B(\mathbf{x})}{I_V(\mathbf{x}) + 1} - \tau_{VB1} \right| + \tau_{VB2} \quad (3.29)$$

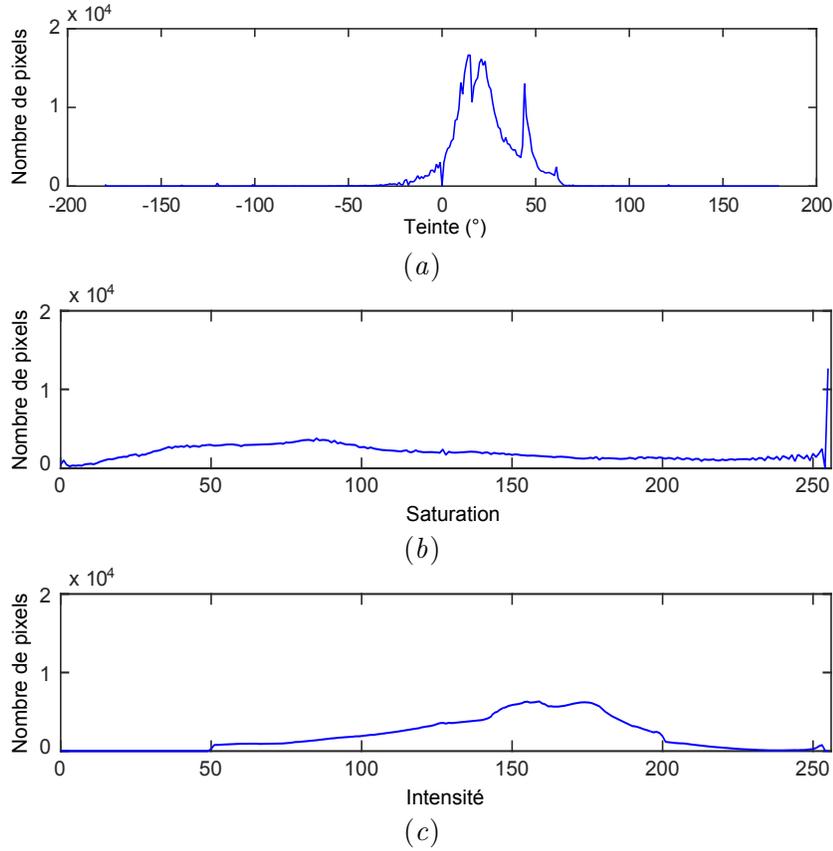


FIGURE 3.10 : Nombre de pixels « feu » de la base d'apprentissage distribués en fonction de l'intensité dans les canaux T , S , et I des pixels /classefeu d'apprentissage : (a) histogramme dans le canal T , (b) histogramme dans le canal S et (c) histogramme dans le canal I .

où τ_{RV1} , τ_{RV2} , τ_{RB1} , τ_{RB2} , τ_{VB1} et τ_{VB2} sont des constantes.

Une recherche par optimisation sur l'ensemble des pixels d'apprentissage a permis d'obtenir les valeurs suivantes : $\tau_{RV1} = 0,515$, $\tau_{RV2} = 0,348$, $\tau_{RB1} = 0,425$, $\tau_{RB2} = 0,425$, $\tau_{VB1} = 0,5$, $\tau_{VB2} = 0,5$.

Ko *et al.* utilise, dans la méthode de détection de pixels de feu décrite dans [31], une règle qui a été présentée dans l'équation (1.4). Elle peut être reformulée de la façon suivante :

$$r_{16}(\mathbf{x}) = p_f(\mathbf{x}) - \tau_K \quad (3.30)$$

où p_f est une fonction présentée dans le cadre de l'équation (1.4) et τ_K est un seuil à déterminer.

La valeur $\tau_K = 3,43 \times 10^{-5}$ a été trouvée en effectuant une recherche par optimisation sur l'ensemble des pixels d'apprentissage.

Celik et Demirel dans [33] proposent huit règles dont les six premières peuvent s'écrire de la manière suivante :

$$r_{17}(\mathbf{x}) = I_Y(\mathbf{x}) - I_{C_b}(\mathbf{x}) \quad (3.31)$$

$$r_{18}(\mathbf{x}) = I_{C_r}(\mathbf{x}) - I_{C_b}(\mathbf{x}) \quad (3.32)$$

$$r_{19}(\mathbf{x}) = I_Y(\mathbf{x}) - \mu_Y \quad (3.33)$$

$$r_{20}(\mathbf{x}) = \mu_{C_b} - I_{C_b}(\mathbf{x}) \quad (3.34)$$

$$r_{21}(\mathbf{x}) = I_{C_r}(\mathbf{x}) - \mu_{C_r} \quad (3.35)$$

$$r_{22}(\mathbf{x}) = |I_{C_b}(\mathbf{x}) - I_{C_r}(\mathbf{x})| - \tau_C \quad (3.36)$$

Avec μ_Y , μ_{C_b} et μ_{C_r} les moyennes de tous les pixels de l'image traitée pour les canaux Y , C_b et C_r ; τ_C une constante à déterminer. Cette dernière a été estimée à $\tau_C = 40$ par optimisation sur l'ensemble des pixels d'apprentissage.

Les deux autres règles de Celik et Demirel utilisent l'équation de trois courbes qui délimitent dans l'espace C_b - C_r les valeurs d'intensité des pixels « feu ». La même étude que celle présentée dans l'article [33] a été menée sur l'ensemble des pixels « feu » de la base d'apprentissage afin de déterminer l'équation de ces fonctions dans le contexte de feux de végétation uniquement. Les résultats obtenus sont présentés dans la figure 3.11.

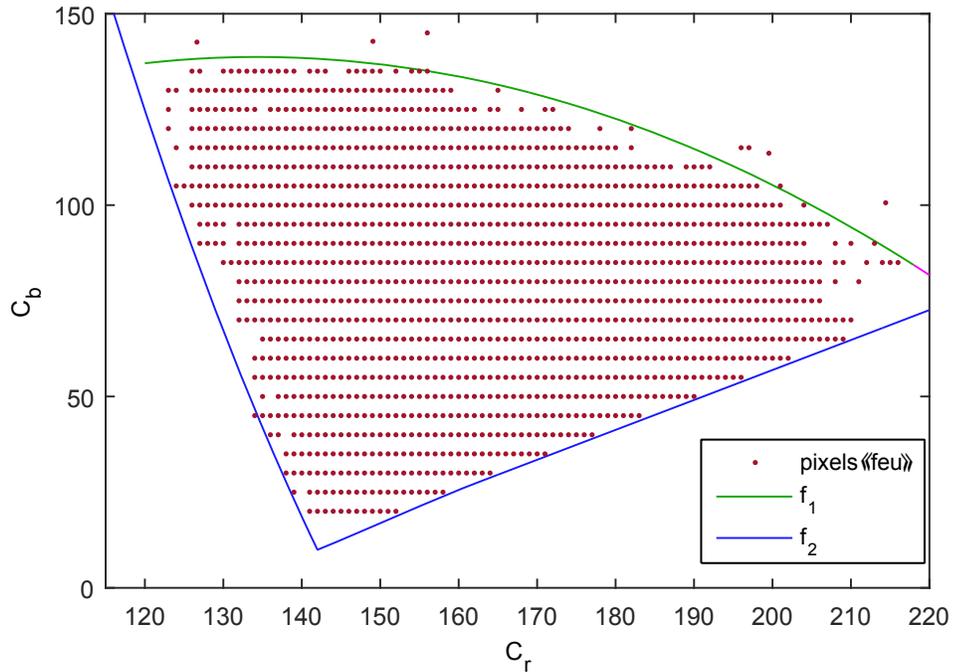


FIGURE 3.11 : Valeurs de C_b et C_r des pixels « feu » de la base d'apprentissage.

Il apparait que le zonage des pixels peut être réalisé à partir de deux fonctions polynômes (f_1 et f_2) dont les équations sont données par :

$$f_1(x) = 7,79 \times 10^{-3}x^2 + 2,10x - 2,25 \quad (3.37)$$

$$f_2(x) = \begin{cases} 4,47 \times 10^{-2}x^2 - 1,94x + 1513,52 & \text{si } x \leq 142 \\ 3,39 \times 10^{-5}x^2 + 0,77x - 98,31 & \text{sinon} \end{cases} \quad (3.38)$$

Les règles sont exprimées par :

$$r_{23}(\mathbf{x}) = f_1(I_{C_r}(\mathbf{x})) - I_{C_b}(\mathbf{x}) \quad (3.39)$$

$$r_{24}(\mathbf{x}) = I_{C_b}(\mathbf{x}) - f_2(I_{C_r}(\mathbf{x})) \quad (3.40)$$

Dans [37], Celik propose des règles sur l'espace $L^*a^*b^*$ qui s'expriment ainsi :

$$r_{25}(\mathbf{x}) = I_{L^*}(\mathbf{x}) - \mu_{L^*} \quad (3.41)$$

$$r_{26}(\mathbf{x}) = I_{a^*}(\mathbf{x}) - \mu_{a^*} \quad (3.42)$$

$$r_{27}(\mathbf{x}) = I_{b^*}(\mathbf{x}) - \mu_{b^*} \quad (3.43)$$

$$r_{28}(\mathbf{x}) = I_{b^*}(\mathbf{x}) - I_{a^*}(\mathbf{x}) \quad (3.44)$$

Où μ_{L^*} , μ_{a^*} et μ_{b^*} sont les moyennes de tous les pixels de l'image traitée dans les canaux L , a^* et b^* .

J. Chen *et al.* présentent dans [40] une méthode qui utilise trois règles identiques aux règles r_2, r_3, r_4 complétées par les deux règles suivantes :

$$r_{29}(\mathbf{x}) = 100 \times I_S(\mathbf{x}) - f_3(I_B(\mathbf{x})) \quad (3.45)$$

$$r_{30}(\mathbf{x}) = f_4(I_B(\mathbf{x})) - 100 \times I_S(\mathbf{x}) \quad (3.46)$$

Où f_3 et f_4 sont des fonctions comparables aux fonctions f_1 et f_2 étudiées précédemment et appliquées au plan (B-S).

Les équations des courbes f_3 et f_4 ont été estimées de la même façon que pour les courbes f_1 et f_2 et sont égales à :

$$f_3(x) = 80 \exp\left(\frac{x}{25}\right) \quad (3.47)$$

$$f_4(x) = 6,5 \times 10^{-3} x^2 - 0,55 x + 100 \quad (3.48)$$

La méthode de détection de pixels de Rossi et Akhloufi [43] se compose de deux phases dont chacune peut être traduite par une règle. La première étape consiste à pré-détecter les pixels « feu » en utilisant une méthode des K-moyennes (en anglais, *K-means*) sur le canal V du système YUV de l'image, avec $k = 2$ afin de séparer les pixels en deux classes : « feu » et non-feu. Liu *et al.* [124] ont prouvé que lorsque $k = 2$, un regroupement par k -moyennes était équivalent à un seuillage d'Otsu. Ce dernier étant plus rapide, il a été utilisé pour réaliser la présélection des pixels « feu » qui se résume par l'application de la règle suivante :

$$r_{31}(\mathbf{x}) = I_V(\mathbf{x}) - \tau_V \quad (3.49)$$

Où τ_V est une constante calculée par la méthode d'Otsu pour chaque image traitée.

La deuxième étape de la méthode de Rossi et Akhloufi consiste à calculer la moyenne d'intensité dans chacun des canaux du système de couleur RVB et de reconsidérer les pixels présélectionnés dans la première phase de la méthode en éliminant ceux dont la couleur est trop éloignée de la couleur moyenne calculée. Cette étape est traduite par la règle suivante :

$$r_{32}(\mathbf{x}) = -\sqrt{(I_R(\mathbf{x}) - \mu_R)^2 + (I_V(\mathbf{x}) - \mu_V)^2 + (I_B(\mathbf{x}) - \mu_B)^2} + \tau_\sigma * \sigma \quad (3.50)$$

où μ_R, μ_V et μ_B sont les moyennes d'intensité dans les canaux R, V et B des pixels vérifiant la règle r_{31} , $\sigma = (\sigma_R, \sigma_V, \sigma_B)$, est l'écart-type moyen maximum des écart-types moyens calculés dans chaque canal des pixels vérifiant la règle r_{31} et τ_σ une constante.

La valeur de τ_σ qui maximise l'efficacité de la méthode de Rossi et Akhloufi a été cherchée par optimisation et vaut $\tau_\sigma = 2,5$.

3.3.2 Méthodes basées sur un apprentissage

Les travaux qui ont été menés dans le cadre de cette thèse concernant la comparaison de performances des méthodes de détection de pixels de feu décrites dans la littérature indiquent que les méthodes de détection par apprentissage automatique obtiennent généralement de meilleurs scores que les autres méthodes (voir partie 3.5).

Dans cette section sera présenté un travail concernant l'amélioration de la méthode de détection de Phillips *et al.* [27] basée sur l'utilisation d'un histogramme gaussien lissé ainsi qu'une nouvelle méthode d'apprentissage par SVM différente de celle de Collumeau *et al.* . [30] (toutes deux présentées dans le chapitre 1).

3.3.2.a Méthode de Phillips *et al.* étendue à plusieurs systèmes de couleur

La méthode de détection de pixels de feu de Phillips *et al.* est basée sur l'utilisation d'un histogramme gaussien lissé obtenu à partir d'informations de couleur de pixels dans l'espace de couleur *RVB*. Il nous a semblé intéressant de tester l'efficacité de cette méthode en considérant les canaux d'autres espaces de couleur et des canaux croisés de systèmes de couleur différents.

Les systèmes de couleur n'ont pas tous les mêmes dimensions car le nombre de valeurs prises dans chaque canal est variable comme ceci est présenté dans le tableau 3.3. Afin de travailler avec un histogramme de même taille quel que soit le système de couleur considéré, toutes les valeurs des espaces de couleur ont été ramenées dans l'intervalle $[0 ; 100]$ en utilisant une interpolation linéaire de sorte d'avoir cent valeurs possibles pour chaque canal, puis en étant arrondies. Le triplet constituées de trois valeurs ainsi obtenues est appelé « couleur normalisée ».

TABLEAU 3.3 : Valeurs minimale et maximale des canaux des différents espaces de couleurs utilisés.

Espace de couleur	<i>RVB</i>			<i>TSI – TSV</i>				<i>YC_bC_r</i>			<i>La*b*</i>		
Canal	<i>R</i>	<i>V</i>	<i>B</i>	<i>T</i>	<i>S</i>	<i>I</i>	<i>V</i>	<i>Y</i>	<i>C_b</i>	<i>C_r</i>	<i>L</i>	<i>a*</i>	<i>b*</i>
Valeur minimale	0	0	0	0	0	0	0	16	16	16	0	-128	-128
Valeur maximale	255	255	255	360	1	255	255	235	240	240	255	127	127

Un histogramme **Hist** est associé à un système de couleur et correspond à un tableau de dimension $100 \times 100 \times 100$ initialisé à zéro. La couleur normalisée d'un pixel ($\mathbf{x}_{C1}, \mathbf{x}_{C2}, \mathbf{x}_{C3}$) permet de localiser une case de ce tableau : **Hist** $[\mathbf{x}_{C1}, \mathbf{x}_{C2}, \mathbf{x}_{C3}]$. La procédure de remplissage de l'historgramme à partir de la couleur normalisée des pixels de la base d'apprentissage est la suivante :

- la couleur de chaque pixel « feu » donne la position centrale de la fonction gaussienne P_f qui doit être additionnée aux valeurs du tableau dans un voisinage de taille $13 \times 13 \times 13$ de la cellule **Hist** $[\mathbf{x}_{C1}, \mathbf{x}_{C2}, \mathbf{x}_{C3}]$. La formule de la fonction P_f est :

$$P_f = \frac{1}{(2\sqrt{2\pi})^3} \exp\left(-\frac{a^2 + b^2 + c^2}{8}\right) \quad (3.51)$$

Avec a, b et c prenant des valeurs dans $[-13 ; 13]$ correspondant à la case du voisinage considérée $[\mathbf{x}_{C1} + a, \mathbf{x}_{C2} + b, \mathbf{x}_{C3} + c]$.

- la couleur de chaque pixel « non-feu » donne la position centrale de la fonction gaussienne P_{nf} qui doit être soustraite aux valeurs du tableau dans un voisinage de taille $13 \times 13 \times 13$ de la cellule **Hist** $[\mathbf{x}_{C1}, \mathbf{x}_{C2}, \mathbf{x}_{C3}]$. La formule de la fonction P_{nf} est :

$$P_{nf} = \frac{1}{(\sqrt{2\pi})^3} \exp\left(-\frac{a^2 + b^2 + c^2}{2}\right) \quad (3.52)$$

Avec a , b et c prenant des valeurs dans $[-13 ; 13]$ correspondant à la case du voisinage considérée $[\mathbf{x}_{C1} + a, \mathbf{x}_{C2} + b, \mathbf{x}_{C3} + c]$.

La méthode de Phillips *et al.* a ainsi été étendue à plusieurs systèmes de couleur (RVB , YC_bC_r , TSV , TSI et $L^*a^*b^*$) ainsi que sur la combinaison de canaux issus de ces espaces de couleur. Les performances des différentes variantes de méthode sont présentées dans la partie 3.5.

3.3.2.b Méthode de détection basée sur les Machines à Vecteurs de Support

Dans cette partie, une nouvelle méthode de détection de pixels de feu développée en collaboration avec le Pr. Moulay Akhoufi est présentée. Elle est basée sur une procédure d'apprentissage de la couleur et de la texture des pixels « feu » et « environnement » par Machines à Vecteurs de Support (soit deux classes).

De la même manière que pour la caractérisation des zones de feu avec ou sans fumée (partie 3.2.2.e), cette méthode fait un apprentissage sur les pixels de régions d'intérêt (ROI) en considérant la représentation spatiale des intensités des trois canaux de l'espace de couleur utilisé. De ce fait, il est impossible d'utiliser des pixels isolés pour réaliser la phase d'apprentissage. Cette dernière a été effectuée à partir de régions d'apprentissage sélectionnées dans dix images de la base CFDB choisies afin de présenter des couleurs de feu différentes et des conditions variées de présence de fumée. Des essais de performances de détection de pixels réalisés pour différentes tailles de ROI ont montré que la la taille de 4×4 pixels était celle qui offrait le meilleur compromis entre robustesse d'apprentissage et vitesse d'exécution.

Le positionnement des régions d'intérêt dans chaque image se fait manuellement par « clic » de souris à partir d'une interface. Une région d'apprentissage composée de $(10 \times 10 \times \text{taille du ROI})$ pixels est créée autour du pixel sélectionné. Il est important que l'utilisateur choisisse des zones caractérisant au mieux le feu et l'environnement (régions du ciel, nuages, végétation) pour augmenter les performances de la méthode de détection. La figure 3.12 présente les différentes zones sélectionnées sur une image d'apprentissage.

Une fois la phase d'apprentissage terminée, le traitement des images se fait de la même façon que pour la caractérisation des zones de feu avec ou sans fumée développée dans la section 3.2.2.e.

Nos travaux ont été menés principalement avec l'espace de couleur RVB . La figure 3.13 présente le résultat de détection de cette méthode pour une image de feu contenant de la fumée. Des travaux ont également été réalisés pour utiliser cette méthode avec le système de couleur $YCrCb$.



FIGURE 3.12 : Exemple de régions d'intérêt sélectionnés dans une image de feu. Les carrés de couleur cyan représentent les zones de l'environnement et les carrés de couleur rouge correspondent aux zones de feu.



FIGURE 3.13 : Exemple de détection de la méthode SVM. (a) Image originale, (b) Pixels détectés par SVM avec élimination des régions de pixels détectées dont la taille est inférieure à 10 % de la taille de la plus grande région.

Les résultats et commentaires de l'évaluation de cette méthode sont présentés dans la partie 3.5.

3.4 Nouvelles méthodes proposées

L'analyse du fonctionnement des méthodes existantes de détection de pixels « feu » et de leurs performances a permis de mettre en exergue leurs limitations et nous proposons dans cette section de nouvelles méthodes de détection.

3.4.1 Nouvelles règles

Cette partie présente deux nouvelles règles qui permettent de détecter des pixels de feu dans des images d'incendie.

La première règle utilise les canaux a^* et b^* du système CIE $L^*a^*b^*$ et prend en compte le fait que les valeurs positives du canal a^* représentent les teintes « rouge » et les valeurs positives du canal b^* représentent les teintes « jaune ». Aussi, les valeurs de $a^* + b^*$ sont élevées pour les pixels « feu ».

La règle proposée consiste à vérifier si la valeur $a^* + b^*$ d'un pixel est supérieure à un seuil t_1 et si tel est le cas, le pixel est labellisé classe « feu ». Cette règle de détection est formulée de la façon suivante :

$$r_{33}(\mathbf{x}) = I_{a^*}(\mathbf{x}) + I_{b^*}(\mathbf{x}) - t_1, \quad (3.53)$$

Avec $t_1 = 32,4$ optimisé sur la base d'apprentissage.

La deuxième règle a été construite en utilisant les canaux de l'espace de couleur RVB . Dans le cas de couleurs de nuance de rouge-orangé, le canal R a des valeurs bien plus élevées que le canal vert ou bleu.

La règle proposée consiste à vérifier que la différence entre la valeur du canal rouge et la valeur minimum entre celle du canal vert et celle du canal bleu est supérieure à un seuil t_2 et si tel est le cas, le pixel est classé « feu ». Cette règle s'écrit de la façon suivante :

$$r_{34} = I_R(\mathbf{x}) - \min(I_V(\mathbf{x}), I_B(\mathbf{x})) - t_2, \quad (3.54)$$

Avec $t_2 = 72$ optimisé sur la base d'apprentissage.

La figure 3.14 montre les résultats de détection de pixels de ces deux règles sur des images présentant des feux de couleur différente et avec des conditions de fumée différente.

3.4.2 Nouvelles méthodes

Cette partie présentent deux nouvelles méthodes de détection de pixels « feu ». La première utilise l'ensemble des règles de détection présentées dans ce chapitre. La deuxième est basée sur une approche jamais utilisée jusqu'à présent dans le cadre de la détection de pixels de feu.

3.4.2.a Détection par vote des règles

Une nouvelle méthode est proposée pour détecter les pixels « feu » à partir des trente-quatre règles de détection énumérées dans ce chapitre. Elle consiste à appliquer toutes les règles à

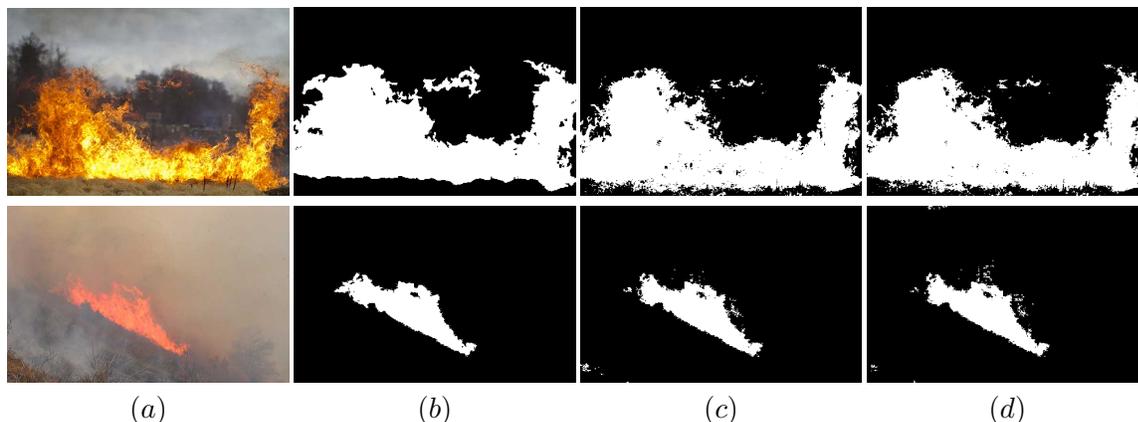


FIGURE 3.14 : Exemples de résultat de détection de pixels obtenus par les règles proposées. (a) Images de feux, (b) vérités terrain (c), pixels « feu » détectés par la règle r_{33} , (d) pixels « feu » détectés par la règle r_{34} .

un même pixel et à le classer en fonction du score obtenu par ce dernier en tenant compte de la labellisation individuelle de chacune des règles. Cette méthode est nommée « méthode par vote » car chaque règle agit comme un élément votant pour un pixel ou non et le pixel est classé « feu » s'il obtient un nombre suffisant de voix. Un pixel est donc classé « feu » si la condition suivante est vérifiée :

$$\left(\sum_{k=1}^{34} \varphi_k(\mathbf{x}) \right) \geq \tau_{MV} \quad (3.55)$$

Avec φ_k définie ainsi :

$$\varphi_k(\mathbf{x}) = \begin{cases} 1 & \text{si } r_k \geq 0 \\ 0 & \text{sinon} \end{cases} \quad (3.56)$$

et τ_{MV} un seuil.

Le valeur de τ_{MV} a été trouvée en calculant la valeur de l'indicateur de performance de détection F-score de la méthode testée sur la base CFDB pour les 34 valeurs possibles du seuil et en sélectionnant celle qui génèrait la plus haute valeur. La valeur $\tau_{MV} = 23$ maximise les performances de cette méthode.

La figure 3.15 présente les résultats de détection de pixels de la méthode par vote sur les mêmes images que celles présentées en figure figure 3.14.

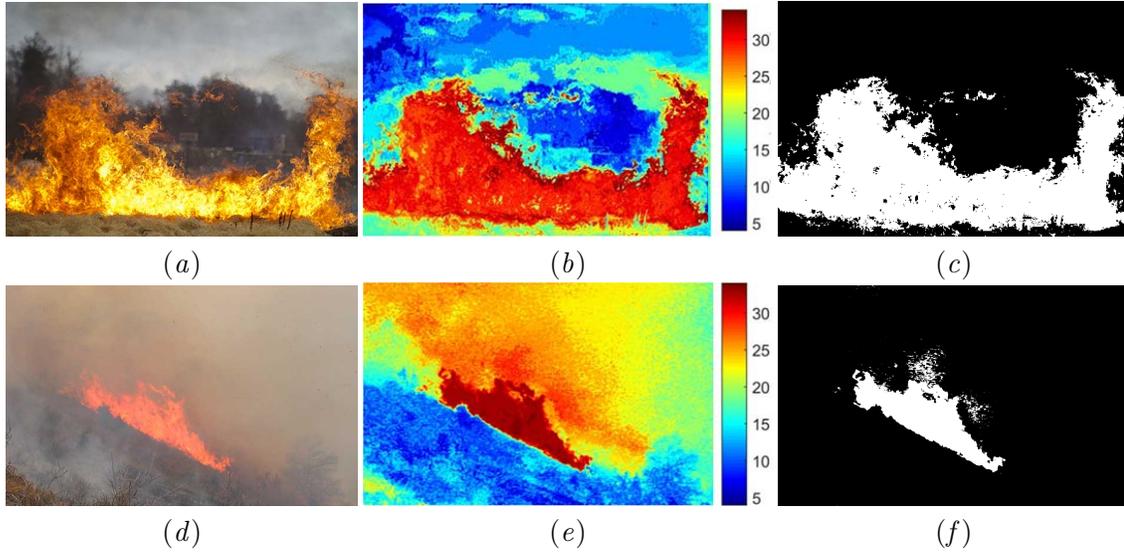


FIGURE 3.15 : Exemples de résultat de détection de pixels obtenus par la méthode par vote. (a) et (d) images originales, (b) et (e) nombre de votes par pixel et (c) et (f) détection finale en gardant les pixels qui vérifient au moins 23 des 34 règles.

3.4.2.b Apprentissage par régression logistique

En collaboration avec le Pr. Turgay Celik de l'Université de Witwatersrand de Johannesburg (Afrique du Sud), deux méthodes de détection de pixels « feu » par apprentissage par régression logistique ont été développées.

La régression logistique est une technique de classification par apprentissage supervisé. La phase d'apprentissage permet, à partir d'un ensemble de pixels, de déterminer la fonction h_ω de sorte que $h_\omega(\mathbf{x})$ soit une bonne prédiction de la classe du pixel \mathbf{x} . Soit $\hat{\mathbf{x}} = [\hat{x}_1, \dots, \hat{x}_d]^\top$ un vecteur de caractéristiques du pixel \mathbf{x} et soit $\omega = [\omega_1, \dots, \omega_d]^\top$ un vecteur de paramètres défini par apprentissage ; la fonction h_ω définie par régression logistique est :

$$h_\omega(\mathbf{x}) = g(\omega^\top \hat{\mathbf{x}}) = \frac{1}{1 + e^{-\omega^\top \hat{\mathbf{x}}}} \quad (3.57)$$

Avec g une fonction dite d'activation et $h(\mathbf{x}) \in [0; 1]$. Cette fonction peut être vue comme la vraisemblance que le pixel \mathbf{x} soit de la classe « feu ».

Le schéma fonctionnel de la régression logistique est présenté en figure 3.16.

Le pixel \mathbf{x} est classé « feu » si la règle suivante est vérifiée :

$$h_\omega(\mathbf{x}) > t_h \quad (3.58)$$

Avec t_h un seuil fixé dans le cadre de cette méthode à $t_h = 0,5$.

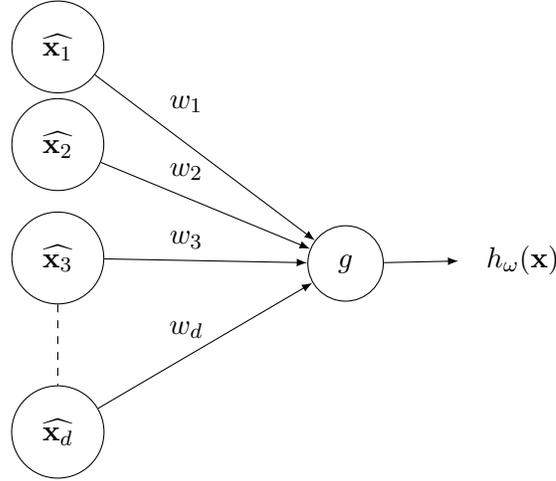


FIGURE 3.16 : Schéma fonctionnel de la régression logistique.

Lors de la phase d'apprentissage, le vecteur de caractéristiques $\widehat{\mathbf{x}}$ est calculé pour les pixels \mathbf{x} « feu » (pour lesquels on définit $\mathcal{W}(\mathbf{x}) = 1$) et pour les pixels \mathbf{x} « non-feu » (pour lesquels on définit $\mathcal{W}(\mathbf{x}) = 0$). Le vecteur ω est estimé à partir des informations obtenues sur l'ensemble de ces pixels en minimisant une fonction d'erreur $J(\omega)$ définie par [125] :

$$J(\omega) = -\frac{1}{N} \sum_{i=1}^N (\mathcal{W}(\mathbf{x}^i) \log(h_\omega(\mathbf{x}^i)) + (1 - \mathcal{W}(\mathbf{x}^i)) \log(1 - h_\omega(\mathbf{x}^i))) \quad (3.59)$$

Avec N le nombre de pixels d'apprentissage et \mathbf{x}^i le $i^{\text{ème}}$ pixel de la base d'apprentissage.

Pour minimiser $J(\mathbf{x})$, l'algorithme de descente en gradient est employé. Cet algorithme consiste à calculer itérativement jusqu'à convergence les valeurs ω_k du vecteur ω par la formule suivante :

$$\omega_k = \omega_k - \alpha \frac{\partial J(\omega)}{\partial \omega_k} \quad (3.60)$$

Avec $\alpha \in]0; 1]$ le taux d'apprentissage, et :

$$\frac{\partial J(\omega)}{\partial \omega_k} = \frac{1}{N} \sum_{i=1}^N (h_\omega(\mathbf{x}^i) - \mathcal{W}(\mathbf{x}^i)) \widehat{\mathbf{x}}_i \quad (3.61)$$

La convergence est atteinte lorsque :

$$\|\omega_c - \omega_{c-1}\| < \epsilon \quad (3.62)$$

Avec ω_c et ω_{c-1} les vecteurs de caractéristiques ω calculés à l'étape courante et précédente, $\|\cdot\|$ la norme euclidienne et ϵ une constante fixée par l'utilisateur.

Les caractéristiques intégrées dans le vecteur $\widehat{\mathbf{x}}$ peuvent être de nature différente. Dans cette thèse, nous avons considéré une première version de méthode basée sur l'utilisation d'un vecteur composé des résultats des règles de détection vues dans la partie 3.3.1 de sorte que :

$$\widehat{\mathbf{x}}_k = r_k \quad \forall k \in [1 ; 34] \quad (3.63)$$

La seconde version de méthode travaille avec des caractéristiques de pixel qui sont les valeurs d'intensité du pixel dans différents canaux de couleur et la différence de ces valeurs. A partir des deux canaux de chrominance C_1 et C_2 des espaces de couleur YC_bC_r , YUV , $L^*a^*b^*$, TSI , et TSV est construit un vecteur $\vartheta_{[C_1, C_2]}$ tel que :

$$\vartheta_{[C_1, C_2]}(\mathbf{x}) = \begin{bmatrix} \mathbf{I}_{C_1}(\mathbf{x}) \\ \mathbf{I}_{C_2}(\mathbf{x}) \\ \mathbf{I}_{C_1}(\mathbf{x}) - \mathbf{I}_{C_2}(\mathbf{x}) \\ \mathbf{I}_{C_1}(\mathbf{x})^2 \\ \mathbf{I}_{C_2}(\mathbf{x})^2 \end{bmatrix} \quad (3.64)$$

et le vecteur de caractéristiques $\widehat{\mathbf{x}}$ définit par :

$$\widehat{\mathbf{x}} = \begin{bmatrix} \vartheta_{[C_b, C_r]}(\mathbf{x}) \\ \vartheta_{[U, V]}(\mathbf{x}) \\ \vartheta_{[a^*, b^*]}(\mathbf{x}) \\ \vartheta_{[T, S_1]}(\mathbf{x}) \\ \vartheta_{[T, S_2]}(\mathbf{x}) \end{bmatrix} \quad (3.65)$$

Avec S_1 et S_2 respectivement la saturation de l'espace TSI et de l'espace TSV .

En supprimant les deux doublons correspondant aux valeurs $\mathbf{I}_T(\mathbf{x})$ et $\mathbf{I}_T(\mathbf{x})^2$, $\widehat{\mathbf{x}}$ est un vecteur de vingt-trois caractéristiques.

La figure 3.17 présente les résultats de détection de pixels de des deux versions de la méthode par régression logistique sur les mêmes images que celles présentées en figure 3.14. On peut remarquer que ces méthodes obtiennent de meilleurs résultats pour la première image ; ceci peut être dû à la présence de fumée dans la seconde image.

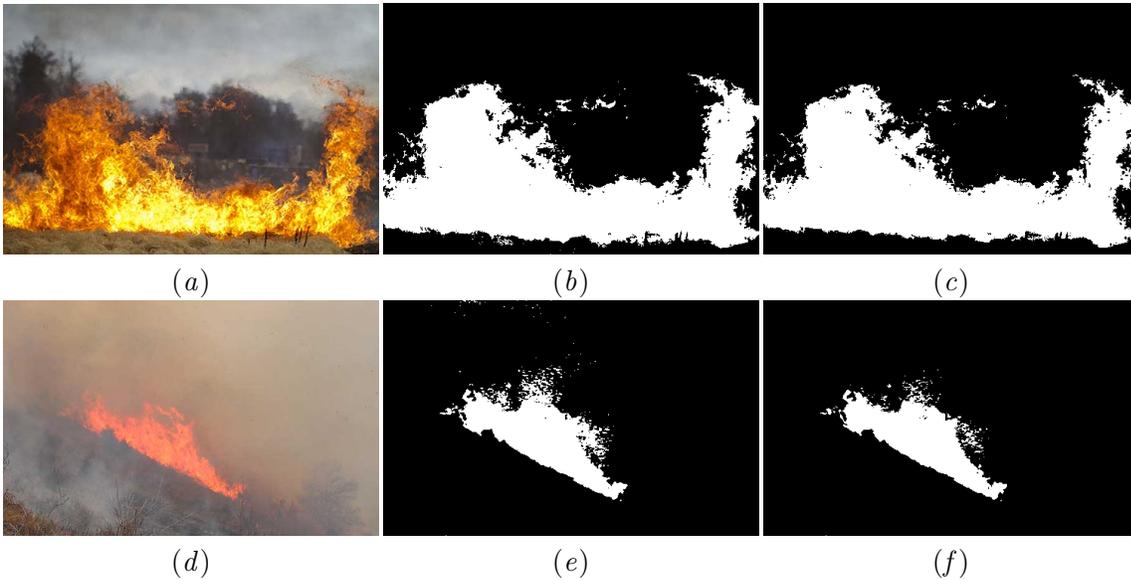


FIGURE 3.17 : Exemples de résultat de détection de pixels obtenus par les deux versions de la méthode par régression logistique. ((a)-(d)) images originales, ((b)-(e)) pixels « feu » détectés par la méthode de régression logistique utilisant les scores des règles et ((c)-(f)) pixels « feu » détectés par la méthode de régression logistique utilisant les intensités de couleur.

3.5 Évaluation des performances des méthodes

Cette section présente les évaluations (quantifiées par la métrique F-Score) et les performances comparées des algorithmes de détection de pixel « feu » décrits dans les parties précédentes.

Tous les algorithmes ont été programmés en Matlab et les évaluations ont été faites sur un ordinateur muni d'un processeur Intel(R) Core™ i7-3632QM CPU à 2,20GHz et muni de 8 Go de mémoire vive.

3.5.1 Évaluation des performances des règles

Dans cette partie, les résultats de l'analyse comparative des performances des trente-quatre règles de détection de pixel « feu » est présentée. Cette analyse a été faite sur deux cent douze millions de pixels « feu » et six cent quatre-vingt-onze millions de pixels « non-feu » (ou « environnement ») des images d'évaluation de la base d'images CFDB.

Les performances des règles ont été étudiées par catégorie de pixel qui sont :

- pour les pixels « feu » : couleur rouge, couleur orange, couleur jaune-blanc, avec fumée, sans fumée, haute texture, faible texture, faible luminosité de l'environnement, moyenne luminosité de l'environnement et haute luminosité de l'environnement
- pour les pixels « non-feu » : faible luminosité, moyenne luminosité et haute luminosité

La catégorie « couleur autre » n'a pas été considérée compte tenu du faible nombre de pixels de

la base qui lui correspond. Lors de la détection des pixels, les pixels déclarés « vrais positifs » (VP) correspondent aux pixels « feu » correctement détectés et les pixels déclarés « vrais négatifs » (VN) correspondent aux pixels « environnement » correctement détectés.

Le tableau 3.4 présente les pourcentages de détection des pixels vrais positifs et des pixels vrais négatifs obtenus pour chaque règle et pour chaque catégorie de pixel. Pour une meilleure lisibilité, les pourcentages ont été présentés en niveau de gris. Plus une case est claire, plus le nombre correspondant est élevé.

Les six premières lignes du tableau donnent les scores de détection des différentes règles par catégorie de pixels « feu » et les trois lignes suivantes indiquent les pourcentages obtenus par ces mêmes règles par catégorie de pixels « non-feu ». Les deux dernières lignes du tableau représentent les taux totaux de pixels vrais positifs et vrais négatifs détectés pour chacune des règles.

La règle r_1 obtient des taux de détection de pixels vrais positifs et vrais négatifs élevés sur l'ensemble des catégories avec un score un peu moins bon pour les pixels « jaune-blanc » et « avec fumée ».

Les règles r_2 , r_3 , r_4 et r_5 ont des taux de détection de pixels vrais positifs très élevés pour les pixels de classe « feu » à l'exception des pixels de couleur rouge pour les règles r_3 et r_5 ainsi que pour la règle r_4 . On peut remarquer, d'autre part, que ces règles ont des taux de détection de pixels vrais négatifs bas, excepté pour la règle r_4 qui obtient des taux élevés pour les pixels de faible et moyenne intensité.

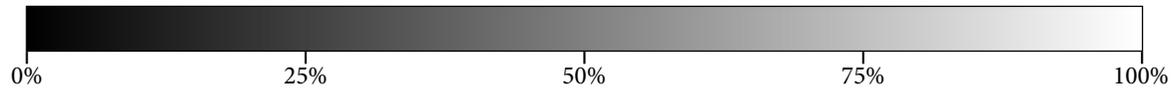
La règle r_6 présente de bons taux de détection de pixels vrais positifs mais des taux de détection de pixels vrais négatifs faibles. La règle r_7 n'est pas robuste aux fumées mais obtient de bons taux de détection de pixels vrais positifs pour toutes les couleurs de feu sans fumée. Cette règle a aussi de faibles taux de détection de pixels vrais négatifs, particulièrement pour les environnements lumineux. La règle r_8 a de bons taux de détection de pixels vrais positifs sauf pour les pixels de couleur rouge avec ou sans fumée. Les taux de détection de pixels faux positifs de cette règle sont bas particulièrement pour les environnements d'intensité faible et moyenne.

Les règles r_9 , r_{10} , et r_{11} ont des taux de détection de pixels vrais positifs élevés mais de faibles taux de détection de pixels vrais négatifs. Ceci s'explique par le fait que les seuils choisis pour ces règles ont été sélectionnés de sorte d'avoir un maximum de détection de pixels vrais positifs.

Les règles r_{12} , r_{13} , r_{14} et r_{15} ont des taux de détection de pixels vrais positifs élevés avec toutefois des scores faibles obtenus par les règles r_{13} et r_{15} pour les pixels de la catégorie « feu jaune-blanc ». Les taux de détection des pixels vrais négatifs sont assez mauvais pour ces quatre règles. On remarque tout de même des scores acceptables pour les pixels « environnement »

TABLEAU 3.4 : Scores des vrais positifs et vrais négatifs pour les 34 règles de détection. La couleur des cases indique l'efficacité d'une méthode : plus la case est blanche, meilleur est le score.

		Phillips		T.H. Chen			Martínez		Horng			Celik RVB				Ko		Celik YC_bC_r								Celik $L^*a^*b^*$				J. Chen		Rossi		Nouvelles règles				
		r_1	r_2	r_3	r_4	r_5	r_6	r_7	r_8	r_9	r_{10}	r_{11}	r_{12}	r_{13}	r_{14}	r_{15}	r_{16}	r_{17}	r_{18}	r_{19}	r_{20}	r_{21}	r_{22}	r_{23}	r_{24}	r_{25}	r_{26}	r_{27}	r_{28}	r_{29}	r_{30}	r_{31}	r_{32}	r_{33}	r_{34}			
Avec Fumée	Rouge																																					
	Orange																																					
	Jaune																																					
Sans Fumée	Rouge																																					
	Orange																																					
	Jaune																																					
I faible																																						
I moyen																																						
I élevé																																						
Tout	VP																																					
	VN																																					



de haute intensité pour les règles r_{13} et r_{14} . La règle r_{12} obtient de bons scores de détection de pixels « environnement » possédant une luminosité faible ou moyenne. Elle est donc adaptée à la détection de pixels de feu dans des environnements sombres (les feux de nuit par exemple). Il est important de rappeler que la règle r_{12} utilise la moyenne des valeurs d'intensité du canal R sur l'image à laquelle appartient le pixel considéré. Elle peut donc produire des résultats différents sur deux pixels ayant la même couleur mais provenant d'images différentes. La règle r_{15} a des taux de détection de pixels vrais négatifs très bas, elle n'est donc pas adaptée à la détection de pixels « feu ».

La règle r_{16} ne semble pas être adaptée à la détection de pixels « feu » car les taux de détection de pixels vrais positifs qu'elle obtient sont médiocres.

Les règles r_{17} , r_{18} , r_{19} , r_{20} , r_{21} , r_{22} , r_{23} et r_{24} ont des bons taux de détection de pixels vrais positifs, notamment pour les pixels de couleur « orange » et « jaune-blanc ». Cependant les règles r_{17} , r_{18} et r_{22} ont des scores plus faibles pour la teinte « rouge ». Les règles r_{20} , r_{21} et r_{22} ont de bons taux de détection de pixels vrais négatifs. Les règles r_{23} et r_{24} ont de très faibles taux de détection de pixels vrais négatifs et de haut taux de détection de pixels vrais positifs. Ces règles sélectionnent donc presque toutes les couleurs de pixels présents dans les images d'évaluation et ne permettent donc pas de bien détecter les pixels « feu » ou même d'améliorer la détection.

Les règles r_{25} , r_{27} et r_{28} obtiennent de bons taux de détection de pixels vrais positifs uniquement pour les pixels de teinte orange et jaune. La règle r_{26} obtiennent de bons taux de détection de pixels vrais positifs uniquement pour les pixels de teinte jaune.

Les règles r_{29} et r_{30} obtiennent de très faibles taux de détection de pixels vrais négatifs. La règle r_{29} détecte presque tous les pixels « feu » et « environnement » et de ce fait elle n'est pas adaptée à la détection de pixels « feu ». Tout comme les règles r_{23} et r_{24} , ces règles sélectionnent une grande partie des couleurs présentes dans la base d'évaluation, elles ne sont donc pas intéressantes pour la détection de pixels « feu ».

Les règles r_{31} et r_{32} obtiennent de bons résultats de détection pour l'ensemble des catégories, particulièrement pour les pixels « sans fumée » et plus spécifiquement ceux de couleur orange.

Enfin, les deux nouvelles règles de détection r_{33} et r_{34} proposées dans cette thèse obtiennent de très bons scores de détection particulièrement pour les pixels « orange ». On remarque qu'ils ont de très bons taux de détection de pixels vrais négatifs.

En conclusion de cette étude relative aux performances comparées des différentes règles de détection de pixels de feu, les règles ayant les meilleurs taux de détection de pixels vrais positifs et vrais négatifs sont les règles r_1 , r_{20} , r_{21} , r_{22} , r_{31} , r_{32} ainsi que les deux nouvelles règles r_{33} et r_{34} . Il est important de souligner que les résultats présentés dans le [tableau 3.4](#) ne donnent pas d'indication sur l'efficacité des méthodes de détection mais des informations

sur le fonctionnement des règles.

3.5.2 Evaluation des performances des différentes versions de la méthode de détection basée sur un apprentissage par histogramme gaussien lissé

Cette partie présente les performances de la méthode de détection basée sur l'utilisation d'un histogramme gaussien lissé et décrite dans la partie 3.3.2.a. Nous avons proposé de considérer cette méthode non plus spécifiquement avec les canaux du système de couleur RVB mais d'étudier ses performances avec des canaux de couleur issus également des systèmes de couleur suivants : YC_bC_r , TSV , TSI et $L^*a^*b^*$. Ces systèmes de couleur présentant quatorze canaux différents, les différentes versions de cette méthode associées aux combinaisons différentes de trois canaux ont été évaluées. Toutes les combinaisons possibles, soit un nombre de 364, ont été considérées.

Les résultats de détection, exprimés sous la forme de la valeur moyenne de la métrique F-score, obtenus par les versions de cette méthode présentant les meilleurs scores sur les images d'évaluation de la base CFDB, sont présentés, par catégorie de pixel, dans le tableau 3.5. Les cinq premières lignes de ce tableau montrent les résultats de la méthode utilisant des canaux d'un même système de couleur ; par exemple les canaux R , V et B du système de couleur RVB . Les autres lignes de ce tableau donnent les scores des versions de cette méthode utilisant des canaux issus de systèmes de couleur différents et obtenant les six meilleurs résultats moyens. Le temps d'exécution de chaque version de méthode inclut les temps d'exécution des actions suivantes : conversion de l'espace de couleur, quantification et normalisation, calcul des probabilités correspondant à chaque valeur de triplet de couleur de pixel par apprentissage, récupération des probabilités associées aux valeurs de triplet du pixel étudié et seuillage. Ce temps est présenté dans la dernière colonne du tableau. Les meilleures performances dans chaque catégorie et par type de triplet de canaux (intra ou inter système) sont présentées en gras. Le meilleur score par catégorie est grisé.

Parmi les cinq systèmes de couleur, le $L^*a^*b^*$ génère les meilleurs résultats dans toutes les catégories de pixels à l'exception de la catégorie « jaune-blanc ». Par contre, le temps d'exécution de la version associée à ce système est très élevé par rapport au temps d'exécution des autres méthodes.

Parmi les versions de méthode utilisant des canaux inter-système de couleur, celle qui est associée aux canaux T , a^* et Cr obtient le meilleur score moyen et dans trois des six catégories de pixel. Son temps d'exécution est élevé. La version utilisant les canaux T , C_b et C_r présente des résultats proches des meilleurs et a un temps d'exécution inférieur d'une demi-seconde par rapport à la meilleure méthode.

Enfin, les meilleurs scores par catégorie (à l'exception du temps d'exécution) sont obtenus par

TABLEAU 3.5 : Valeurs de F-Score des méthodes de détection par histogramme tridimensionnel testées sur la base d'images CFDB.

Canaux	Rouge	Orange	Jaune-Blanc	Sans fumée	Avec fumée	Faible Int.	Int. moyenne	Haute Int.	Moyenne	Écart-type	Temps (s)
RVB	0,785	0,807	0,734	0,791	0,806	0,819	0,819	0,775	0,798	0,175	0,253
YC_bC_r	0,791	0,814	0,743	0,800	0,811	0,819	0,822	0,786	0,805	0,175	0,416
TSV_2	0,774	0,798	0,734	0,781	0,798	0,816	0,812	0,763	0,789	0,180	0,477
TS_2I	0,770	0,796	0,728	0,777	0,798	0,818	0,810	0,759	0,787	0,181	0,491
$L^*a^*b^*$	0,796	0,820	0,734	0,806	0,816	0,819	0,827	0,793	0,811	0,174	0,953
Ta^*C_r	0,820	0,858	0,669	0,836	0,846	0,793	0,849	0,840	0,841	0,146	1,168
$TCrb^*$	0,817	0,855	0,662	0,829	0,846	0,796	0,845	0,836	0,837	0,149	1,182
TC_bC_r	0,815	0,855	0,672	0,829	0,847	0,799	0,846	0,834	0,837	0,148	0,528
$C_bC_r b^*$	0,818	0,853	0,679	0,829	0,846	0,799	0,845	0,834	0,837	0,151	1,064
SC_rL	0,798	0,815	0,740	0,804	0,812	0,827	0,823	0,790	0,808	0,176	1,029
RIV_2	0,764	0,811	0,774	0,788	0,807	0,803	0,814	0,780	0,797	0,175	0,266

des versions utilisant des canaux inter-système.

3.5.3 Évaluation des performances de la nouvelle méthode basée sur un apprentissage par machines à vecteurs de support

La nouvelle méthode de détection de pixels proposée dans cette thèse est basée sur l'emploi de machines à vecteurs de support qui nécessitent un apprentissage sur des zones de pixels et non pas uniquement sur des pixels isolés. Idéalement, un travail comparable à celui effectué sur la méthode par histogramme gaussien lissé aurait dû être réalisé afin d'obtenir des résultats sur plusieurs espaces de couleurs. Par manque de temps, ceci n'a pas été effectué et donc il n'est pas possible d'évaluer les performances de cette nouvelle méthode sur différents espaces de couleur sur la base d'images CFDB.

Une petite étude a été effectuée sur les espaces RVB et YC_bC_r sur neuf séquences d'images de feux de végétation acquises à 2 Hz et choisies car elles présentent des caractéristiques d'image différentes. Dans une même séquence de propagation de feu, les images ont des caractéristiques similaires au niveau des couleurs de l'environnement, des textures et des couleurs du feu. Ces propriétés des séquences de feu de propagation se prêtent particulièrement à l'utilisation de la méthode utilisant les SVM décrite dans ce chapitre. Pour appliquer l'algorithme de détection sur une séquence, nous avons donc choisi d'effectuer l'apprentissage sur la première image de propagation.

Le tableau 3.6 présente les moyennes et écarts types des pourcentages de pixels appartenant

aux différentes catégories. Pour les catégories « rouge », « orange », « jaune-blanc » et « fumée », la colonne intitulée μ indique le pourcentage moyen de pixels des images d'une même séquence classés dans cette catégorie et celle intitulée σ donne leur écart-type. Concernant la catégorie « texture », μ et σ représentent la moyenne et l'écart type de l'entropie locale moyenne de chaque pixel classé « feu ». Pour la catégorie « luminosité », μ et σ représentent la moyenne et l'écart type de la luminosité moyenne des pixels classés « environnement ». Enfin, la variable N représente le nombre d'images de chaque séquence. Au total, cent-cinquante-six images ont été segmentées manuellement et utilisées pour réaliser cette évaluation de performance.

TABLEAU 3.6 : Statistiques des caractéristiques d'images des neuf séquences utilisées

Sequene	N	Rouge		Orange		Jaune-Blanc		Fumée		Texture		Luminosité	
		μ	σ	μ	σ	μ	σ	μ	σ	μ	σ	μ	σ
1	19	67,6	5,0	8,8	4,4	0,5	0,6	35,1	7,5	4,5	0,2	27,1	2,3
2	31	25,4	7,7	45,4	2,8	6,5	4,4	7,0	2,5	5,1	0,1	26,5	1,3
3	13	62,3	5,0	19,9	3,7	3,2	2,4	22,7	6,7	4,9	0,1	32,1	2,0
4	7	20,3	6,8	7,4	2,3	13,6	7,1	8,2	4,5	5,5	0,0	35,0	4,9
5	20	29,1	1,0	51,1	21,0	15,9	7,3	3,6	2,2	5,3	0,1	41,9	2,6
6	8	40,8	2,2	41,9	19,8	6,4	4,3	5,5	3,1	5,4	0,1	31,2	4,4
7	31	49,4	1,3	29,6	9,7	7,4	3,4	11,2	5,2	5,5	0,1	36,3	1,1
8	18	47,0	1,0	40,4	8,0	8,9	2,4	2,2	1,5	5,4	0,1	38,6	1,3
9	9	1,1	0,3	53,7	11,3	2,3	3,4	7,5	3,9	4,9	0,2	21,2	0,1

Ce tableau montre que les couleurs de feu, le taux de fumée qui se superpose sur les pixels « feu », ainsi que le niveau de texture et la luminosité de l'environnement sont différents d'une séquence à l'autre. Les séquences 4, 5 et 6 présentent des écarts types relatif à toutes les catégories plus élevés que les autres séquences. Cela signifie que les attributs des pixels des images de ces séquences varient plus dans ces séquences que dans les autres.

Dans chacune des neuf séquences étudiées, cinq zones d'intérêt du feu et dix zones d'intérêt de l'environnement, appartenant à la première image de la séquence faisant apparaître un feu établi, ont été choisies et ceci pour réaliser l'étape d'apprentissage. Un plus grand nombre de zones d'intérêt de pixels « environnement » a été considéré car les couleur et texture de ce type de région sont plus variées que celles des régions de pixels « feu ». Le tableau 3.7 présente les valeurs de F-Score obtenues sur les neuf séquences d'images par la méthode SVM utilisant les espaces RVB et YC_bC_r . L'image ayant servi pour l'apprentissage n'est pas considérée dans le calcul de la valeur du F-score moyen. La dernière colonne du tableau correspond aux moyennes des scores obtenus sur toutes les images.

TABLEAU 3.7 : Comparaison de la méthode SVM sur deux espaces de couleurs sur neuf séquences d’images de propagation

Séquence d’images	1	2	3	4	5	6	7	8	9	Tout
<i>RVB</i>	0.90	0.92	0.93	0.87	0.88	0.87	0.89	0.88	0.88	0.89
<i>YC_bC_r</i>	0.90	0.91	0.92	0.87	0.87	0.86	0.89	0.90	0.89	0.89

La méthode proposée obtient des scores similaires dans les deux espaces de couleur. On remarque que les séquences 4, 5 et 6 ont des scores généralement plus bas que ceux des autres séquences. Il est probable que ceci vienne du fait que ces séquences ont des variations de couleurs plus importantes que les autres.

3.5.4 Comparaison des performances des méthodes

Les onze méthodes de détection de pixels « feu » de la littérature détaillées dans le chapitre 1 ainsi que les cinq nouvelles méthodes proposées dans cette thèse ont été évaluées sur les quatre cent cinquante images de feu de végétation de la base CFDB. Les apprentissages et le calcul des constantes et des seuils nécessaires à certains algorithmes ont été réalisés en utilisant les pixels présentés dans la figure 3.9. L’apprentissage de la méthode utilisant les machines à vecteurs de support a quant à lui été effectué à partir de dix images choisies parmi les images d’apprentissage sur lesquelles ont été sélectionnées neuf régions de 1600 pixels « feu » et dix-huit régions de 1600 pixels « environnement » par image.

Les résultats des différentes méthodes de détection, exprimés sous la forme de la valeur moyenne de la métrique F-score, sont présentés, par catégorie d’image, dans le tableau 3.8. Les catégories relatives à la texture n’ont pas été présentées dans ce tableau car une seule méthode utilise les textures pour la détection de pixels.

Les onze premières lignes de ce tableau montrent les résultats des méthodes de la littérature référencées par le nom de son premier auteur et éventuellement une information de date ou de système de couleur si le même auteur a développé plusieurs méthodes. Les cinq lignes suivantes présentent les résultats des cinq nouvelles méthodes référencées de la façon suivante : la méthode par histogramme gaussien lissé utilisant l’espace Ta^*C_r est notée « HGL Ta^*C_r », celle utilisant les machines à vecteurs de support est mentionnée par « SVM », la méthode par vote des règles est notée « Vote » et les méthodes de régression logistique utilisant les canaux de couleur et sur les règles sont respectivement référencées par « RL Couleurs » et « RL Règles ». Les meilleures performances des méthodes de la littérature et des nouvelles méthodes par catégorie d’image sont présentées en gras. Le meilleur score par catégorie est grisé.

Parmi les méthodes de la littérature, la méthode référencée « Rossi » obtient le meilleur score dans six des huit catégories et le meilleur score moyen. Son temps d’exécution est de 0,379

seconde, et donc 3,24 plus petit que le temps d'exécution moyen des méthodes de la littérature qui s'élève à 1,23. Les méthodes référencées « Phillips » et « Martínez » obtiennent également de bons scores moyens.

L'analyse des performances de chaque méthode par catégorie conduit aux observations suivantes :

Parmi les méthodes de la littérature, la méthode référencée « Rossi » obtient le meilleur score dans six des huit catégories et le meilleur score moyen. Son temps d'exécution est de 0,379 seconde, et est donc 3,24 fois plus petit que le temps d'exécution moyen des méthodes de la littérature qui s'élève à 1,23 seconde. Les méthodes référencées « Phillips » et « Martínez » obtiennent également de bons scores moyens.

L'analyse des performances de chaque méthode par catégorie conduit aux observations suivantes :

- la méthode « Phillips » présente des scores similaires dans les différentes catégories d'image avec toutefois une valeur de F-score moindre dans la catégorie « jaune-blanc ». On peut également remarquer que la performance de détection décroît légèrement à mesure que l'intensité lumineuse augmente. On peut dire que ses performances sont dépendantes de la luminosité.

- la méthode « Chen 2004 » obtient ses meilleurs scores (toutefois moyens) dans les catégories « jaune-blanc » et « faible intensité » ce qui la conditionnerait à un usage plutôt ciblé sur des images obtenues de nuit ou d'intérieur.

- la méthode « Martínez » présente de bonnes performances dans la catégorie « orange ». Ses performances sont dépendantes de l'intensité.

- la méthode « Horng » obtient de très faibles résultats dans toutes les catégories. Son score moyen est égal à 0,116.

- la méthode « Celik RVB » présente de faibles performances dans toutes les catégories. Son score moyen est égal à 0,345.

- la méthode « Celik YC_bC_r » obtient des résultats moyens avec ses meilleurs scores obtenus dans les catégories « jaune-blanc » (0,709) et « intensité moyenne » (0,727)

- la méthode « Ko » présente de faibles scores dans toutes les catégories, avec un score bas dans la catégorie « jaune-blanc » égale à 0,277. Son score moyen est égal à 0,545

- la méthode « Chen 2010 » obtient des performances très similaires à celles de « Chen 2004 ». Son score moyen est égal à 0,657

- la méthode « Rossi » obtient de bonnes performances et son score moyen est égal à 0,808. Toutefois elle n'est pas robuste pour des images de classe « jaune blanc ». De plus, il apparaît que les scores sont moins bons lorsque la luminosité est faible. Ceci peut s'expliquer par le fait que les zones de feux sont le plus souvent de couleur jaune ou blanc dans les images présentant une luminosité faible car lorsque les caméras sont en mode automatique, le temps d'ouverture de la caméra s'adapte à l'environnement.

- la méthode « Celik $L^*a^*b^*$ » n'obtient pas de très bonnes performances (score moyen égale à 0,542) surtout pour les images de classe « jaune-blanc ». De plus cette méthode est dépendante

TABLEAU 3.8 : Valeurs de F-Score des méthodes de détection testées sur la base d'images CFDB.

Méthode	Rouge	Orange	Jaune-Blanc	Sans fumée	Avec fumée	Faible Int.	Int. moyenne	Haute Int.	Moyenne	Écart-type	Temps (s)
Phillips [27]	0,759	0,776	0,650	0,759	0,775	0,791	0,788	0,742	0,766	0,183	0,561
Chen 2004 [41]	0,604	0,678	0,698	0,622	0,701	0,775	0,718	0,585	0,660	0,217	0,313
Martínez [23]	0,682	0,788	0,598	0,720	0,788	0,801	0,772	0,726	0,753	0,176	0,060
Hornig [38]	0,171	0,100	0,021	0,170	0,056	0,007	0,103	0,144	0,116	0,124	0,466
Celik <i>RVB</i> [28]	0,198	0,411	0,146	0,213	0,490	0,307	0,405	0,291	0,345	0,234	0,135
Celik <i>YC_bC_r</i> [33]	0,647	0,680	0,709	0,684	0,660	0,667	0,727	0,621	0,673	0,209	0,168
Ko [31]	0,547	0,559	0,277	0,550	0,539	0,378	0,582	0,534	0,545	0,238	0,114
Chen 2010 [40]	0,601	0,675	0,698	0,619	0,698	0,777	0,716	0,581	0,657	0,219	0,137
Rossi [43]	0,822	0,815	0,602	0,804	0,813	0,752	0,817	0,808	0,808	0,177	0,379
Celik <i>L*a*b*</i> [37]	0,344	0,624	0,400	0,473	0,617	0,623	0,593	0,480	0,542	0,219	5,666
Collumeau [30]	0,696	0,645	0,404	0,703	0,589	0,481	0,618	0,704	0,648	0,236	5,571
HGL <i>Ta*C_r</i>	0,820	0,858	0,669	0,836	0,846	0,793	0,849	0,840	0,841	0,146	1,168
SVM	0,592	0,707	0,755	0,653	0,707	0,691	0,711	0,645	0,679	0,223	8,853
Vote	0,768	0,790	0,727	0,772	0,793	0,817	0,798	0,760	0,782	0,196	2,694
RL Couleurs	0,810	0,842	0,755	0,827	0,834	0,821	0,842	0,821	0,830	0,163	3,077
RL Règles	0,758	0,765	0,699	0,760	0,762	0,820	0,766	0,747	0,761	0,217	2,909

de la luminosité.

- la méthode « Collumeau » obtient des scores moyens (score moyen égal à 0,648) et sa performance la plus faible apparaît dans la catégorie « jaune-blanc ». On peut remarquer que ses performances augmentent avec luminosité de l'environnement et qu'elles sont sensibles à la fumée.

Parmi les nouvelles méthodes, la méthode « « HGL Ta^*C_r » » obtient le meilleur score dans six des huit catégories et le meilleur score moyen égal à 0,841. Toutefois cette méthode n'est pas robuste pour la détection de pixels « jaune-blanc » (score égal à 0,669) et son score dans la catégorie « faible intensité » est moyen (0,793). L'analyse des performances par catégorie des autres nouvelles méthodes conduit aux observations suivantes :

- la méthode « SVM » a le score moyen égal à 0,79 et c'est le plus bas des méthodes proposées. Elle génère toutefois le meilleur score de toutes les méthodes dans la catégorie « jaune-blanc » (0,755). Ses performances sont légèrement meilleures que la méthode « Collumeau » qui utilise également les machines à vecteurs de support, particulièrement dans la catégorie « jaune-blanc ».

- la méthode « Vote » présente des scores élevés dans l'ensemble des catégories et son score moyen est égal à 0,782. Sa performance la plus basse est obtenue dans la catégorie « jaune-blanc » (0,727). On peut noter que ses performances dépendantes de la luminosité de l'environnement. D'autre part, elle présente de meilleures performances que la plupart des méthodes de la littérature et notamment dans la catégorie « jaune-blanc ».

- la méthode « RL Couleurs » obtient le deuxième meilleur score moyen de toutes les nouvelles méthodes proposées (0,830) et de toutes les méthodes. Elle obtient dans la catégorie « jaune-blanc » le meilleur score de toutes les méthodes étudiées (0,755).

- Enfin la méthode « RL Règles » obtient un score moyen égale à 0,761 et est donc moins performante la méthode « RL Couleurs ». Son plus faible score apparaît dans la catégorie « jaune-blanc » (0,699) et ses performances dépendent de la luminosité de l'environnement.

En résumé, la méthode « « HGL Ta^*C_r » » obtient le meilleur score moyen de toutes les méthodes présentées (0,841) et les cinq meilleurs scores dans les huit catégories étudiées. La plupart des méthodes ont leur score le plus faible dans la catégorie « jaune-blanc » et il n'existe donc pas de méthode de détection performante pour tout type de feu. Les cinq méthodes proposées ont un score moyen égal à 0,779 qui est donc bien supérieur au score moyen des méthodes de la littérature qui est de 0,592.

Enfin, si on se rapporte au [tableau 1.4](#) (page 31) qui présente les indices de qualité de Hafiane lus dans différents articles présentant les performances de certaines méthodes de détection, on peut remarquer que les méthodes obtenant les meilleures performances ne sont pas les mêmes que celles présentées dans le [tableau 3.8](#). Par exemple, la méthode « Rossi » présente les moins bons scores dans ce tableau alors que dans notre étude elle est la meilleure des méthodes de la littérature. Ceci est sûrement dû au fait que les images utilisées dans les précédents travaux

pour l'évaluation des différentes méthodes n'étaient pas assez hétérogènes et nombreuses.

3.6 Conclusion

Dans ce chapitre, une étude relative aux performances des méthodes de détection de pixels « feu » de la littérature et de nouveaux algorithmes proposés dans le cadre de cette thèse a été présentée. Le travail qui a été mené à apporter plusieurs contributions scientifiques dont :

- la construction d'une base de cinq-cents images caractérisées de feux de végétation afin d'assurer la robustesse des résultats obtenus
- la proposition de deux nouvelles règles de détection de pixels « feu »
- l'analyse détaillée des performances des règles de détection (de la littérature et proposées) par catégorie de pixel
- la proposition de cinq nouvelles méthodes de détection de pixels « feu »
- l'analyse détaillée des performances des méthodes de détection (de la littérature et proposées) par catégorie d'image

Ces travaux ont montré qu'il n'existait pas une méthode de détection performante pour tout type d'image de feux de végétation.

Enfin, ces travaux n'ont pas permis de résoudre la problématique de la détection de pixels « feu » dans des images présentant une fumée épaisse rendant invisible les zones de flamme. L'utilisation d'images acquises dans le spectre infrarouge et le spectre visible doit être envisagée pour dépasser cette difficulté.

Chapitre 4

Détection multimodale de pixels

4.1	Introduction	101
4.2	Base d'images multimodales	102
4.2.1	Triplet d'images	103
4.2.2	Recalage des images	103
4.2.3	Caractérisation des images	106
4.2.4	Création d'un ensemble d'images d'apprentissage	106
4.2.5	Synthèse d'images avec différents types de fumée	107
4.3	Fusion de données : principes théoriques	109
4.3.1	Méthode statistique	110
4.3.2	Méthode du Modèle des Croyances Transférables (MCT)	113
4.4	Méthodes de fusion appliquées à la détection de pixels « feu »	115
4.4.1	Probabilités conditionnelles	116
4.4.2	Détection multimodale par fusion de probabilités	123
4.4.3	Détection multimodale par régression logistique	130
4.5	Évaluation des performances des algorithmes de détection de pixels « feu » et comparaison de leurs résultats	134
4.5.1	Évaluation des performances des algorithmes de détection de pixels « feu » sur la base d'images multimodales et comparaison de leurs résultats	135
4.5.2	Évaluation de la robustesse aux fumées des algorithmes de détection travaillant sur des images du visible et multimodales	138
4.6	Conclusion	141

4.1 Introduction

Dans le chapitre précédent, il a été vu que la détection de pixels « feu » dans des images du visible était complexe du fait des couleurs multiples et inhomogènes des flammes, de la présence ou non de fumée ainsi que de l'intensité lumineuse de l'environnement. La plus grande difficulté est la présence de fumée masquant tout ou partie du feu dans les images et le rendant impossible à détecter. Tous ces problèmes n'apparaissent pas dans les images de feu acquises dans le domaine de l'infrarouge ; de façon anecdotique de petites zones de feu peuvent être rendues invisibles du fait de la présence de fumée de suie dans des images du proche-infrarouge.

Une des difficultés d'exploitation des zones de feu segmentées dans des images de l'infrarouge rencontrée par la communauté scientifique travaillant sur la problématique des feux de forêt résulte de la différence entre les zones de feu qui apparaissent dans le visible et celles de l'infrarouge ; ceci étant dû au fait que les éléments qui émettent dans l'infrarouge ne sont pas identiques à ceux rayonnant dans le visible. Aussi, les chercheurs continuent d'utiliser comme référence des mesures acquises à partir d'images du visible. En tenant compte de l'ensemble de ces éléments, il est apparu intéressant d'exploiter conjointement les informations issues des domaines de l'infrarouge et du visible pour développer un algorithme efficace de détection de pixels « feu ». Les travaux développés dans cette thèse ont été menés avec des images obtenues simultanément dans les spectres du visible et du proche-infrarouge que nous appellerons aussi images multimodales.

La section 4.2 décrit la base d'images que nous avons construite afin de tester et d'évaluer les méthodes de détection multimodale de pixels « feu ».

La section 4.3 présente les principes théoriques de deux des trois approches de combinaison de données utilisables pour la détection multimodale. La troisième méthode repose sur la régression logistique qui a été présentée au chapitre 3.

La section 4.4 détaille les nouvelles méthodes développées dans cette thèse en se basant sur les probabilités conditionnelles, la fusion de probabilités et la régression logistique.

Enfin, la section 4.5 présente les performances des différentes méthodes de détection multimodale testées sur notre base d'images et les compare à celles obtenues sur cette même base par des méthodes monomodales.

4.2 Base d'images multimodales

La base d'images multimodales est composée d'images acquises simultanément dans le domaine du visible et dans le domaine du proche-infrarouge par des caméras JAI AD80-GE munies d'une focale de 6 mm. Quatorze brûlages ont été réalisés lors de quatre journées d'expériences durant les mois de décembre, mars et juin 2012 et 2014. Les propagations de feu ont été effectuées sur une plateforme de 10 m sur 10 m recouverte de différents combustibles (bruyère, genêt, frisure de bois). En moyenne, 25 images par journée d'expérience ont été sélectionnées pour intégrer la base : chaque image du domaine du visible sélectionnée possède une zone de feu distinguable sans équivoque à l'œil nu avec des couleurs et une texture différentes des autres images de la base. Cent images du visible ont ainsi été retenues ainsi que les images simultanément acquises dans le domaine du proche-infrarouge.

4.2.1 Triplet d'images

La base d'images est composée d'un ensemble de triplets d'images : à chaque image dans le domaine du visible lui correspond sa vérité terrain, établie manuellement par un expert, et l'image acquise simultanément dans le proche-infrarouge. La figure 4.1 présente deux triplets d'images de la base d'images multimodales.

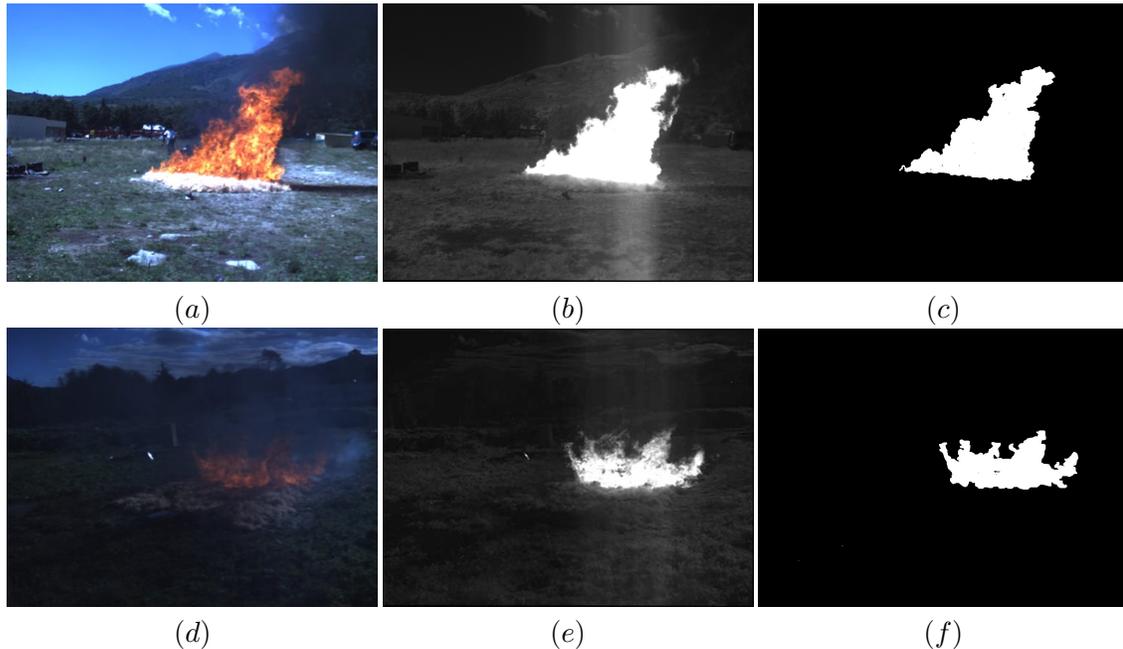


FIGURE 4.1 : Exemple d'images multimodales. (a) et (d) images prises dans le spectre visible ; (b) et (e) images prises dans le spectre proche-infrarouge ; (c) et (f) vérités terrain.

4.2.2 Recalage des images

L'exploitation conjointe d'informations issues des domaines du proche-infrarouge et du visible nécessite que les images multimodales soient superposables. Toutes les images de cette étude ont été prises avec des caméras JAI AD80-GE dans lesquelles un même faisceau lumineux est décomposé en deux rayons, chacun impactant un capteur différent. De très légers décalages entre les images prises dans le visible et dans le proche-infrarouge sont présents. Ces décalages sont de l'ordre du pixel, ils ne sont donc pas visibles à l'œil nu sur une image mais peuvent causer de légères erreurs lors d'une détection multimodale.

Afin d'obtenir des images superposables, un recalage des images a été effectué de sorte que le pixel \mathbf{x} de deux images prises simultanément dans le visible et le proche-infrarouge se rapporte au même point de la scène. L'opération de recalage d'image peut se résumer en trois étapes [126] :

- Mise en correspondance de points de la scène

- Estimation de la transformation géométrique
- Transformation de l'image et ré-échantillonnage

Notons que dans le cas traité, le recalage est plus simple qu'un recalage multimodal tel que l'on peut trouver dans la littérature [127, 128] puisque les images à recaler proviennent de la même caméra et ont la même taille. Pour que le recalage soit précis, il faut que les correspondances de points le soient aussi. Pour cela, une grille de calibrage (telle que celle décrite dans la partie 6.3) a été utilisée afin d'avoir des points de correspondances à partir desquels la transformation géométrique a été estimée. La fonction `detectCheckerboardPoints` de la boîte à outil *Computer Vision System* de `Matlab` a été employée pour la détection automatique de ces points. La sélection des coins de la grille sur une vingtaine de paires d'images a permis d'établir plus de 700 correspondances de points pour chaque caméra utilisée.

L'estimation de la transformation géométrique se fait en utilisant les n correspondances de pixels entre les images du visible et du proche-infrarouge. Des matrices P_1 et P_2 de dimension $n \times 3$ sont créées à partir des coordonnées (x,y) de l'ensemble des coordonnées des points des images respectivement du visible et du proche-infrarouge. Leur première ligne représente les coordonnées ligne, leur deuxième ligne contient les coordonnées colonne et la troisième ligne n'est composée que de 1. La transformation H telle que $P_1H = P_2$ est cherchée. Une solution pour estimer H , décrite par Hartley et Zisserman [129], est l'algorithme de transformation linéaire directe normalisée. Cet algorithme consiste à calculer la décomposition en valeurs singulières d'une matrice A construite ainsi :

$$A = \begin{pmatrix} A_1 \\ \dots \\ A_n \end{pmatrix}; A_i = \begin{pmatrix} \mathbf{O} & -P_2(i,3)P_1(i) & P_2(i,2)P_1(i) \\ P_2(i,3)P_1(i) & \mathbf{O} & -P_2(i,1)P_1(i) \\ -P_2(i,2)P_1(i) & P_2(i,1)P_1(i) & \mathbf{O} \end{pmatrix} \quad (4.1)$$

Avec $\mathbf{O} = (0,0,0)$ et $P_1(i) = (P_1(i,1), P_1(i,2), P_1(i,3))$. A se présente donc sous la forme UDV^T avec D une matrice diagonale contenant des coefficients diagonaux positifs, ordonnés dans l'ordre décroissant. U et V sont des matrices unitaires, c'est-à-dire qu'elles vérifient $U^T U = Id$ et $V^T V = Id$ telle que Id soit la matrice identité. La matrice H correspond au réarrangement en matrice 3×3 de la dernière colonne de la matrice V .

La dernière étape du recalage est la transformation d'une des deux images (visible ou proche-infrarouge) pour qu'elle puisse être superposée à l'autre image. La première approche pour transformer une image consiste à calculer, pour chaque pixel de l'image, le pixel transformé par la matrice H . Le problème de cette technique est que les coordonnées obtenues sont rarement des coordonnées entières. De plus, des pixels sans information peuvent apparaître dans l'image construite comme cela est illustré pour une transformation par étirement dans le schéma présenté en figure 4.2.

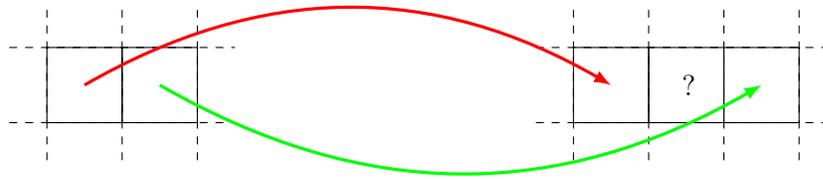


FIGURE 4.2 : Transformation directe d'image par étirement. Il n'y a aucune information sur le pixel central dans l'image transformée

Il existe une méthode qui permet de garantir que tous les pixels de l'image transformée aient un correspondant (voir figure 4.3). C'est une transformation inverse qui nécessite que la matrice H puisse être inversée. Ses différentes étapes sont les suivantes : dans un premier temps, une image composée uniquement de pixels de couleur noire est créée (I_t). Dans notre application, cette image est choisie de taille égale à l'image à transformer (I_d). Les coordonnées de chaque pixel de I_t sont ensuite transformées par H^{-1} de sorte d'obtenir les coordonnées correspondantes dans l'image originale I_d . Dans la figure 4.3, on remarque que le pixel central, dont on ignorait la valeur dans la figure 4.2, a bien un correspondant dans l'image originale. Dans un deuxième temps, les valeurs de couleur des pixels correspondants dans I_d sont copiées dans l'image à transformer I_t . Cependant, comme illustré dans la figure, les coordonnées d'un pixel obtenues par H^{-1} peuvent ne pas être des valeurs entières et se trouver donc entre plusieurs pixels. Il se pose alors le problème de la couleur à donner au pixel dans I_t . Lorsque ce cas se présente, une interpolation des couleurs des pixels voisins de celui calculé est utilisée. Dans le cas de la transformation présentée dans la figure 4.3 la couleur associée au pixel de la flèche bleue sera une interpolation de la couleur des pixels correspondant aux pixels des flèches rouge et verte. Enfin, si la coordonnée calculée n'appartient pas à l'image originale, la couleur de pixel est laissée à zéro.

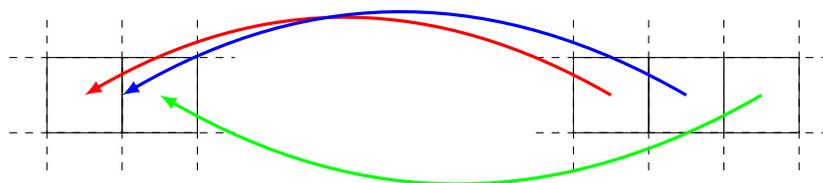


FIGURE 4.3 : Transformation inverse d'image par étirement.

La méthode d'interpolation couleur que nous avons utilisée est l'interpolation bilinéaire. Elle consiste à faire une moyenne pondérée des couleurs des quatre pixels les plus proches du pixel recherché. La pondération associée à la couleur de chaque pixel voisin est fonction de la distance entre le pixel recherché et le pixel voisin. Prenons le cas d'une image en niveaux de gris. Si le pixel recherché dans l'image I_d a des coordonnées (u,v) qui ne sont pas entières, on

définit d_u et d_v les distances en ligne et en colonne des coordonnées entières inférieures. Par exemple si $(u,v) = (40,2 - 120,6)$, on a $d_u = 0,2$ et $d_v = 0,6$. La valeur d'intensité qui sera affectée au pixel \mathbf{x} de l'image I_t est calculée par la formule suivante :

$$I_t(\mathbf{x}) = \begin{pmatrix} d_u & 1 - d_u \end{pmatrix} \begin{pmatrix} I_d(u - d_u, v - d_v) & I_d(u + 1 - d_u, v - d_v) \\ I_d(u - d_u, v + 1 - d_v) & I_d(u + 1 - d_u, v + 1 - d_v) \end{pmatrix} \begin{pmatrix} d_v \\ 1 - d_v \end{pmatrix} \quad (4.2)$$

Avec $I_d(a,b)$ la valeur de l'image I au pixel ayant les coordonnées (a,b) . Pour des images couleurs il faut appliquer ces calculs sur les canaux R , V et B .

Dans cette étude, les images prises dans le domaine proche-infrarouge ont été transformées pour se superposer aux images prises dans le domaine visible.

4.2.3 Caractérisation des images

De la même manière que pour la base d'images CFDB du chapitre précédent, les images du spectre visible de cette nouvelle base sont caractérisées selon la couleur dominante du feu (rouge, orange, jaune-blanc), la présence de fumée (avec fumée, sans fumée) et l'intensité de l'environnement du feu (intensité élevée, moyenne, faible). Les images prises dans le domaine du proche-infrarouge sont catégorisées en fonction de l'intensité moyenne de l'arrière-plan. Si l'intensité moyenne des pixels de l'arrière-plan est supérieure à un seuil τ_{Int} alors l'attribut « intensité élevée » est donné sinon c'est l'attribut « intensité faible » qui est associé. Le seuil a été empiriquement choisi à $\tau_{Int} = 60$. Les images de la figure 4.1 ont été caractérisées avec cette méthode et l'image (a) possède comme caractéristiques : couleur principale rouge, sans fumée, intensité d'environnement moyenne et l'image (d) a comme caractéristiques : couleur principale rouge, avec fumée, intensité d'environnement faible. L'environnement des deux images prises dans le proche-infrarouge est caractérisé avec « intensité faible ».

Le tableau 4.1 donne la répartition par catégorie des 100 images utilisées dans la base multimodale. Les images de couleur principale rouge sont les plus nombreuses. Ceci est dû à la caméra utilisée ainsi qu'à ses réglages. En effet, les feux de ces expériences ont aussi été pris par des appareils photographiques classiques et ajoutés à la base CFDB, et pour ces images, les couleurs principales ont été classées principalement comme orange. On peut donc en conclure que les caractéristiques colorimétriques des feux pris par ces caméras ne sont pas les mêmes que celles des feux de la base CFDB. Il est donc nécessaire de faire une base d'apprentissage spécifique à cette base d'images pour laquelle toutes les images ont été prises avec le même modèle de caméra.

4.2.4 Création d'un ensemble d'images d'apprentissage

Pour construire l'ensemble des images d'apprentissage de la base multimodale, plusieurs triplets d'image ont été sélectionnés. Toutes les images sélectionnées pour l'apprentissage pro-

TABLEAU 4.1 : Répartition par catégorie des images de la base multimodale

Spectre	Base		Nombre d'images
			100
Visible	Couleur	Rouge	87
		Orange	6
		Jaune blanc	7
	Fumée	Avec	14
		Sans	86
	Luminosité de l'environnement	Faible	25
		Moyenne	56
Haute		19	
IR	Intensité	Faible	84
		Haute	16

viennent des acquisitions faites au cours d'une même journée d'expérience. Ce choix a été fait pour que les images d'apprentissage ne soient pas issues des mêmes expériences que les images d'évaluation. Les expériences de juin 2012 ont été choisies pour leur diversité (combustibles différents, feux avec et sans fumée, différents niveaux d'intensité lumineuse). En tout, 24 triplets d'images ont été sélectionnés pour construire cet ensemble d'apprentissage. L'ensemble d'évaluation est donc constitué de soixante-seize triplets d'images.

4.2.5 Synthèse d'images avec différents types de fumée

Les images de feux les plus difficiles à segmenter dans le domaine du visible sont celles qui présentent des zones de flammes recouvertes de fumée. Ce sont aussi celles qui se rapprochent le plus des images acquises lors de feux à grande échelle. La figure 4.4 donne un exemple de telles images obtenues lors d'un brûlage dirigé sur la commune de Solaro (Haute Corse).



FIGURE 4.4 : Exemple de feu masqué par de la fumée dans des conditions réelles. Brûlage dirigé sur la commune de Solaro (Haute Corse) : (a) image acquise dans le spectre visible, (b) image acquise dans le spectre infrarouge.

Avec de telles images il est impossible de délimiter à partir des images obtenues dans le visible le contour des zones de feu et d'établir leur vérité terrain. Aussi, pour évaluer la robustesse aux fumées des algorithmes de détection de pixels de feu, nous avons créé de nouvelles images en fusionnant des images de fumée avec des images de feu sans fumée.

La figure 4.5 présente les trois images de fumée de feux qui ont été utilisées pour synthétiser de nouvelles images. Les fumées sont de couleur blanche, grise et noire. La figure 4.5 présente également les histogrammes associées à ces images de fumée.

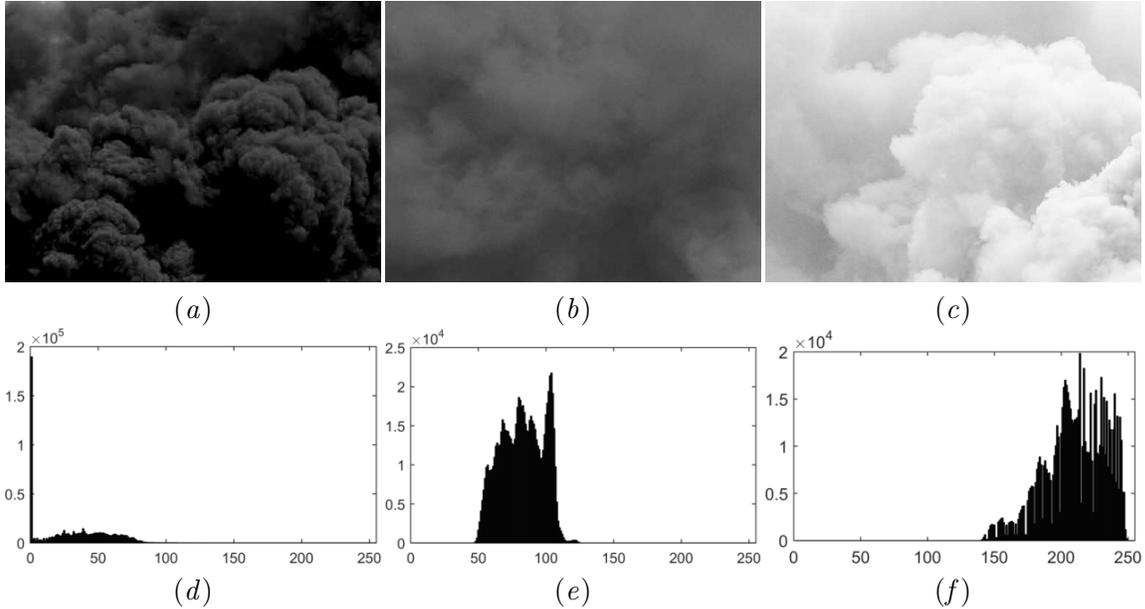


FIGURE 4.5 : Images de fumée utilisées pour construire les images de feu avec fumée. (a) fumée noire et (d) son histogramme associé; (b) fumée grise et (e) son histogramme associé; (c) fumée blanche et (f) son histogramme associé.¹

L'intensité $I_{f_s}(\mathbf{x})$ de chaque pixel \mathbf{x} de l'image synthétisée est obtenue en utilisant la formule suivante :

$$I_{f_s}(\mathbf{x}) = \alpha I_f(\mathbf{x}) + (1 - \alpha)I_s(\mathbf{x}) \quad (4.3)$$

où α est un facteur de transparence compris entre 0 et 1, $I_f(\mathbf{x})$ est l'intensité du pixel \mathbf{x} dans l'image de feu et $I_s(\mathbf{x})$ est l'intensité du pixel \mathbf{x} dans l'image de fumée.

Huit niveaux de transparence de fumée ont été utilisés pour générer, à partir d'une image sans fumée, huit nouvelles images avec des niveaux différents d'opacité de fumée. Le premier niveau ($\alpha = 1$) correspond à une transparence totale, donc à l'image originale. Le deuxième

1. Pour la figure 4.5, les images sont des recadrages de fumée d'images de feux qui ont été récupérées sur Internet. Crédits : (a) U.S. Navy, (b) Hila Shaked (Creative commons), (c) U.S. Department of Agriculture.

niveau ($\alpha = 0,75$) ainsi que le troisième niveau ($\alpha = 0,5$) correspondent à des niveaux de transparence pour lesquels le feu est encore clairement visible (voir figure 4.6). Les niveaux 3 à 6 de transparence ($\alpha = 0,25$, $\alpha = 0,20$, $\alpha = 0,10$, et $\alpha = 0,05$) correspondent à des niveaux élevés de fumée et le dernier niveau ($\alpha = 0$) donne une opacité totale de la fumée, donc l'image de fumée originale.

La figure 4.6 présente les images générées à partir d'une image originale contenant une zone de feu sans fumée fusionnée avec l'image présentée en figure 4.5(a) et en appliquant des niveaux d'opacité $\alpha = 0,5$, $\alpha = 0,25$ et $\alpha = 0,1$.

Les images avec fumée synthétisées ne sont pas insérées dans la base d'images proche-infrarouge/visible ; elles sont uniquement utilisées pour évaluer la robustesse des algorithmes aux fumées dans la partie 4.5.2. Pour chaque algorithme, seules les images ayant permis les meilleures performances de détection de pixels « feu » dans le domaine du visible sont utilisées pour synthétiser des images avec fumée suivant la procédure décrite précédemment. De la sorte, la robustesse aux fumées est évaluée pour chaque algorithme à partir d'un ensemble d'images pour lequel son score de détection de pixels « feu » dans le visible est le meilleur.



FIGURE 4.6 : Exemple de feu masqué synthétiquement par de la fumée. (a) image originale, (b) image modifiée avec $\alpha = 0,5$, (c) image modifiée avec $\alpha = 0,25$ et (d) image modifiée avec $\alpha = 0,1$.

4.3 Fusion de données : principes théoriques

La fusion de données se définit comme « la combinaison de données ou d'informations pour estimer ou prédire l'état d'entités » [130]. Dans nos travaux, les données sont acquises avec des systèmes d'imagerie. On parlera donc de fusion d'images. On différenciera la fusion de résultats qui combine les résultats d'algorithmes comme par exemple les scores de classification des images utilisées, et la fusion de valeurs qui combine les valeurs des pixels des images. La fusion de valeurs est généralement utilisée pour enrichir une image ou créer une image complémentaire, et chaque image contient des informations non contenues dans l'autre. La fusion de résultats elle, est utilisée pour améliorer les performances de reconnaissance. La détection de pixels « feu » étant un problème de classification, l'approche choisie est donc celle de la fusion de résultats. La fusion de valeurs sera utilisée dans le chapitre suivant pour créer des images « feu » plus détaillées afin d'améliorer l'appariement stéréoscopique.

Le principe général de détection de pixels par fusion multimodale de résultats est illustré en figure 4.7. A chaque pixel \mathbf{x} d'une image obtenue à partir d'une source $s = 1, \dots, n$ correspond un vecteur de données \mathbf{x}_s . Ces données peuvent être par exemple l'intensité, la couleur, la température ou la profondeur associée au pixel. Chaque vecteur de données \mathbf{x}_s associé à chaque source s est classé séparément et génère un résultat ; par exemple le résultat peut être « intensité forte » ou « température haute ». Les résultats de toutes les sources sont ensuite fusionnés pour obtenir la classe du pixel qui peut être par exemple « feu » ou « environnement » comme cela sera illustré dans la partie 4.4.2.

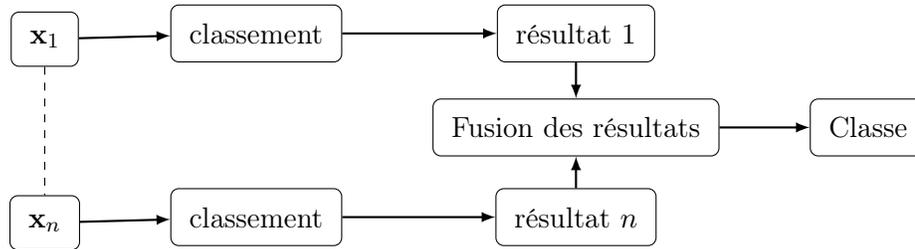


FIGURE 4.7 : Principe général de la fusion de résultats pour un pixel \mathbf{x} ayant des données obtenues à partir de n sources.

Les principales méthodes de fusion de résultats de la littérature utilisent des approches telles que les méthodes statistiques [131], la théorie de croyance de Dempster-Shafer [132] ou les réseaux de neurones [133]. La régression logistique, développée dans le chapitre précédent, est une méthode employant les réseaux de neurones que nous utiliserons aussi pour la fusion multimodale.

Nous présentons dans cette partie la fusion de données par approche statistique et par modèle de croyance transférable. Dans la section suivante, nous décrirons comment appliquer ces fusions de résultats dans le cadre de la détection multimodale de feu.

4.3.1 Méthode statistique

Lee *et al.* [134] ont été les premiers à utiliser une méthode statistique pour le classement de pixels. Cette partie reprend les théories énoncées dans [134] qui seront utilisées pour la détection de pixels « feu ».

Soit n sources de données qui fournissent les mesures \mathbf{x}_s , $s = 1, \dots, n$ pour chacun des pixels de l'image \mathbf{I} . Le but est de classer les pixels de \mathbf{I} en M classes d'information notées w_j , $j = 1, \dots, M$ à partir des données \mathbf{x}_s . Ces classes sont définies par l'utilisateur de la méthode. La donnée \mathbf{x}_s ne permet pas nécessairement à elle seule de classer le pixel \mathbf{x} dans une des classes d'information w_j .

Considérons maintenant chaque source de données séparément. Pour chacune des sources s des

données, il est possible de répartir les pixels dans des classes de données $d_{s,i}$, $i = 1, \dots, m_s$ avec m_s le nombre de classes de données pour le capteur s . Supposons maintenant que les classes d'informations w_j soient liées aux classes de données pour chaque source par une fonction d'appartenance spécifique de la source que nous noterons $f(w_j|d_{s,i}(\mathbf{x}_s))$. Cette fonction mesure la force de l'association de la classe de donnée $d_{s,i}$ avec la classe d'information w_j et peut aussi être influencée par la valeur \mathbf{x}_s (c'est à dire que l'on peut choisir que la classe de donnée d'un pixel ne dépende uniquement que de la classe d'information ou bien qu'elle dépende aussi de la valeur du pixel). Enfin, on définit une fonction globale d'appartenance à la classe d'information w_j comme étant une fonction collectant les différentes forces associées à w_j pour chaque source de données :

$$F_j = F_j\{f(w_j|d_{s,i}(\mathbf{x}_s)); \alpha_s | i = 1, \dots, m_s; s = 1, \dots, n\} \quad (4.4)$$

Avec α_s un facteur de fiabilité associé à la source s .

Le pixel \mathbf{x} est classé dans la classe d'information w_* qui maximise la fonction globale d'appartenance :

$$w_* = \arg \max_j F_j \quad (4.5)$$

Le schéma fonctionnel de la figure 4.8 récapitule la fusion de résultats par la méthode statistique présentée. Pour chaque valeur des pixels $\mathbf{x}_1, \dots, \mathbf{x}_n$ la fonction d'appartenance est calculée pour toutes les valeurs $i \in [1, \dots, m_s]$ et $j \in [1, \dots, M]$. Les fonctions globales d'appartenance de chaque classe d'information sont ensuite calculées à partir des fonctions d'appartenances calculées. La classe d'information w_* choisie est celle dont la fonction globale d'appartenance est maximale.

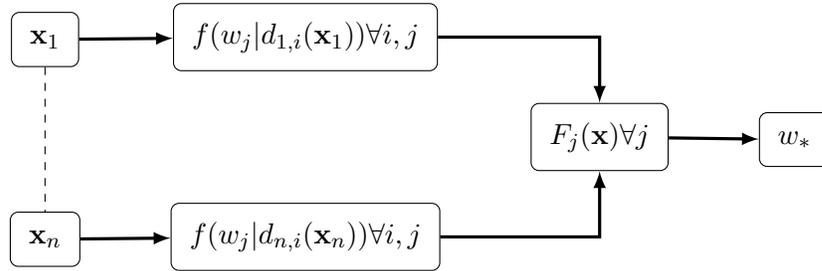


FIGURE 4.8 : Schéma fonctionnel de la fusion de résultats par méthode statistique.

Dans [134] la fonction globale d'appartenance est définie par la probabilité *a posteriori* des sources jointes :

$$F_j(\mathbf{x}) = p(w_j|\mathbf{x}_1, \dots, \mathbf{x}_n) \quad (4.6)$$

Soit deux évènements A et B , le théorème de Bayes stipule que $p(A|B) = \frac{p(B|A)p(A)}{p(B)}$, on a donc :

$$F_j(\mathbf{x}) = \frac{p(w_j)p(\mathbf{x}_1, \dots, \mathbf{x}_n|w_j)}{p(\mathbf{x}_1, \dots, \mathbf{x}_n)} \quad (4.7)$$

En considérant les sources indépendantes, on peut écrire :

$$\begin{aligned} p(\mathbf{x}_1, \dots, \mathbf{x}_n|w_j) &= \prod_{s=1}^n p(\mathbf{x}_s|w_j) \\ &= \prod_{s=1}^n \frac{p(w_j|\mathbf{x}_s)p(\mathbf{x}_s)}{p(w_j)} \end{aligned}$$

La fonction globale d'appartenance devient donc :

$$\begin{aligned} F_j(\mathbf{x}) &= \frac{p(w_j)p(\mathbf{x}_1, \dots, \mathbf{x}_n|w_j)}{p(\mathbf{x}_1, \dots, \mathbf{x}_n)} \\ &= \frac{p(w_j)}{p(\mathbf{x}_1, \dots, \mathbf{x}_n)} \prod_{s=1}^n \frac{p(w_j|\mathbf{x}_s)p(\mathbf{x}_s)}{p(w_j)} \\ &= \frac{H(\mathbf{x})}{p(w_j)^{n-1}} \prod_{s=1}^n p(w_j|\mathbf{x}_s) \end{aligned}$$

Avec

$$H(\mathbf{x}) = \frac{1}{p(\mathbf{x}_1, \dots, \mathbf{x}_n)} \prod_{s=1}^n p(\mathbf{x}_s) \quad (4.8)$$

Le terme $H(\mathbf{x})$ est le même quel que soit j . Il n'intervient donc pas dans la recherche du maximum de F_j et il peut être remplacé par $H(\mathbf{x}) = 1$. La fonction globale d'appartenance peut donc s'exprimer ainsi :

$$F_j(\mathbf{x}) = \frac{1}{p(w_j)^{n-1}} \prod_{s=1}^n p(w_j|\mathbf{x}_s) \quad (4.9)$$

Considérons la probabilité qu'une classe d'information soit vérifiée sachant la valeur observée par la source comme la fonction d'appartenance à cette classe pour la source considérée.

Comme cette fonction est dépendante des classes de données de la source, elle peut donc s'écrire :

$$p(w_j|\mathbf{x}_s) = \sum_{i=1}^{m_s} p(w_j, d_{s,i}|\mathbf{x}_s)/p(\mathbf{x}_s) \quad (4.10)$$

En utilisant le théorème de Bayes on peut écrire cette formule sous deux formes qui peuvent être utilisées selon les cas traitées :

$$p(w_j|\mathbf{x}_s) = \sum_{i=1}^{m_s} p(w_j|d_{s,i}, \mathbf{x}_s)p(d_{s,i}|\mathbf{x}_s) \quad (4.11)$$

ou

$$p(w_j|\mathbf{x}_s) = \sum_{i=1}^{m_s} \frac{p(\mathbf{x}_s|w_j, d_{s,i})p(d_{s,i}|\mathbf{x}_s)p(w_j)}{p(\mathbf{x}_s)} \quad (4.12)$$

Les méthodes pour calculer les termes des équations (4.11) et (4.12) dépendent du contexte de l'utilisation de la fusion. Nous définirons ces termes pour la fusion appliquée à la détection de feu dans la sous-section 4.4.1.

Dans [134], il est proposé de donner plus d'importance à une source en fonction de la qualité du capteur en réécrivant la fonction $F_j(\mathbf{x})$ de l'équation (4.9) ainsi :

$$F_j(\mathbf{x}) = \frac{1}{p(w_j)^{n-1}} \prod_{s=1}^n p(w_j|\mathbf{x}_s)^{\alpha_s} \quad (4.13)$$

Avec α_s le facteur de contribution de la source s . α_s est choisi en fonction de l'incertitude u_s du capteur s :

$$\alpha_i u_i = \alpha_j u_j \quad \forall i, j \in [1 ; n] \quad (4.14)$$

4.3.2 Méthode du Modèle des Croyances Transférables (MCT)

Le Modèle des Croyances Transférables (MCT) développé par Smets *et al.* en 1994 [135] permet la représentation et la combinaison des connaissances. Il se base sur des fonctions de croyance fournies par différentes sources d'information pour déterminer la classe $w_j \in \Omega$ d'un système à partir d'observations \mathbf{x}_s issues de n sources. Son principe est donc semblable à la méthode probabiliste présentée précédemment. Il se distingue de cette dernière par la manière de calculer les probabilités d'appartenance aux classes d'information. Une description très

complète de ce modèle est proposée par Ramasso *et al.* dans [136]. Nous présentons dans cette partie les concepts théoriques importants pour notre application.

La théorie des fonctions de croyance introduite par Dempster en 1968 [137] et Shafer en 1976 [138] permet de modéliser la connaissance par une fonction de masse de croyance m^Ω définie par :

$$\begin{aligned} m^\Omega &: 2^\Omega \rightarrow [0; 1] \\ B &\mapsto m^\Omega(B) \end{aligned}$$

Avec $2^\Omega = \{\emptyset, \{w_1\}, \{w_2\}, \{w_1, w_2\}, \{w_3\}, \dots, \{w_1, w_2, w_3\}, \dots, \{w_1, \dots, w_M\}\}$, l'espace qui rassemble tous les sous-ensembles possibles des classes et des unions de classes de Ω et M le nombre de classes de Ω . Les distributions de masse de croyance pour une source s sont définies comme l'ensemble des masses de croyance des éléments $A \in 2^\Omega$ tel que :

$$\sum_{A \in 2^\Omega} m_s^\Omega(A) = 1 \quad (4.15)$$

Pour chaque élément A de 2^Ω , $m_s^\Omega(A)$ représente la part de croyance de la source s pour cet élément. Contrairement au modèle de Dempster et Shafer, Smets considère le concept du monde ouvert [139] pour les MCT qui admet qu'une masse non nulle peut être attribuée à l'ensemble vide. Les masses de \emptyset et de Ω correspondent respectivement à la masse conflictuelle et la masse d'ignorance, c'est à dire la part de conflit des sources et le taux de croyance attribué au fait de ne pas connaître la classe d'un pixel.

Dans le cadre de la fusion de résultats issus des données de sources distinctes, il existe plusieurs opérateurs de fusion de distribution de masses développées pour le MCT [136]. Nous présentons ici la règle conjonctive de combinaison (CRC) notée \odot qui permet de combiner deux distributions de masses distinctes m_1^Ω et m_2^Ω de la façon suivante :

$$m_{1 \odot 2}^\Omega(A) = m_1^\Omega \odot m_2^\Omega = \sum_{B \cap C = A} m_1^\Omega(B) m_2^\Omega(C) \quad (4.16)$$

A partir de ces connaissances une décision peut être prise. Dans le cadre des MCT, elle se base sur la distribution de probabilités pignistiques notée BetP. La transformée pignistique consiste à répartir la masse d'une classe A sur les hypothèses contenues dans A :

$$\text{BetP} \{m^\Omega\} (A) = \sum_{B \in A} \frac{m^\Omega(B)}{|B|(1 - m^\Omega(\emptyset))} \quad (4.17)$$

Avec $|B|$ le nombre d'éléments dans B . La décision est ensuite prise en choisissant la classe maximisant la probabilité pignistique :

$$w_* = \arg \max_{w_j \in \Omega} \text{BetP} \{m^\Omega\} (w_j) \quad (4.18)$$

Le théorème de Bayes peut être utilisé en extension du modèle des croyances transférables pour calculer les fonctions de masse à partir de vraisemblances. En effet, il permet de calculer des probabilités à partir de données *a priori*. Cette extension, appelée Théorème Bayésien Général, associe une croyance à chaque élément de 2^Ω à partir de plausibilités $pl[w_j](\mathbf{x}_s)$. Dans de nombreux cas, la croyance conditionnelle de \mathbf{x}_s sachant w_j peut être remplacée par une fonction de probabilité $pl[w_j](\mathbf{x}_s) = p(\mathbf{x}_s|w_j)$ [140]. D'après les équations de Smets [141], on a pour tout $A \in \Omega$ le calcul des masses *a posteriori* qui s'exprime ainsi :

$$m[\mathbf{x}_s](A) = \prod_{w_j \in A} pl[w_j](\mathbf{x}_s) \prod_{w_j \in \bar{A}} (1 - pl[w_j](\mathbf{x}_s)) = \prod_{w_j \in A} p(\mathbf{x}_s|w_j) \prod_{w_j \in \bar{A}} (1 - p(\mathbf{x}_s|w_j)) \quad (4.19)$$

4.4 Méthodes de fusion appliquées à la détection de pixels « feu »

Dans cette partie, nous considérons qu'un pixel \mathbf{x} peut appartenir à l'une des deux classes suivantes : la classe « feu » notée \mathcal{F} (correspondant à w_2) et la classe « environnement » (ou « non-feu ») notée \mathcal{E} (correspondant à w_1). $\Omega = \{\mathcal{F}; \mathcal{E}\}$ est l'ensemble de ces deux classes. Chaque pixel est classé à partir de deux modalités différentes : la valeur d'intensité \mathbf{x}_1 produite par le capteur travaillant dans le domaine du proche-infrarouge ($\mathbf{x}_1 = \mathbf{I}_{IR}(\mathbf{x})$) et le vecteur de valeurs \mathbf{x}_2 produit par le capteur travaillant dans le domaine du visible. Ce vecteur contient l'intensité du pixel dans les canaux rouge, vert et bleu : $\mathbf{x}_2 = (\mathbf{I}_R(\mathbf{x}), \mathbf{I}_V(\mathbf{x}), \mathbf{I}_B(\mathbf{x}))$.

Les probabilités conditionnelles $p(w_i|\mathbf{x}_s)$ calculées dans chaque domaine spectral sont fusionnées. La fonction de probabilité $p(\cdot)$ prend ses valeurs dans $[0; 1]$. De manière générale, la classe $w_i(\mathbf{x})$ est associée à un pixel si sa probabilité $p(w_i(x))$ est supérieure à 50% (*i.e.* $p(w_i(x)) > \frac{1}{2}$).

Cette section est organisée de la façon suivante : le calcul des probabilités conditionnelles à partir des différentes modalités est d'abord présenté. Une partie est ensuite consacrée à la description de l'adaptation pour la détection de pixels « feu » de la méthode probabiliste et de la méthode du Modèle des Croyances Transférables. Enfin, un dernier paragraphe est dédié à la présentation d'une méthode de détection multimodale utilisant une méthode de régression logistique.

4.4.1 Probabilités conditionnelles

Cette partie présente le calcul de la probabilité $p(w_i|\mathbf{x}_s)$ qu'un pixel \mathbf{x} appartienne à la classe $w_i \in \Omega$ sachant la donnée \mathbf{x}_s pour le canal $s \in \{1,2\}$.

Comme évoqué précédemment, la valeur $\frac{1}{2}$ représente la frontière entre un état et un autre. Ainsi, si la décision de la classe d'un pixel doit se faire uniquement avec les données d'un capteur s , on aura :

$$w_i(\mathbf{x}) = \begin{cases} \mathcal{F} & \text{si } p(w_i|\mathbf{x}_s) > \frac{1}{2} \\ \mathcal{E} & \text{sinon} \end{cases} \quad (4.20)$$

4.4.1.a Probabilités conditionnelles appliquées aux images du spectre proche-infrarouge

Cette section présente le calcul des probabilités conditionnelles des pixels d'une image d'appartenir à une classe connaissant leur valeur d'intensité dans le spectre proche-infrarouge. Nous définissons les probabilités d'un pixel \mathbf{x} d'être « feu » (respectivement « environnement ») sachant \mathbf{x}_1 par : $p(\mathcal{F}|\mathbf{x}_1)$ (respectivement $p(\mathcal{E}|\mathbf{x}_1)$).

Dans l'infrarouge thermique, la démarcation entre le feu et l'environnement est nette, c'est à dire que l'intensité des pixels des zones de feu est bien plus élevée que celles des pixels de l'environnement. Ceci est dû au fait des grandes différences de rayonnement infrarouge entre le feu et l'arrière-plan. Dans le proche-infrarouge, il en est de même sauf en présence de réflexions du soleil ou du feu lui-même sur des surfaces comme de l'eau, des vitres ou le sol [54]. Comme pour le spectre visible, un mauvais réglage des caméras (temps d'ouverture, mise au point, ouverture du diaphragme,...) peut compliquer l'étape de détection du feu. Nous considérerons donc dans la suite du document que les caméras sont bien réglées pour l'observation du feu, c'est à dire que les pixels du feu ont une plus grande intensité que les pixels de l'environnement dans les images du proche-infrarouge.

Nous faisons la supposition que plus un pixel a une grande intensité dans le proche-infrarouge, plus la probabilité qu'il appartienne à la classe « feu » est grande. Sa probabilité conditionnelle d'appartenir à la classe « feu » est calculée en considérant l'écart entre son intensité \mathbf{x}_1 et un certain seuil t . Comme $\mathbf{x}_1 \in [0; 255]$, l'écart $a = \mathbf{x}_1 - t$ peut prendre comme valeur minimum $a_{min} = -t$ et comme valeur maximum $a_{max} = 255 - t$. Afin de ramener un élément $a \in [a_{min}; a_{max}]$ dans $[0; 1]$ tel que pour tout $a > 0$ alors $r(a) > \frac{1}{2}$, la fonction $r : [a_{min}; a_{max}] \mapsto [0; 1]$ définie $\forall a \in [a_{min}; a_{max}]$ est employée :

$$r(a) = \begin{cases} \frac{a}{-2(a_{min})} + \frac{1}{2} & \text{si } a \leq 0 \\ \frac{a}{2(a_{max})} + \frac{1}{2} & \text{sinon} \end{cases} \quad (4.21)$$

La probabilité conditionnelle d'un pixel d'appartenir à la classe « feu » sachant sa valeur \mathbf{x}_1 acquise par le capteur proche-infrarouge est donnée par :

$$p(\mathcal{F}|\mathbf{x}_1) = r(\mathbf{x}_1 - t) = \begin{cases} \frac{\mathbf{x}_1 - t}{2t} + \frac{1}{2} & \text{si } \mathbf{x}_1 \leq t \\ \frac{\mathbf{x}_1 - t}{2(255 - t)} + \frac{1}{2} & \text{sinon} \end{cases} \quad (4.22)$$

A titre d'exemple, la figure 4.9 présente les valeurs $p(\mathcal{F}|\mathbf{x}_1)$ pour $t = 230$.

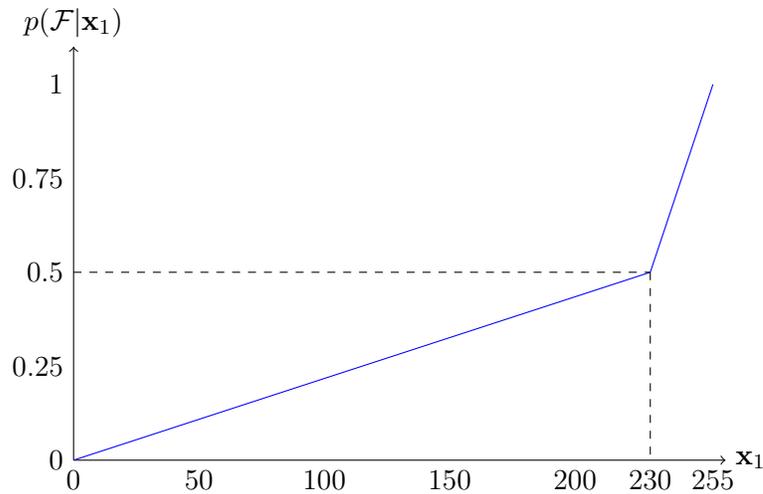


FIGURE 4.9 : Probabilité conditionnelle de la classe « feu » sachant la valeur de pixel dans le spectre proche-infrarouge pour un seuil $t = 230$.

Plusieurs algorithmes de la littérature tels que le seuillage d'Otsu [53], la méthode de Ridler et Calvard [44], l'algorithme de Kittler et Illingworth [58], le seuillage de Martínez [59] ont été testés pour l'estimation du seuil t dans le cas d'images du proche-infrarouge de notre base. Aucun ne nous a semblé satisfaisant, particulièrement lorsque des pixels de l'environnement avaient des valeurs d'intensité élevées.

Nous proposons dans le cadre de cette thèse un nouvel algorithme de sélection de seuil. Sa particularité est qu'il est adaptatif, c'est à dire que la valeur du seuil dépend de l'image à traiter. Le calcul du seuil est basé sur un modèle gaussien représenté par une moyenne μ et un

écart type σ estimés à partir des niveaux d'intensité de tous les pixels de l'image sur laquelle se fait le seuillage. Le seuil t est choisi ainsi :

$$t = \begin{cases} \mu + 2\sigma & \text{si } \mu + 2\sigma \leq 255 \\ 255 & \text{sinon} \end{cases} \quad (4.23)$$

La figure 4.10 présente la détection de pixels « feu » par les différents algorithmes de seuillage cités précédemment dans deux images du proche-infrarouge. Le classement des pixels a été réalisé à partir de la règle (4.20) en utilisant la probabilité conditionnelle définie par la formule (4.22).

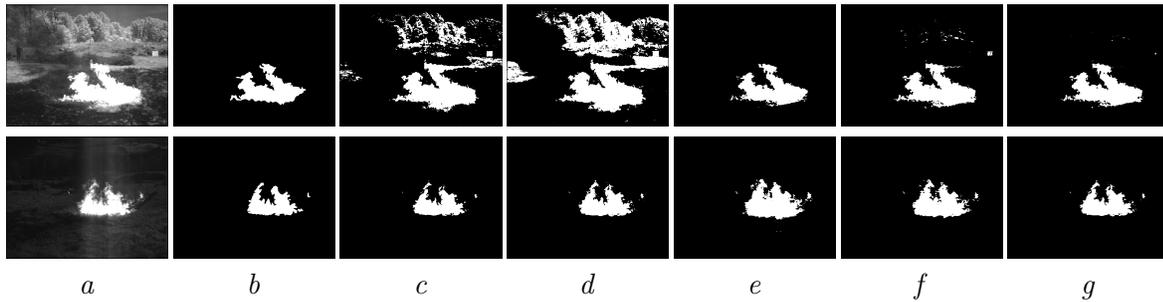


FIGURE 4.10 : Résultats de détection de pixels « feu » dans des images du proche-infrarouge : (a) Images originales du proche-infrarouge; (b) Vérité terrain; (c) Seuillage d'Otsu; (d) Seuillage de Ridler et Calvard; (e) Seuillage de Kittler et Illingworth; (f) Seuillage de Martínez; et (g) Seuillage proposé.

Dans cette figure, il apparaît que l'image supérieure (Image 1) est à l'origine de nombreux pixels faux-positifs détectés par les algorithmes de seuillage d'Otsu, de Ridler et Calvard et l'algorithme de Martínez. L'algorithme de seuillage de Kittler et Illingworth détecte quant à lui des pixels faux-positifs pour l'image inférieure (Image 2). Les valeurs de F-score obtenues par ces méthodes de détection sont présentées dans le tableau 4.2. La méthode de seuillage adaptatif obtient des scores proches du meilleur score des autres méthodes pour les deux images tandis que les autres méthodes présentent des scores moyens ($< 0,85$) ou mauvais ($< 0,7$) pour au moins une des deux images.

TABLEAU 4.2 : Valeurs de F-score obtenues par les différentes méthodes de seuillage

	Otsu	Ridler et Calvard	Kittler et Illingworth	Martínez	Seuillage adaptatif
Image 1	0,574	0,402	0,913	0,857	0,891
Image 2	0,917	0,903	0,783	0,846	0,909

La probabilité que le pixel \mathbf{x} appartienne à la classe « environnement » sachant sa valeur \mathbf{x}_1 est définie comme étant la probabilité de ne pas appartenir à la classe « feu », c'est à dire :

$$p(\mathcal{E}|\mathbf{x}_1) = 1 - p(\mathcal{F}|\mathbf{x}_1) \quad (4.24)$$

4.4.1.b Probabilités conditionnelles appliquées aux images du spectre visible

Il existe dans la littérature plusieurs méthodes permettant d'obtenir la probabilité d'un pixel d'appartenir à la classe « feu » sachant la valeur donnée par le capteur visible [31, 27]. Les techniques d'apprentissage sont les plus efficaces pour une bonne détection de pixels « feu » mais nécessitent d'avoir une base d'apprentissage qui contient des images de feux de même caractéristiques que celles traitées (voir chapitre 3).

Dans le cadre de cette thèse, il est proposé de réaliser un apprentissage de la couleur du feu non pas sur une base d'images mais sur les pixels de l'image à traiter aux positions correspondant aux pixels classés « feu » dans l'image du proche-infrarouge. L'avantage de cette méthode est que la couleur du feu apprise n'est pas générique mais spécifique à l'image traitée.

La procédure s'effectue en deux étapes : dans une première phase, une détection de pixels « feu » par la méthode de seuillage adaptatif présenté dans la partie précédente est appliquée à l'image proche-infrarouge. Les pixels contenus dans la zone rectangulaire (dite « boîte ») englobant les régions de feu ainsi détectées sont classés suivant qu'ils sont détectés comme des pixels « feu » ou « environnement ». Dans une deuxième étape, les pixels de l'image couleur situés aux positions des pixels identifiés comme appartenant au feu dans l'image proche-infrarouge sont utilisés pour faire un apprentissage de la couleur du feu en utilisant un histogramme gaussien lissé tridimensionnel dans l'espace *RVB*. Les autres pixels permettent de faire un apprentissage de la couleur de l'environnement. L'espace *RVB* a été choisi car c'est l'espace dans lequel sont enregistrées les images et qu'il est donc plus rapide de traiter les images dans cet espace. La figure 4.11 présente le schéma des étapes suivies pour le calcul des probabilités conditionnelles dans le spectre visible.

La méthode d'apprentissage par histogramme tridimensionnel présentée dans le chapitre 3 permet de générer une probabilité d'appartenance à la classe « feu » pour chaque couleur de l'espace *RVB*. Cette probabilité a des valeurs positives (respectivement négatives) si la couleur appartient probablement à la classe « feu » (respectivement à la classe « environnement »). Ici, les pixels utilisés pour générer l'histogramme tridimensionnel sont uniquement ceux contenus dans la boîte englobante. Soit $HGL(\mathbf{x}_2)$ la valeur de l'histogramme gaussien lissé pour la couleur du pixel \mathbf{x}_2 , alors la probabilité conditionnelle que le pixel \mathbf{x} appartienne au feu sachant la valeur de \mathbf{x}_2 donnée par le capteur visible est :

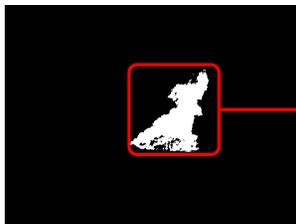
$$p(\mathcal{F}|\mathbf{x}_2) = r(HGL(\mathbf{x}_2)) \quad (4.25)$$

Avec la fonction $r()$ définie par l'équation (4.21) et $HGL(\mathbf{x}_2) \in [a_{min}; a_{max}]$ avec les valeurs

Spectre proche-infrarouge



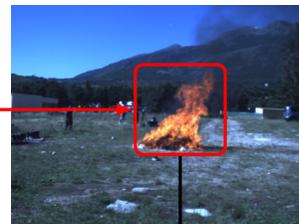
Seuillage adaptatif



Spectre visible



Boîte englobante



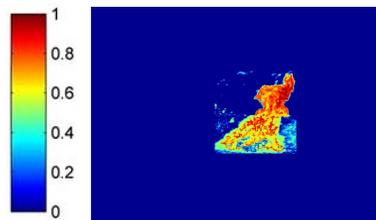
Sélection de pixels



Apprentissage de la couleur



Détection *HGL*



$p(\mathcal{F}|\mathbf{x}_2)$

FIGURE 4.11 : Schéma fonctionnel de la méthode pour calculer les probabilités conditionnelles associées au spectre visible.

des bornes données par :

$$a_{min} = \min_{a \in [0;255]^3} HGL(a) \qquad a_{max} = \max_{a \in [0;255]^3} HGL(a)$$

La figure 4.12 présente un exemple des valeurs de probabilité conditionnelle $p(\mathcal{F}|\mathbf{x}_2)$ obtenues à partir des pixels (dont la composante bleue de \mathbf{x}_2 vaut 100) contenus dans la boîte englobante présentée dans la figure 4.11. On peut remarquer d'une part qu'un grand nombre de couleurs ne sont pas contenues dans la zone de l'image délimitée par la boîte et ont donc une probabilité de 0,5. D'autre part, les couleurs qui ont une grande composante rouge ont une grande probabilité d'appartenir au feu. Enfin, la probabilité $p(\mathcal{F}|\mathbf{x}_2)$ a sa plus faible valeur pour $\mathbf{x}_2 = (63; 42; 100)$ (couleur mauve sombre) ce qui signifie que cette couleur correspond très probablement à une couleur de l'environnement.

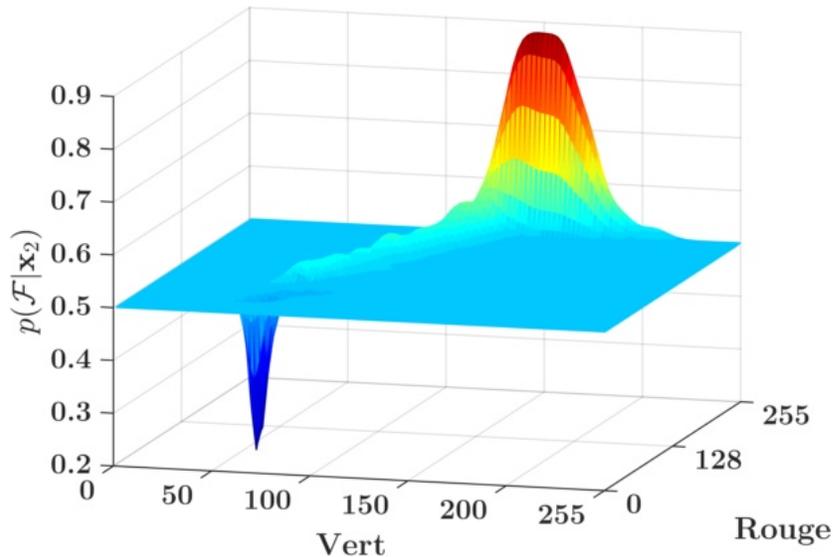


FIGURE 4.12 : Probabilité conditionnelle $p(\mathcal{F}|\mathbf{x}_2)$ obtenue à partir des pixels (dont la composante bleue de \mathbf{x}_2 vaut 100) contenus dans la boîte englobante présentée dans la figure 4.11

Une fois que l'apprentissage est effectué, les pixels de l'image du spectre visible contenus dans la boîte englobant les pixels détectés dans le proche-infrarouge sont classés en suivant la règle (4.20). La figure 4.13 montre le résultat de ce classement de pixels pour l'image couleur présentée dans la figure 4.11

Le temps correspondant à l'étape d'apprentissage de cette méthode peut être diminué en réduisant le nombre de pixels utilisés pour réaliser cette action mais s'accompagne d'une baisse de performance de la méthode de détection de pixels. Les images de la base ont une taille de



FIGURE 4.13 : Classement de pixels pour l'image couleur présentée dans la figure 4.11. (a) Classement de pixels correspondant à $p(\mathcal{F}|\mathbf{x}_2) > \frac{1}{2}$ et (b) vérité terrain.

768 × 1024. La méthode de détection a été évaluée sur les soixante-seize images de test de la base avec trois autres tailles d'image : 255 × 300, 150 × 200 et 112 × 150. Le tableau 4.3 présente les valeurs moyennes de F-score et du temps d'exécution par image obtenues en fonction de la taille des images testées. Il apparaît clairement que la performance et le temps d'exécution de la méthode de détection diminue avec la taille de l'image. L'image originale (768 × 1024) permet d'obtenir le meilleur score mais conduit à un temps d'exécution de 5,248 s par image traitée. L'image de taille 255 × 300 permet d'avoir des scores meilleurs que dans le spectre proche-infrarouge (voir tableau 4.7) seul tout en ayant un temps d'exécution de l'ordre d'une demi seconde en moyenne pour une image. Cette taille d'image pour l'apprentissage a donc été retenue et sera utilisée dans la suite du document.

TABLEAU 4.3 : Valeurs moyenne de F-score obtenues par la méthode de détection par apprentissage pour différentes tailles d'image.

	Sur image 768 × 1024	Sur image 255 × 300	Sur image 150 × 200	Sur image 112 × 150
Moyenne	0,912	0,908	0,905	0,903
Écart-type	0,029	0,030	0,033	0,032
Temps (s)	5,248	0,477	0,242	0,161

La probabilité que le pixel \mathbf{x} appartienne à la classe « environnement » sachant sa valeur \mathbf{x}_2 est définie comme étant la probabilité de ne pas appartenir à la classe « feu » :

$$p(\mathcal{E}|\mathbf{x}_2) = 1 - p(\mathcal{F}|\mathbf{x}_2) \quad (4.26)$$

4.4.2 Détection multimodale par fusion de probabilités

La détection multimodale par fusion de probabilités consiste à combiner les probabilités associées à chaque pixel des images acquises dans le visible et dans le proche-infrarouge. Trois méthodes ont été étudiées pour combiner ces données. La première consiste à calculer une moyenne pondérée des probabilités obtenues dans chaque domaine spectral, la seconde est une approche statistique et la troisième utilise le modèle des croyances transférables.

L'emploi de ces méthodes suppose que les images acquises dans le domaine du visible sont exploitables, c'est à dire que les zones de feu sont discernables toutes ou en grande partie. Un seuil de 70% de pixels « feu » non recouverts de fumée a été choisi empiriquement au deçà duquel les images acquises dans le domaine du visible sont considérées comme inexploitables et la détection des zones de feu s'effectue uniquement à partir d'un traitement monomodal des images acquises dans le proche-infrarouge. Pour chaque couple d'images du visible et du proche-infrarouge, le pourcentage de pixels recouverts de fumée est estimé à partir du nombre de pixels « feu » détectés par seuillage adaptatif dans l'image du proche-infrarouge et du nombre de pixels « feu » détectés par les règles de Horng r_6 , (3.23) r_7 , (3.24) et r_8 (3.25) dans l'image du visible. La méthode de Horng a été choisie pour l'obtention de cet indicateur car parmi les méthodes sensibles aux fumées elle obtient les meilleurs scores de détection sur la base multimodale (voir section 4.5). En faisant la supposition que les pixels détectés dans les images du visible sont tous des pixels « feu » sans fumée, le pourcentage de pixels sans fumée est calculé en effectuant le quotient du nombre de pixels détectés dans le visible par le nombre de pixels détectés dans le proche-infrarouge.

4.4.2.a Méthode par moyenne pondérée

La première méthode de fusion proposée consiste à effectuer une moyenne pondérée des probabilités calculées à partir des images du visible et du proche-infrarouge selon la formule suivante :

$$p(\mathcal{F}|\mathbf{x}) = \alpha p(\mathcal{F}|\mathbf{x}_2) + (1 - \alpha)p(\mathcal{F}|\mathbf{x}_1) \quad (4.27)$$

Avec α une valeur comprise entre 0 et 1.

A partir de $p(\mathcal{F}|\mathbf{x})$ un pixel est classé « feu » si :

$$p(\mathcal{F}|\mathbf{x}) > \frac{1}{2} \quad (4.28)$$

Deux méthodes différentes ont été utilisées pour déterminer la valeur de α .

La première est la méthode d'optimisation décrite dans la section 3.3.1. Elle a été employée pour déterminer la valeur de α qui génère le plus haut score moyen de détection de pixels « feu » sur l'ensemble des images d'apprentissage de la base. La valeur déterminée est $\alpha = 0,629$.

La deuxième procédure est basée sur l'équation (4.14) qui intègre l'incertitude des capteurs. Les valeurs de l'équation ont été estimées à partir de la moyenne des résultats de F-score obtenus sur l'ensemble des images d'apprentissage par les méthodes monomodales de détection travaillant respectivement sur les images du proche-infrarouge (μ_1) et du visible (sans utiliser la boîte englobante) (μ_2) présentées en section 4.4.1. Les valeurs u_1 et u_2 ont ensuite été calculées en effectuant les opérations $u_1 = 1 - \mu_1$ et $u_2 = 1 - \mu_2$. Les valeurs moyennes de F-score obtenues sur notre base d'apprentissage ont été $\mu_1 = 0,908$ et $\mu_2 = 0,794$. Dans le cas d'une moyenne pondérée, on a $\alpha_2 = 1 - \alpha_1$. En utilisant l'équation (4.14), on obtient :

$$\begin{aligned} \alpha_1 u_1 = (1 - \alpha_1) u_2 &\Leftrightarrow \frac{1 - \alpha_1}{\alpha_1} = \frac{u_1}{u_2} \\ &\Leftrightarrow \frac{1}{\alpha_1} - 1 = \frac{u_1}{u_2} \\ &\Leftrightarrow \frac{1}{\alpha_1} = \frac{u_1}{u_2} + 1 = \frac{u_1 + u_2}{u_2} \\ &\Leftrightarrow \alpha_1 = \frac{u_2}{u_1 + u_2} \end{aligned}$$

Ce qui donne $\alpha_1 = 0,691$ pour notre étude.

La méthode par moyenne pondérée a été testée les valeurs de α obtenues par les deux méthodes exposées précédemment ainsi que pour $\alpha = 0,5$. Ce dernier poids correspond à une simple moyenne des probabilités conditionnelles des deux modalités, sans favoriser une modalité plutôt qu'une autre.

Lors des essais effectués sur la base d'image, la méthode avec $\alpha = 0,5$ obtient des résultats légèrement supérieurs. Une explication possible est que la base de test employée contient des images dont les caractéristiques diffèrent légèrement de la base d'apprentissage. Une base d'apprentissage contenant plus d'images permettrait de déterminer une valeur de α plus robuste.

Dans le reste du document seule la méthode avec une valeur $\alpha = 0,5$ est utilisée. Elle est nommée méthode de détection multimodale par moyenne des probabilités conditionnelles.

La figure 4.14 montre les résultats de cette nouvelle méthode de détection appliquée aux couples d'images du visible et du proche-infrarouge présentés en figure 4.1. La figure 4.1 (a) présente une zone de feu contenant un pourcentage de pixels non recouverts de fumée supérieur à 70% ; une moyenne des probabilités conditionnelles du spectre du visible et du spectre du proche-infrarouge est donc effectuée et les pixels sélectionnés sont présentés en figure 4.14 (d). La figure 4.1 (d) ne valide pas, quant à elle, la condition relative au pourcentage de pixels « feu » non recouverts de fumée. Seules les probabilités conditionnelles du spectre proche-infrarouge sont alors considérées pour sélectionner les pixels « feu ». La valeur de F-Score

obtenue pour la première image est de 0,958 et est supérieure à celle obtenue pour la détection dans le spectre infrarouge seule (0,936) et que pour la détection dans le spectre visible seule (0,952).

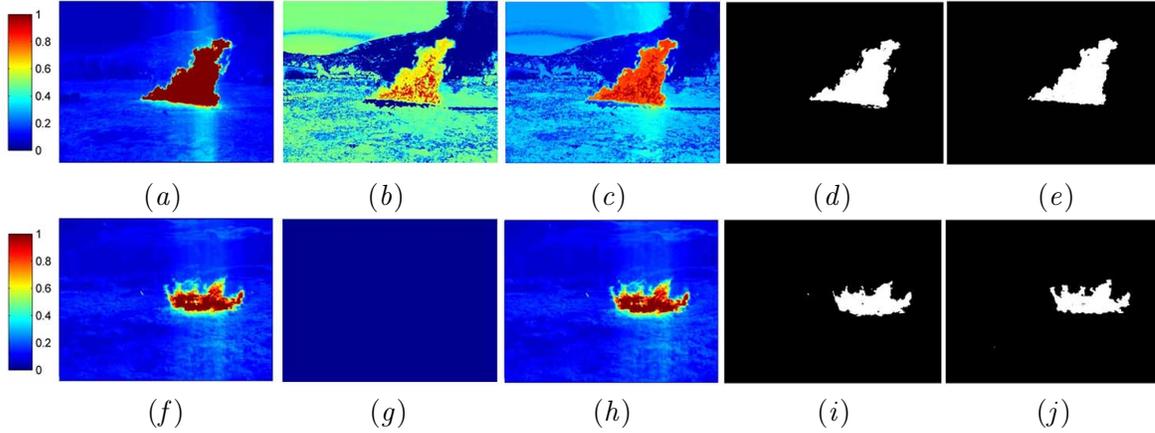


FIGURE 4.14 : Détection multimodale par moyenne pour les images de la figure 4.1. (a) et (f) $p(\mathcal{F}|\mathbf{x}_1)$, (b) et (g) $p(\mathcal{F}|\mathbf{x}_2)$, (c) et (h) $p(\mathcal{F}|\mathbf{x})$ (d) et (i) pixels détectés, (e) et (j) vérité terrain.

4.4.2.b Méthode statistique

Le deuxième algorithme de détection multimodale de pixels « feu » proposé dans cette thèse est basé sur la méthode statistique de fusion de données présentée en section 4.3.1. En utilisant les notations déjà employées dans cette section, nous considérons dans cette partie :

- un système composé de deux capteurs optiques ($n = 2$), l'un travaille dans le domaine du proche-infrarouge et correspond à la source $s = 1$, l'autre travaille dans le domaine du visible et correspond à la source $s = 2$; les deux sources sont supposées indépendantes.
- Deux classes d'informations : w_1 correspond à la classe « feu » et w_2 à la classe « environnement » ; donc $M = 2$.
- À chaque capteur correspond deux classes de sources : $d_{s,1}$ correspondant à la classe « feu » du capteur s et $d_{s,2}$ correspondant à la classe « environnement » du capteur s .

Le tableau 4.4 récapitule les noms des différentes classes utilisées dans le cadre de cette méthode.

La probabilité conditionnelle sur les classes d'information sachant les données des sources $p(w_j|\mathbf{x}_s)$ est calculée en utilisant l'équation (4.11) et s'écrit donc :

$$p(w_j|\mathbf{x}_s) = \sum_{i=1}^2 p(w_j|d_{s,i},\mathbf{x}_s)p(d_{s,i}|\mathbf{x}_s)$$

TABLEAU 4.4 : Noms associés aux classes d'information et de source.

Classe d'information	Nom
w_1	« feu »
w_2	« environnement »
Classe de source	Nom
$d_{1,1}$	« feu » dans le spectre proche-infrarouge
$d_{1,2}$	« environnement » dans le spectre proche-infrarouge
$d_{2,1}$	« feu » dans le spectre visible
$d_{2,2}$	« environnement » dans le spectre visible

Le calcul de $p(w_j|d_{s,i},\mathbf{x}_s)$ est réalisé en considérant que la probabilité conditionnelle d'appartenir à une classe d'information est uniquement dépendante de la classe de source, ce qui implique que la valeur \mathbf{x}_s n'a pas d'influence dans le calcul de cette probabilité. On a donc :

$$p(w_j|d_{s,i},\mathbf{x}_s) = p(w_j|d_{s,i}) \quad (4.29)$$

De plus, il est posé que la probabilité d'appartenir à une classe d'information sachant la classe de source est égale à 1 si les indices des classes sont égaux et qu'elle est égale à 0 si les indices des classes sont différents (voir tableau 4.5).

TABLEAU 4.5 : Valeurs des $p(w_j|d_{s,i})$ choisies pour la détection de pixels de classe « feu ».

	$d_{1,1}$	$d_{1,2}$	$d_{2,1}$	$d_{2,2}$
w_1	1	0	1	0
w_2	0	1	0	1

La probabilité conditionnelle de classes d'information sachant les données des sources s'écrit :

$$\begin{aligned} p(w_j|\mathbf{x}_s) &= \sum_{i=1}^2 p(w_j|d_{s,i},\mathbf{x}_s)p(d_{s,i}|\mathbf{x}_s) \\ &= \sum_{i=1}^2 p(w_j|d_{s,i})p(d_{s,i}|\mathbf{x}_s) \\ &= p(d_{s,j}|\mathbf{x}_s) \end{aligned}$$

Elle est donc égale à la probabilité conditionnelle de classe de source et correspond à $p(\mathcal{F}|\mathbf{x}_s)$ et $p(\mathcal{E}|\mathbf{x}_s)$ définies dans la partie 4.4.1.

Comme cela a été présenté dans les sections précédentes, les probabilités conditionnelles associées à la source proche-infrarouge $s = 1$ sont calculées à partir de l'écart $(\mathbf{x}_1 - t)$ ramené à

une valeur entre 0 et 1 alors que celles associées à la source visible $s = 2$ sont estimées à partir de la valeur de l'histogramme gaussien lissé ramenée également à une valeur entre 0 et 1. Le tableau 4.6 présente les formules de calcul de $p(w_j|\mathbf{x}_s)$ pour les deux sources. r est la fonction définie dans l'équation (4.21), t est le seuil calculé pour l'image du spectre proche-infrarouge par la règle (4.23) et $HGL(\mathbf{x}_2)$ est la valeur de l'histogramme gaussien lissé pour la couleur \mathbf{x}_2 (voir partie 4.4.1.b).

TABLEAU 4.6 : Formule de calcul de $p(w_j|\mathbf{x}_s)$ en fonction des sources et des classes.

	\mathbf{x}_1	\mathbf{x}_2
w_1	$r(\mathbf{x}_1 - t)$	$r(HGL(\mathbf{x}_2))$
w_2	$1 - r(\mathbf{x}_1 - t)$	$1 - r(HGL(\mathbf{x}_2))$

Les fonctions globales d'appartenance aux classes d'information sont estimées par la formule (4.13). Deux fonctions globales d'appartenance correspondant chacune à une classe sont à calculer :

$$F_1(\mathbf{x}) = \frac{1}{p(w_1)} p(w_1|\mathbf{x}_1)^{\alpha_1} p(w_1|\mathbf{x}_2)^{\alpha_2} \quad (4.30)$$

et

$$F_2(\mathbf{x}) = \frac{1}{p(w_2)} p(w_2|\mathbf{x}_1)^{\alpha_1} p(w_2|\mathbf{x}_2)^{\alpha_2} \quad (4.31)$$

Leurs calculs nécessitent de déterminer, en plus des valeurs des probabilités conditionnelles, les valeurs de $p(w_j)$ et de α_s pour les deux capteurs.

Dans cette thèse, $p(w_j)$ est défini comme la moyenne des $p(w_s|\mathbf{x}_s)$ quelle que soit la valeur de \mathbf{x}_s . Pour cela les $p(w_s|\mathbf{x}_1)$ sont sommés pour les 256 valeurs de \mathbf{x}_1 possibles (car $\mathbf{x}_1 \in [0; 255]$). Les $p(w_s|\mathbf{x}_2)$ sont sommés à ce résultat pour toutes les valeurs de \mathbf{x}_2 . Il y en a 256^3 possibles du fait que \mathbf{x}_2 est un vecteur de valeurs rouge, verte et bleue pouvant chacune prendre leurs valeurs dans $[0; 255]$. On divise ensuite le résultat par le nombre de combinaisons possibles c'est à dire par $256 + 256^3$:

$$p(w_j) = \frac{\sum_{\mathbf{x}_1^i \in \mathcal{X}_1} p(w_s|\mathbf{x}_1^i) + \sum_{\mathbf{x}_2^i \in \mathcal{X}_2} p(w_s|\mathbf{x}_2^i)}{256 + 256^3} \quad (4.32)$$

Comme dans l'article de Lee *et al.* [134], le facteur de contribution de la source la plus fiable est fixé à 1. On a donc $\alpha_1 = 1$ et à l'aide de la formule (4.14), on obtient :

$$\alpha_2 = \frac{\alpha_1 u_1}{u_2} \quad (4.33)$$

Ce qui donne $\alpha_2 = 0,447$ pour notre étude.

Pour finir, la classe d'information \mathcal{W} attribuée au pixel \mathbf{x} est celle qui maximise la fonction globale d'appartenance :

$$\mathcal{W}(\mathbf{x}) = \begin{cases} w_1 & \text{si } F_1(\mathbf{x}) > F_2(\mathbf{x}) \\ w_2 & \text{sinon} \end{cases} \quad (4.34)$$

La figure 4.15 présente les valeurs $F_1(\mathbf{x})$ et $F_2(\mathbf{x})$ pour les images de la figure 4.1 (a) et (b) ainsi que la détection finale et la vérité terrain. La valeur du F-Score obtenue pour cette image est de 0,956 et est meilleure que pour la détection dans le spectre infrarouge seule (0,936) et que pour la détection dans le spectre visible seule (0,952).

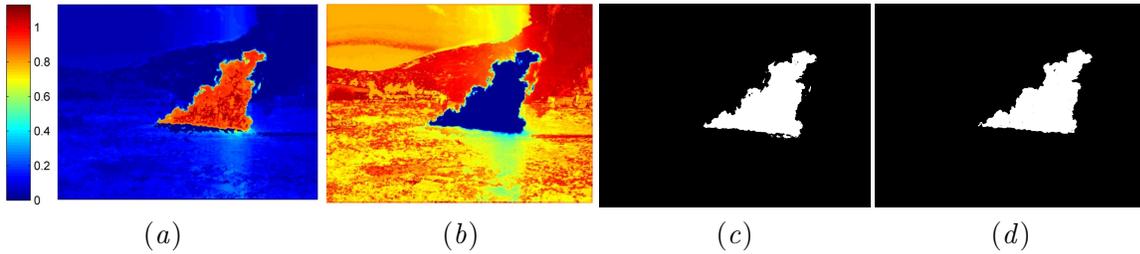


FIGURE 4.15 : Détection par méthode statistique. (a) valeur de la fonction globale d'appartenance à la classe « feu » ($F_1(\mathbf{x})$) pour chaque pixel de l'image, (b) valeur de la fonction globale d'appartenance à la classe « environnement » ($F_2(\mathbf{x})$) pour chaque pixel de l'image, (c) détection finale et (d) vérité terrain.

4.4.2.c Méthode du Modèle des Croyances Transférables (MCT)

On considère que chaque pixel de l'image peut prendre deux états : « feu » ($w_1 = \mathcal{F}$) et « environnement » ($w_2 = \mathcal{E}$) qui sont les éléments de l'ensemble $\Omega = \{\mathcal{F}, \mathcal{E}\}$ appelé état d'ignorance. Dans la théorie de Dempster-Shafer (ou théorie de crédibilité) cet ensemble d'états est élargi à $2^\Omega = \{\emptyset, \{\mathcal{F}\}, \{\mathcal{E}\}, \{\Omega\}\}$ [138]. La classe d'un pixel \mathbf{x} est estimée à partir du MCT selon les phases suivantes :

La première étape de cette méthode de détection consiste à calculer les probabilités $p(\mathbf{x}_s|w_j)$ à partir des $p(w_j|\mathbf{x}_s)$ données dans la partie 4.4.1 en appliquant le théorème de Bayes :

$$p(\mathbf{x}_s|w_j) = \frac{p(\mathbf{x}_s)p(w_j|\mathbf{x}_s)}{\sum_{\mathbf{x}_s^i \in \mathcal{X}_s} p(\mathbf{x}_s^i)p(w_j|\mathbf{x}_s^i)} \quad (4.35)$$

\mathcal{X}_s étant l'ensemble des N_s valeurs que peut prendre le pixel \mathbf{x} pour la source s .

N'ayant pas de connaissances *a priori* sur la valeur d'un pixel, la distribution des probabilités *a priori* pour la valeur des pixels est fixée uniforme. On a donc $p(\mathbf{x}_s^i) = \frac{1}{N_s}, \forall \mathbf{x}_s^i \in \mathcal{X}_s$. Ce qui donne :

$$p(\mathbf{x}_s|w_j) = \frac{\frac{1}{N_s}p(w_j|x_s)}{\sum_{\mathbf{x}_s^i \in \mathcal{X}_s} \frac{1}{N_s}p(w_j|\mathbf{x}_s^i)} = \frac{p(w_j|\mathbf{x}_s)}{\sum_{\mathbf{x}_s^i \in \mathcal{X}_s} p(w_j|\mathbf{x}_s^i)} \quad (4.36)$$

Les masses *a posteriori* peuvent alors être calculées à partir de ces probabilités en utilisant le théorème de Bayes généralisé selon l'équation (4.19) :

$$m^\Omega[\mathbf{x}_s](\emptyset) = \prod_{w_j \in \Omega} (1 - p(\mathbf{x}_s|w_j)) = (1 - p(\mathbf{x}_s|w_1)) (1 - p(\mathbf{x}_s|w_2)) \quad (4.37)$$

$$m^\Omega[\mathbf{x}_s](w_1) = p(\mathbf{x}_s|w_1) (1 - p(\mathbf{x}_s|w_2)) \quad (4.38)$$

$$m^\Omega[\mathbf{x}_s](w_2) = (1 - p(\mathbf{x}_s|w_1)) p(\mathbf{x}_s|w_2) \quad (4.39)$$

$$m^\Omega[\mathbf{x}_s](\Omega) = \prod_{w_j \in \Omega} p(\mathbf{x}_s|w_j) = p(\mathbf{x}_s|w_1)p(\mathbf{x}_s|w_2) \quad (4.40)$$

Il est facilement vérifiable que la somme des masses pour une source donnée est bien égale à 1.

Dans la deuxième étape, les masses jointes $m_{1 \odot 2}^\Omega[\mathbf{x}_s](w_j)$ sont calculées à partir des masses (4.37) à (4.40) en utilisant la formule (4.16) :

$$\begin{aligned} m_{1 \odot 2}^\Omega(\emptyset) &= \sum_{B \cap C = \emptyset} m^\Omega[\mathbf{x}_1](B)m^\Omega[\mathbf{x}_2](C) \\ &= m^\Omega[\mathbf{x}_1](\emptyset)(m^\Omega[\mathbf{x}_2](\emptyset) + m^\Omega[\mathbf{x}_2](w_1) + m^\Omega[\mathbf{x}_2](w_2) + m^\Omega[\mathbf{x}_2](\Omega)) + \\ &\quad m^\Omega[\mathbf{x}_1](w_1)(m^\Omega[\mathbf{x}_2](\emptyset) + m^\Omega[\mathbf{x}_2](w_2)) + \\ &\quad m^\Omega[\mathbf{x}_1](w_2)(m^\Omega[\mathbf{x}_2](\emptyset) + m^\Omega[\mathbf{x}_2](w_1)) + m^\Omega[\mathbf{x}_1](\Omega)m^\Omega[\mathbf{x}_2](\emptyset) \end{aligned} \quad (4.41)$$

$$m_{1 \odot 2}^\Omega(w_1) = m^\Omega[\mathbf{x}_1](w_1)(m^\Omega[\mathbf{x}_2](w_1) + m^\Omega[\mathbf{x}_2](\Omega)) + m^\Omega[\mathbf{x}_1](\Omega)m^\Omega[\mathbf{x}_2](w_1) \quad (4.42)$$

$$m_{1 \odot 2}^\Omega(w_2) = m^\Omega[\mathbf{x}_1](w_2)(m^\Omega[\mathbf{x}_2](w_2) + m^\Omega[\mathbf{x}_2](\Omega)) + m^\Omega[\mathbf{x}_1](\Omega)m^\Omega[\mathbf{x}_2](w_2) \quad (4.43)$$

$$m_{1 \odot 2}^\Omega(\Omega) = m^\Omega[\mathbf{x}_1](\Omega)m^\Omega[\mathbf{x}_2](\Omega) \quad (4.44)$$

Les probabilités pignistiques permettent de prendre une décision sur la classe des pixels. La probabilité que le pixel \mathbf{x} soit dans un état $w_j \in \Omega$ connaissant les données qui lui sont associées dans le spectre proche-infrarouge et visible est donnée par l'équation (4.17). Dans le contexte de notre étude, cette probabilité s'écrit :

$$\text{BetP}\{m_{1\odot 2}^{\Omega}\}(w_j) = \frac{m_{1\odot 2}^{\Omega}(w_j)}{1 - m_{1\odot 2}^{\Omega}(\emptyset)} + \frac{m_{1\odot 2}^{\Omega}(\Omega)}{2(1 - m_{1\odot 2}^{\Omega}(\emptyset))} \quad (4.45)$$

Le pixel \mathbf{x} est classé « feu » si :

$$\text{BetP}\{m_{1\odot 2}^{\Omega}\}(\mathcal{F}) > \text{BetP}\{m_{1\odot 2}^{\Omega}\}(\mathcal{E}) \quad (4.46)$$

Ceci s'écrit également :

$$\frac{m_{1\odot 2}^{\Omega}(w_1)}{1 - m_{1\odot 2}^{\Omega}(\emptyset)} + \frac{m_{1\odot 2}^{\Omega}(\Omega)}{2(1 - m_{1\odot 2}^{\Omega}(\emptyset))} > \frac{m_{1\odot 2}^{\Omega}(w_2)}{1 - m_{1\odot 2}^{\Omega}(\emptyset)} + \frac{m_{1\odot 2}^{\Omega}(\Omega)}{2(1 - m_{1\odot 2}^{\Omega}(\emptyset))} \quad (4.47)$$

Ce qui se réduit à :

$$m_{1\odot 2}^{\Omega}(w_1) > m_{1\odot 2}^{\Omega}(w_2) \quad (4.48)$$

La décision de la classe du pixel \mathbf{x} peut donc se faire uniquement à l'aide de la masse jointe des deux sources pour la classe « feu » et « environnement » *i.e.* les équations (4.42) et (4.43).

La figure 4.16 présente les images des masses jointes et la détection finale pour les pixels des images du visible et du proche-infrarouge de la figure 4.1 (a) et (b) et la vérité terrain. Le F-Score obtenu pour cette image est de 0,953 et est meilleur que celui obtenu par la méthode de détection travaillant uniquement sur les images du spectre proche-infrarouge seul (0,936) et par la méthode de détection travaillant uniquement sur les images du spectre visible seule (0,952).

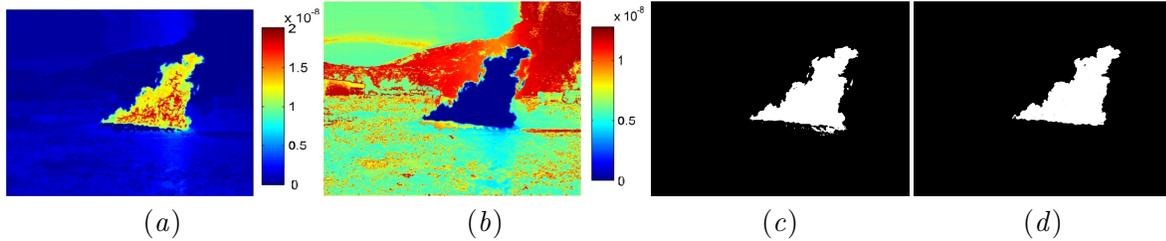


FIGURE 4.16 : Détection multimodale par MCT. (a) valeurs des masses jointes pour la classe « feu » pour chaque pixel de l'image, (b) valeurs des masses jointes pour la classe « environnement » pour chaque pixel de l'image, (c) détection finale et (d) vérité terrain.

4.4.3 Détection multimodale par régression logistique

En collaboration avec le Professeur Turgay Celik de l'Université de Witwatersrand de Johannesburg (Afrique du Sud), deux architectures différentes d'apprentissage par régression logistique ont été développées pour la détection multimodale de pixels « feu ». Ces méthodes nécessitent la construction de vecteurs de caractéristiques créés à partir d'informations estimées en utilisant des images du spectre du proche-infrarouge ($\hat{\mathbf{x}}_1$) et du spectre du visible ($\hat{\mathbf{x}}_2$).

Le vecteur $\widehat{\mathbf{x}}_1$ contient quatre caractéristiques et est construit de la façon suivante :

$$\widehat{\mathbf{x}}_1 = \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_1 - \mu_I \\ \mathbf{x}_1^2 \\ (\mathbf{x}_1 - \mu_I)^2 \end{bmatrix} \quad (4.49)$$

Avec \mathbf{x}_1 la valeur du pixel \mathbf{x} de l'image I dans le spectre proche-infrarouge et μ_I la moyenne des valeurs d'intensités de tous les pixels de l'image du spectre proche infrarouge.

Le vecteur $\widehat{\mathbf{x}}_2$ est identique au vecteur utilisé par la méthode monomodale de détection de pixels « feu » présentée dans le chapitre 3. Il est obtenu en utilisant les équations (3.64) et (3.65) et contient 23 caractéristiques. Sa structure est la suivante :

$$\widehat{\mathbf{x}}_2 = \begin{bmatrix} I_{C_b}(\mathbf{x}) \\ I_{C_r}(\mathbf{x}) \\ I_{C_b}(\mathbf{x}) - I_{C_r}(\mathbf{x}) \\ I_{C_b}(\mathbf{x})^2 \\ I_{C_r}(\mathbf{x})^2 \\ I_U(\mathbf{x}) \\ I_{V_2}(\mathbf{x}) \\ I_U(\mathbf{x}) - I_{V_2}(\mathbf{x}) \\ I_U(\mathbf{x})^2 \\ I_{V_2}(\mathbf{x})^2 \\ I_{a^*}(\mathbf{x}) \\ I_{b^*}(\mathbf{x}) \\ I_{a^*}(\mathbf{x}) - I_{b^*}(\mathbf{x}) \\ I_{a^*}(\mathbf{x})^2 \\ I_{b^*}(\mathbf{x})^2 \\ I_T(\mathbf{x}) \\ I_{S_1}(\mathbf{x}) \\ I_T(\mathbf{x}) - I_{S_1}(\mathbf{x}) \\ I_T(\mathbf{x})^2 \\ I_{S_1}(\mathbf{x})^2 \\ I_{S_2}(\mathbf{x}) \\ I_T(\mathbf{x}) - I_{S_2}(\mathbf{x}) \\ I_{S_2}(\mathbf{x})^2 \end{bmatrix} \quad (4.50)$$

$I_{C_b}(\mathbf{x})$ et $I_{C_r}(\mathbf{x})$ étant la valeur d'intensité du pixel dans les deux canaux de chrominance du système de couleur YC_bCr , $I_U(\mathbf{x})$ et $I_{V_2}(\mathbf{x})$ la valeur d'intensité du pixel dans les deux

canaux de chrominance du système de couleur YUV , $I_{a^*}(\mathbf{x})$ et $I_{b^*}(\mathbf{x})$ la valeur d'intensité du pixel dans les canaux de chrominance du système de couleur La^*b^* , $I_T(\mathbf{x})$ la valeur d'intensité du pixel dans le canal teinte du système de couleur TSI , $I_{S_1}(\mathbf{x})$ la valeur d'intensité du pixel dans le canal teinte du système de couleur TSI et $I_{S_2}(\mathbf{x})$ la valeur d'intensité du pixel dans le canal teinte du système de couleur TSV .

La première architecture proposée est schématisée dans la figure 4.17. Elle diffère de celle présentée au chapitre 3 uniquement par le vecteur de caractéristiques employé qui est ici une combinaison des vecteurs de caractéristiques des deux modalités : $\hat{\mathbf{x}} = [\hat{\mathbf{x}}_0, \hat{\mathbf{x}}_1, \hat{\mathbf{x}}_2]^\top$, avec $\hat{\mathbf{x}}_0 = 1$ une constante unitaire.

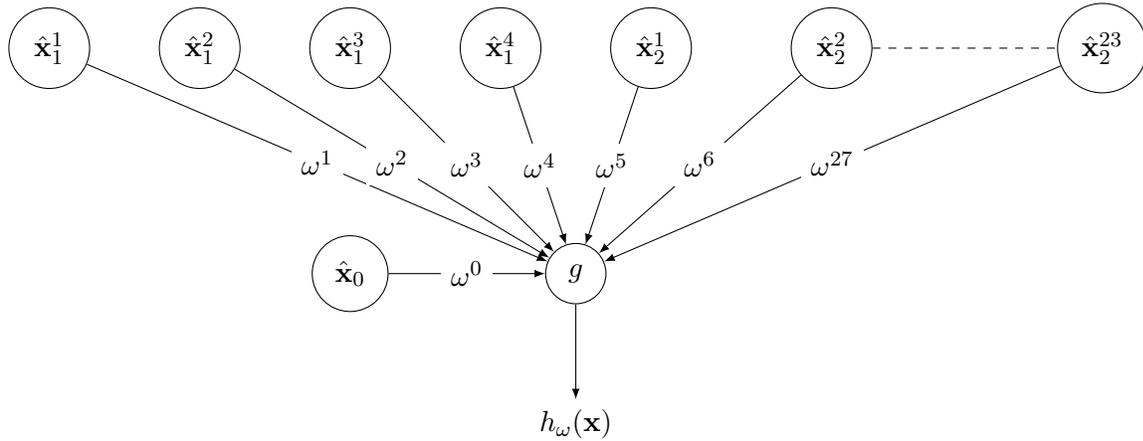


FIGURE 4.17 : Schéma fonctionnel de l'architecture 1 de régression logistique.

$\omega = [\omega^0, \omega^1, \dots, \omega^{27}]$ est un vecteur de 28 éléments (23 du spectre visible, 4 du spectre proche-infrarouge et la constante unitaire). Une étape d'apprentissage permet de l'estimer. De façon analogue à la méthode présentée au chapitre 3, le vecteur ω est obtenu par descente en gradient selon l'équation (3.60) en minimisant la fonction d'erreur $J(\omega)$ définie dans l'équation (3.59).

La fonction $h_\omega(\mathbf{x}) = g(\omega^T \hat{\mathbf{x}})$ produit une valeur comprise entre 0 et 1. Elle correspond à la probabilité que le pixel \mathbf{x} appartienne à la classe « feu ». Un pixel est labellisé « feu » si la valeur de $h_\omega(\mathbf{x})$ est supérieure à un seuil t_h :

$$h_\omega(\mathbf{x}) > t_h \quad (4.51)$$

Le seuil t_h est estimé par optimisation par recherche de motifs (*pattern search* [123]) sur la base d'apprentissage. Sa valeur a été estimée à 0,92 sur notre base multimodale d'apprentissage.

La figure 4.18 présente les valeurs de la fonction de probabilité $h_\omega(\mathbf{x})$ de l'architecture 1, le résultat du classement des pixels et la vérité terrain pour les images des domaines visible et proche-infrarouge de la figure 4.1 (a). Cette méthode de détection qui utilise des informations multimodales obtient pour cette image une valeur de F-Score égale à 0,938. Ce résultat est

meilleur que celui obtenu par la même méthode n'utilisant que des informations exclusivement issues des images du spectre-infrarouge (0,936) mais inférieur à celui obtenu par l'utilisation des probabilités conditionnelles dans le spectre visible $p(\mathcal{F}|\mathbf{x}_2) > \frac{1}{2}$ (0,952).

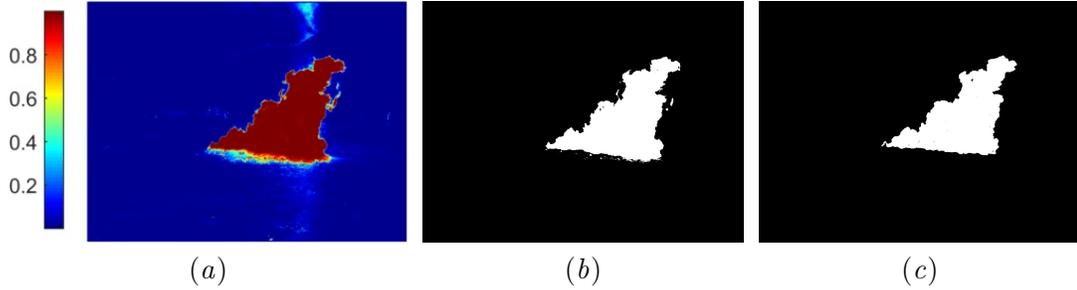


FIGURE 4.18 : Détection multimodale par Régression Logistique (Architecture 1). (a) valeurs des $h_\omega(\mathbf{x})$ pour chaque pixel \mathbf{x} de l'image, (b) détection finale et (c) vérité terrain.

Cette première architecture de régression logistique fusionne toutes les caractéristiques des pixels de la même manière, sans différencier les types de source dont elles sont issues. Une deuxième architecture (nommée architecture 2) qui réalise cette différence a été étudiée et est présentée en figure 4.19.

Dans un premier niveau, les données issues d'une même source sont fusionnées entre elles avec un vecteur ω_i différent pour chaque type de données : $\omega_1 = [\omega_1^1, \omega_1^2, \dots, \omega_1^{23}]$ et $\omega_2 = [\omega_2^1, \omega_2^2, \dots, \omega_2^4]$. Deux vecteurs augmentés $\tilde{\mathbf{x}}_1 = [\hat{\mathbf{x}}_1^0, \hat{\mathbf{x}}_1]^\top$ et $\tilde{\mathbf{x}}_2 = [\hat{\mathbf{x}}_2^0, \hat{\mathbf{x}}_2]^\top$ sont construits avec $\hat{\mathbf{x}}_1^0 = \hat{\mathbf{x}}_2^0 = \hat{\mathbf{x}}^0 = 1$.

Dans un deuxième niveau, les données issues des différents capteurs sont fusionnées. Le vecteur $\omega = \{\omega^k, \omega_1^l, \omega_2^m\}$ contient les paramètres d'apprentissage avec $k = 0, 1, 2, l = 0, \dots, 4$ et $m = 0, \dots, 23$.

Les paramètres d'apprentissage ω_i^j sont estimés itérativement de la façon suivante :

$$\begin{aligned}\omega_0 &= \omega_0 - \alpha \frac{1}{N} \sum_{i=1}^N \delta_i \hat{\mathbf{x}}_0 \\ \omega_1 &= \omega_1 - \alpha \frac{1}{N} \sum_{i=1}^N \delta_i g_1^i \\ \omega_2 &= \omega_2 - \alpha \frac{1}{N} \sum_{i=1}^N \delta_i g_2^i\end{aligned}\tag{4.52}$$

$$\omega_1^l = \omega_1^l - \alpha \frac{1}{N} \sum_{i=1}^N \delta_i \omega_1 g_1^i (1 - g_1^i) \hat{\mathbf{x}}_1^l\tag{4.53}$$

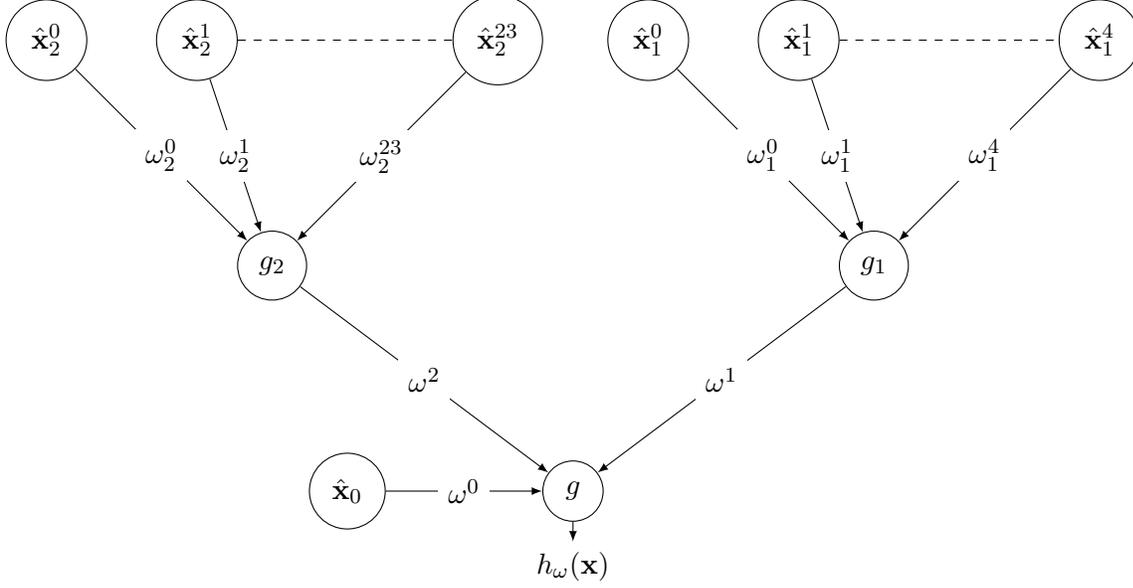


FIGURE 4.19 : Schéma fonctionnel de l'architecture 2.

$$\omega_2^m = \omega_2^m - \alpha \frac{1}{N} \sum_{i=1}^N \delta_i \omega_2 g_2^i (1 - g_2^i) \hat{\mathbf{x}}_2^m \quad (4.54)$$

avec, $g_1^i = g(\omega_1^\top \tilde{\mathbf{x}}_1^i)$, $g_2^i = g(\omega_2^\top \tilde{\mathbf{x}}_2^i)$, $\omega_1 = [\omega_1^0, \dots, \omega_1^{d_1}]$ et $\omega_2 = [\omega_2^0, \dots, \omega_2^{d_2}]$. Le même critère de convergence que pour l'architecture 1 est utilisé. Comme pour l'architecture 1, l'architecture 2 produit une valeur dans $[0; 1]$ représentant la probabilité que le pixel appartienne à la classe « feu ». Cette valeur peut être ensuite utilisée pour classer les pixels à l'aide d'un seuil. Ici, le seuil t_h de la règle de classement (1,73) est le même que pour l'architecture 1 : $t_h = 0,92$ les valeurs des $h_\omega(\mathbf{x})$ pour l'architecture 2, le classement des pixels et la vérité terrain correspondant à la figure 4.1 (a).

La valeur de F-Score obtenue par cette méthode pour l'image de la figure 4.1 (a) est de 0,938. Elle est supérieure à la valeur obtenue par la même méthode n'utilisant que des informations exclusivement issues des images du spectre-infrarouge (0,952) mais inférieure à celle obtenue par la méthode utilisant des probabilités conditionnelles dans le spectre visible $p(\mathcal{F}|\mathbf{x}_2) > \frac{1}{2}$ (0,952).

4.5 Évaluation des performances des algorithmes de détection de pixels « feu » et comparaison de leurs résultats

Afin de comparer les performances de tous les algorithmes de détection de pixels « feu » présentés dans cette thèse, un même ensemble d'images appartenant à la base d'images multimodales a été utilisé pour leur évaluation. Les résultats de cette étude sont présentés dans la première partie de cette section. De plus, les performances de ces algorithmes sur des images de syn-

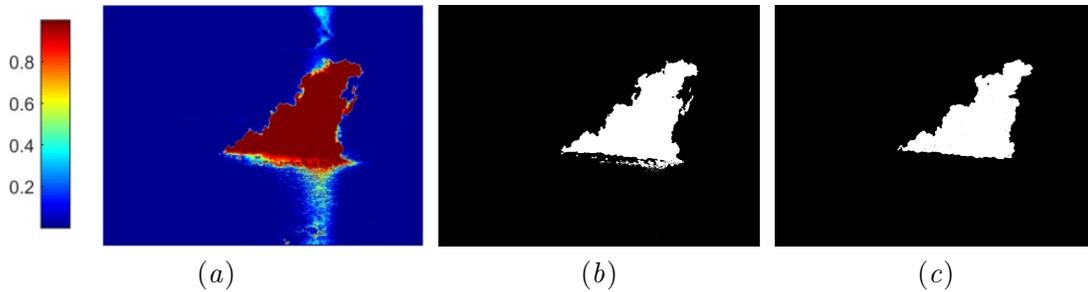


FIGURE 4.20 : Détection multimodale par Régression Logistique (Architecture 2). (a) valeurs des $h_{\omega}(\mathbf{x})$ pour chaque pixel \mathbf{x} de l'image, (b) détection finale et (c) vérité terrain.

thèse contenant de la fumée (dont la méthode de création est décrite dans la section 4.2.5) ont également été évaluées et sont montrées dans la deuxième partie de cette section.

4.5.1 Évaluation des performances des algorithmes de détection de pixels « feu » sur la base d'images multimodales et comparaison de leurs résultats

Tous les algorithmes de détection décrits dans ce chapitre et le chapitre précédent ont été évalués sur la base de test contenant soixante-seize images multimodales.

Les algorithmes évalués uniquement sur les images du visible de cette base sont les algorithmes de : Phillips [27], Chen 2004 [41], Martínez [23], Horng [38], Celik RVB [28], Celik YCbCr [33], Ko [31], Chen 2010 [40], Rossi [43], Celik Lab [37], Collumeau [30] ainsi que les nouvelles méthodes proposées dans cette thèse c'est à dire la détection bayésienne (Bayésienne), la détection par machine à vecteurs de support (SVM), la détection par vote (Vote) et la détection par régression logistique utilisant uniquement des informations couleurs (RL. Couleur) et celle utilisant les critères de détection (RL. Critères). Les apprentissages des algorithmes de Phillips et de ceux utilisant la régression logistique ont été faits sur les 24 images de la base d'apprentissage. Pour les détections de Collumeau et par machine à vecteurs de support, l'apprentissage a été fait en choisissant 15 régions de 1600 pixels « feu » et 30 régions de 1600 pixels de « environnement » par image sur trois images de la base d'apprentissage sélectionnées parmi les trois séries d'images gardées pour l'apprentissage et dans lesquelles les zones de feu sont aisément visibles.

Les méthodes de détection testées uniquement sur les images acquises dans le domaine du proche-infrarouge de cette base sont les algorithmes : d'Otsu [53], de Ridler et Calvard [44], de Kittler et Illingworth [58], de Martínez 2001 [59], et de la méthode basée sur un seuillage adaptatif qui a été développée dans le cadre de cette thèse (Adaptatif). Nous avons aussi testé une technique de segmentation par contours actifs [142] pour détecter le feu. Tous ces algorithmes sont présentés dans la sous-section 1.2.2 du chapitre 1.

Enfin, les cinq techniques de détection multimodales présentées dans ce chapitre ont été éva-

luées sur les soixante-seize images multimodales de test de la base.

Les valeurs de F-score obtenues par les méthodes évaluées sont calculées en utilisant la vérité terrain des images de la base de test. Elles sont présentées dans le [tableau 4.7](#). Le temps de calcul associé à chaque algorithme correspond à son temps moyen d'exécution calculé en considérant les soixante-seize images d'évaluation.

4.5.1.a Comparaison des performances des méthodes de détection travaillant uniquement sur des images du spectre du visible

Les scores des algorithmes utilisant les machines à vecteurs de support (Collumeau *et al.* et SVM) sont les meilleurs. Les méthodes de Phillips *et al.* et de Collumeau *et al.* obtiennent un F-score supérieur à 0,75 avec en temps d'exécution inférieur à une seconde. Si on compare ces scores avec ceux du chapitre 3, les méthodes de Phillips *et al.*, Horng *et al.*, Celik *RVB*, Collumeau *et al.* ainsi que les méthodes SVM et de Vote obtiennent de meilleurs résultats sur cette base. Les autres méthodes ont des résultats plus faibles. Pour les méthodes d'apprentissage, cette amélioration peut s'expliquer par le fait que les images de cette base sont plus homogènes que celles de la base CFDB car elles ont été toutes acquises avec le même type de caméras.

4.5.1.b Comparaison des performances des méthodes de détection travaillant uniquement sur des images du spectre du proche-infrarouge

La méthode de seuillage adaptatif obtient un score moyen supérieur aux autres méthodes de seuillage à égalité avec la méthode de contours actifs (cette dernière ayant un temps d'exécution plus grand). De plus, elle est robuste aux changements d'intensité dans les images (les scores de faible et haute intensité pour les images du spectre proche-infrarouge sont très proches (0,906 et 0,902)).

4.5.1.c Comparaison des performances des méthodes de détection travaillant uniquement sur des images du visible et des méthodes de détection travaillant uniquement sur des images du spectre du proche-infrarouge

Les scores des algorithmes de détection travaillant sur des images du proche-infrarouge sont globalement plus élevés que ceux travaillant sur des images du spectre visible. L'algorithme de Collumeau *et al.* et la méthode par SVM ont toutefois des scores comparables à ceux des méthodes de détection dans le proche-infrarouge. Les seuillages adaptatif et par contours actifs obtiennent des meilleurs scores que toutes les détections du spectre visible. D'autre part, les temps d'exécution des méthodes de détection travaillant dans le domaine du proche-infrarouge sont plus petits que ceux des autres méthodes. Pour la détection des pixels « feu », l'exploitation uniquement d'images du spectre proche-infrarouge paraît donc plus performante que l'exploitation uniquement d'images du spectre visible (avec des vérités terrain établies à partir d'images du visible).

TABLEAU 4.7 : Valeurs de F-Score des méthodes de détection testées sur la base d'images multimodales.

Spectre	Méthode	Visible								Infrarouge		Tout		
		Rouge	Orange	Jaune-Blanc	Faible Int.	Int. moyenne	Haute Int.	Sans fumée	Avec fumée	Faible Int.	Haute Int.	Moyenne	Écart-type	Temps (s)
Visible	Phillips [27]	0,805	0,669	0,620	0,799	0,805	0,745	0,794	0,788	-	-	0,793	0,101	0,594
	Chen 2004 [41]	0,453	0,894	0,569	0,117	0,687	0,568	0,558	0,098	-	-	0,479	0,309	0,155
	Martínez 2002 [23]	0,428	0,746	0,282	0,340	0,525	0,390	0,482	0,240	-	-	0,441	0,168	0,015
	Hornig [38]	0,554	0,912	0,660	0,193	0,778	0,723	0,653	0,204	-	-	0,576	0,312	0,116
	Celik <i>RVB</i> [28]	0,659	0,856	0,378	0,623	0,747	0,489	0,705	0,454	-	-	0,662	0,190	0,053
	Celik <i>YC_bC_r</i> [33]	0,374	0,860	0,435	0,033	0,598	0,533	0,473	0,054	-	-	0,401	0,314	0,072
	Ko [31]	0,393	0,730	0,111	0,124	0,622	0,302	0,468	0,092	-	-	0,404	0,277	0,050
	Chen 2010 [40]	0,459	0,896	0,576	0,122	0,693	0,576	0,564	0,102	-	-	0,485	0,309	0,056
	Rossi [43]	0,724	0,809	0,534	0,724	0,743	0,668	0,736	0,665	-	-	0,724	0,073	3,552
	Celik <i>L*a*b*</i> [37]	0,227	0,590	0,220	0,082	0,352	0,252	0,283	0,064	-	-	0,246	0,182	2,362
	Collumeau [30]	0,870	0,914	0,649	0,874	0,896	0,764	0,881	0,795	-	-	0,866	0,088	0,826
	Bayésienne	0,432	0,864	0,484	0,141	0,633	0,545	0,531	0,094	-	-	0,456	0,280	0,163
	SVM	0,862	0,932	0,673	0,861	0,893	0,765	0,879	0,771	-	-	0,861	0,096	1,221
	Vote	0,807	0,891	0,719	0,805	0,824	0,776	0,827	0,726	-	-	0,809	0,083	1,345
	RL. Couleurs	0,842	0,699	0,494	0,866	0,834	0,723	0,822	0,841	-	-	0,826	0,094	1,492
RL Critères	0,755	0,852	0,633	0,825	0,717	0,741	0,750	0,792	-	-	0,757	0,142	1,119	
Infrarouge	Otsu [53]	-	-	-	-	-	-	0,881	0,894	0,888	0,830	0,883	0,117	0,002
	Ridler [44]	-	-	-	-	-	-	0,830	0,827	0,859	0,830	0,848	0,223	0,050
	Kittler [58]	-	-	-	-	-	-	0,857	0,860	0,854	0,895	0,857	0,120	0,019
	Martínez 2001 [59]	-	-	-	-	-	-	0,838	0,798	0,827	0,887	0,831	0,061	0,029
	Contours actifs	-	-	-	-	-	-	0,907	0,904	0,907	0,891	0,906	0,035	2,761
	Adaptatif	-	-	-	-	-	-	0,908	0,897	0,906	0,902	0,906	0,042	0,009
Multimodal	Moyenne	0,917	0,933	0,865	0,913	0,921	0,909	0,920	0,898	0,917	0,910	0,916	0,026	0,499
	Stat	0,915	0,934	0,869	0,912	0,921	0,901	0,919	0,895	0,915	0,907	0,915	0,032	0,526
	MCT	0,915	0,930	0,883	0,914	0,919	0,908	0,919	0,899	0,917	0,901	0,915	0,025	0,517
	RL. Arch.1	0,884	0,868	0,859	0,870	0,889	0,888	0,891	0,841	0,881	0,903	0,882	0,047	1,541
	RL. Arch.2	0,906	0,919	0,863	0,898	0,915	0,894	0,907	0,901	0,906	0,905	0,906	0,030	1,619

4.5.1.d Comparaison des performances des méthodes multimodales de détection

Les trois méthodes de fusion de probabilités conditionnelles obtiennent des scores similaires et supérieurs aux scores des autres méthodes monomodales avec de meilleures valeurs de F-Score sur la plupart des catégories d'images et les meilleurs écarts types. Leur temps d'exécution est d'environ une demi-seconde. Les méthodes par régression logistique obtiennent des performances inférieures aux autres méthodes de cette catégorie. Ceci peut s'expliquer par le fait que les images utilisées pour l'apprentissage contiennent plus de fumée que celles de l'évaluation et sont en nombre restreint. Enfin, un point important est à remarquer : la méthode de fusion par moyenne a un score des plus élevés et un temps d'exécution le plus petit.

4.5.1.e Comparaison des performances des méthodes multimodales de détection avec celles des autres méthodes

Les moyennes des résultats obtenus par les méthodes de détection multimodales par fusion de probabilité sont meilleures que celles des méthodes travaillant uniquement avec des images du spectre du visible ou du spectre du proche-infrarouge. Les méthodes par régression logistique sont celles pour lesquelles les valeurs de F-score sont les plus basses. Ceci peut s'expliquer par le faible nombre d'images d'apprentissage et la présence de plus de fumées dans les images d'apprentissage que de test. Enfin, les algorithmes multimodaux semblent plus robustes aux fumées (colonnes « avec fumée » et « sans fumée » du [tableau 4.7](#)) que les autres algorithmes. La sous-section suivante présente le travail de quantification de cette robustesse qui a été réalisé pour les algorithmes travaillant sur des images du visible et sur des images multimodales.

4.5.2 Évaluation de la robustesse aux fumées des algorithmes de détection travaillant sur des images du visible et multimodales

La robustesse aux fumées des algorithmes de détection travaillant sur des images du visible et multimodales a été évaluée en utilisant des images de synthèse présentant différents types et densités de fumée.

La méthode utilisée pour synthétiser de telles images a été présentée dans la partie [4.2.5](#). Pour chaque algorithme, les quinze images de la base de test ayant obtenues les meilleurs scores de détection ont été retenues pour cette évaluation. Ce choix d'images permet de tester la robustesse de chaque algorithme en utilisant uniquement les images qui lui ont permis d'obtenir ses meilleurs scores de détection. A partir de chaque image sélectionnée, il est généré trois séries de huit images en combinant l'image de départ avec les trois images différentes de fumée présentées en [figure 4.5 \(a\)](#), [figure 4.5 \(b\)](#) et [figure 4.5 \(c\)](#) pour des niveaux différents de transparence. Les huit niveaux de transparence de fumée correspondant à $\alpha = 1$ (niveau de transparence total) , $\alpha = 0,75$, $\alpha = 0,5$ (niveaux de transparence où le feu est encore clairement visible dans l'image) $\alpha = 0,25$, $\alpha = 0,20$, $\alpha = 0,10$, $\alpha = 0,05$ (niveaux de transparence correspondant à des niveaux élevés de fumée) $\alpha = 0$ (opacité totale de la fumée). Pour cette

dernière valeur, seule l'image de fumée est utilisée. Pour chaque algorithme, la moyenne des scores de détection est calculée pour chaque densité et couleur de fumée. Le score obtenu pour chaque niveau de transparence est ensuite divisé par le score obtenu pour les images sans fumée ($\alpha = 1$). Ceci permet d'avoir le taux de diminution de performance par catégorie d'image sans tenir compte de la performance intrinsèque de l'algorithme.

Le [tableau 4.8](#) présente les ratios des scores en fonction de α et par type et densité de fumée des méthodes de détection travaillant sur des images du visible et multimodales.

4.5.2.a Robustesse aux fumées des algorithmes travaillant avec uniquement des images du spectre visible

Seuls les scores des algorithmes ayant obtenus les dix meilleures performances de détection sur les images du spectre visible de la base multimodale sont présentés. Les méthodes travaillant uniquement sur des images du visible ne peuvent pas détecter des pixels « feu » quand $\alpha = 0$ car les zones de feu sont totalement masquées. D'une manière générale, la fumée blanche produit des identifications de pixels faux positifs car la couleur de fumée se confond avec la couleur du feu. Les algorithmes les plus robustes aux fumées sont la méthode de régression logistique sur les espaces de couleur ainsi que la méthode de Rossi *et al.* qui obtiennent des ratios élevés y compris avec une très faible transparence de fumée. La méthode de détection par vote sur les critères est quant à elle très robuste aux fumées blanches. Les autres méthodes obtiennent des ratios faibles et sont donc à proscrire dans des cas d'images présentant des zones de feu avec fumée. Enfin il serait possible de rendre plus robustes les méthodes par apprentissage travaillant sur des images du spectre visible en réalisant un apprentissage de feu sur des images présentant de la fumée dense.

4.5.2.b Robustesse aux fumées des algorithmes travaillant avec des images multimodales

Tous les algorithmes sont robustes aux fumées avec toutefois une nette diminution des scores lorsque la fumée est totalement opaque. La méthode du modèle des croyances transférables est la plus sensible aux fumées et ses performances diminuent fortement lorsque la transparence de fumée est inférieure à 0,10. Enfin, l'architecture 1 de la régression logistique est très robuste aux fumées contrairement à l'architecture 2 qui obtient de moins bons scores lorsque la fumée est totalement opaque.

Toutes ces évaluations ont été faites en utilisant la fusion des données sans tester au préalable le taux de superposition du feu par la fumée.

TABLEAU 4.8 : Ratio des scores en fonction de la transparence de fumée pour les algorithmes de détection travaillant sur des images du visible et multimodales. N : fumée noire, G : fumée grise et B : fumée blanche.

Méthodes du spectre visible															
Phillips [27]			Horng [38]			Celik RGB [28]			Rossi [43]			Collumeau [30]			
α	N	G	B	N	G	B	N	G	B	N	G	B	N	G	B
0,75	1,00	1,02	1,02	0,95	1,00	0,99	0,99	0,99	0,98	1,00	1,00	1,00	1,00	0,99	0,97
0,50	0,91	1,01	0,88	0,80	0,97	0,90	0,98	0,98	0,86	0,98	1,00	0,99	0,99	0,96	0,86
0,25	0,70	0,84	0,54	0,27	0,85	0,48	0,90	0,84	0,15	0,95	0,98	0,97	0,92	0,83	0,42
0,20	0,59	0,74	0,35	0,16	0,78	0,05	0,83	0,74	0,01	0,92	0,96	0,94	0,89	0,76	0,25
0,10	0,22	0,31	0,22	0,01	0,23	0,00	0,47	0,12	0,00	0,91	0,93	0,93	0,68	0,36	0,00
0,05	0,02	0,04	0,13	0,00	0,00	0,00	0,10	0,00	0,00	0,93	0,96	0,96	0,38	0,00	0,00
0,00	0,00	0,00	0,09	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
Bayesienne			SVM			Vote			RL. Couleur			RL. Critères			
α	N	G	B	N	G	B	N	G	B	N	G	B	N	G	B
0,75	0,92	0,92	0,92	1,00	0,99	0,98	0,99	1,00	1,00	1,00	1,00	0,98	1,03	1,03	1,02
0,50	0,44	0,44	0,44	0,98	0,96	0,88	0,98	0,98	1,00	1,01	1,00	0,99	0,99	0,98	0,94
0,25	0,13	0,13	0,13	0,90	0,83	0,51	0,91	0,90	1,00	1,00	1,01	1,01	0,85	0,74	0,61
0,20	0,04	0,04	0,04	0,85	0,74	0,35	0,86	0,86	0,97	1,00	1,00	1,01	0,76	0,62	0,44
0,10	0,00	0,00	0,00	0,63	0,37	0,00	0,62	0,62	0,96	0,99	0,99	0,99	0,31	0,13	0,00
0,05	0,00	0,00	0,00	0,28	0,00	0,00	0,27	0,32	0,94	0,98	0,98	0,96	0,00	0,00	0,00
0,00	0,00	0,00	0,00	0,02	0,00	0,00	0,00	0,00	0,00	0,07	0,00	0,00	0,00	0,00	0,00
Méthodes multimodales															
Moyenne			Méthode statistique			Méthode MCT			RL. Architecture 1			RL. Architecture 2			
α	N	G	B	N	G	B	N	G	B	N	G	B	N	G	B
0,75	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	0,99	0,99	1,00	1,00	1,00	1,00
0,50	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	0,97	0,99	1,00	1,00	1,00	1,00
0,25	0,99	0,99	0,99	0,98	0,98	0,98	0,98	0,97	0,98	0,95	1,00	1,00	1,00	1,00	0,99
0,20	0,98	0,98	0,98	0,98	0,98	0,98	0,97	0,96	0,96	0,93	1,00	1,00	1,00	1,00	0,99
0,10	0,97	0,97	0,97	0,96	0,96	0,96	0,88	0,89	0,90	0,91	1,00	1,00	0,98	0,99	0,98
0,05	0,96	0,96	0,96	0,94	0,94	0,95	0,69	0,72	0,77	0,90	1,00	1,00	0,95	0,96	0,96
0,00	0,76	0,83	0,74	0,48	0,64	0,38	0,08	0,11	0,09	1,00	1,00	1,00	0,56	0,56	0,56

4.5.2.c Robustesse aux fumées des algorithmes de fusion multimodale de probabilités conditionnelles avec calcul au préalable du pourcentage de superposition des zones de feu par de la fumée

Il a été montré que l'exploitation d'une image du domaine du visible dont les zones de feux sont masquées par de la fumée fait diminuer les performances des algorithmes de détection utilisant des images multimodales. Aussi, une procédure a été développée afin d'utiliser des images du visible et du proche-infrarouge que si les zones de feu sont recouvertes de fumée avec un pourcentage de superposition inférieur à 70% ; dans le cas contraire, seules les images du proche-infrarouge sont utilisées. Le tableau 4.9 présente les ratios des algorithmes de fusion de probabilités conditionnelles. Le fait de basculer sur l'infrarouge uniquement permet de garder de bons résultats quel que soit le niveau de fumée. En effet, les algorithmes obtiennent de bons résultats même pour des fumées opaques.

TABLEAU 4.9 : Ratio des scores des algorithmes de détection multimodaux avec évaluation préalable du taux de superposition de fumée en fonction de la transparence de fumée.

α	Moyenne			Méthode Statistique			Méthode MCT		
	N	G	B	N	G	B	N	G	B
0,75	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00
0,50	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00
0,25	1,00	1,00	0,99	1,00	1,00	0,99	0,99	0,99	0,98
0,20	1,00	0,99	0,99	1,00	0,99	0,98	0,99	0,98	0,98
0,10	1,00	1,00	0,98	1,00	1,00	0,98	0,99	0,99	0,96
0,05	1,00	1,00	0,98	1,00	1,00	0,98	0,99	0,99	0,97
0,00	1,00	1,00	1,00	1,00	1,00	1,00	0,99	0,99	0,99

L'étude présentée dans cette section a permis de montrer que les méthodes de détection basées sur l'exploitation conjointes d'images du visible et du proche-infrarouge multimodales obtiennent de meilleurs résultats que les méthodes n'exploitant qu'un seul type d'image. La méthode par moyenne des probabilités conditionnelles qui est une méthode simple permettant d'obtenir de bons scores avec un temps d'exécution court est choisie pour réaliser l'étape de détection de pixels « feu » du système présentée dans cette thèse.

4.6 Conclusion

Dans ce chapitre, une nouvelle approche pour la détection de pixels « feu » consistant à fusionner les données extraites d'images du visible et du proche-infrarouge a été présentée. Une base de 100 images multimodales a été utilisée pour évaluer les performances de toutes les méthodes de détection étudiées dans cette thèse. Les résultats obtenus ont tout d'abord permis de confirmer que les méthodes de détection travaillant uniquement sur les images du proche-infrarouge sont plus efficaces (particulièrement en présence de fumée se superposant aux zones

de feu dans les images), plus simples à mettre en œuvre et plus rapides que les méthodes travaillant sur les images du visible. Lorsque les fumées ne masquent pas totalement la zone de feu dans les images, les résultats ont montré que les méthodes de détection multimodales développées permettent d'obtenir les meilleurs résultats. Enfin, la robustesse aux fumées des différents algorithmes de détection utilisés dans cette thèse a été testée et les détections multimodales proposées sont les plus robustes à tout type et toute densité de fumée. Ce chapitre permet d'établir l'intérêt de l'utilisation conjointe des images du spectre visible et du proche infrarouge pour la détection des pixels « feu ». Une méthode simple et rapide consistant à moyennner les probabilités conditionnelles d'appartenance au feu sachant la valeur des pixels des deux modalités a notamment montré de bonnes performances de détection sur la base d'image.

Chapitre 5

Optimisation par fusion d'images multimodales du nombre de points appariés

5.1	Introduction	143
5.2	État de l'art de la fusion d'images	144
5.2.1	Approches classiques	145
5.2.2	Approches multirésolutions	146
5.3	Appariement d'images multimodales	155
5.3.1	Détection et appariement de points d'intérêt	156
5.3.2	Fusion de données d'images pour l'appariement	161
5.3.3	Concaténation de correspondances	163
5.4	Evaluation de l'efficacité d'appariement de points d'intérêt en fonction du type d'image	163
5.5	Conclusion	172

5.1 Introduction

L'estimation de caractéristiques géométriques de feux par stéréovision nécessite plusieurs étapes clefs qui sont : la détection des pixels de feu, la détection de points d'intérêt dans les zones de pixels de feu, l'appariement des points d'intérêt entre les images stéréoscopiques, le calcul de points 3D par triangulation et enfin l'estimation des caractéristiques géométriques des zones de feu.

Ce chapitre porte sur la recherche dans les images stéréoscopiques des couples de pixels correspondants aux mêmes points de feu projetés. Pour obtenir une reconstruction tridimensionnelle précise du feu, il est nécessaire d'avoir de nombreuses correspondances avec une précision au pixel près, réparties sur toute la zone de feu.

L'appariement stéréoscopique est une problématique centrale dans le domaine de la vision par ordinateur depuis une vingtaine d'années et de nombreuses recherches ont été faites sur ce su-

jet. Les thèses de Gales [143] et de Galindo [144] présentent les différentes méthodes existantes sur le sujet. Les solutions proposées dépendent du contexte de l'utilisation de l'appariement, de la précision recherchée et du temps alloué pour l'exécution de la procédure. Dans le cadre de cette thèse, nous nous sommes intéressés à des méthodes d'appariement non dense (c'est à dire que tous les points des zones de feu détectées ne sont pas appariés) en nous focalisant davantage sur leurs performances que sur leur temps d'exécution. Nous avons toutefois considéré le temps d'exécution comme un facteur important dans le développement de nouvelles méthodes afin qu'elles puissent être adaptées et exécutées en temps réel. L'acquisition simultanée d'images stéréoscopiques du domaine du visible et du proche-infrarouge donne la possibilité de développer de nouveaux moyens pour obtenir cette double condition. En effet, les correspondances peuvent être cherchées dans les images du spectre du visible et/ou du proche-infrarouge ainsi que sur des images fusionnées à l'aide des images des deux spectres.

Ce chapitre présente dans la section 5.2 les techniques de fusion de valeurs de pixels de la littérature. Ces méthodes sont utilisées pour obtenir une image avec plus de détails que les images seules.

La section 5.3 décrit les algorithmes de recherche d'appariements utilisés dans le cadre de cette thèse.

Enfin, la section 5.4 est consacrée à l'étude d'efficacité de ces méthodes sur des images de feu. Cette partie permettra de mettre en évidence les meilleures méthodes pour l'appariement des points du feu avec le système utilisé.

5.2 État de l'art de la fusion d'images

Dans le chapitre précédent, la fusion multimodale de résultats a été utilisée pour améliorer la détection de pixels de feu. Lors de la présentation des techniques de fusion de résultats, la fusion de valeurs de pixels d'image (plus communément appelée fusion d'images) a été énoncée. Les techniques de fusion d'images permettent d'obtenir une nouvelle image à partir de plusieurs images de la même scène. Le but est d'obtenir une image combinant les informations importantes contenues dans chaque image. Les techniques de fusion d'images, dépendent de l'objectif recherché. Pohl *et al.* [145] ont recensé plusieurs objectifs courants dans la fusion d'images, à savoir :

- Substituer des données ou caractéristiques manquantes sur une image avec des données extraites sur l'autre image
- Compléter l'ensemble des données pour améliorer leur classement
- Appliquer des corrections géométriques améliorées
- Augmenter la résolution spatiale d'une image

Il existe plusieurs façons de fusionner les valeurs de pixels d'image. Les algorithmes choisis pour réaliser cette action dépendent du type d'application et du type d'images [146]. Dans le cadre de cette thèse, l'objectif est de fusionner des images de même dimension acquises dans les spectres du visible et du proche-infrarouge pour obtenir une image avec plus de détails. Les méthodes de fusion d'images de même dimension sont très bien exposées dans le livre de Mitchell [147], dans la thèse d'Akhloufi [148] ainsi que dans les articles de Zeng *et al.* [146] et Sadjadi *et al.* [149]. Ces techniques sont décrites dans cette section. Nous considérerons que les images multimodales ont été recalées de la même façon que dans le chapitre précédent.

5.2.1 Approches classiques

5.2.1.a Superposition linéaire par pondération

La plus simple des méthodes de fusion d'images consiste à attribuer une pondération à chaque image. La superposition pondérée entre deux images peut s'exprimer ainsi [145] :

$$I_f = A(\alpha I_1 + (1 - \alpha)I_2) + B \quad (5.1)$$

Où I_f est l'image fusionnée, I_1 et I_2 sont les images à fusionner, α est un paramètre de pondération et A et B sont des facteurs d'échelle. La superposition linéaire par le calcul de moyenne [148] se fait avec $\alpha = 0,5$, $A = 1$ et $B = 0$.

5.2.1.b Analyse en Composante Principale

L'Analyse en Composante Principale (ACP) est une technique linéaire de réduction de dimensionnalité qui permet de diminuer l'espace dans lequel sont représentées les données. La fusion d'images utilisant l'ACP consiste à réaliser une moyenne pondérée des images avec des coefficients de pondération calculés par ACP [150]. Ces coefficients sont obtenus à partir des composantes principales ayant les plus grandes valeurs propres. Pour calculer ces composantes principales, il faut dans un premier temps générer la matrice de corrélation des images à fusionner. Ensuite, les vecteurs propres de cette matrice sont estimés ainsi que leurs valeurs propres. Les composantes principales sont les vecteurs propres qui ont les plus grandes valeurs propres.

Soit une matrice X composée de n colonnes qui contiennent les intensités des pixels des n images auxquelles a été soustrait le niveau moyen d'intensité μ calculé en considérant toutes les images et V_m le vecteur propre maximal de la matrice de corrélation $X^T X$ alors l'image fusionnée I_f est calculée en utilisant la formule :

$$I_f = X|V_m||V_m^T| - \mu \quad (5.2)$$

L'équation (5.2) permet d'obtenir un vecteur colonne correspondant à l'image fusionnée. Ce vecteur est ensuite mis en forme pour avoir les mêmes dimensions des images qui ont été utilisées pour la fusion.

5.2.2 Approches multirésolutions

Les approches multirésolutions pour la fusion d'images sont principalement utilisées pour faire ressortir les détails des images. Pour une meilleure compréhension, toutes les explications des algorithmes de fusion multirésolutions seront données pour deux images car c'est le nombre d'images qui est utilisé dans le cadre de cette thèse. Toutefois, les algorithmes présentés peuvent être utilisés sur plus de deux images.

La première étape consiste à calculer les décompositions multirésolutions pour les deux images [149]. Cette opération décompose chaque image en détails et approximation. Les détails d'une image correspondent aux endroits importants d'une image comme les contours d'un objet. Les hautes valeurs (positives ou négatives) d'une image de détails sont obtenues par les pixels dont l'intensité est très différente de celle de leurs voisins. L'approximation d'une image est une image de taille deux fois inférieure à l'image traitée et dont l'intensité des pixels est obtenue à partir de celle des pixels des images de départ. Cette image est construite en parallèle à l'image de détails de sorte que l'on puisse retrouver l'image de départ par une opération inverse qui dépend de la décomposition choisie. En calculant les détails et l'approximation de l'image d'approximation obtenue, un nouveau niveau de décomposition est généré. Cette opération est répétée n fois, permettant ainsi d'obtenir n images de détails et une image d'approximation pour chacune des images; n étant le niveau de décomposition choisi par l'utilisateur. Plusieurs méthodes de décomposition multirésolution sont présentées dans cette section. Elles permettent de générer différents types de détails et d'approximations pour une image. La même méthode avec le même niveau de décomposition doit être appliquée sur les deux images pour pouvoir les fusionner.

La deuxième étape consiste à fusionner les images de détails et d'approximation des deux images considérées. Ici encore, plusieurs méthodes existent pour combiner ces données. Enfin, la décomposition inverse est appliquée aux images de détails et approximation combinées pour obtenir l'image fusionnée. La figure 5.1 présente le schéma des opérations de la fusion multirésolution de deux images.

Plusieurs méthodes existent pour fusionner les images de détails et d'approximations.

La première méthode consiste à choisir pour chaque pixel à fusionner, celui des deux images qui a la plus grande intensité en valeur absolue (les détails pouvant être négatifs).

Une autre règle de combinaison consiste à appliquer une moyenne pondérée des images à combiner en choisissant un poids $\omega(\mathbf{x})$ adapté à chaque pixel \mathbf{x} . L'image d'approximation ou de détails fusionnée $\widehat{\mathbf{I}}_f$ est calculée ainsi :

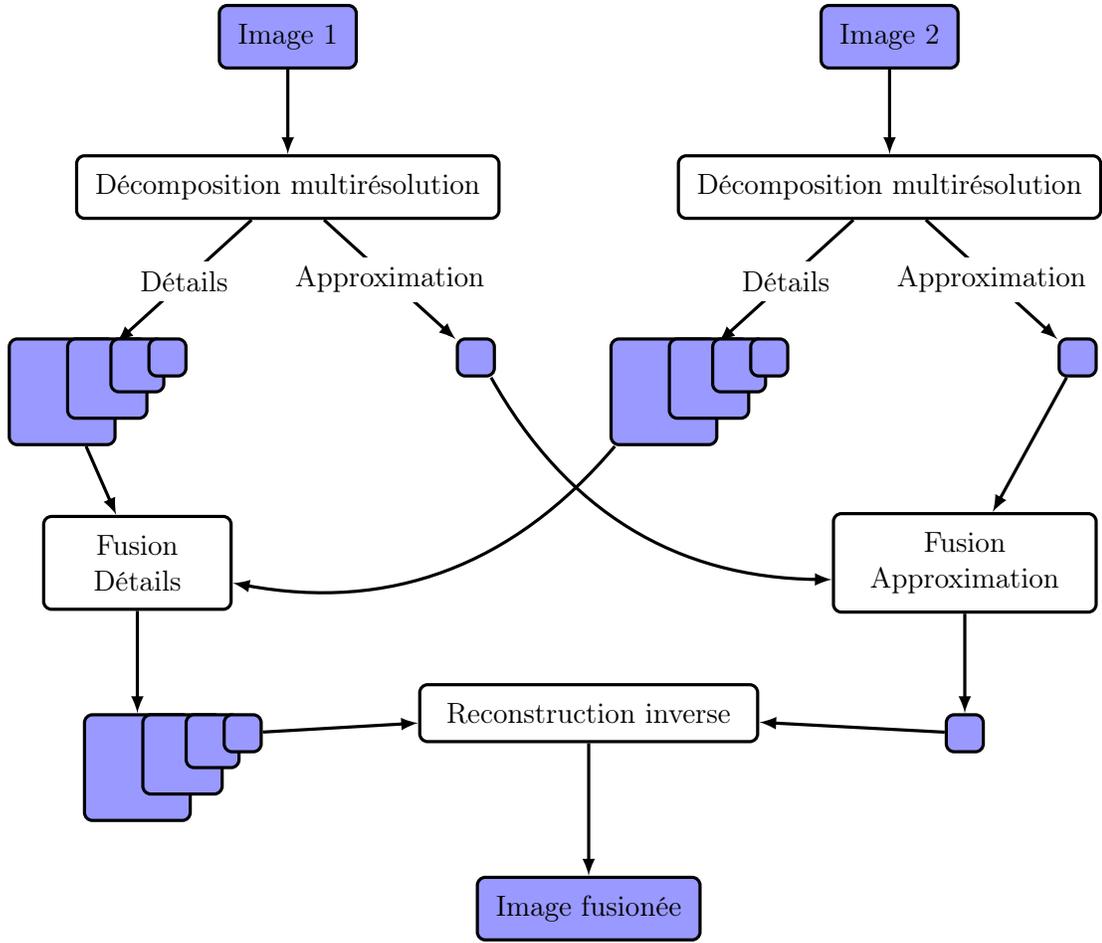


FIGURE 5.1 : Schéma des opérations réalisées pour la fusion multirésolution de deux images.

$$\widehat{\mathbf{I}}_f(\mathbf{x}) = \omega(\mathbf{x})\widehat{\mathbf{I}}_1(\mathbf{x}) + (1 - \omega(\mathbf{x}))\widehat{\mathbf{I}}_2(\mathbf{x}) \quad (5.3)$$

Avec $\widehat{\mathbf{I}}_1$ et $\widehat{\mathbf{I}}_2$ les décompositions multirésolutions des images à fusionner.

Les formules pour calculer les poids optimums $\omega(\mathbf{x})$ sont décrites par Burt *et al.* dans [151]. Soit N_v la taille de fenêtre de voisinage pour un pixel \mathbf{x} (généralement égale à 3 x 3 ou 5 x 5 pixels) et soit $A_{\mathbf{I}}(\mathbf{x})$ la moyenne pondérée du voisinage du pixel \mathbf{x} de coordonnées (m,n) dans l'image \mathbf{I} définie ainsi :

$$A_{\mathbf{I}}(\mathbf{x}) = \sum_{i=1}^{N_v} \sum_{j=1}^{N_v} s(i,j)\widehat{\mathbf{I}}(m+i, n+j) \quad (5.4)$$

Avec s une fonction de poids telle que :

$$\sum_{i=1}^{N_v} \sum_{j=1}^{N_v} s(i, j) = 1 \quad (5.5)$$

En considérant $M_{I_1 I_2}(\mathbf{x})$ la corrélation normalisée moyennée sur le voisinage de \mathbf{x} des images à fusionner :

$$M_{I_1 I_2}(\mathbf{x}) = \frac{2 \sum_{i=1}^{N_v} \sum_{j=1}^{N_v} s(i, j) \widehat{I}_1(m+i, n+j) \widehat{I}_2(m+i, n+j)}{A_{I_1}(\mathbf{x})^2 + A_{I_2}(\mathbf{x})^2} \quad (5.6)$$

Le poids $\omega(\mathbf{x})$ est défini de la façon suivante :

$$\omega(\mathbf{x}) = \begin{cases} \frac{1}{2} \left(\frac{M_{I_1 I_2}(\mathbf{x}) - \alpha}{1 - \alpha} \right) & \text{si } M_{I_1 I_2}(\mathbf{x}) \geq \alpha \\ 0 & \text{sinon} \end{cases} \quad (5.7)$$

Avec α un seuil fixé par l'utilisateur.

Dans certains cas, comme pour les images de modalités différentes, l'image obtenue par la fusion des images approximations peut contenir moins d'informations que l'une des deux images d'approximation. De ce fait, une troisième méthode est de ne pas combiner les images d'approximation et de choisir l'image d'approximation d'une seule des deux images.

Plusieurs processus de décomposition multirésolution peuvent être utilisés pour la fusion d'image. Nous avons choisi d'utiliser quatre méthodes parmi les plus les utilisées compte tenu de leurs bonnes performances [148] :

- Pyramide laplacienne
- Pyramide de ratios
- Pyramide de contraste
- Transformée en ondelettes discrète

5.2.2.a Pyramide laplacienne

Le processus de fusion par décomposition en pyramide laplacienne est décrit par Burt *et al.* [152]. La figure 5.2 présente le schéma explicatif de cet algorithme sur trois niveaux de décomposition pour l'image Lenna.

La première étape de la décomposition par pyramide laplacienne d'une image est la création d'une pyramide d'images d'approximation de cette image. L'image originale est sous-échantillonnée successivement afin d'obtenir une série d'images de taille différente. Après

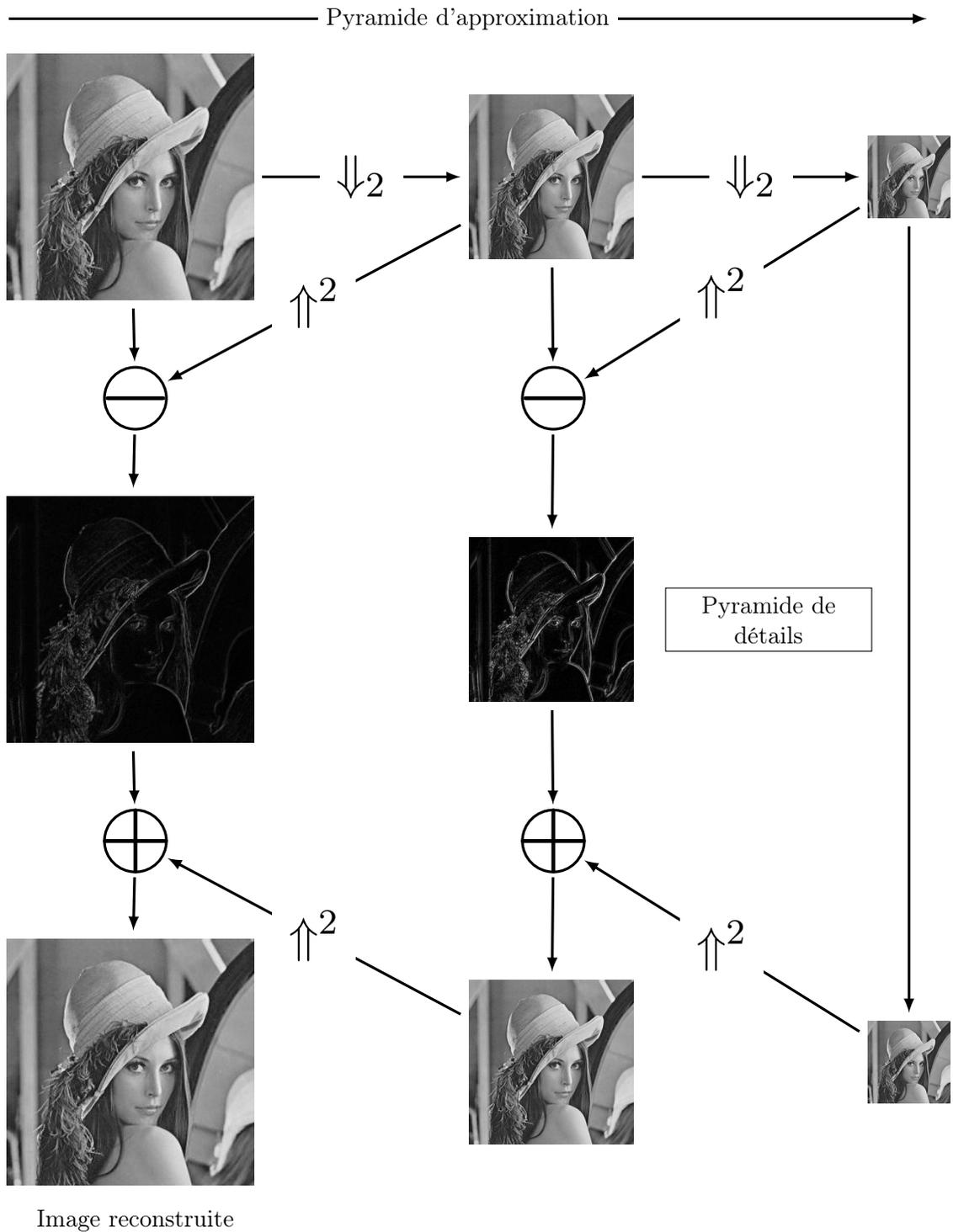


FIGURE 5.2 : Représentation de l'algorithme de la pyramide laplacienne sur 3 niveaux de décomposition pour l'image de Lenna. Les opérateurs \Downarrow_2 et \Uparrow^2 symbolisent respectivement les opérations de réduction et d'expansion d'image par un facteur 2 et les opérateurs \ominus et \oplus symbolisent respectivement les opérations de soustraction et d'addition d'images.

chaque étape de sous-échantillonnage, la taille de l'image est divisée par deux. Cette pyramide d'images d'approximation est aussi appelée pyramide Gaussienne car l'opération de réduction de l'image comprend un filtrage Gaussien de l'image suivi d'une décimation de facteur deux. Le filtrage Gaussien est fait par une convolution entre l'image et un noyau w défini dans [152] par :

$$w = \left[\frac{1}{16}, \frac{1}{4}, \frac{3}{8}, \frac{1}{4}, \frac{1}{16} \right] \quad (5.8)$$

Une pyramide d'images de détails est également calculée. Dans un premier temps, chaque image réduite est agrandie d'un facteur deux par interpolation bilinéaire. Les images obtenues ainsi sont soustraites aux images du niveau supérieur (qui ont donc la même taille). La soustraction de deux images se fait en effectuant une soustraction des intensités pour chaque pixel des images. La figure 5.3 présente les images de détail laplacien pour l'image de Lenna sur 4 niveaux de résolution.

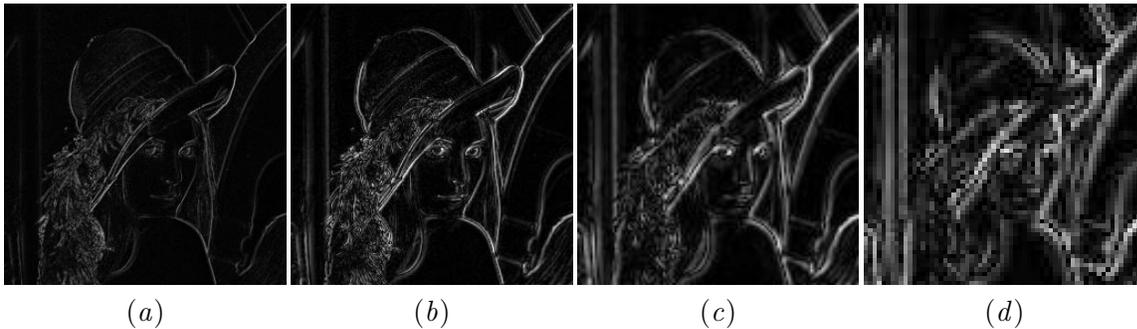


FIGURE 5.3 : Exemple de détails laplaciens sur 4 niveaux pour l'image de Lenna ; (a) détails du niveau 0, (b) du niveau 1, (c) du niveau 2 et (d) du niveau 3.

À partir de la pyramide de détails et de la plus petite image de la pyramide Gaussienne, il est possible de retrouver exactement l'image originale en appliquant l'opérateur inverse (ici l'addition). La plus petite image d'approximation est agrandie d'un facteur deux par interpolation bilinéaire puis additionnée à la plus petite image de détails pour obtenir l'image d'approximation du niveau supérieur. Cette image est ensuite utilisée avec l'image des détails du niveau supérieur pour reconstruire l'image d'approximation suivante et ainsi de suite jusqu'à retrouver l'image originale.

Pour fusionner deux images avec la décomposition en pyramide laplacienne, les pyramides d'images de détails ainsi que les images de la plus petite échelle des pyramides Gaussiennes sont combinées par une des règles de combinaison vues précédemment.

5.2.2.b Pyramide de ratios

Le principe de la fusion pyramidale de ratio est identique à celui de la fusion pyramidale laplacienne. La seule différence réside dans l'opérateur utilisé pour le calcul de la pyramide de détails qui est une division (et non pas une soustraction) et dont l'opérateur inverse associé est une multiplication. La figure 5.4 présente les images de détails de ratios pour l'image de Lenna sur 4 niveaux de résolution.



FIGURE 5.4 : Exemple de détails de ratios sur 4 niveaux pour l'image de Lenna ; (a) détails du niveau 0, (b) du niveau 1, (c) du niveau 2 et (d) du niveau 3.

5.2.2.c Pyramide de contraste

Le principe de la fusion pyramidale de contraste est également identique à celui de la fusion pyramidale laplacienne. Il diffère uniquement au niveau de l'opérateur de soustraction qui est remplacé par un calcul de contraste symbolisé par \odot et défini ainsi :

$$a \odot b = \frac{a - b}{b} \quad (5.9)$$

La figure 5.5 présente les images de détails de contraste pour l'image de Lenna sur 4 niveaux de résolution.

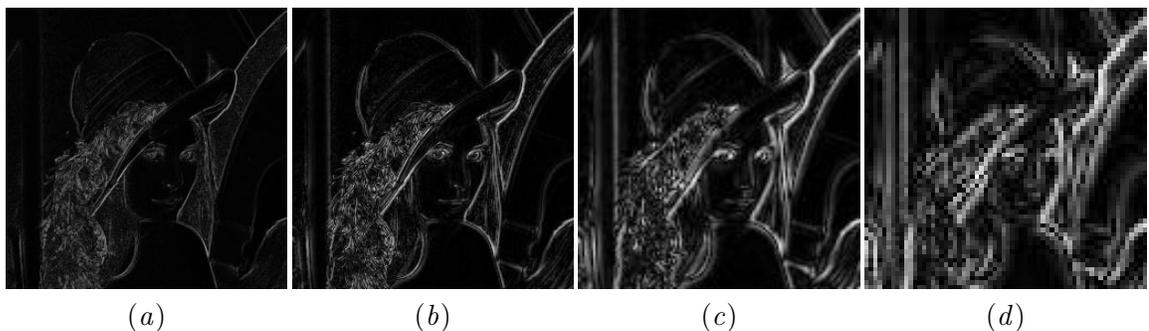


FIGURE 5.5 : Exemple de détails de contraste sur 4 niveaux pour l'image de Lenna ; (a) détails du niveau 0, (b) du niveau 1, (c) du niveau 2 et (d) du niveau 3.

5.2.2.d Transformée en ondelettes discrète

La décomposition en ondelettes discrète [153] permet de projeter une image dans un espace d'approximation (App) et dans un espace de détails qui est composé des détails horizontaux (DH), verticaux (DV) et diagonaux (DD). La décomposition en ondelettes discrète se fait par des convolutions de l'image par des filtres définis par les ondelettes. Rappelons qu'une ondelette est une fonction ψ qui vérifie les conditions suivantes :

$$\int_{-\infty}^{+\infty} |\psi(\omega)| d\omega < +\infty \qquad \int_{-\infty}^{+\infty} |\psi(\omega)|^2 d\omega < +\infty \qquad (5.10)$$

Pour chaque type d'ondelettes des filtres de décomposition passe-bas (B_D) et passe-haut (H_D) peuvent être calculés. Pour décomposer une image en ondelettes, les lignes de l'image originale sont convoluées avec les filtres de décomposition. Ensuite, une opération de réduction de lignes est appliquée aux images (\downarrow_2^l). Les colonnes des deux images obtenues sont ensuite convoluées avec les filtres B_D et H_D , puis une réduction sur les colonnes (\downarrow_2^c) est appliquée pour obtenir quatre images du niveau inférieur. Pour appliquer la décomposition sur un autre niveau, le même processus est appliqué sur l'image projetée dans l'espace des approximations, c'est à dire l'image qui a été convoluée deux fois avec le filtre passe-bas. Le schéma de la figure 5.6 représente les étapes de la décomposition en ondelettes telle qu'elle vient d'être décrite. Les exposants des images d'approximation et de détails correspondent au niveau de décomposition.

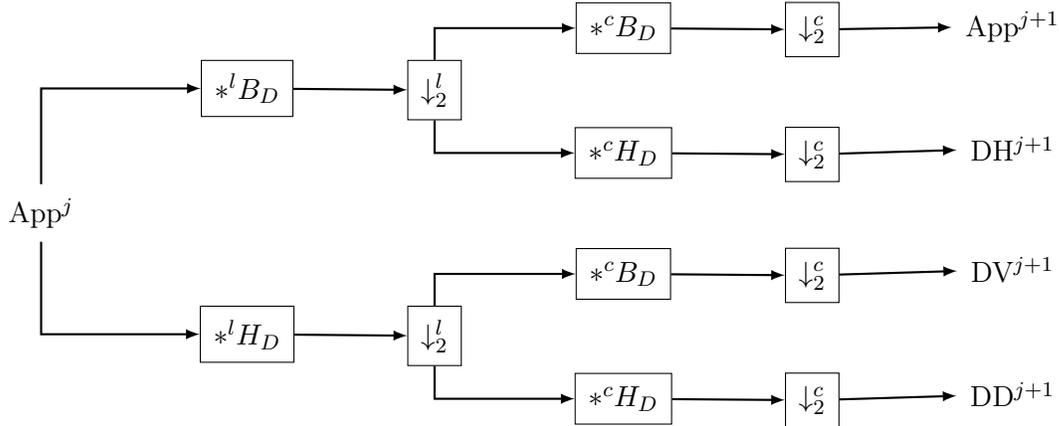


FIGURE 5.6 : Schéma de la décomposition en ondelettes discrète sur deux dimensions.

Il existe de nombreuses ondelettes différentes, nous avons choisi d'en décrire quatre parmi les plus usuelles : Haar, Debauchies, Coiflet et Meyer.

L'ondelette de Haar discrète, (ou de Debauchies 1) est la plus simple des ondelettes. Elle est représentée par la fonction suivante :

$$\psi(\omega) = \begin{cases} 1 & \text{si } 0 \leq \omega < \frac{1}{2} \\ -1 & \text{si } \frac{1}{2} \leq \omega < 1 \\ 0 & \text{sinon} \end{cases} \quad (5.11)$$

Les filtres passe-bas et passe-haut de l'ondelette de Haar utilisés pour la décomposition sont les suivants :

$$B_D^{Haar} = \left[\frac{1}{\sqrt{2}} ; \frac{1}{\sqrt{2}} \right] \quad H_D^{Haar} = \left[-\frac{1}{\sqrt{2}} ; \frac{1}{\sqrt{2}} \right] \quad (5.12)$$

La figure 5.7 présente un exemple de décomposition de Haar de l'image de Lenna par ondelettes avec un schéma explicatif.

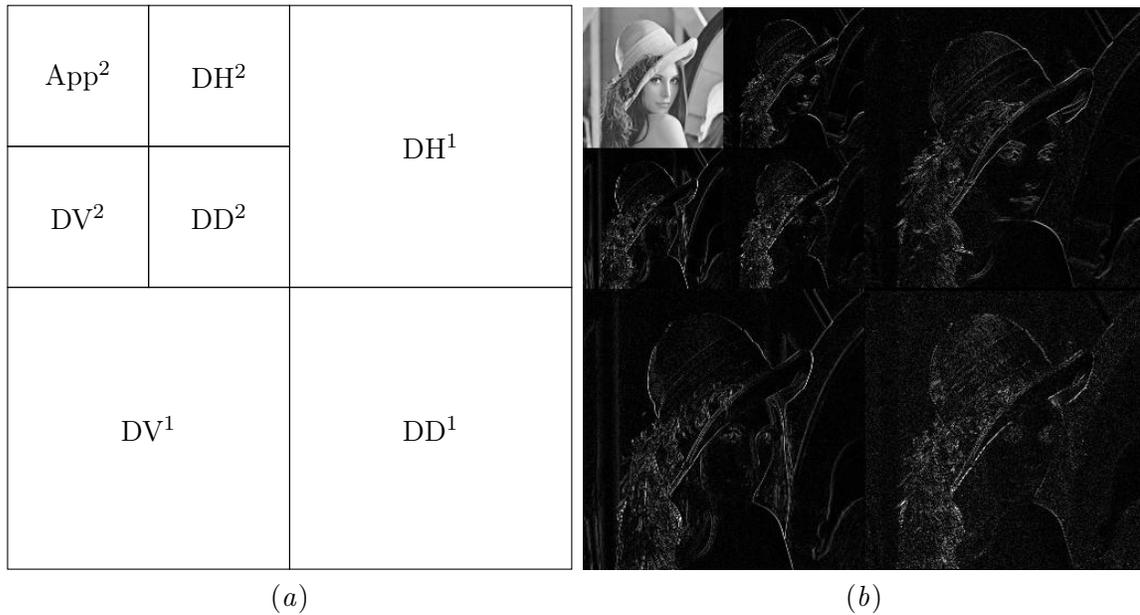


FIGURE 5.7 : Exemple d'une image décomposée par ondelettes. (a) Schéma d'une décomposition d'image par ondelettes de niveau 2 avec les numéros en indice correspondant au niveau de décomposition ; App, l'espace des approximations ; DH, DV et DD respectivement les espaces de détails horizontaux, verticaux et diagonaux (b) exemple de décomposition en ondelettes sur l'image de Lenna.

La transformée en ondelettes discrète par les ondelettes de Debauchies est l'une des plus utilisées en fusion [148]. Elle a la particularité de générer des détails parcimonieux (grand nombre de moments nuls) lors de la décomposition d'images. Parmi les ondelettes de Debauchies celle d'indice deux est une des plus utilisées. Les filtres passe-bas et passe-haut de cette dernière employés pour la décomposition sont les suivants :

$$\begin{aligned} B_D^{Db2} &= [-0,1294 ; 0,2241 ; 0,8365 ; 0,4830] \\ H_D^{Db2} &= [-0,4830 ; 0,8365 ; -0,2241 ; -0,1294] \end{aligned} \quad (5.13)$$

L'ondelette de Coiflet d'indice 1 permet d'avoir les filtres passe-bas et passe-haut suivants :

$$\begin{aligned} B_D^{Coif1} &= [-0,0157 ; -0,0727 ; 0,3849 ; 0,8526 ; 0,3379 ; -0,0727] \\ H_D^{Coif1} &= [0,0727 ; 0,3379 ; -0,8526 ; 0,3849 ; 0,0727 ; -0,0157] \end{aligned} \quad (5.14)$$

L'ondelette de Meyer est définie ainsi :

$$\psi(\omega) = \begin{cases} \frac{1}{\sqrt{2\pi}} \sin\left(\frac{\pi}{2}\nu\left(\frac{3|\omega|}{2\pi} - 1\right)\right) e^{\frac{j\omega}{2}} & \text{si } \frac{2\pi}{3} \leq |\omega| < \frac{4\pi}{3} \\ \frac{1}{\sqrt{2\pi}} \cos\left(\frac{\pi}{2}\nu\left(\frac{3|\omega|}{4\pi} - 1\right)\right) e^{\frac{j\omega}{2}} & \text{si } \frac{4\pi}{3} \leq |\omega| < \frac{8\pi}{3} \\ 0 & \text{sinon} \end{cases} \quad (5.15)$$

Une version discrète de cette ondelette est utilisée afin de pouvoir l'appliquer à la fusion d'image. La toolbox *wavelet* de **Matlab** permet de générer les filtres de décomposition discrets de cette ondelette.

La figure 5.8 présente les images de détails horizontaux du premier niveau de décomposition l'image de Lenna pour les quatre ondelettes présentées.

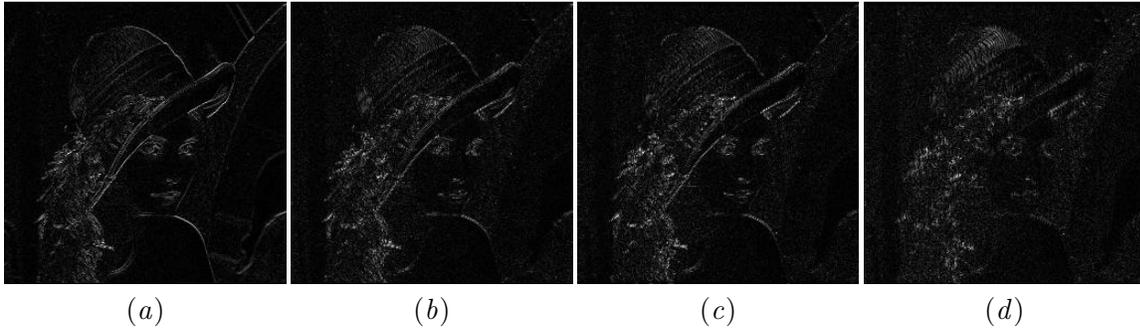


FIGURE 5.8 : Exemple de détails horizontaux du premier niveau de décomposition en ondelettes de l'image de Lenna; (a) ondelette de Haar, (b) ondelette de Debauchies 2, (c) ondelette de Coiflet et (d) ondelette de Meyer.

Les espaces d'approximation et de détails sont fusionnés en utilisant les règles de combinaison vues précédemment. Pour obtenir l'image fusionnée finale, une transformée en ondelettes inverse est appliquée aux résultats des fusions d'approximation et de détails. La reconstruction inverse suit le schéma de la figure 5.9 avec \uparrow_2^l et \uparrow_2^c correspondant respectivement aux expansions de facteur deux des lignes et des colonnes et B_R et H_R respectivement les filtres de reconstruction passe-bas et passe-haut. Pour chaque type d'ondelettes, ces filtres sont calculés ainsi :

$$B_R(\omega) = B_D(t_f - \omega) \qquad H_R(\omega) = H_D(t_f - \omega) \qquad (5.16)$$

Avec t_f la taille du filtre considéré.

Par exemple, pour l'ondelette de Haar, $t_f = 2$ et les filtres de reconstruction sont les suivants :

$$B_R^{Haar} = \left[\frac{1}{\sqrt{2}} ; \frac{1}{\sqrt{2}} \right] \qquad H_R^{Haar} = \left[\frac{1}{\sqrt{2}} ; -\frac{1}{\sqrt{2}} \right] \qquad (5.17)$$

Enfin, l'image est rognée car le filtrage augmente la taille de l'image en ajoutant des colonnes et des lignes. La coupure se fait donc de sorte que l'image soit de la taille désirée pour le niveau en retirant les pixels des bordures extérieures de l'image.

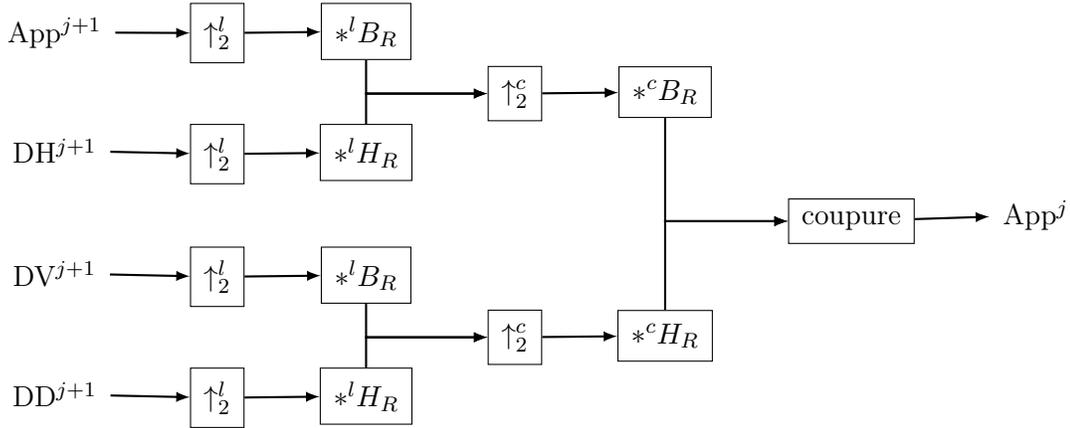


FIGURE 5.9 : Schéma de la transformation inverse en ondelettes discrètes sur deux dimensions.

5.3 Appariement d'images multimodales

Cette section est dédiée à l'utilisation de la fusion d'images multimodales dans le but d'optimiser le nombre de points appariés entre des images stéréoscopiques ; c'est à dire que l'on cherche à obtenir un maximum de points de correspondance tout en garantissant un faible nombre de fausses correspondances.

La première partie de cette section décrit l'algorithme de recherche de correspondances utilisé dans cette thèse et qui peut s'appliquer aussi bien aux images du spectre visible, du proche-infrarouge ainsi qu'aux images fusionnées. Une deuxième partie est consacrée à la fusion d'images du visible et du proche-infrarouge permettant d'obtenir un grand nombre de points appariés. Enfin une dernière partie présente une méthode de concaténation des points de correspondance trouvés afin d'augmenter leur nombre.

Afin de pouvoir estimer les caractéristiques géométriques d'un feu, seuls les points de correspondance appartenant aux zones de flammes doivent être recherchés dans les images. Nous avons vu dans le chapitre précédent que la détection multimodale par moyenne des probabilités donnait les meilleurs résultats de détection de zones de feu. Cette détection est donc appliquée afin de ne considérer que les pixels de feu des images du spectre visible et proche infrarouge lors de la recherche des points de correspondance. Les images du spectre visible sont converties en niveau de gris afin de rendre les opérations de fusion et de recherche de correspondances plus simples et rapides.

5.3.1 Détection et appariement de points d'intérêt

Le calcul des caractéristiques géométriques d'un feu s'effectue à partir de ses points 3D estimés par stéréovision. L'évaluation de chaque point tridimensionnel nécessite d'identifier dans les deux images stéréoscopiques (quel que soit la modalité de l'image) le couple de points gauche-droite ($\mathbf{x}^g, \mathbf{x}^d$) appariés provenant de la projection d'un même point tridimensionnel.

Lors de l'appariement, le correspondant le plus probable de chaque pixel de l'image gauche est cherché dans l'image droite. Il est possible de sélectionner ces points manuellement mais cette approche est longue et peu précise. Une recherche automatique de correspondance est donc préférable pour calculer les points tridimensionnels d'un objet. Ce processus peut se découper en plusieurs étapes :

- Détection de points d'intérêt
- Extraction du descripteur des points d'intérêt
- Appariement des points d'intérêt par comparaison de leur descripteur

5.3.1.a Détection de points d'intérêt

Un point d'intérêt dans une image correspond à une double discontinuité de la fonction d'intensité ; Ce peut être par exemple des jonctions de lignes correspondant à des textures ou à des coins d'objet. Il existe de nombreuses méthodes de détection de points d'intérêt dans une image dont les plus connues sont : l'algorithme FAST (*Feature From Accelerate Segment Test*) qui s'appuie sur la comparaison entre le niveau de gris du point courant et celui de certains points avoisinants le point traité [154], la détection de caractéristiques par la méthode de Harris [155] qui calcule la matrice de second ordre obtenue à partir des pixels voisins du pixel considéré, ou la détection par la méthode de la valeur propre minimale [156] qui calcule le minimum des valeurs propres de la matrice de Harris.

Parmi toutes les méthodes que nous avons testées sur des images de feu, la méthode de Harris, qui est simple, rapide et qui fait partie des meilleures méthodes de détection [143], a donné les meilleurs résultats et nous l'avons retenue. Cette méthode consiste à calculer la matrice des moments du second ordre, appelée matrice de Harris et permet de faire apparaître les coins

par autocorrélation indépendamment de la direction. Cette matrice, notée H , se définit pour chaque pixel \mathbf{x} de l'image \mathbf{I} ainsi :

$$H_{\mathbf{x}} = \sigma_d^2 \begin{bmatrix} \mathbf{I}_a(\mathbf{x}) & \mathbf{I}_c(\mathbf{x}) \\ \mathbf{I}_c(\mathbf{x}) & \mathbf{I}_b(\mathbf{x}) \end{bmatrix} \quad (5.18)$$

Avec les images $\mathbf{I}_a, \mathbf{I}_b$ et \mathbf{I}_c calculées de la façon suivante :

$$I_a = g_{\sigma_d} * \left(\frac{\partial g_{\sigma_d} * \mathbf{I}}{\partial l} \right)^2 ; I_b = g_{\sigma_i} * \left(\frac{\partial g_{\sigma_d} * \mathbf{I}}{\partial c} \right)^2 ; I_c = g_{\sigma_i} * (I_a I_b) \quad (5.19)$$

Avec g_{σ_d} et g_{σ_i} des gaussiennes d'écart types σ_d et σ_i , $\frac{\partial}{\partial l}$ la dérivée d'une image sur les lignes et $\frac{\partial}{\partial c}$ la dérivée d'une image sur les colonnes. Comme dans [157], nous avons fixé $\sigma_d = 1$ et $\sigma_i = 1,5$.

La fonction de force des points d'intérêt f_H est ensuite définie comme le rapport du déterminant de la matrice de Harris sur sa trace [158] :

$$f_{H_l}(\mathbf{x}) = \frac{\det H_l(\mathbf{x})}{\text{tr } H_l(\mathbf{x})} = \frac{\lambda_1 \lambda_2}{\lambda_1 + \lambda_2} \quad (5.20)$$

Avec λ_1 et λ_2 les valeurs propres de la matrice H .

Les points d'intérêt \mathbf{x} d'une image ont une valeur $f_{H_l(\mathbf{x})}$ supérieure à celle des autres pixels de son environnement. Pour chaque pixel de l'image, un bloc de voisinage de taille 3×3 pixels est considéré et si le pixel correspond au maximum local de ce voisinage alors il est considéré comme point d'intérêt.

La figure 5.10 présente les points détectés par la méthode de Harris sur une image de feu acquise dans le domaine du visible et convertie en niveau de gris. Dans cet exemple, 1073 points sont détectés.

5.3.1.b Descripteurs des points d'intérêt

Cette étape consiste à décrire les pixels entourant les points d'intérêt. Généralement les descripteurs sont invariants aux transformations géométriques et lumineuses. De nombreuses techniques d'extraction de descripteurs ont été proposées dans la littérature : méthode par invariance par échelle (*Scale-invariant feature transform* : SIFT) [159], méthode d'histogramme des gradients d'orientations et d'emplacement (*Gradient location-orientation histogram* : GLOH) [160], méthode de descripteurs basés sur les filtres [161], méthode de hachage par analyse discriminante linéaire (*Linear Discriminant Analysis Hashing* : LDAHash) [162] ou encore les

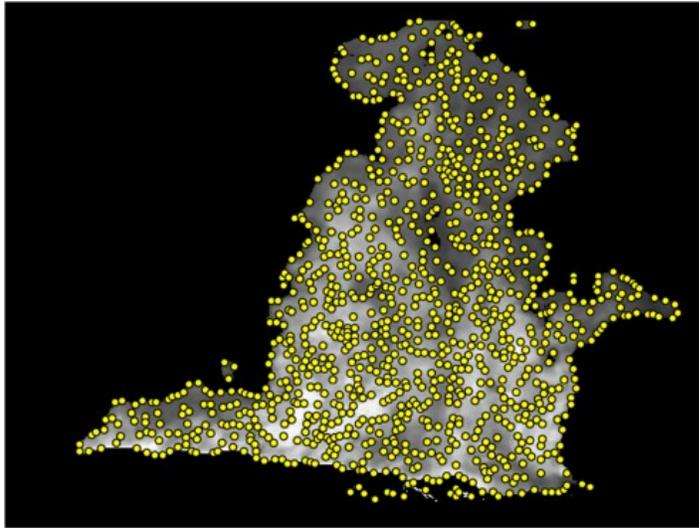


FIGURE 5.10 : Détection de points d'intérêts Harris sur une image de feu.

descripteurs basés sur la couleur [163]. Dans cette thèse, nous utiliserons une forme simple de descripteurs adaptés de ceux proposés par Brown *et al.* [157].

Les descripteurs de Brown *et al.* sont sous la forme de matrices 8×8 obtenues à partir de l'intensité des pixels situés autour de chaque point d'intérêt. Chaque matrice est obtenue en considérant dans un premier temps une zone de 41×41 pixels située autour de chaque pixel d'intérêt puis en l'échantillonnant afin de travailler sur un groupement de 8×8 pixels. Cet échantillonnage revient à laisser quatre pixels entre chaque pixel sélectionné. Pour éviter des effets de repliement de spectre, les pixels sont échantillonnés en faisant une interpolation, c'est à dire en faisant une moyenne pondérée des pixels voisins. Afin d'être robuste aux changements de luminosité entre les images contenant les points à appairer, le bloc B de 8×8 pixels est normalisé de la manière suivante :

$$\overline{\mathbf{x}^j} = \frac{\mathbf{x}^j - \mu_B}{\sigma_B} \quad (5.21)$$

Avec \mathbf{x}^j un pixel appartenant au bloc de pixels B , μ_B et σ_B la moyenne et l'écart-type de l'intensité des pixels du bloc B .

La figure 5.11 montre un exemple de ce descripteur pour un point d'intérêt de l'image de feu présenté dans la figure 5.10.

Il est possible de rendre les descripteurs invariants par rotation et robustes aux changements d'échelle en orientant le bloc de pixels selon l'orientation du gradient du point d'intérêt et en calculant les descripteurs de 8×8 sur différentes tailles d'images.

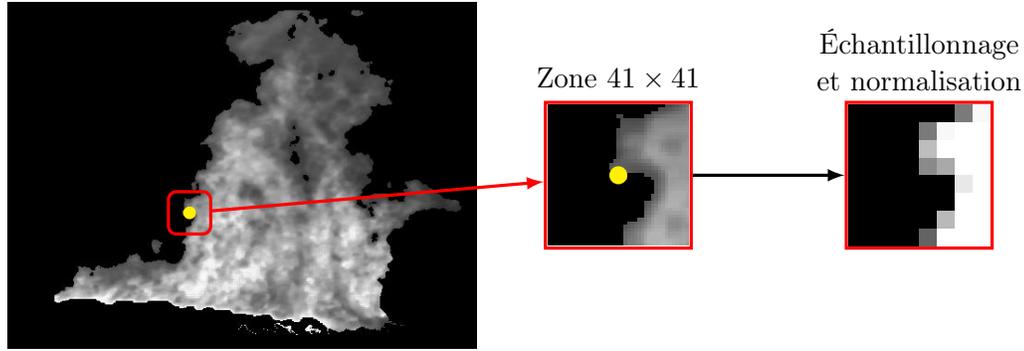


FIGURE 5.11 : Exemple de descripteur pour un point d'intérêt particulier.

Dans le contexte de cette thèse, les images stéréoscopiques qui contiennent les points à appairer sont de même échelle et ont approximativement la même orientation. Aussi les dernières opérations sur les descripteurs décrites dans ce paragraphe n'ont pas été appliquées afin de diminuer le temps d'exécution du programme d'appariement.

5.3.1.c Appariement des points d'intérêt par comparaison de leur descripteur

Les étapes précédentes permettent d'obtenir dans chaque image stéréoscopique des points d'intérêt avec leur descripteur associé. L'appariement des points d'intérêt nécessite une mesure de similarité de leur descripteur. Un grand nombre de méthodes de mesure de similarité ont été répertoriées par Chambon et Couzil [164]. Parmi les méthodes ayant un bon ratio temps d'exécution-nombre de bons appariements, nous avons choisi la méthode dite de « Somme des distances au carré normalisées à moyenne nulle » (*Zero mean Normalised Sum of Squared Differences*, ZNSSD). Le coefficient de corrélation qui mesure la similitude entre deux descripteurs est défini par :

$$\text{ZNSSD}(D_1, D_2) = \frac{\| (D_1 - \overline{D_1}) - (D_2 - \overline{D_2}) \|_2}{\sqrt{\|D_1\|_2 \|D_2\|_2}} \quad (5.22)$$

Où D_1 et D_2 sont les descripteurs à comparer, $\|\cdot\|_2$ est la norme euclidienne et $\overline{D_i}$ est la moyenne des valeurs du descripteur D_i .

Plus ce coefficient est petit, plus grande est la similitude.

Pour chaque point d'intérêt de l'image prise comme référence, la ligne épipolaire correspondante dans l'autre image est calculée ainsi qu'une zone située à ± 2 pixels autour de cette ligne. Seuls les points d'intérêt candidats situés dans cette région de l'image sont considérés pour l'appariement et le calcul du coefficient ZNSSD de la paire de points. Le point d'intérêt générant le coefficient ZNSSD le plus élevé est apparié. La figure 5.12 présente deux cas montrant une droite épipolaire, la zone de recherche associée et un point candidat à l'appariement. Dans

cet exemple, le point rouge n'est pas retenu pour le calcul du coefficient ZNSSD alors que le point bleu est retenu.

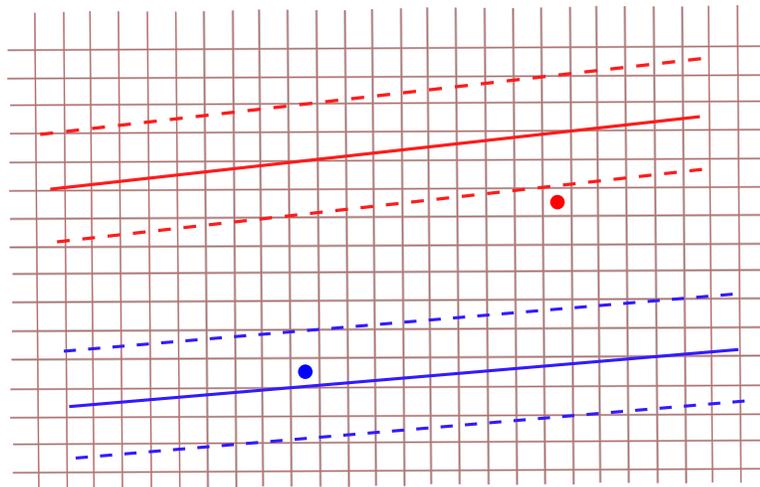


FIGURE 5.12 : Recherche de point autour d'une ligne épipolaire. Le point rouge se trouve à plus de deux pixels de distance de la droite épipolaire considérée alors que le point bleu se trouve dans la zone de sélection. Le point rouge n'est donc pas considéré pour l'appariement contrairement au point bleu.

La figure 5.13 présente les correspondances trouvées pour une paire non rectifiée d'images stéréoscopiques de feu dans le spectre du visible. Dans cet exemple, 643 correspondances ont été trouvées en 0,41 s. Une vérification manuelle a permis d'affirmer que les correspondances trouvées dans cet exemple sont justes.

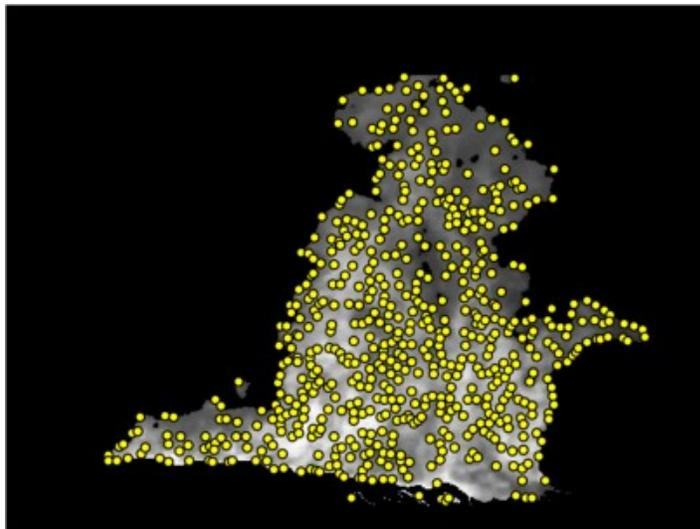


FIGURE 5.13 : Exemple de points appariés dans une image de feu du domaine du visible. Les points jaunes représentent les points appariés.

5.3.2 Fusion de données d'images pour l'appariement

La fusion de données d'images multimodales peut générer une image avec plus de détails. Ceci permet d'obtenir des descripteurs plus détaillés et donc plus caractéristiques des points d'intérêt auxquels ils sont associés. L'étape d'appariement des descripteurs est donc améliorée lorsque les images sont plus détaillées. Dans cette partie, nous présentons à partir d'un exemple d'image de feu, l'intérêt pour l'appariement de points d'intérêt de fusionner des images acquises simultanément dans le spectre du visible et du proche infra-rouge. Les résultats de fusion pour les différentes techniques présentées en section 5.2 sont montrés.

Les méthodes de fusion par moyenne et par analyse en composantes principales ont été considérées ainsi que les quatre méthodes suivantes de fusion utilisant des approches multirésolutions : pyramide laplacienne, pyramide de ratios, pyramide de contraste et transformée en ondelettes discrète (Haar, Debauchies 2, Coiflet et Meyer). Pour les approches multirésolutions, il existe une technique pour combiner les détails et une autre pour combiner les approximations (voir partie 5.2.2). Deux méthodes ont été testées pour combiner les détails. La première technique consiste à garder le maximum des valeurs absolues des détails, garantissant ainsi d'obtenir une image finale avec beaucoup de détails. La seconde méthode est la moyenne pondérée (équation (5.3)). Concernant la fusion des approximations, les deux méthodes citées précédemment ainsi que la sélection uniquement de l'image du spectre visible ont été considérées. Le choix de la sélection unique de l'image du spectre visible lors de l'étape de fusion des approximations permet de garder une bonne hétérogénéité des niveaux d'intensité de l'image.

La figure 5.14 présente une image d'un feu acquise simultanément dans le domaine du visible (*a*) et dans le domaine proche-infrarouge (*b*) ainsi que les images fusionnées obtenues par fusion par moyenne (*c*), analyse en composantes principales (*d*), pyramide laplacienne (*e*), pyramide de ratios (*f*), pyramide de contraste (*g*), transformée en ondelettes discrète de Haar (*h*), transformée en ondelettes discrète de Debauchies 2 (*i*), transformée en ondelettes discrète de Coiflet (*j*) et transformée en ondelettes discrète de Meyer (*k*). Les images (*a'*), (*b'*), (*c'*), (*d'*), (*e'*), (*f'*), (*g'*), (*h'*), (*i'*), (*j'*) et (*k'*), montrent respectivement le contenu de la zone délimitée par un carré de taille 90×90 pixels dans les images (*a*), (*b*), (*c*), (*d*), (*e*), (*f*), (*g*), (*h*), (*i*), (*j*) et (*k*). Ce carré est représenté en rouge dans l'image (*a*). Ces parties grossies permettent de mieux distinguer les détails qui apparaissent dans les images fusionnées par rapport aux images originales. Dans les images fusionnées, le feu est plus texturé que dans l'image du spectre proche-infrarouge (figure 5.14 (*a*)) et les contours ressortent plus que dans l'image du spectre visible (figure 5.14 (*b*)).

On remarque que les résultats de la fusion par moyenne et de la fusion par analyse en composantes principales sont très proches. Les images fusionnées par des méthodes utilisant des approches multirésolutions ont plus de détails que les images obtenues par moyenne ou par analyse en composantes principales. Les techniques de fusion par méthode pyramidale font

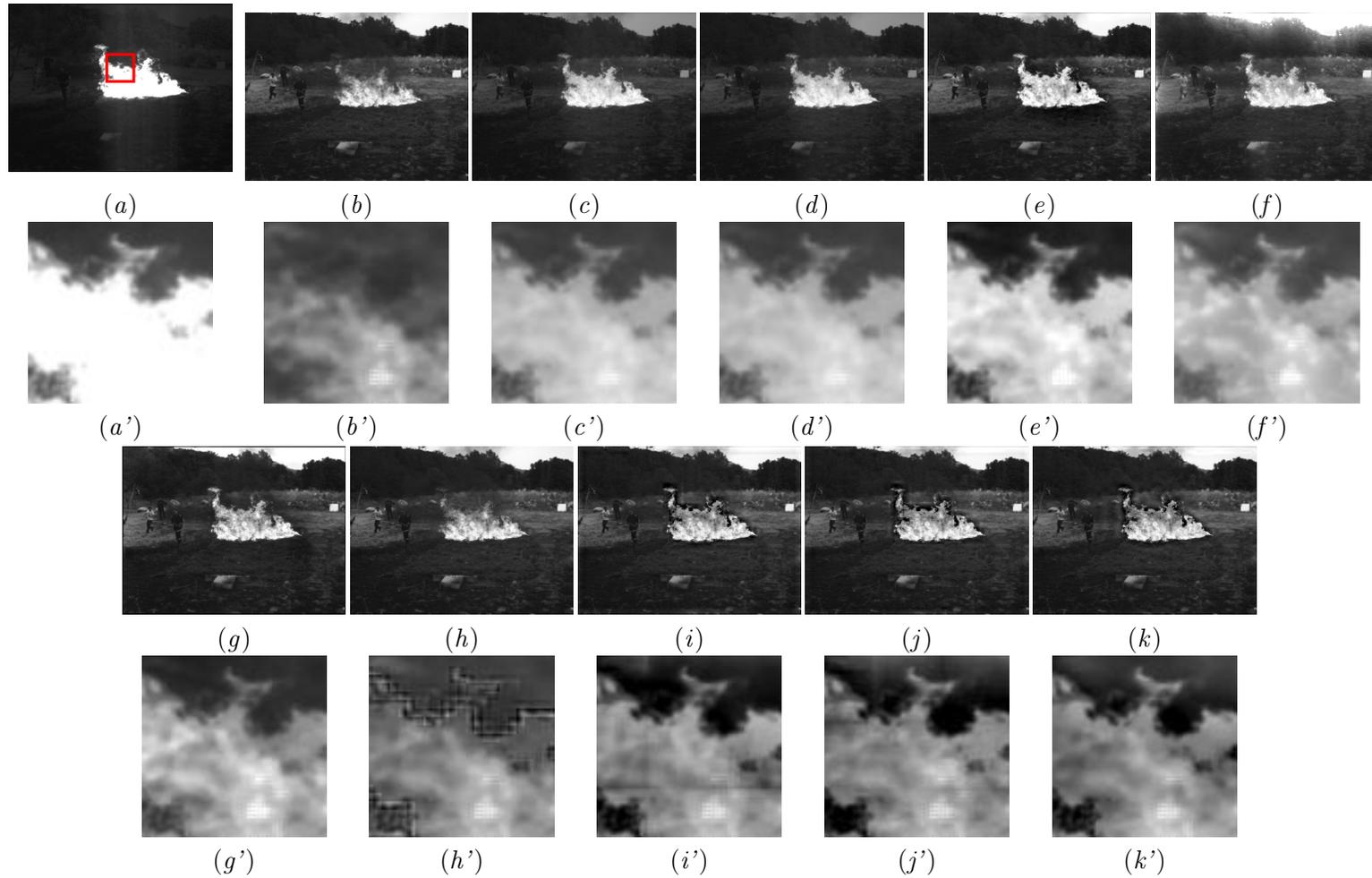


FIGURE 5.14 : Exemples de fusion d'images de feu; (a) image du spectre proche-infrarouge (b) image du spectre visible, (c) image fusionnée par moyenne, (d) image fusionnée par analyse en composantes principales (e) image fusionnée par pyramide laplacienne (f) image fusionnée par pyramide de ratios, (g) image fusionnée par pyramide de contraste, et (h), (i), (j), et (k) images fusionnées par transformée en ondelettes discrète de Haar, Debauchies 2, Coiflet et Meyer. Les détails ont été fusionnés en choisissant le maximum des valeurs absolues et les approximations en choisissant uniquement les approximations du visible. Les images (a')-(k') sont un agrandissement respectif des images (a)-(k).

bien ressortir les détails des images des deux spectres. Concernant la méthode par décomposition en ondelettes de Haar, on observe l'apparition d'artéfacts au niveau des contours du feu. Ces artéfacts sont courants dans la fusion d'image par décomposition en ondelettes [165, 166] et leur cause n'a jamais été clairement identifiée. De légers artéfacts pour les autres décompositions en ondelettes peuvent également être observés mais les résultats sont meilleurs visuellement.

D'une manière générale, les techniques de fusion d'images permettent d'obtenir des images plus intéressantes pour la recherche de points d'intérêt en faisant ressortir les détails et les contours du feu, particulièrement pour les fusions multirésolutions.

5.3.3 Concaténation de correspondances

Les méthodes de détection et d'appariement de points présentées précédemment peuvent s'appliquer aussi bien à des images du spectre du visible, du spectre du proche-infrarouge qu'aux images fusionnées. Nous proposons dans le cadre de cette thèse de considérer tous les points appariés obtenus dans les différents types d'image (images du spectre visible avec ceux des images du spectre proche-infrarouge, appariements des images fusionnées avec ceux des images du spectre visible et/ou proche-infrarouge, les appariements obtenus à partir de plusieurs types d'images fusionnées) afin d'augmenter le nombre final de points considérés pour calculer les points tridimensionnels. Ceci est réalisé par une méthode que nous appelons « concaténation » et que nous présentons dans cette section.

Soit deux ensembles P_A et P_B de points appariés ($\mathbf{x}_g, \mathbf{x}_d$). La première étape de la procédure de concaténation consiste à calculer un nouvel ensemble P obtenu en réalisant un ET logique entre les ensembles P_A et P_B ($P = P_A \cap P_B$) de sorte à éliminer les doublons d'éléments entre les deux ensembles P_A et P_B . La deuxième étape de la procédure consiste à ne conserver parmi tous les couples de points de P qui ont en commun un pixel \mathbf{x}_g ou un pixel \mathbf{x}_d uniquement celui pour lequel le coefficient ZNSSD est le plus petit.

La concaténation de N ensembles de points s'effectue en réalisant dans un premier temps la concaténation de deux ensembles de points puis en renouvelant la procédure de concaténation autant de fois que nécessaire en considérant comme ensemble à concaténer l'ensemble de points obtenu à l'étape précédente et un nouvel ensemble.

5.4 Evaluation de l'efficacité d'appariement de points d'intérêt en fonction du type d'image

Une étude a été réalisée pour évaluer l'efficacité d'appariement des points d'intérêt en fonction du type d'image (visible, proche-infrarouge et fusionnée). L'ensemble des images utilisées pour cette étude contient les 100 images acquises simultanément dans le proche-infrarouge et

le visible (constituant la base multimodale) ainsi que leurs images stéréoscopiques correspondantes.

Le nombre de points appariés dans chaque paire d'images étant dépendant du nombre de pixels constituant la zone de feu détectée, les performances des méthodes d'appariement sont présentées sous la forme du pourcentage du nombre de pixels appariés par rapport au nombre de pixels « feu ».

Cette partie présente, les résultats de l'étude d'efficacité d'appariement en considérant, dans une première partie, les images du spectre visible et du spectre proche-infrarouge, puis dans une deuxième partie, les images fusionnées (avec les différents types de fusion énoncés dans ce chapitre), et enfin, en concaténant les appariements obtenus dans les différents types d'image.

Le tableau 5.1 montre les pourcentages de points appariés par catégorie d'image en considérant séparément les images acquises dans le visible et dans le proche-infrarouge.

TABLEAU 5.1 : Pourcentage de pixels de feux appariés et temps d'exécution moyens pour les paires d'images du spectre visible et proche-infrarouge.

		Spectre visible <i>V</i>	Spectre proche-infrarouge <i>N</i>
Spectre visible	Rouge	0,65 %	-
	Orange	1,07 %	-
	Jaune-Blanc	0,14 %	-
	Texturé	0,78 %	-
	Non-texturé	0,56 %	-
	Faible intensité	0,78 %	-
	Intensité moyenne	0,64 %	-
	Haute intensité	0,43 %	-
	Sans fumée	0,64 %	-
	Avec fumée	0,62 %	-
N	Faible intensité	-	0,41 %
	Haute intensité	-	0,07 %
	Moyenne	0,64 %	0,35 %
	Temps (s)	0,379	0,342

Le feu étant plus texturé dans les images du spectre du visible que dans les images du proche-infrarouge, le pourcentage de points appariés dans les images du visible (0,64%) est supérieur à celui estimé dans les images du spectre proche-infrarouge (0,35%). De plus, parmi les images acquises dans le visible, celles qui sont texturées permettent d'obtenir un meilleur score d'appariement (0,78% au lieu de 0,56%). Ceci apparaît également indirectement en considérant la couleur principale des feux. En effet, les images de couleur principale jaune-blanc sont moins

texturées dans la zone de feu et donc génèrent moins de points appariés (0,14%) que les images ayant une autre couleur principale (1,07% pour orange et 0,65% pour rouge). Cette couleur jaune-blanc est due à un temps d'exposition élevé de l'appareil de prise d'image. On peut également remarquer que, quel que soit le type d'image, plus l'intensité lumineuse de l'environnement est grand, moins il y a de points de correspondance trouvés. Les détails du feu apparaissent de meilleure façon lorsque l'environnement est sombre. Il est donc préférable de choisir un temps d'exposition de la caméra de façon à être légèrement sous-exposé pour une meilleure vision du feu.

Concernant les temps d'exécution, la recherche d'appariements dans les images du proche-infrarouge est plus rapide car il y a moins de points d'intérêt dans l'image et donc moins de descripteurs à construire et à comparer.

Le tableau 5.2 présente les pourcentages de points appariés obtenus en fonction de la méthode de fusion utilisée pour combiner les images du visible et du proche-infrarouge. Pour chaque méthode de fusion multirésolution, les scores sont donnés en fonction du type de fusion de détails et d'approximation. Concernant la fusion des détails, « Max » représente la fusion par maximum des valeurs absolue et « Pond » représente la fusion par pondération des valeurs. Concernant la fusion des approximations, « Max » et « Pond » représentent les mêmes types de fusion que ceux utilisés pour les détails et « Vis » représente le cas pour lequel seule l'approximation du spectre visible est gardée pour la reconstruction de l'image.

Les méthodes de fusion ne permettent pas d'obtenir des images qui ont, en moyenne, plus de points appariés que les images du visible. Une explication à ceci est que les images traitées sont composées de la zone de feu segmentée à partir de l'image originale et d'un fond noir correspondant à l'environnement. De ce fait, les contours de feu sont très marqués dans tout type d'image et la différence du nombre de points appariés en fonction des images est due principalement à la texture des zones de feu très présente dans les images du visible.

En ce qui concerne les temps d'exécution, les plus bas sont obtenus avec les méthodes de fusion classiques. Il peut être remarqué aussi que le temps d'exécution des méthodes de fusion par décomposition en ondelettes sont plus grands que ceux des méthodes de décompositions pyramidales.

Enfin, les images générées par la méthode de pyramide de contraste utilisant une fusion de détails par maximum des valeurs absolues et uniquement l'image du spectre visible pour obtenir son image d'approximation obtiennent le meilleur pourcentage de points appariés des méthodes de fusion. Cette fusion sera notée F_1 dans la suite de cette section. Les pourcentages d'obtention de points appariés, détaillés par catégorie d'images fusionnées par la méthode F_1 , sont présentés dans le tableau 5.3 et peuvent être comparés avec ceux obtenus par les méthodes travaillant uniquement sur des images monomodales (également présentés dans ce tableau). On peut constater que les images générées par la méthode F_1 obtiennent en moyenne

TABLEAU 5.2 : Pourcentage d'appariements de points et temps d'exécution moyens obtenus en fonction de la méthode de fusion utilisée pour combiner les images du visible et du proche-infrarouge. Les méthodes de fusion sont Max : maximum des valeurs absolues ; Pond : pondération des valeurs ; Vis : spectre visible

Méthode	Fusion Détails	Fusion Approximations	Fusion (F)	Temps (s)
Moyenne	-	-	0,54 %	0,386
ACP	-	-	0,49 %	0,469
Pyramide laplacienne	Max	Vis	0,57 %	0,632
		Max	0,42 %	0,613
		Pond	0,49 %	0,621
	Pond	Vis	0,37 %	0,686
		Max	0,34 %	0,673
		Pond	0,34 %	0,685
Pyramide de ratios	Max	Vis	0,52 %	0,620
		Max	0,33 %	0,596
		Pond	0,40 %	0,612
	Pond	Vis	0,57 %	0,675
		Max	0,49 %	0,666
		Pond	0,56 %	0,676
Pyramide de contraste	Max	Vis	0,62 %	0,627
		Max	0,42 %	0,611
		Pond	0,51 %	0,622
	Pond	Vis	0,35 %	0,686
		Max	0,32 %	0,672
		Pond	0,35 %	0,683
Décomposition en ondelettes de Haar	Max	Vis	0,46 %	1,563
		Max	0,38 %	1,563
		Pond	0,43 %	1,574
	Pond	Vis	0,29 %	1,645
		Max	0,33 %	1,627
		Pond	0,34 %	1,628
Décomposition en ondelettes de Debauchies	Max	Vis	0,44 %	1,569
		Max	0,35 %	1,555
		Pond	0,43 %	1,569
	Pond	Vis	0,32 %	1,631
		Max	0,32 %	1,624
		Pond	0,34 %	1,625
Décomposition en ondelettes de Coiflet	Max	Vis	0,46 %	1,573
		Max	0,39 %	1,561
		Pond	0,45 %	1,578
	Pond	Vis	0,29 %	1,636
		Max	0,30 %	1,632
		Pond	0,31 %	1,631
Décomposition en ondelettes de Meyer	Max	Vis	0,52 %	2,645
		Max	0,37 %	2,655
		Pond	0,42 %	2,646
	Pond	Vis	0,36 %	2,733
		Max	0,32 %	2,737
		Pond	0,33 %	2,762

un pourcentage de points appariés plus bas que les images du visible. Toutefois, les images de feu fusionnées présentant les caractéristiques telles que couleur jaune-blanc, feux non-texturés, environnement de faible intensité, avec fumée, ont un pourcentage d'appariement plus élevé que celles du visible. La fusion d'images multimodales a donc un intérêt particulier dans ces types d'images. Les images de feux réels contenant généralement plus de fumée que les images utilisées pour cette étude, ces techniques d'appariement d'images fusionnées sont donc importantes pour l'appariement d'images de feu de végétation.

TABLEAU 5.3 : Pourcentages de pixels de feux appariés par catégorie d'image obtenus pour des images fusionnées par la méthode F_1 et pour des images couleur du spectre du visible (V).

	V	F_1
Rouge	0,65 %	0,64 %
Orange	1,07 %	0,82 %
Jaune-Blanc	0,14 %	0,28 %
Texturé	0,78 %	0,72 %
Non-texturé	0,56 %	0,57 %
Faible intensité	0,78 %	0,86 %
Intensité moyenne	0,64 %	0,59 %
Haute intensité	0,43 %	0,41 %
Sans fumée	0,64 %	0,61 %
Avec fumée	0,62 %	0,72 %
Faible intensité	0,70 %	0,69 %
Haute intensité	0,28 %	0,27 %
Moyenne	0,64 %	0,62 %

Le tableau 5.4 présente les pourcentages des pixels de feu appariés estimés après concaténation des points appariés à partir de différents types d'images tel que cela a été décrit précédemment. Dans ce tableau, le signe \sqcap représente la concaténation, E_F l'ensemble des points appariés à partir d'images fusionnées, E_N l'ensemble des points appariés à partir d'images du proche-infrarouge et E_V l'ensemble des points appariés à partir d'images du visible.

Les pourcentages sont toujours meilleurs lorsque la concaténation des appariements obtenus dans les images fusionnées est faite avec les points appariés à partir d'images du spectre visible plutôt qu'à partir d'images du spectre proche-infrarouge. Ceci s'explique par le fait qu'il y a un nombre moins important de points appariés obtenus à partir des images du spectre proche-infrarouge qu'à partir des images du spectre visible. De plus, les scores sont encore meilleurs lorsque les points appariés issus des trois types d'images sont concaténés.

De plus, la fusion par ondelettes de Meyer avec une fusion des détails par maximum de valeur absolue et en gardant ne gardant que les approximations du spectre visible obtient les meilleurs scores pour la concaténation $E_F \sqcap E_V$ ainsi que pour la concaténation $E_F \sqcap E_V \sqcap E_N$. Nous noterons F_2 cette méthode ainsi que E_{F_2} l'ensemble des points appariés obtenus à partir d'images générées par cette fusion. Pour la concaténation $E_F \sqcap E_N$, c'est l'ensemble obtenu

TABLEAU 5.4 : Pourcentage d'appariements pour les concaténations d'appariement. E_V correspond aux appariements d'images du spectre visible, E_N aux appariements d'images du spectre proche-infrarouge et F aux appariements d'images fusionnées. Le signe \sqcap correspond à la concaténation.

Méthode	Fusion Détails	Fusion Approximations	$E_F \sqcap E_V$	$E_F \sqcap E_N$	$E_F \sqcap E_V \sqcap E_N$
Moyenne	-	-	0,84 %	0,65 %	0,94 %
ACP	-	-	0,81 %	0,65 %	0,95 %
Pyramide laplacienne	Max	Vis	0,98 %	0,70 %	1,07 %
		Max	0,89 %	0,52 %	0,97 %
		Pond	0,94 %	0,59 %	1,01 %
	Pond	Vis	0,90 %	0,50 %	0,99 %
		Max	0,85 %	0,43 %	0,91 %
		Pond	0,86 %	0,43 %	0,92 %
Pyramide de ratios	Max	Vis	0,92 %	0,67 %	1,06 %
		Max	0,80 %	0,48 %	0,93 %
		Pond	0,86 %	0,54 %	0,99 %
	Pond	Vis	0,86 %	0,72 %	1,00 %
		Max	0,87 %	0,64 %	1,01 %
		Pond	0,87 %	0,71 %	1,00 %
Pyramide de contraste	Max	Vis	0,93 %	0,76 %	1,04 %
		Max	0,89 %	0,56 %	1,00 %
		Pond	0,93 %	0,64 %	1,04 %
	Pond	Vis	0,86 %	0,48 %	0,96 %
		Max	0,82 %	0,43 %	0,91 %
		Pond	0,86 %	0,45 %	0,93 %
Décomposition en ondelettes de Haar	Max	Vis	0,91 %	0,69 %	1,06 %
		Max	0,87 %	0,50 %	0,96 %
		Pond	0,90 %	0,57 %	1,00 %
	Pond	Vis	0,86 %	0,52 %	1,01 %
		Max	0,83 %	0,43 %	0,91 %
		Pond	0,86 %	0,45 %	0,94 %
Décomposition en ondelettes de Debauchies	Max	Vis	0,93 %	0,67 %	1,08 %
		Max	0,85 %	0,48 %	0,96 %
		Pond	0,91 %	0,57 %	1,01 %
	Pond	Vis	0,89 %	0,54 %	1,02 %
		Max	0,84 %	0,41 %	0,90 %
		Pond	0,86 %	0,44 %	0,93 %
Décomposition en ondelettes de Coiflet	Max	Vis	0,94 %	0,69 %	1,09 %
		Max	0,88 %	0,50 %	0,96 %
		Pond	0,92 %	0,57 %	1,01 %
	Pond	Vis	0,85 %	0,51 %	0,99 %
		Max	0,81 %	0,40 %	0,89 %
		Pond	0,83 %	0,43 %	0,91 %
Décomposition en ondelettes de Meyer	Max	Vis	1,00 %	0,73 %	1,14 %
		Max	0,86 %	0,49 %	0,96 %
		Pond	0,90 %	0,55 %	1,00 %
	Pond	Vis	0,92 %	0,55 %	1,04 %
		Max	0,82 %	0,41 %	0,90 %
		Pond	0,85 %	0,43 %	0,92 %

par la méthode F_1 qui obtient les meilleurs scores. Cet ensemble sera noté E_{F_1}

La concaténation $E_{F_1} \sqcap E_V$ obtient donc un score inférieur à celui de $E_{F_2} \sqcap E_V$ alors que l'ensemble E_{F_1} est composé de plus d'appariements que la concaténation E_{F_2} . Cette observation implique que le nombre d'appariements contenus dans E_{F_2} mais pas dans E_V est plus grand que le nombre d'appariements contenus dans E_{F_1} mais pas dans E_V . Cette observation peut s'expliquer par les artefacts générés par cette méthode de fusion qui permettent de créer des détails n'existant dans aucun des spectres ; en effet, les artefacts se créent de la même manière dans les mêmes zones des images pour les vues de gauche et de droite.

On remarque par ailleurs que la fusion des détails par maximum des valeurs absolue en gardant les approximations du spectre visible permet toujours d'obtenir les meilleurs scores pour chaque type de fusion. Ce type de fusion de détails et d'approximation est donc à privilégier pour la suite.

Dans ce tableau les indications de temps d'exécution ne sont pas données car la concaténation des points d'appariement prend un temps total de l'ordre de la somme des temps des différentes recherches d'appariement combinées. En effet, le temps de calcul de la concaténation des points est négligeable par rapport au temps de recherche d'appariements. Par exemple, le temps d'exécution pour obtenir des appariements issus de la concaténation de ceux obtenus à partir d'images du spectre visible et à partir d'images fusionnées par moyenne est de l'ordre de 77 centièmes de seconde (0,38 + 0,39).

Le tableau 5.5 présente le pourcentage de points appariés par catégorie d'image et par type de concaténation pour les méthodes ayant obtenues les meilleurs scores dans le tableau 5.4. Dans ce tableau, $E_V \sqcap E_N$ est l'ensemble des points obtenus par concaténation de l'ensemble des pixels appariés extraits depuis des images du visible avec l'ensemble des pixels appariés obtenus depuis des images du proche-infrarouge, $E_{F_2} \sqcap E_N$ est l'ensemble des points obtenus par concaténation de l'ensemble E_{F_2} avec l'ensemble des pixels appariés obtenus depuis des images du proche-infrarouge, $E_{F_1} \sqcap E_N$ est l'ensemble des points obtenus par concaténation de l'ensemble E_{F_1} avec l'ensemble des pixels appariés extraits depuis des images du proche-infrarouge, et $E_{F_2} \sqcap E_V \sqcap E_N$ est l'ensemble des points obtenus par concaténation de l'ensemble E_{F_2} avec l'ensemble des pixels appariés obtenus depuis des images du proche-infrarouge et avec l'ensemble des pixels appariés extraits depuis des images du visible.

La concaténation $E_{F_2} \sqcap E_V$ donne de meilleurs résultats que $E_V \sqcap E_N$ ou $E_{F_1} \sqcap E_N$. La fusion des trois types d'image permet d'obtenir des ratios de plus de 1 % pour la plupart des catégories. Comme pour les méthodes sans concaténation, les meilleurs scores sont obtenus sur des feux de couleur principale orange, texturés et lorsque l'environnement est de faible intensité.

Étant donné que les meilleurs scores sont obtenus avec la concaténation de trois ensembles,

TABLEAU 5.5 : Pourcentage de pixels de feu appariés par catégorie pour des concaténations particulières. $E_V \sqcap E_N$ représente la concaténation des points appariés obtenus à partir d’images du visible et d’images du proche-infrarouge ; $E_{F_2} \sqcap E_V$ représente la concaténation de l’ensemble E_{F_2} avec l’ensemble des points appariés obtenus à partir d’images du visible ; $E_{F_1} \sqcap E_N$ représente la concaténation de l’ensemble E_{F_1} avec l’ensemble des points appariés obtenus à partir d’images du proche-infrarouge ; et $E_{F_2} \sqcap E_V \sqcap E_N$ représente la concaténation de l’ensemble E_{F_2} avec l’ensemble des points appariés obtenus à partir d’images du proche-infrarouge.

	$E_V \sqcap E_N$	$E_{F_2} \sqcap E_V$	$E_{F_1} \sqcap E_N$	$E_{F_2} \sqcap E_V \sqcap E_N$
Rouge	0,84 %	1,02 %	0,75 %	1,14 %
Orange	1,43 %	1,58 %	1,09 %	1,78 %
Jaune-Blanc	0,50 %	0,27 %	0,55 %	0,60 %
Texturé	1,04 %	1,17 %	0,89 %	1,34 %
Non-texturé	0,74 %	0,92 %	0,69 %	1,04 %
Faible intensité	1,10 %	1,38 %	1,04 %	1,58 %
Intensité moyenne	0,80 %	0,96 %	0,70 %	1,06 %
Haute intensité	0,66 %	0,64 %	0,58 %	0,82 %
Sans fumée	0,83 %	0,97 %	0,73 %	1,10 %
Avec fumée	0,99 %	1,23 %	0,93 %	1,44 %
Faible intensité	0,95 %	1,12 %	0,85 %	1,28 %
Haute intensité	0,31 %	0,40 %	0,29 %	0,42 %
Moyenne	0,85 %	1,00 %	0,76 %	1,14 %

on peut aussi penser à concaténer les points issus de fusion différente. Cette procédure peut avoir un temps d’exécution élevé longue selon le nombre de fusions utilisées et produit de moins en moins de nouveaux points à chaque nouvelle concaténation. Le tableau 5.6 présente le pourcentage de points appariés en fonction du nombre de concaténations de fusion utilisées. L’ordre des fusions utilisées est choisi en fonction du score de la colonne $E_F \sqcap E_V \sqcap E_N$ du tableau 5.4 ; ce qui correspond à l’ordre suivant :

- décomposition en ondelettes de Meyer
- décomposition en ondelettes de Coiflet
- décomposition en ondelettes de Debauchies 2
- décomposition par pyramide laplacienne
- décomposition par pyramide de ratios
- décomposition en ondelettes de Haar
- décomposition par pyramide de contraste
- analyse en composantes principales
- moyenne des intensités

Pour chaque fusion multirésolution, les détails sont fusionnés par maximum des valeurs absolues et l’approximation du spectre visible est gardée. En effet, il a été vu dans le tableau 5.2

que ce type de fusion obtenait toujours les meilleurs scores. Lorsque les ensembles de tous les types de fusion sont concaténés, les scores obtenus sont de l'ordre de 2 %.

TABLEAU 5.6 : Pourcentage de pixels appariés pour la concaténation des points obtenus par fusion.

Nombre Fusions	0	1	2	3	4	5	6	7	8	9
Ratios (%)	0,85	1,14	1,32	1,47	1,60	1,77	1,89	1,97	2,02	2,03
Temps (s)	0,73	3,37	4,93	6,49	7,11	7,72	9,28	9,90	10,36	10,74

La figure 5.15 (a) présente les correspondances trouvées pour les images de la figure 5.13 en utilisant la concaténation des points appariés des deux spectres avec ceux de la fusion par méthode de décomposition par ondelettes de Meyer. Pour le deuxième type de concaténation, 918 correspondances ont été trouvées en 3,42 s. Soit 275 points de correspondance de plus que pour les appariements dans le spectre visible seul mais en 6 fois plus de temps.

Une vérification manuelle a permis de déterminer que seules quatre correspondances étaient fausses. Elles correspondent à des points détectés à la base du feu qui ont beaucoup de similitudes et qui se situent sur la même ligne épipolaire. Ces faux appariements peuvent être évités en amont en ne considérant pour la recherche d'appariements que les pixels appartenant à de petites zones de feu. La figure 5.15 (b) présente aussi les points de correspondance obtenus par la méthode de concaténation. Dans ce cas, 1460 points de correspondance ont été trouvés en 12,04 s avec une dizaine de faux appariements.

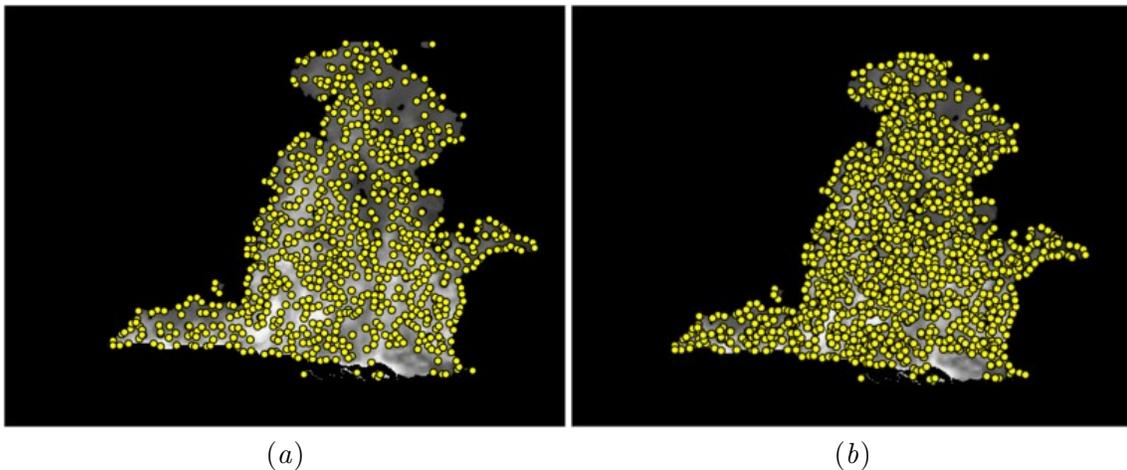


FIGURE 5.15 : Concaténation des correspondances de points pour une image de feu (a) pour les points trouvés dans le spectre visible, proche-infrarouge et dans l'image fusionnée par décomposition en ondelette et (b) concaténation des points de correspondance de tous les types de fusion.

Les méthodes de recherche de points décrites précédemment ont de bons résultats pour des images dans lesquelles le feu n'est pas totalement masqué par de la fumée. Si, lors de l'étape de détection de pixels, l'image du spectre visible est considérée comme ayant trop de fumée

pour être traitée, celle-ci n'est pas considérée non plus pour effectuer la recherche de points de correspondances. Dans ce cas-là, les appariements sont recherchés seulement dans les images du spectre proche-infrarouge. Cela évite de reconstruire d'éventuelles textures de fumée.

5.5 Conclusion

Dans ce chapitre, nous avons présenté une nouvelle approche pour la recherche de correspondances rendue possible grâce à l'aspect multimodal du système de stéréovision. En effet, nous avons proposé de fusionner les valeurs des pixels des images des deux spectres pour obtenir de nouveaux détails dans les images. Il a été montré que pour certaines catégories d'images, certains types de fusion d'images multimodales permettaient d'obtenir plus de points appariés que dans les images du spectre visible. Nous avons aussi proposé de concaténer les appariements obtenus à partir des différents spectres ainsi que des images fusionnées. Le choix de la méthode à utiliser peut se faire sur la base du temps d'exécution alloué à cette étape. Si il n'y a pas de contraintes de temps, la meilleure méthode est celle qui concatène tous les types d'appariements, sinon il est préférable de choisir la méthode concaténant les appariements de la fusion par pyramide laplacienne avec les appariements obtenus dans le spectre visible seul. Le premier choix a obtenu dans notre étude un temps d'exécution moyen de plus de dix secondes alors que le second permet d'obtenir des appariements en une seconde en moyenne. Dans la suite de ce document, le premier type de concaténation d'appariements est utilisé.

Ces méthodes s'appliquent particulièrement bien pour les feux de végétation mais il est aussi envisageable de les appliquer à d'autres types d'objet qui ont des caractéristiques différentes dans les deux spectres (visages, silhouettes, végétaux, ...).

L'amélioration de la précision des méthodes de recherche de correspondances proposées pourraient être obtenue en réalisant une recherche d'appariements subpixelique [167, 168] ou des méthodes de mise en correspondance par propagation de germes [169, 170, 143].

Les méthodes présentées dans ce chapitre permettent d'optimiser le nombre de paires de points appariés dans des images de feux de végétation. Grâce à ces appariements, des points tridimensionnels sont calculés à partir desquels les caractéristiques géométriques des feux sont estimées. L'optimisation du nombre de points permet une reconstruction plus dense du feu et donc plus précise.

Chapitre 6

Estimation de caractéristiques géométriques de fronts de feu à partir d'images stéréoscopiques

6.1	Introduction	173
6.2	Principes théoriques de la stéréovision	174
6.2.1	Modèle géométrique de la caméra	174
6.2.2	Estimation des paramètres intrinsèques d'une caméra	176
6.2.3	Système stéréoscopique	179
6.2.4	Estimation de la matrice essentielle	182
6.2.5	Reconstruction 3D des points appariés	184
6.3	Estimation des paramètres intrinsèques et extrinsèques d'un système de vision	185
6.4	Reconstruction tridimensionnelle de points de feu	188
6.5	Estimation des caractéristiques géométriques d'un feu	188
6.5.1	Estimation du plan de base	190
6.5.2	Calcul des caractéristiques géométriques	191
6.6	Conclusion	198

6.1 Introduction

Ce chapitre présente les algorithmes utilisés et développés pour l'estimation de caractéristiques géométriques de feu à partir de points tridimensionnels obtenus depuis des images stéréoscopiques d'un système de vision positionné devant le front de feu par rapport à sa direction de propagation. Les données suivantes sont ainsi calculées : la position des points au sol du front et leur vitesse, la direction principale de propagation du front de feu, sa largeur, sa hauteur en fonction de sa largeur, sa longueur, son inclinaison ainsi que sa surface.

La section 6.2 décrit les principes théoriques de la stéréovision et la procédure d'obtention de points tridimensionnels.

La section 6.3 présente la procédure de calibrage utilisée pour l'estimation des paramètres intrinsèques et extrinsèques des deux systèmes de vision considérés dans cette thèse.

La section 6.4 explique une procédure qui a été développée afin d'éliminer les points tridimensionnels situés trop loin du centre des points 3D et considérés comme erronés

La section 6.5 présente les méthodes d'estimation de caractéristiques géométriques de feu à partir de leurs points tridimensionnels estimés.

6.2 Principes théoriques de la stéréovision

Cette section décrit les principes théoriques permettant la reconstruction tridimensionnelle d'un objet à partir de plusieurs vues décalées de cet objet. Un des ouvrages de référence dans le domaine de la vision 3D est le livre de Hartley et Zisserman [129] et cette section en est largement inspirée. Dans une première partie, le modèle géométrique d'une caméra est décrit. Une deuxième partie est consacrée aux calculs des paramètres de ce modèle. Dans une troisième partie, le modèle géométrique d'un système de stéréovision est présenté. Une quatrième partie est dédiée aux calculs des paramètres de ce système. Enfin, une dernière partie est consacrée aux calculs permettant d'obtenir des points tridimensionnels à partir des coordonnées de points appariés dans des images stéréoscopiques.

6.2.1 Modèle géométrique de la caméra

Une caméra est définie par son centre optique O_c , son axe optique z_c (perpendiculaire au plan image et passant par le point O_c), le centre de l'image x_c (qui est l'intersection entre le plan image et l'axe optique), sa focale f_c (qui est la distance entre les points O_c et x_c). Ces paramètres sont appelés « paramètres intrinsèques » de la caméra. On note (u,v) les coordonnées image d'un pixel \mathbf{x}_p correspondant à un point P de coordonnées (X_c, Y_c, Z_c) dans le repère caméra (figure 6.1) [129].

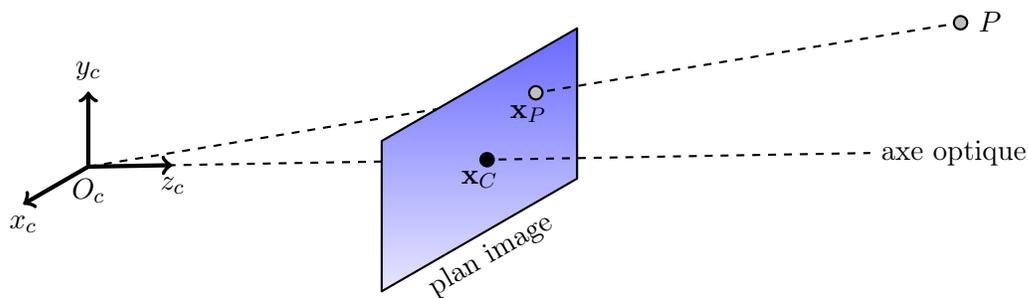


FIGURE 6.1 : Modèle géométrique d'une caméra. O_c est le centre optique, x_c le centre de l'image, P est un point du plan et \mathbf{x}_p est la projection du point P sur le plan image.

Tout point de l'espace P de coordonnées (X_c, Y_c, Z_c) exprimées dans le repère $(\vec{x}_c, \vec{y}_c, \vec{z}_c)$ est projeté dans le plan image en un point \mathbf{x}_P de coordonnées (u, v) selon l'équation :

$$\begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = \begin{bmatrix} \alpha_u & \gamma & u_0 & 0 \\ 0 & \alpha_v & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{pmatrix} X_c \\ Y_c \\ Z_c \\ 1 \end{pmatrix} \quad (6.1)$$

Avec :

- (u_0, v_0) les coordonnées en pixel de x_C
- α_u et α_v les distances focales en pixel sur les directions u et v
- γ un paramètre différent de zéro lorsque le repère image n'est pas orthogonal.

On a $\alpha_u = f_c m_u$ et $\alpha_v = f_c m_v$ avec m_u et m_v les facteurs d'échelle de l'image dans la direction de u et de v ; ils représentent le nombre de pixels par unité de distance dans les directions u et v .

Le modèle de la caméra est dit « sténopé » si on considère $m_u = m_v = 1$. Dans le cas d'une caméra CCD, la matrice de calibrage K_c est définie par :

$$K_c = \begin{bmatrix} \alpha_u & \gamma & u_0 \\ 0 & \alpha_v & v_0 \\ 0 & 0 & 1 \end{bmatrix} \quad (6.2)$$

Cette matrice contient les paramètres intrinsèques de la caméra qui définissent son modèle géométrique.

La transformation qui lie le repère de la caméra $(\vec{x}_c, \vec{y}_c, \vec{z}_c)$ à celui de la scène est une transformation rigide qui se compose d'une rotation et d'une translation. En considérant que le point P de l'espace de coordonnées (X_c, Y_c, Z_c) dans le repère de la caméra a pour coordonnées (X, Y, Z) dans le repère de la scène, on a la relation suivante :

$$\begin{pmatrix} X_c \\ Y_c \\ Z_c \end{pmatrix} = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix} \begin{pmatrix} X \\ Y \\ Z \end{pmatrix} + \begin{pmatrix} t_x \\ t_y \\ t_z \end{pmatrix} \quad (6.3)$$

Avec une homogénéisation des coordonnées, la relation (6.3) peut s'écrire sous la forme :

$$\begin{pmatrix} X_c \\ Y_c \\ Z_c \\ 1 \end{pmatrix} = \begin{bmatrix} \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \\ 0 & 0 & 0 \end{bmatrix} & \begin{pmatrix} t_x \\ t_y \\ t_z \\ 1 \end{pmatrix} \end{bmatrix} \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix} = \begin{bmatrix} R_c & \mathbf{t}_c \\ 0 & 1 \end{bmatrix} \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix} \quad (6.4)$$

Avec R_c une matrice de rotation et \mathbf{t}_c un vecteur de translation, qui sont appelés « paramètres extrinsèques » de la caméra.

En combinant les équations (6.1) et (6.4), la relation qui lie les coordonnées du point P dans le repère de la scène à celui de son représentant dans l'image s'écrit :

$$\begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = \begin{bmatrix} K_c & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} R_c & \mathbf{t}_c \\ 0 & 1 \end{bmatrix} \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix} = M_c \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix} \quad (6.5)$$

M_c est la matrice de projection associée à la caméra.

La sous-section suivante présente les calculs pour estimer les paramètres intrinsèques d'une caméra.

6.2.2 Estimation des paramètres intrinsèques d'une caméra

Les paramètres intrinsèques d'une caméra sont estimés à partir de correspondances établies entre les coordonnées 3D de points d'un plan (P^i) et les coordonnées en pixel de leurs projections dans l'image de la caméra (x_c^i). Ceci est réalisé en utilisant une mire constituée de motifs dont les caractéristiques géométriques et les positions sont précisément connues. Les paramètres sont obtenus en déterminant l'homographie H_c qui minimise $\|\mathbf{x}_c^i - H_c P^i\|^2$. Pour cela, plusieurs étapes sont nécessaires.

Une normalisation des coordonnées P^i et c^i est faite dans une première étape. Cette normalisation consiste à appliquer une transformation de similarité S_1 aux points \mathbf{x}_c^i et une transformation S_2 aux points P^i . Chaque transformation de similarité se compose d'une translation, qui déplace les points de sorte que leur origine corresponde à leur barycentre, et d'une mise à l'échelle qui agit de sorte que la distance moyenne des points à l'origine soit égale à $\sqrt{2}$. Cette mise à l'échelle dite « isotrope » a la particularité de générer un point moyen de coordonnées $(1, 1, 1)^\top$. En posant $\tilde{\mathbf{x}}_i = S_1 \mathbf{x}_c^i = (\tilde{u}_i, \tilde{v}_i, \tilde{w}_i)$ et $\tilde{P}_i = S_2 P^i$, à partir de l'équation (6.5), les équations suivantes sont obtenues [129] :

$$\begin{bmatrix} \mathbf{0} & -\tilde{w}_i \tilde{P}_i & \tilde{v}_i \tilde{P}_i \\ \tilde{w}_i \tilde{P}_i & \mathbf{0} & -\tilde{u}_i \tilde{P}_i \end{bmatrix} \tilde{\mathcal{H}} = A \tilde{\mathcal{H}} = 0 \quad (6.6)$$

Avec :

$$\tilde{\mathcal{H}} = \left[\begin{array}{cccccccccccc} h_{11} & h_{12} & h_{13} & h_{14} & h_{21} & h_{22} & h_{23} & h_{24} & h_{31} & h_{32} & h_{33} & h_{34} \end{array} \right]^\top \quad (6.7)$$

Où h_{ab} est l'élément de la matrice \tilde{H}_c d'indice de ligne a et de colonne b (matrice de projection normalisée). $\tilde{\mathcal{H}}$ est donc une matrice colonne composée de tous les éléments de \tilde{H}_c . Estimer $\tilde{\mathcal{H}}$ permet donc d'estimer \tilde{H}_c .

A partir de l'équation (6.6) appliquée à douze correspondances minimum, il est possible de déterminer les éléments de la matrice \tilde{H}_c . Pour optimiser la performance de l'estimation, une contrainte sur \tilde{H}_c peut être rajoutée telle que la suivante [129] :

$$\|\tilde{H}_c\| = 1 \quad (6.8)$$

En fixant cette contrainte, une solution de (6.6) peut être obtenue en décomposant la matrice A en valeurs singulières (en anglais, *singular value decomposition* : SVD) [171]. Ainsi, si la matrice A peut s'écrire $A = UDV^\top$ avec D une matrice diagonale dont les éléments sont arrangés dans l'ordre décroissant, alors $\tilde{\mathcal{H}}$ correspond à la dernière colonne de V .

La solution ainsi obtenue est une estimation de \tilde{H}_c qu'il faut améliorer en réduisant l'erreur géométrique, c'est à dire en réduisant l'écart entre les points \mathbf{x}_c^i et leur estimation $\tilde{\mathbf{x}}_c^i = \tilde{H}_c P^i$ donné par :

$$\sum_i \|\mathbf{x}_c^i - \tilde{\mathbf{x}}_c^i\|^2 \quad (6.9)$$

Pour cela, la technique itérative de Levenberg-Marquardt [172] est utilisée avec comme valeur initiale l'estimation de \tilde{H}_c obtenue grâce à $\tilde{\mathcal{H}}$. La matrice \tilde{H}_c qui vérifie l'équation suivante est ainsi recherchée :

$$\min_{\tilde{H}_c} \sum_i \|\mathbf{x}_c^i - \tilde{H}_c P^i\|^2 \quad (6.10)$$

La matrice \tilde{H}_c est ensuite dénormalisée et génère la matrice H_c selon la formule :

$$H_c = S_1^{-1} \tilde{H}_c S_2 \quad (6.11)$$

La matrice H_c s'écrit sous la forme :

$$\left[\begin{array}{ccc} \mathbf{h}_c^1 & \mathbf{h}_c^2 & \mathbf{h}_c^3 \end{array} \right] \quad (6.12)$$

Et l'on peut écrire [173] :

$$\begin{bmatrix} \mathbf{h}_c^1 & \mathbf{h}_c^2 & \mathbf{h}_c^3 \end{bmatrix} = \lambda K_c M_c \quad (6.13)$$

Avec \mathbf{h}_c^i la colonne i de la matrice \tilde{H}_c , λ un facteur d'échelle et M_c une matrice définie à partir des paramètres extrinsèques de la caméra introduits dans l'équation (6.4) :

$$M_c = \begin{bmatrix} r_{11} & r_{12} & t_x \\ r_{21} & r_{22} & t_y \\ r_{31} & r_{32} & t_z \end{bmatrix} \quad (6.14)$$

Il n'est pas nécessaire d'estimer cette matrice pour calculer les paramètres intrinsèques contenues dans la matrice K_c . Comme les deux premières colonnes de M_c sont deux colonnes d'une matrice de rotation, elles sont orthogonales. Cette propriété permet d'obtenir les deux contraintes suivantes à partir de l'équation (6.13) :

$$\mathbf{h}_c^{1\top} K_c^{-\top} K_c^{-1} \mathbf{h}_c^{2\top} = 0 \quad (6.15)$$

$$\mathbf{h}_c^{1\top} K_c^{-\top} K_c^{-1} \mathbf{h}_c^{1\top} = \mathbf{h}_c^{2\top} K_c^{-\top} K_c^{-1} \mathbf{h}_c^{2\top} \quad (6.16)$$

Pour estimer les paramètres intrinsèques qui constituent la matrice de calibrage K_c à partir de la matrice H_c , une nouvelle matrice B définie par l'équation (6.17) est introduite :

$$B = K_c^{-\top} K_c^{-1} \quad (6.17)$$

La matrice B étant symétrique, un vecteur de six éléments \mathbf{b} contenant les différentes valeurs de B est utilisé :

$$\mathbf{b} = \begin{bmatrix} B_{11} & B_{12} & B_{22} & B_{13} & B_{23} & B_{33} \end{bmatrix}^\top$$

Avec ces notations, il est possible d'écrire l'égalité suivante :

$$\mathbf{h}_c^i{}^\top B \mathbf{h}_c^j{}^\top = \mathbf{v}_{ij}^\top \mathbf{b} \quad \forall i, j \in \{1, 2, 3\} \quad (6.18)$$

Avec :

$$\mathbf{v}_{ij} = \begin{bmatrix} h_{i1}h_{j1} & h_{i1}h_{j2} + h_{i2}h_{j1} & h_{i2}h_{j2} & h_{i3}h_{j1} + h_{i1}h_{j3} & h_{i3}h_{j2} + h_{i2}h_{j3} & h_{i3}h_{j3} \end{bmatrix}^\top$$

Grâce à cette écriture, les contraintes (6.15) et (6.16) peuvent être mises sous la forme de la relation matricielle suivante :

$$\begin{bmatrix} \mathbf{v}_{12}^\top \\ (\mathbf{v}_{12} - \mathbf{v}_{22})^\top \end{bmatrix} \mathbf{b} = \mathbf{V}\mathbf{b} = 0 \quad (6.19)$$

Les valeurs des six paramètres intrinsèques ($u_0, v_0, \lambda, \alpha_u, \alpha_v$ et γ) sont déterminées en résolvant l'équation (6.19). Afin d'obtenir au moins six équations différentes, il est nécessaire de réaliser des correspondances de points ($\mathbf{x}_i \Leftrightarrow P_i$) sur au moins trois images de la mire prises depuis la même caméra.

Une solution au vecteur \mathbf{b} est ensuite obtenue de la même manière que pour l'estimation de $\tilde{\mathcal{H}}$ grâce à la décomposition de la matrice \mathbf{V} en valeurs singulières (SVD). Le vecteur obtenu permet d'obtenir la matrice B à un facteur d'échelle λ près. Il est possible d'estimer les paramètres intrinsèques situés dans la matrice K_c à partir de la matrice B en utilisant les relations suivantes [173] :

$$\begin{aligned} v_0 &= \frac{B_{12}B_{13} - B_{11}B_{23}}{B_{11}B_{22} - B_{12}^2} \\ \lambda &= B_{33} - \frac{B_{13}^2 + v_0(B_{12}B_{13} - B_{11}B_{23})}{B_{11}} \\ \alpha_u &= \sqrt{\frac{\lambda}{B_{11}}} \\ \alpha_v &= \sqrt{\frac{\lambda B_{11}}{B_{11}B_{22} - B_{12}^2}} \\ \gamma &= -\frac{B_{12}\alpha_u^2\alpha_v}{\lambda} \\ u_0 &= \frac{\gamma v_0}{\alpha_v} - \frac{B_{13}\alpha_u^2}{\lambda} \end{aligned} \quad (6.20)$$

6.2.3 Système stéréoscopique

Un système de stéréovision est composé de deux caméras définies chacune par leur paramètres intrinsèques. La figure 6.2 présente un schéma d'un tel système avec O_g et O_d les centres optiques des deux caméras, $\overrightarrow{O_g z_g}$ et $\overrightarrow{O_d z_d}$ leur axe optique et P un point de l'espace qui se projette en \mathbf{x}_g dans le plan caméra de gauche et en \mathbf{x}_d dans le plan caméra de droite. On notera X_g et X_d les coordonnées tridimensionnelles des projections du point P exprimées dans le repère caméra.

La relation entre les repères des deux caméras est définie par une transformation homogène. Dans ce document, cette transformation est fixée comme étant celle qui permet de passer du

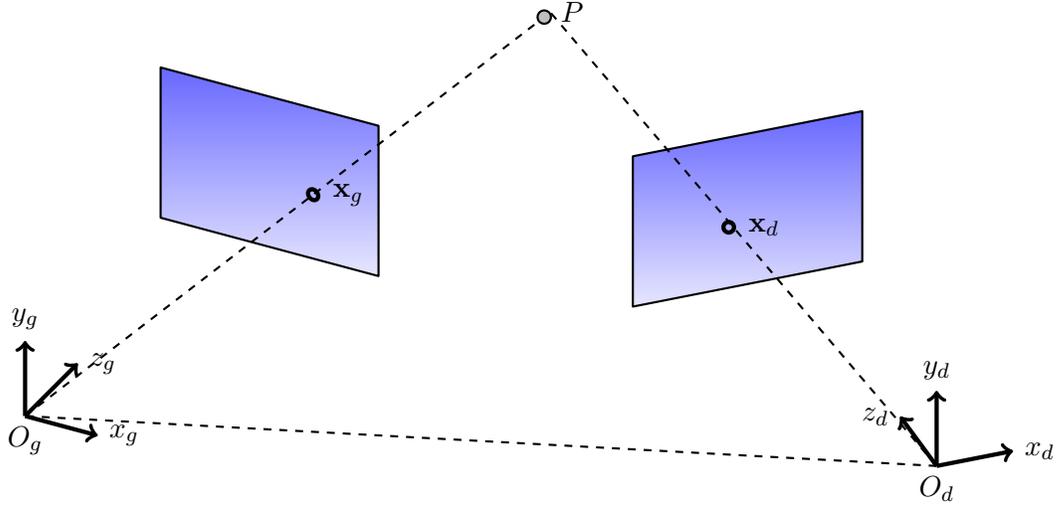


FIGURE 6.2 : Système de stéréovision

repère de la caméra droite à celui de la caméra gauche. Elle est représentée par une matrice homogène M de dimension 4×4 définie comme suit :

$$M = \begin{bmatrix} R & \mathbf{t} \\ 0 & 1 \end{bmatrix} \quad (6.21)$$

Où $\mathbf{t} = [t_1 \ t_2 \ t_3]^\top$ est un vecteur représentant la translation entre les deux centres optiques et R est une matrice de rotation 3×3 représentant la rotation à appliquer au repère de la caméra droite pour l'orienter comme celui de la caméra gauche. \mathbf{t} et R sont appelés paramètres extrinsèques.

On note $[\mathbf{t}]_\times$ la matrice antisymétrique de \mathbf{t} définie par :

$$[\mathbf{t}]_\times = \begin{bmatrix} 0 & -t_3 & t_2 \\ t_3 & 0 & -t_1 \\ -t_2 & t_1 & 0 \end{bmatrix} \quad (6.22)$$

La relation qui lie les coordonnées (X_g) aux coordonnées (X_d) est donnée par :

$$X_g = MX_d \quad (6.23)$$

La géométrie épipolaire est un modèle mathématique de géométrie qui décrit les relations géométriques d'une paire stéréoscopique. La matrice fondamentale, notée F , contient toute l'information nécessaire sur cette géométrie. Elle permet de relier les coordonnées \mathbf{x}_g aux

coordonnées \mathbf{x}_d selon la relation dite « de contrainte épipolaire » qui s'exprime de la façon suivante :

$$\mathbf{x}_d^\top F \mathbf{x}_g = 0 \quad (6.24)$$

En considérant les matrices de calibrage K_g et K_d des caméras gauche et droite, la matrice fondamentale peut se mettre sous la forme [129] :

$$F = K_d^{-\top} [\mathbf{t}]_{\times} R K_g^{-1} \quad (6.25)$$

D'autre part, à tout point de l'image de gauche \mathbf{x}_g correspond un ensemble de possibilités de projection dans l'image de droite \mathbf{x}_d en fonction de la position du point de l'espace P (voir figure 6.3). La droite épipolaire notée l_d représente cet ensemble et l'on a :

$$l_d = F \mathbf{x}_g \quad (6.26)$$

La droite épipolaire l_g d'un point \mathbf{x}_d dans l'image gauche est définie ainsi :

$$l_g = F^\top \mathbf{x}_d \quad (6.27)$$

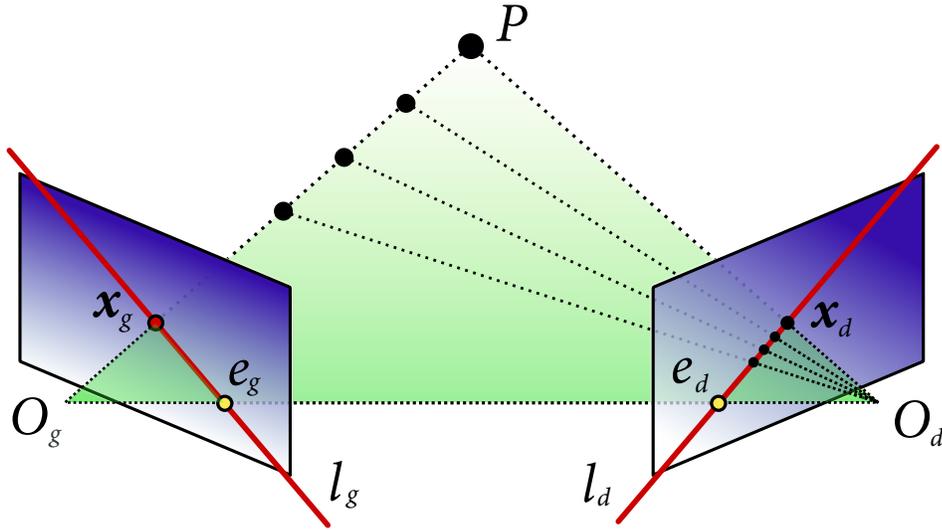


FIGURE 6.3 : Géométrie épipolaire.

L'épipôle d'un plan image associé à une caméra (noté e_g pour la caméra gauche et e_d pour la caméra droite) est le point d'intersection de toutes les lignes épipolaires de ce plan. Les épipôles e_g et e_d vérifient la propriété suivante [129] :

$$Fe_g = e_d^\top F = 0 \quad (6.28)$$

En considérant les relations suivantes :

$$\mathbf{x}'_g = K_g \mathbf{x}_g \mathbf{x}'_d{}^\top = \mathbf{x}_d{}^\top K_d^{-\top} \quad (6.29)$$

La contrainte épipolaire se simplifie de la manière suivante :

$$\mathbf{x}'_d{}^\top E \mathbf{x}'_g = 0 \quad (6.30)$$

Avec $E = [\mathbf{t}]_\times R$ appelée matrice essentielle.

En conclusion, la calibration d'un système de stéréovision permet d'estimer la matrice R et le vecteur \mathbf{t} à partir desquels la matrice essentielle E est construite. Cette matrice lie les coordonnées images des deux pixels issus du même point de l'espace P et situés dans les deux images du système de stéréovision.

6.2.4 Estimation de la matrice essentielle

L'estimation de la matrice essentielle se fait principalement selon deux méthodes. La première, développée par Longuet-Higgins [174] calcule cette matrice à partir de huit correspondances de points dans les images stéréoscopiques. La seconde méthode, développée par Nistér [175], permet d'estimer la matrice essentielle à partir de seulement cinq correspondances. Cette partie décrit la solution proposée par Nistér.

L'équation (6.30) est réécrite de la façon suivante :

$$\tilde{\mathbf{x}}^\top \tilde{E} = 0 \quad (6.31)$$

Avec :

$$\tilde{\mathbf{x}} = \begin{bmatrix} u_g u_d & v_g u_d & u_d & u_g v_d & v_g v_d & v_d & u_g & v_g & 1 \end{bmatrix}^\top \quad (6.32)$$

$$\tilde{E} = \begin{bmatrix} E_{11} & E_{12} & E_{13} & E_{21} & E_{22} & E_{23} & E_{31} & E_{32} & E_{33} \end{bmatrix}^\top \quad (6.33)$$

À partir des cinq paires de points appariés, cinq vecteurs $\tilde{\mathbf{x}}$ sont construits et concaténés pour créer une matrice de dimension 5×9 . Cette dernière est ensuite décomposée sous la forme du produit d'une matrice orthogonale et d'une matrice triangulaire supérieure (décomposition

QR). Les quatre dernières lignes de la matrice orthogonale sont remodelées en quatre matrices X , Y , Z et W de dimension 3×3 .

La matrice essentielle peut s'exprimer à partir de ces quatre matrices de la façon suivante :

$$E = xX + yY + zZ + W \quad (6.34)$$

Avec x , y et z des scalaires à déterminer.

En considérant les contraintes suivantes [129, 176] :

$$\det(E) = 0 \quad (6.35)$$

$$E E^\top E = \frac{1}{2} \text{tr}(E E^\top) E = 0 \quad (6.36)$$

En y insérant l'équation (6.34), et en réalisant une élimination de Gauss-Jordan, un système de dix équations en x , y et z est obtenu. Le réarrangement de ces équations conduit à l'obtention d'une matrice B de dimension 3×3 contenant des polynômes en z . Comme le vecteur $\begin{bmatrix} x & y & 1 \end{bmatrix}^\top$ est un vecteur nul de B , les valeurs qui annulent le déterminant de B (qui est un polynôme de degré dix) sont estimées. On pose :

$$\langle n \rangle \equiv \det(B) \quad (6.37)$$

Après une normalisation de $\langle n \rangle$ de sorte que $n_{10} = 1$, les racines sont obtenues en cherchant les valeurs propres de la matrice compagnon suivante :

$$\begin{bmatrix} -n_9 & -n_8 & \dots & -n_0 \\ 1 & & & \\ & \ddots & & \\ & & & 1 \end{bmatrix} \quad (6.38)$$

Pour chaque racine z , les variables x et y peuvent être obtenues en utilisant le système d'équation défini par B . La matrice essentielle est ensuite obtenue à partir de l'équation (6.34). Les valeurs x et y sont alors calculées à partir des racines z grâce au système d'équations définies dans C . La matrice essentielle E est enfin obtenue grâce à l'équation (6.34).

À partir de la matrice E , il est alors possible de déterminer R et \mathbf{t} en considérant le théorème énoncé dans [129, 171] selon lequel si la matrice E est décomposée en valeurs singulières et que l'on peut écrire $E = U \text{diag}(1,1,0) V^\top$ avec U et V des matrices choisies de sorte que leur déterminant soit strictement positif, alors :

$$\mathbf{t} = \pm t_u \text{ avec } t_u = \begin{bmatrix} u_{13} & u_{23} & u_{33} \end{bmatrix} \quad (6.39)$$

$$R = UDV^\top \text{ ou } R = U D^\top V^\top \quad (6.40)$$

Avec :

$$D = \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (6.41)$$

La matrice M définie dans (6.21) a donc quatre solutions possibles qui satisfont la contrainte épipolaire (6.30), et une de ces solutions correspond à la vraie matrice M . Le fait que les points de la scène doivent se trouver devant les caméras est imposé pour trouver la bonne estimation de M . Une seule correspondance de points est suffisante pour déterminer la solution. Il suffit pour cela de calculer la triangulation tridimensionnelle du point P correspondant en utilisant chacune des quatre solutions possibles et de vérifier que P se trouve dans le champ de vue des deux caméras.

En conclusion, cette méthode permet de calculer la matrice essentielle à partir de cinq points de correspondance. Ce même calcul peut être fait avec un plus grand nombre de points afin d'améliorer la précision de l'estimation de la matrice E .

6.2.5 Reconstruction 3D des points appariés

La reconstruction tridimensionnelle de points appariés entre deux images stéréoscopiques s'effectue par une méthode dite de triangulation [129] dont voici les détails :

Soit M_g la matrice de projection de la caméra gauche et M_d celle de la caméra droite. Les relations suivantes peuvent être écrites $\forall P \in \mathbb{R}^3$:

$$\mathbf{x}_g = M_g P \qquad \mathbf{x}_d = M_d P \quad (6.42)$$

Les équations (6.42) donnent un système de quatre équations. Si \mathbf{x}_g , \mathbf{x}_d , M_g et M_d sont connus, les coordonnées tridimensionnelles (X_p, Y_p, Z_p) du point P sont obtenues par la résolution des quatre équations.

Dans cette thèse, les paramètres intrinsèques et extrinsèques du système de stéréovision sont estimés par calibrage. Les coordonnées tridimensionnelles $P = (X_g, Y_g, Z_g, 1)$ du point P dans le repère de la caméra gauche peuvent être calculées en résolvant les équations (6.42) avec :

$$M_g = K_g \qquad M_d = K_d \begin{bmatrix} R & \mathbf{t} \end{bmatrix} \quad (6.43)$$

K_g et K_d étant les matrices de calibrage des caméras droite et gauche et $\begin{bmatrix} R & \mathbf{t} \end{bmatrix}$ étant la matrice de transformation du repère de la caméra droite vers le repère de la caméra gauche.

Lorsque les caméras ne sont pas calibrées, il est possible de calculer des matrices de projection à partir de la matrice fondamentale F de la façon suivante [129] :

$$M_g = \begin{bmatrix} Id & \mathbf{0} \end{bmatrix} \qquad M_d = \begin{bmatrix} [e_d]_{\times} F & e_d \end{bmatrix} \qquad (6.44)$$

Avec :

- Id : la matrice identité de taille 3×3
- $\mathbf{0} = \begin{bmatrix} 0 & 0 & 0 \end{bmatrix}^T$
- e_d : l'épipôle de l'image droite,
- $[e_d]_{\times}$ la matrice antisymétrique de e_d

6.3 Estimation des paramètres intrinsèques et extrinsèques d'un système de vision

Les paramètres intrinsèques et extrinsèques d'un système de vision sont estimés en réalisant une procédure de calibrage au cours de laquelle l'acquisition d'une série d'images d'une mire (généralement entre dix et vingt) présentant des positions, orientations et inclinaisons différentes est réalisée. La figure 6.4 présente trois exemples de mire dont la plus utilisée qui est un damier, est montrée en (a).

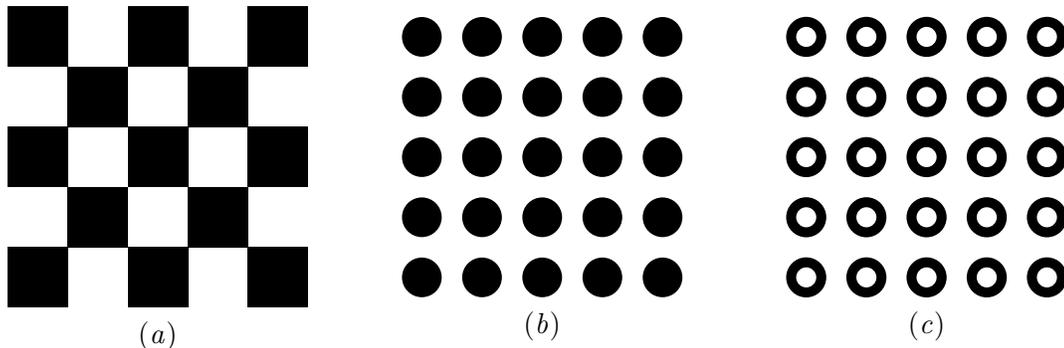


FIGURE 6.4 : Mires de calibrage utilisées dans la littérature : (a) Damier, (b) cercles et (c) anneaux.

Data *et al.* ont présenté dans [177], une étude relative à la précision du calibrage de caméra obtenue pour les différentes mires existantes. Leur conclusion est que la mire formée d'anneaux est celle qui permet d'avoir les meilleures performances. Des essais que nous avons réalisés en extérieur dans des conditions réelles de feux dirigés montrent toutefois que la mire damier est celle qui permet d'obtenir des points d'intérêt avec la plus grande facilité. Dans le cadre de

cette thèse, deux mires d'un mètre de côté ont été utilisées. L'une, en damier, se compose de carrés noirs et blancs de 15 cm de côté. L'autre contient des anneaux de cercle de diamètre extérieur de 5 cm et de diamètre intérieur de 2,5 cm.

Les paramètres intrinsèques et extrinsèques d'un système de stéréovision sont estimés en utilisant les méthodes présentées dans la section précédente. Bouguet [178] a développé sous MATLAB une boîte à outils contenant les programmes de l'ensemble de ces méthodes. Depuis sa version 2014b, `Matlab` propose dans une de ses boîtes à outils intitulée *Computer Vision System* de nouvelles fonctions et applications pour le calibrage de systèmes de vision à partir d'images de damier.

Des améliorations de performance des méthodes présentées dans cette section ont été proposées, notamment par Data *et al.* qui présentent dans [177] une méthode de raffinement d'estimation basée sur un processus itératif dont voici le principe :

Dans une première phase, les paramètres intrinsèques de la caméra sont estimés en utilisant les algorithmes « classiques » implémentés par exemple dans les boîtes à outils citées précédemment.

Dans une deuxième phase, l'estimation des paramètres intrinsèques est améliorée par un processus qui permet une meilleure sélection des points d'intérêt détectés sur l'image de la mire. Ce processus est réalisé plusieurs fois jusqu'à convergence de l'algorithme en suivant les étapes :

- Les images de la mire sont transformées à partir des paramètres intrinsèques estimés afin de supprimer les distorsions et de les projeter dans un plan dit « canonique » qui permet d'obtenir des gradients orthogonaux pour les images des mires (voir figure 6.8). C'est-à-dire que les gradients entre les lignes horizontales et verticales de la mire sont orthogonaux dans cette nouvelle image.
- Les points d'intérêt sont recherchés sur cette image et sont re-projetés sur l'image originale.
- les paramètres intrinsèques sont une nouvelle fois calculés en utilisant les nouveaux points d'intérêt obtenus à partir du plan canonique.

Les résultats présentés dans [177] suggèrent que cette méthode diminue d'un facteur deux l'erreur de calibrage. En règle générale, les résultats convergent dès la deuxième itération. Cette méthode est donc plus longue qu'un calibrage « classique » mais permet d'obtenir des données plus précises et nécessite un temps d'exécution égale à environ deux à trois fois le temps d'exécution des anciennes méthodes.

Les résultats de l'évaluation de performance de cette procédure n'ayant été présentés que sur des objets situés à de courtes distances et pour un système mono-caméra, un travail d'évaluation sur de grandes distances et pour un système de stéréovision a été réalisé dans le cadre de cette thèse en utilisant la salle de métrologie de l'Université Laval. Cette salle a une longueur de plus de 30 mètres et possède un mur constitué de plusieurs cibles dont les

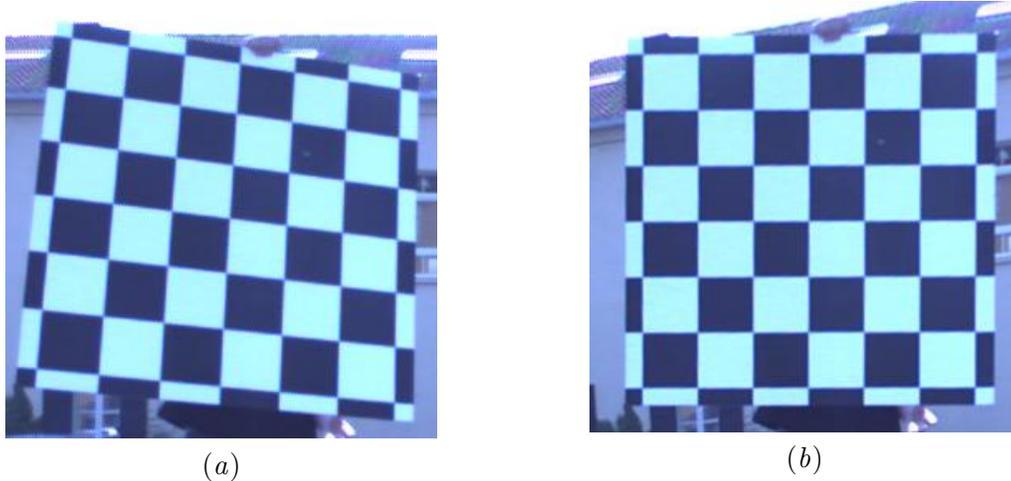


FIGURE 6.5 : Projection canonique d'une image de calibration : (a) image de calibration originale et (b) sa projection canonique faite grâce aux paramètres de la caméra.

positions sont référencées (voir figure 6.6).

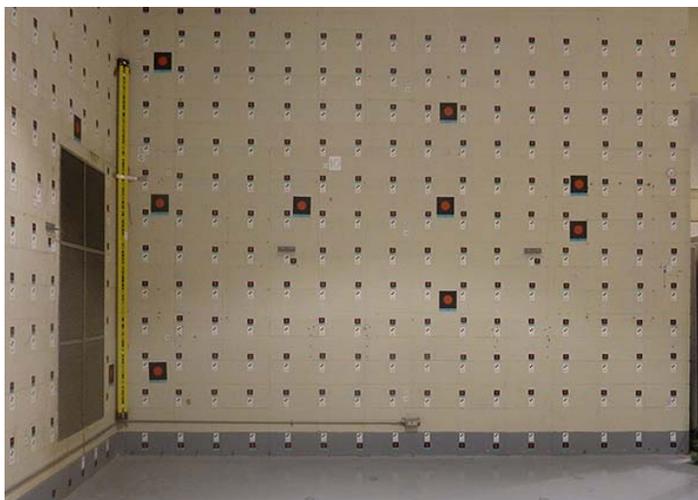


FIGURE 6.6 : Mur de la salle de métrologie. La position relative de chaque point est référencée.

Les erreurs de reconstruction ont été estimées en comparant les distances entre les points réels et celles calculées à partir des points reconstruits. Les tableau 6.1 et tableau 6.2 présentent ces erreurs.

TABEAU 6.1 : Focale et distance inter-caméra mesurées et estimées.

	Focale	Distance inter-caméras
Mesures	8 mm	520 mm
Estimations	8,2 mm	516 mm

Les résultats obtenus confirment que la reconstruction est d'autant moins précise que la dis-

TABLEAU 6.2 : Erreurs de reconstruction à différentes distances.

Erreur à 10 m	Erreur à 20 m	Erreur à 30 m
9 mm	26 mm	102 mm

tance est grande. Les erreurs à 30 mètres sont acceptables dans le contexte de l'étude des feux de végétation. La procédure de calibrage du système stéréoscopique utilisant une méthode itérative décrite dans [177] a donc été validée pour la réalisation de nos travaux.

6.4 Reconstruction tridimensionnelle de points de feu

La figure 6.7 donne une illustration du processus proposé dans cette thèse permettant d'obtenir des points 3D d'un feu à partir d'images stéréoscopiques multimodales. L'ultime traitement de cette méthode est l'obtention de points tridimensionnels de feu calculés par triangulation à partir des points appariés dans les images stéréoscopiques.

La dernière image de cette figure montre un ensemble de points tridimensionnels dont certains ont été colorés en vert parce qu'ils ont été identifiés comme étant erronés. La procédure suivante permet de les supprimer :

- La moyenne (μ_d) et l'écart type (σ_d) de la distance (d_P) entre les points 3D (P) obtenus et le centre du repère tridimensionnel (qui est situé au centre de la caméra gauche du système) ont été calculées
- Parmi l'ensemble des points 3D, ceux vérifiant la condition suivante sont considérés comme erronés et éliminés :

$$|d_P - \mu_d| \geq s * \sigma_d \quad (6.45)$$

Avec s une constante.

$s = 3$ a été fixée à partir du traitement des points de dix feux différents réalisés sur un terrain de 100 m^2 .

6.5 Estimation des caractéristiques géométriques d'un feu

Cette section présente les méthodes qui ont été développées pour l'estimation de caractéristiques de feu (telles que la position au sol du front, sa vitesse, sa largeur, sa hauteur, sa longueur, son inclinaison et sa surface) à partir de points 3D obtenus par stéréovision. Les traitements présentés dans cette partie ont été développés pour des points tridimensionnels obtenus à partir d'images stéréoscopiques d'un système de vision positionné devant le front de feu. Ils correspondent à une amélioration des méthodes présentées dans [179].

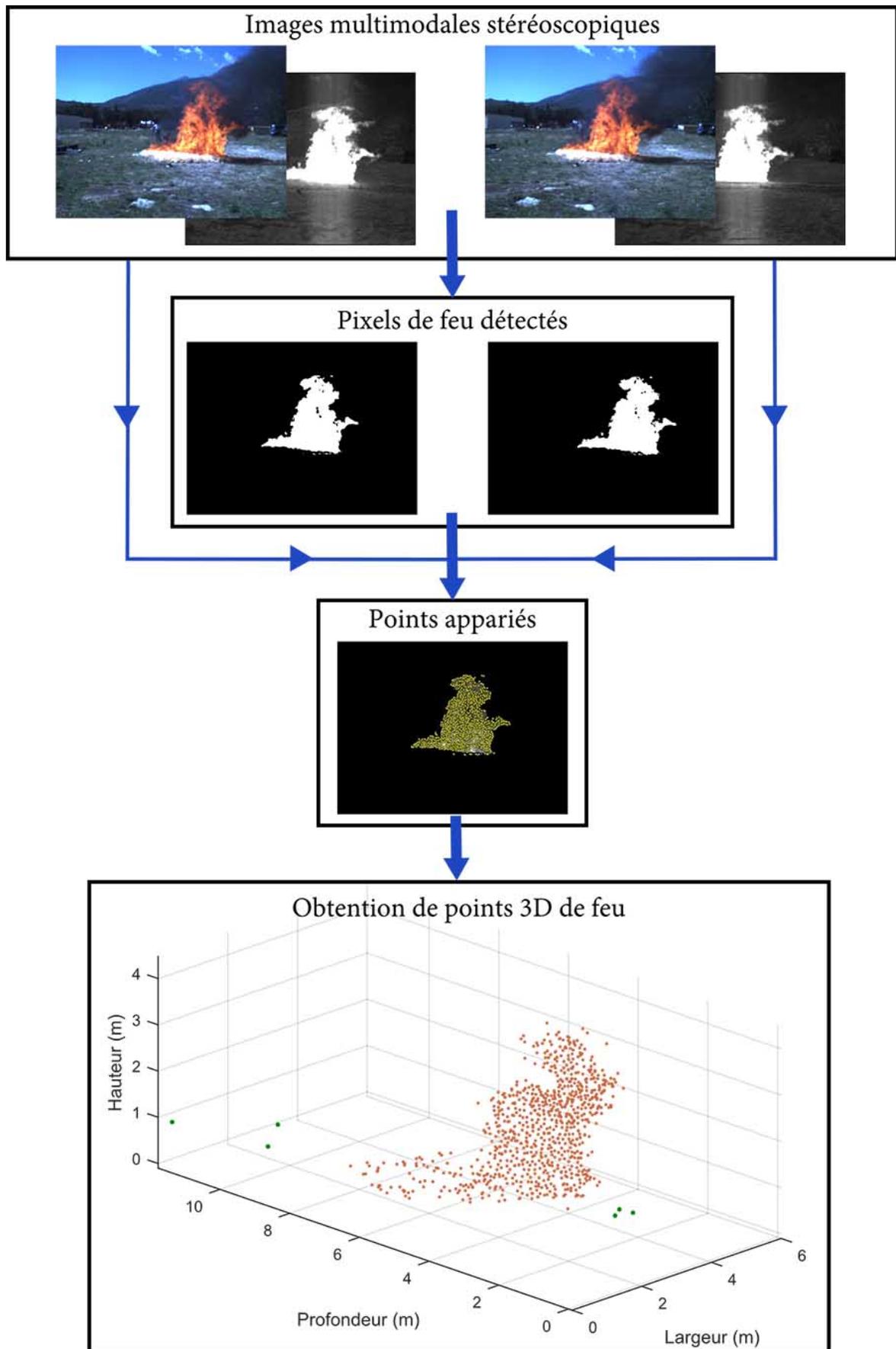


FIGURE 6.7 : Obtention de points 3D d'un feu par stéréovision.

6.5.1 Estimation du plan de base

Les caractéristiques géométriques de feu sont exprimées par rapport à un plan de base qui peut être soit la limite du végétal si le système de vision est positionné devant le front de feu, soit la surface du terrain si le système de vision est positionné à l'arrière du front de feu. La figure 6.8 montre un exemple d'images prises dans ces deux configurations.



FIGURE 6.8 : Vue avant et arrière d'une propagation. (a) vue avant et (b) vue arrière.

Le plan de base est estimé en considérant les points 3D de feu les plus bas calculés à partir d'images successives (par exemple trois cents images correspondant à cinq minutes de propagation) et renouvelées par tranche d'images (par exemple cent-vingt images correspondant à deux minutes de propagation). La sélection des points 3D de feu les plus bas s'effectue de la façon suivante :

Les points appariés de l'image de gauche du système de stéréovision sont triés en fonction de leurs coordonnées afin de pouvoir sélectionner dans chaque colonne de l'image, ceux, s'ils existent, qui ont l'indice de ligne le plus grand (c'est à dire ceux situés au bas de l'image). Parmi tous les points 2D sélectionnés, certains, particulièrement ceux positionnés aux extrémités de la zone de segmentation de feu, sont les plus bas de leur colonne mais ne se situent pas au niveau du plan de base. Pour les éliminer, la moyenne (μ_p) et l'écart type (σ_p) des indices de ligne (v_p) des points retenus sont calculées et les points situés à une certaine distance de la moyenne de position de ligne μ_p sont éliminés. Seuls les points vérifiant la condition suivante sont retenus :

$$v_p > \mu_p - 2\sigma_p \quad (6.46)$$

La figure 6.9 présente les points de sol d'une zone de feu identifiés parmi les points appariés.

L'équation mathématique du plan de base est déterminée en minimisation la somme des distances quadratiques entre ce plan et les points 3D retenus. La méthode utilisée est la suivante :

- dans une première étape, un point 3D moyen (P_M) est calculé en effectuant une moyenne

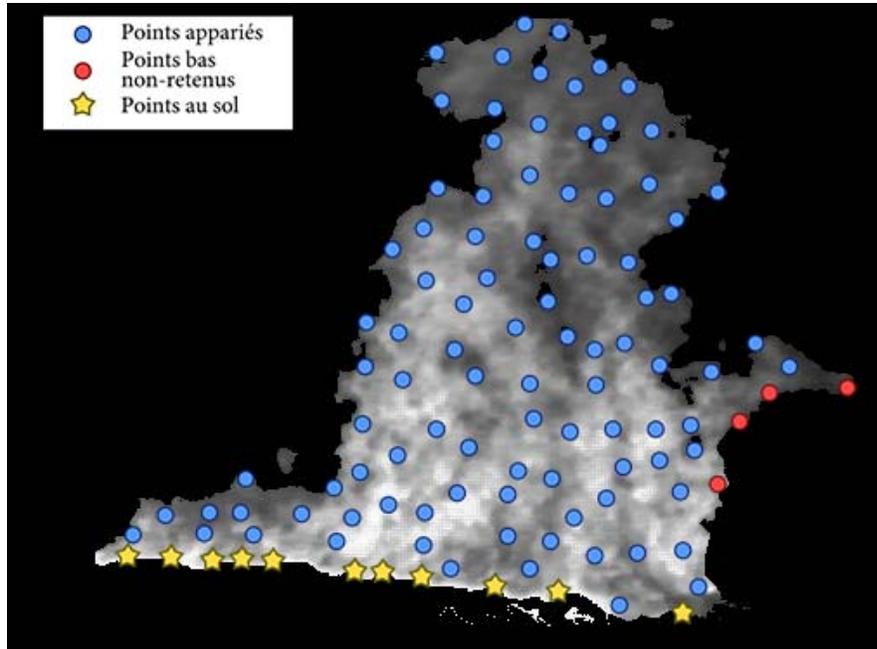


FIGURE 6.9 : Points appariés d'un feu détectés comme étant au sol.

sur l'ensemble des n points bas du feu,

- dans une deuxième étape, une matrice G de dimension $3 \times n$ est construite de sorte que chaque ligne contienne les coordonnées d'un point bas auquel a été soustrait les coordonnées du point moyen,
- Dans une troisième étape, les vecteurs propres droits v_G de la matrice $G^T G$ vérifiant l'équation (6.47) sont estimés.

$$G^T G v_G = \lambda_G v_G \quad (6.47)$$

Avec λ_G la valeur propre de v_G .

Comme $G^T G$ est une matrice 3×3 symétrique, les deuxième et troisième vecteurs propres droits sont orthogonaux et représentent des vecteurs directeurs du plan recherché. Le plan de base est donc déterminé à partir de ces deux vecteurs et du point moyen (P_M).

La figure 6.10 présente le plan de base estimé en utilisant des points 3D bas identifiés à partir d'images acquises par un système de vision positionné à l'arrière du front de feu.

6.5.2 Calcul des caractéristiques géométriques

6.5.2.a Transformation des points 3D

Les caractéristiques géométriques d'un feu doivent être estimées par rapport au plan de base. Ceci nécessite d'appliquer aux points 3D de feu une transformation afin de ramener à zéro la

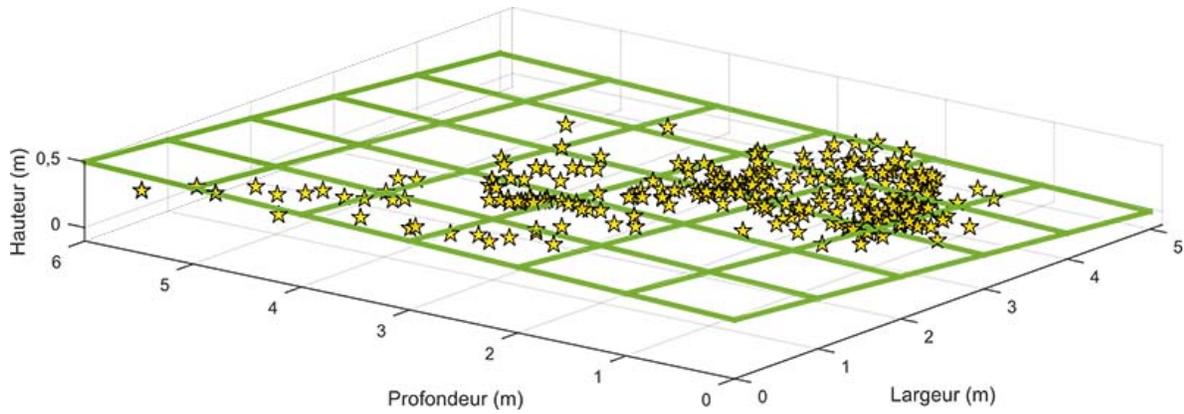


FIGURE 6.10 : Estimation d'un plan de propagation à partir des points 3D bas d'un feu.

hauteur des points les plus bas. Cette transformation se compose d'une translation, qui place le point moyen P_M à $(0 \ 0 \ 0)$, et d'une rotation. La rotation est estimée grâce à l'analyse en composantes principales d'une matrice de dimension 3×3 contenant trois points du plan de base générant un point moyen égal à $(0 \ 0 \ 0)$. La matrice $R_P = V\sqrt{\Lambda}$ transforme les points du plan base de sorte que leur troisième coordonnées (hauteur) soit nulle.

6.5.2.b Position de la ligne de front

La ligne de front est la position du front de feu sur le plan de base. Elle est obtenue en projetant les points 3D du bas du front de feu sur le plan de base et en les interpolant par une courbe de Bézier [180]. La figure 6.11 donne un exemple de ligne de front ainsi obtenue.

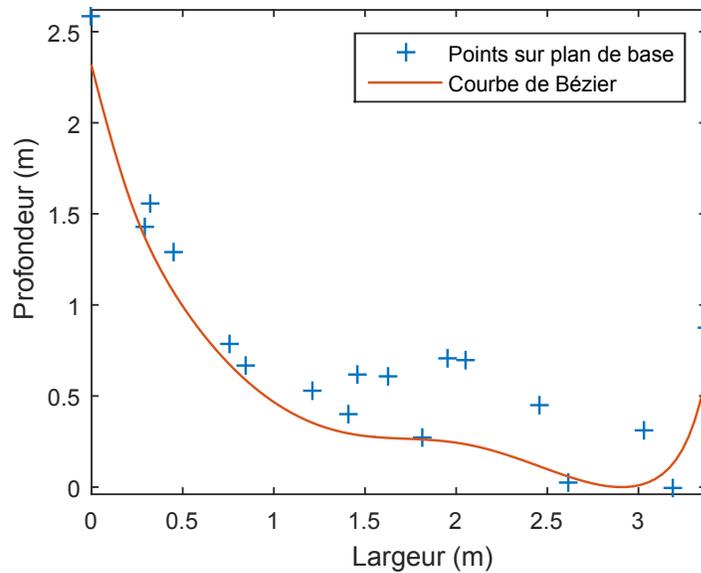


FIGURE 6.11 : Position du front de feu sur le plan de base et courbe de Bézier associée.

6.5.2.c Vitesse des points de la ligne de front

Les courbes de position du front de feu sur le plan de base obtenues à différents instants sont utilisées pour estimer la vitesse des points du front. Cette vitesse est le rapport de la distance qui sépare deux points équivalents situés sur deux fronts successifs divisée par l'intervalle de temps entre les deux instants d'acquisition des images à partir desquelles les courbes ont été calculées. Pour calculer cette grandeur, les courbes de position sont traitées deux à deux. La normale à la première courbe au point considéré coupe la courbe suivante en un point dit « équivalent ». La figure 6.12 présente les normales estimées pour trois points du front avec les points équivalent du front suivant. Dans cette figure, un temps de cinq secondes a été choisi pour une meilleure visualisation. Les vitesses obtenues à partir des distances estimées sont affichées pour les trois points considérés.

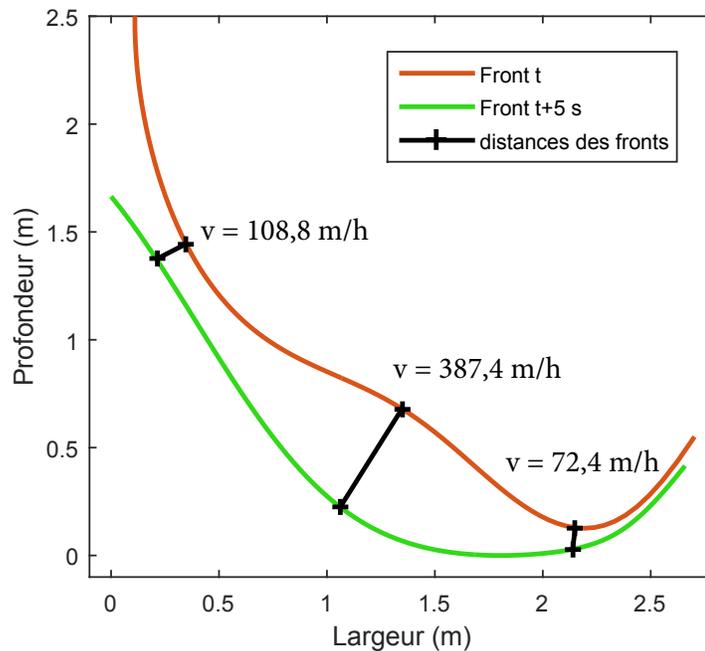


FIGURE 6.12 : Estimation des vitesses de trois points au sol.

Les physiciens de l'UMR CNRS 6134 SPE - Université de Corse qui travaillent sur la modélisation de la propagation des feux de forêt sont intéressés par la vitesse de certains points caractéristiques de la ligne de front de feu. Ces points sont aisés à identifier sur une courbe numérique mais parfois compliqués à positionner sur une courbe réelle par un non-spécialiste. Aussi, dans le cadre de cette thèse, une interface a été développée afin de donner la possibilité à un utilisateur de positionner par clic les points dont il veut l'estimation de vitesse.

6.5.2.d Direction principale

La direction principale prise par un feu au cours de sa propagation correspond à l'axe moyen selon lequel il se propage. L'estimation de cette direction est réalisée à partir des courbes de

Bézier représentant les lignes de front successives telles que celles présentées en figure 6.13. Le nombre de ces courbes considérées dans le cadre de cette thèse est de trois cents (correspondant à cinq minutes de propagation) et le nombre de courbes renouvelées par tranche est de cent-vingt (correspondant à deux minutes de propagation) mais ces nombres sont modifiables.

La procédure d'estimation de la direction principale est la suivante : Les courbes de Bézier sont considérées deux à deux et sur chacune d'elle cent points équirépartis sont numérotés. Tous les vecteurs reliant les points de même numéro des paires de courbe sont ensuite calculés et la moyenne de ces vecteurs donne la direction principale de propagation.

La figure 6.13 présente la direction principale estimée d'une propagation de feu.

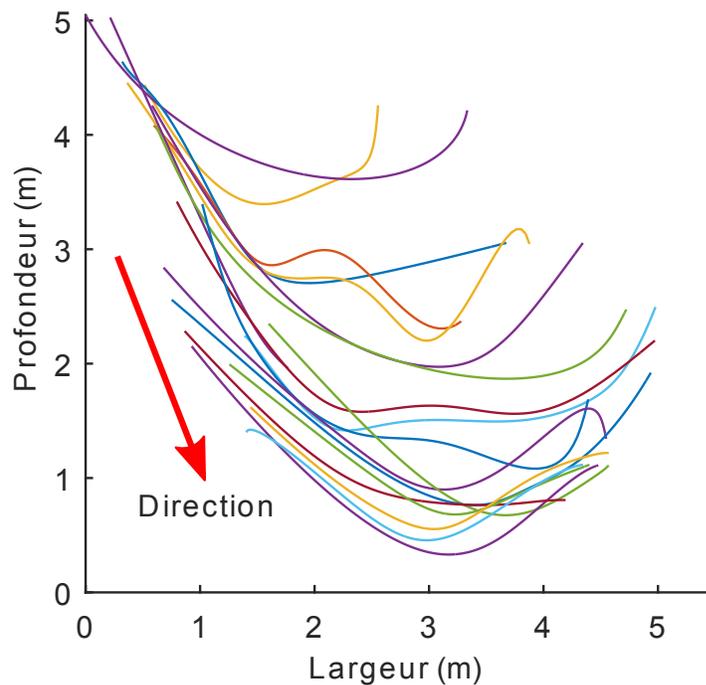


FIGURE 6.13 : Estimation de la direction de propagation. Les courbes représentent les différentes positions au sol estimées au cours de la propagation et la flèche rouge est la direction calculée à partir de ces positions.

6.5.2.e Hauteur, largeur, longueur et inclinaison

Les mesures de la hauteur, largeur, longueur et inclinaison du front de feu sont plus aisées lorsque la direction principale du feu est colinéaire à l'axe de profondeur. Aussi une rotation autour de l'axe de hauteur avec un angle égale à l'angle entre le vecteur de direction principale du front et le vecteur de profondeur est appliquée à l'ensemble des points 3D (ce qui correspond à une deuxième transformation pour ces points).

La hauteur du front de feu correspond à la hauteur du point 3D ayant la composante de hauteur maximale. La figure 6.14 présente la visualisation de la hauteur d'un front de feu.

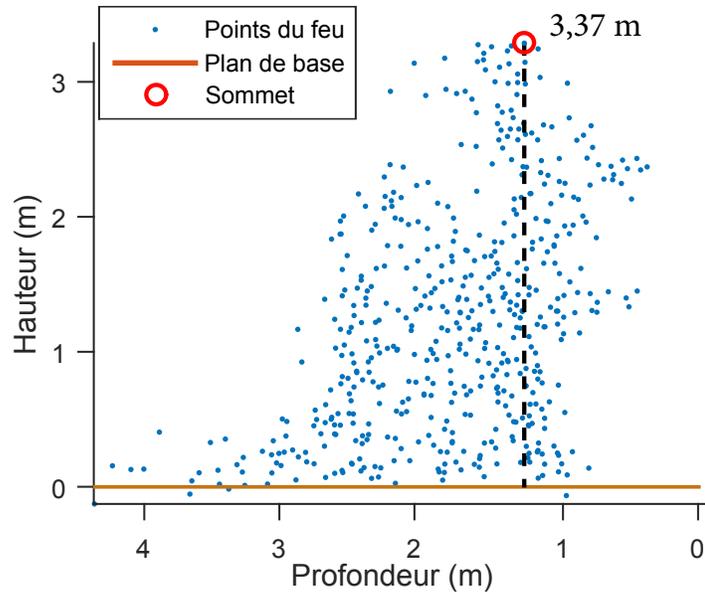


FIGURE 6.14 : Estimation automatique de la hauteur d'un front de feu.

La largeur du front de feu est la différence de coordonnées de largeur des points situés aux deux extrémités de la courbe de position du front. La figure 6.15 visualise cette distance pour un exemple de feu réel.

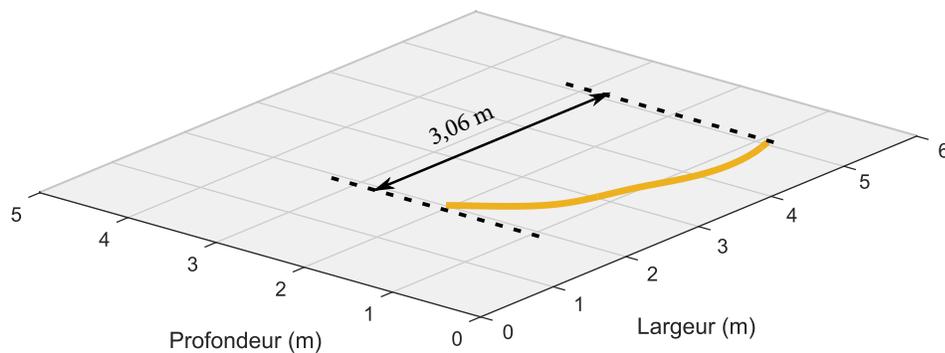


FIGURE 6.15 : Estimation automatique de la largeur d'un front de feu.

La longueur du front de feu est la distance entre le point le plus haut du front de feu et le point le plus avancé sur la ligne de front. L'angle d'inclinaison représente l'angle compris entre la verticale au sol et le segment de longueur de flamme. La figure 6.16 montre un exemple de détermination de ces mesures et présente avec un triangle vert le point le plus avancé de la ligne de front, avec un rond rouge le point le plus haut du front de feu, la longueur du front est matérialisée par un trait fuchsia et la verticale au sol par un trait noir.

Les points détectés automatiquement correspondant au point le plus haut du front de feu et au point le plus avancé de la ligne de front ne sont pas toujours ceux qu'un expert aurait choisis compte tenu de configurations particulières comme une partie de flamme qui se détache du

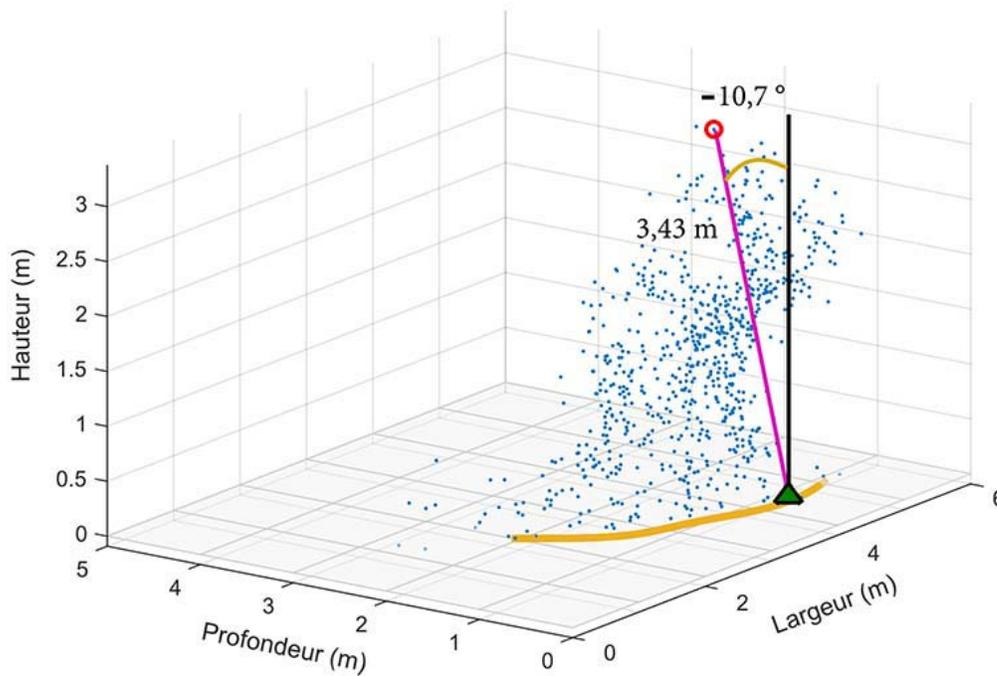


FIGURE 6.16 : Estimation automatique de la longueur et de l'inclinaison d'un front de feu.

front par exemple. Aussi, une interface graphique a été développée afin de laisser la possibilité à un utilisateur de sélectionner des points qui correspondent. La figure 6.17 montre cette interface qui permet la saisie des points et donne les valeurs de hauteur, longueur et inclinaison calculées à partir de ces derniers.

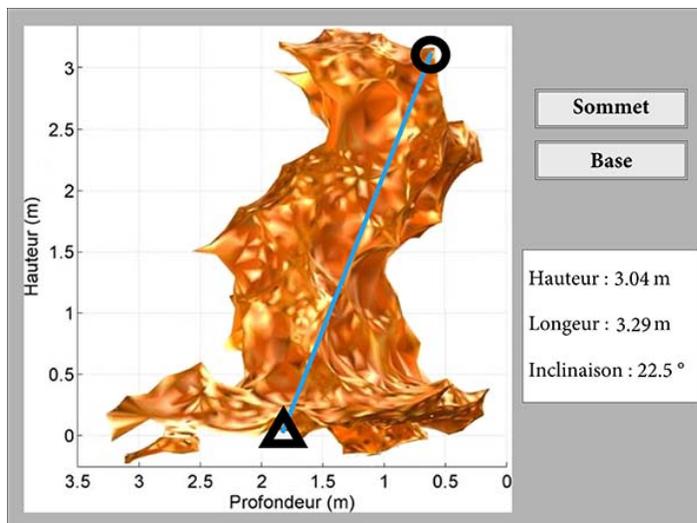


FIGURE 6.17 : Interface graphique permettant la saisie de points nécessaires au calcul de la hauteur, longueur et inclinaison d'un front de feu.

6.5.2.f Profil de la hauteur du front

Le profil de la hauteur du front de feu est obtenu de la façon suivante : les points 3D de feu sont projetés dans le plan (largeur-hauteur) et une triangulation de Delaunay [181] est appliquée afin d'obtenir une surface fermée et d'en déterminer son contour. Les points de ce contour considérés comme faisant partie du haut de la flamme sont situés sur la portion de ligne reliant les extrémités gauche et droite des points du feu passant par le point correspondant au sommet de la flamme.

La figure 6.18 donne le profil de la hauteur du front de feu présenté dans la figure 6.9.

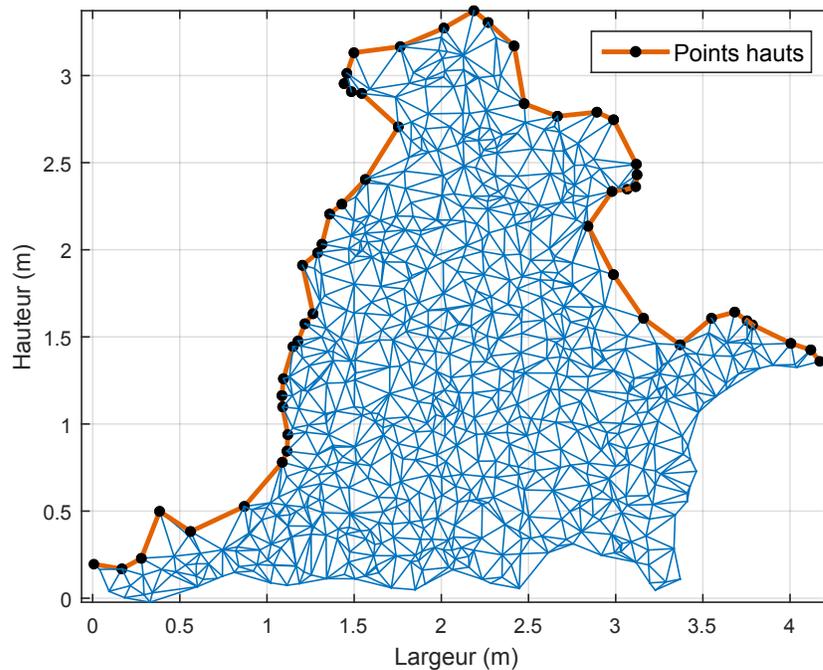


FIGURE 6.18 : Triangulation de Delaunay appliquée aux points du feu permettant la détermination du profil de hauteur du front.

6.5.2.g Surface

La surface d'un front de feu est l'aire de sa partie avant du front. Une triangulation de Delaunay appliquée aux points 3D de feu permet d'obtenir une surface constituée de triangles. Ceux qui ne sont pas masqués par d'autres triangles et qui sont orientés vers le système de vision constituent la partie avant du front à partir desquels le calcul d'aire est effectué. Ces derniers présentent la propriété suivante : la demi-droite colinéaire à l'axe de la profondeur et dont l'origine est le centre du triangle n'intercepte aucun autre triangle.

La figure 6.19 présente les triangles retenus pour la triangulation pour l'image de feu présentée dans la figure 6.9. On voit que la partie arrière du feu n'est pas prise en compte dans les calculs. La somme des aires des triangles affichés est de 21 m^2 .

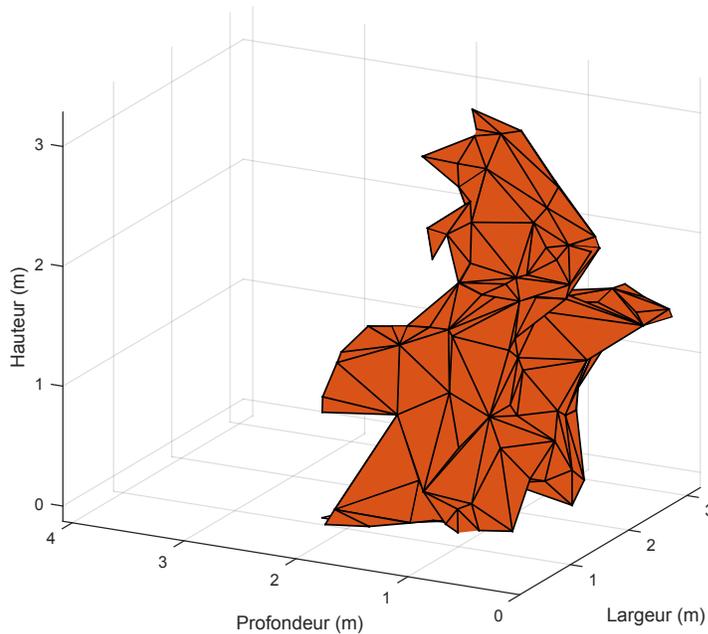


FIGURE 6.19 : Estimation de la surface d'un front de feu.

6.6 Conclusion

Ce chapitre a présenté dans une première partie les méthodes permettant l'estimation des caractéristiques géométriques d'un front de feu à partir d'images stéréoscopiques. Les principes théoriques de la reconstruction tridimensionnelle ont été décrits dans une deuxième partie. Une troisième partie a été dédiée aux méthodes d'estimation des paramètres intrinsèques et extrinsèques d'un système de stéréovision. Une technique itérative permettant l'amélioration des estimations a été présentée ainsi que son évaluation sur de grandes distances. La reconstruction tridimensionnelle de points de feu a été développée dans une cinquième partie. La dernière partie de ce chapitre a été dédiée aux méthodes permettant l'estimation des caractéristiques géométriques de front feu à partir d'un nuage de points 3D.

Les méthodes d'estimation qui ont été présentées ont été développées pour travailler à partir d'un nuage de points obtenu depuis un seul système de stéréovision. Cette configuration ne permet pas d'obtenir certaines caractéristiques géométriques telles que l'épaisseur du front ou son volume car la vision du feu est partielle et génère des erreurs de précision d'estimation lorsque la distance entre le feu et le système de vision est grande. L'utilisation de plusieurs systèmes de stéréovision permet de résoudre ces problèmes.

Chapitre 7

Estimation de caractéristiques géométriques de fronts de feux à partir de plusieurs systèmes de stéréovision

7.1	Introduction	199
7.2	Recalage de points tridimensionnels obtenus depuis différents systèmes de vision	200
7.2.1	Matériel utilisé	200
7.2.2	Détermination de la matrice de passage entre les repères des différents systèmes de stéréovision	202
7.3	Reconstruction tridimensionnelle de la forme de feu	208
7.4	Calcul de caractéristiques géométriques	210
7.4.1	Périmètre du feu	211
7.4.2	Surface du végétal participant à la combustion	211
7.4.3	Direction principale de propagation	212
7.4.4	Lignes de front	212
7.4.5	Épaisseur	212
7.4.6	Volume	214
7.5	Incertitude de mesure	214
7.6	Conclusion	216

7.1 Introduction

Dans le chapitre précédent, les possibilités d'estimation de caractéristiques géométriques d'un feu à partir d'images stéréoscopiques ont été montrées. Il a été souligné que des images acquises uniquement depuis un seul point de vue donnent une vision incomplète du feu et ne permettent pas de mesurer certaines caractéristiques. Ce chapitre présente l'estimation des caractéristiques telles que l'aire du végétal participant à la combustion, le périmètre, l'épaisseur et le volume d'un feu à partir d'un dispositif composé de plusieurs systèmes de stéréovision.

La section 7.2 présente le dispositif de vision complet et la méthode de recalage utilisée afin de projeter les points 3D issus de différents systèmes de stéréovision dans un repère commun.

La section 7.4 détaille la méthode d'exploitation des points tridimensionnels issus de différents systèmes de vision et les calculs d'estimation de nouvelles caractéristiques géométriques du feu.

Enfin, la section 7.5 présente le contexte expérimental de prise d'images à partir de deux systèmes de stéréovision réalisées sur deux objets en mouvement de forme constante et de dimensions connues et les résultats d'estimation de leurs caractéristiques.

7.2 Recalage de points tridimensionnels obtenus depuis différents systèmes de vision

Les points tridimensionnels d'un objet obtenus depuis un système de stéréovision sont exprimés dans un repère positionné dans le dispositif de vision. Afin d'exploiter simultanément les points issus de différents systèmes de stéréovision, il est nécessaire de les projeter tous dans un repère commun, c'est ce qui est réalisé par un recalage.

Dans la littérature il existe plusieurs méthodes de recalage de points tridimensionnels. L'algorithme itératif du point le plus proche (en anglais, *Iterative Closest Points (ICP)*) [104] est le plus utilisé. Il cherche à minimiser la distance entre les points de deux nuages tridimensionnels issus de deux vues. Ses performances sont d'autant meilleures que le nombre de points communs aux deux vues est grand. Dans le cadre de cette thèse, seuls deux systèmes de stéréovision sont employés et, afin d'estimer le plus grand nombre possible de caractéristiques, ils sont positionnés de façon à obtenir des vues complémentaires du feu, ce qui ne permet pas d'avoir beaucoup de points en commun issus des différents systèmes. Cette méthode ne peut donc pas s'appliquer dans le contexte de cette thèse. Nous proposons de recalculer les points tridimensionnels à partir des informations de position, d'orientation et d'inclinaison de chaque système de vision. La première partie de cette section décrit le matériel utilisé pour obtenir ces données. La seconde partie détaille les méthodes développées qui optimisent le recalage des points.

7.2.1 Matériel utilisé

Deux systèmes de stéréovision identiques à celui décrit précédemment sont utilisés. Ils sont positionnés en fonction de la topologie du terrain et des possibilités de placement. Dans la mesure du possible, un système est placé en position frontale par rapport à la direction supposée de propagation du feu

Le récepteur GPS Leica Viva CS10, avec le système RTK permettant l'obtention d'une précision centimétrique en temps réel, est utilisé pour déterminer la position de la caméra gauche de chaque système de stéréovision (position du repère des points 3D générés par ce dispositif). La précision d'un tel dispositif dépend de son contexte d'utilisation ; dans le cas de nos expérimentations, elle a varié entre deux et cinq centimètres. Il a été choisi d'obtenir les coordonnées de points GPS dans le repère cartésien global WGS84 ayant pour origine le centre de la terre.

La carte de navigation du drone AR Parrot 2.0 équipée d'une centrale inertielle (en anglais, *inertial measurement unit* (IMU)) est utilisée pour obtenir l'orientation et l'inclinaison des systèmes de vision. La figure 7.1 montre le dispositif qui a été construit et son positionnement sur la caméra de gauche de chaque système. La figure 7.2 donne un schéma présentant les axes associés aux angles de roulis (noté α_i), de tangage (noté β_i), et de lacet (noté γ_i) qui sont utilisés pour définir l'orientation des caméras.

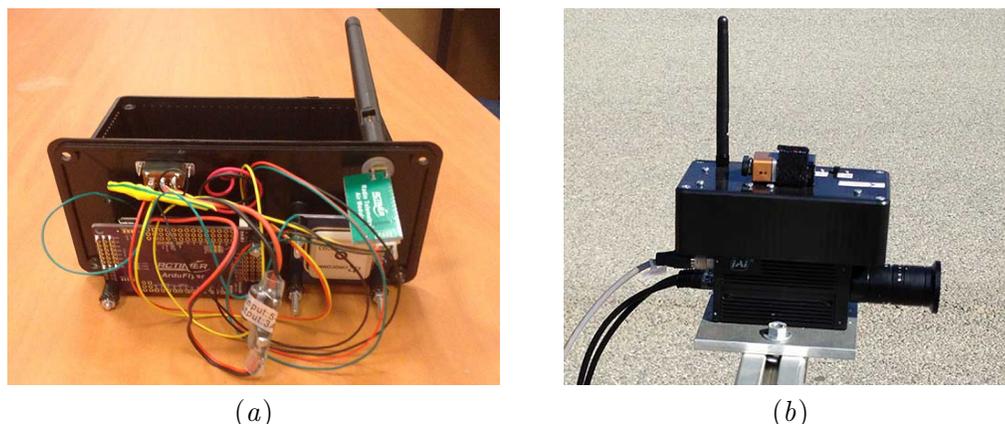


FIGURE 7.1 : Positionnement de la carte de navigation de l'AR Parrot 2.0. (a) Intérieur du dispositif construit et (b) mise en place de l'appareil sur la caméra gauche d'un système de stéréovision.

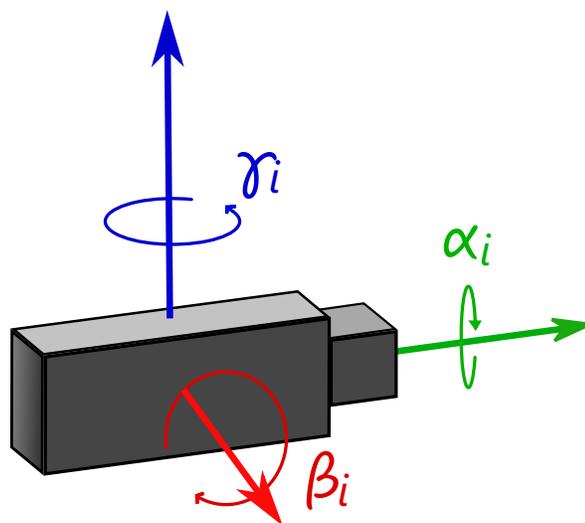


FIGURE 7.2 : Angles de roulis, de tangage et de lacet d'une caméra.

Les données d'orientation et d'inclinaison sont acquises pendant une durée qui est généralement de l'ordre de deux minutes et les moyennes de ces valeurs sont utilisées par la procédure de recalage des points 3D. La figure 7.3 présente un rendu d'acquisition.

Le principal avantage de la carte de navigation du drone AR Parrot 2.0 par rapport à d'autres

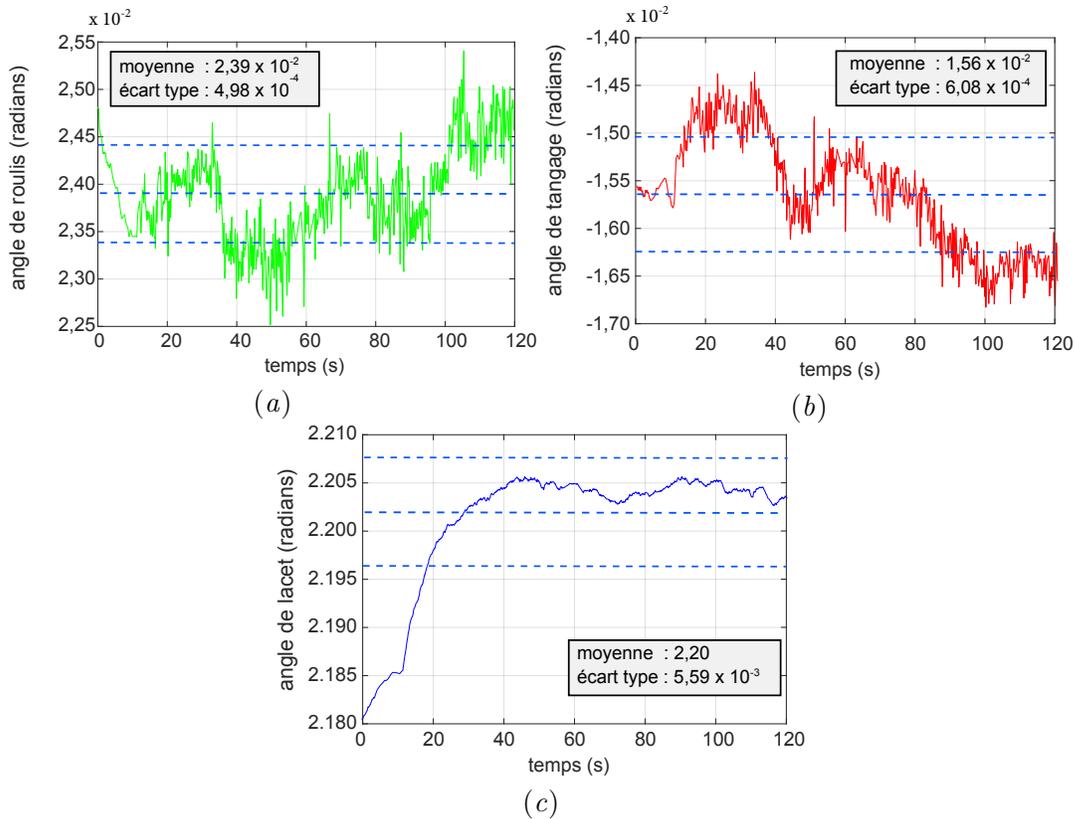


FIGURE 7.3 : Exemple de données acquises par la carte de navigation sur 120 secondes. (a) angles de roulis (b) angles de tangage et (c) angles de lacet. Ces données ont été acquises sur un système immobile

systèmes présents sur le marché qui permettent l’obtention des données d’orientation et d’inclinaison est son prix (qui est de l’ordre de soixante euros). Cela a permis de fixer un dispositif sur chacune des caméras et est compatible avec l’idée de pouvoir fournir des équipements de prix accessibles à des professionnels de la lutte contre les feux.

Les données de position, d’orientation et d’inclinaison des deux systèmes de stéréovision sont acquises pour chaque nouveau contexte expérimental.

7.2.2 Détermination de la matrice de passage entre les repères des différents systèmes de stéréovision

Cette section décrit une méthode de calibrage du dispositif composé de deux systèmes de stéréovision. Cette méthode est réalisée en trois étapes :

Dans une première étape, les points 3D qui sont isolés ou qui se trouvent à une grande distance de l’origine du repère local du système de vision qui la génère sont éliminés afin de ne garder que ceux permettant d’avoir une reconstruction précise. Ceci permet toutefois d’avoir des points tridimensionnels de feu tout au long de la propagation car généralement lorsque la distance entre le feu et un système de vision augmente, elle diminue avec l’autre système de

vision.

Dans une deuxième phase, les données de position, orientation et inclinaison des dispositifs enregistrées lors des expérimentations sont utilisées pour déterminer la matrice de transformation permettant de passer d'un repère local au repère de référence.

Dans une troisième étape, cette matrice est utilisée pour reprojeter les points 3D dans les images et améliorer l'estimation des données en réalisant une optimisation par recherche de motif.

La première partie de la méthode proposée consiste à éliminer les points 3D de feu qui sont incertains. Pour évaluer cette incertitude, une probabilité est associée à chaque point issu d'un système de stéréovision. Cette probabilité tient compte de l'erreur théorique de profondeur et de la distance entre ce point et les autres points tridimensionnels issus du même système.

Pour chaque point 3D, une première probabilité est déterminée à partir de l'erreur théorique de profondeur (Δz_X d'équation (2.2)) ramenée à une valeur comprise dans l'intervalle $[0 ; 1]$ par la fonction r d'équation (4.21). L'erreur théorique est calculée en considérant la profondeur estimée du point, les paramètres intrinsèques et extrinsèques du système de vision et une erreur de disparité égale à un pixel. La fonction r est utilisée avec les valeurs a_{min} et a_{max} fixées respectivement à -300 et 300 . La valeur absolue de ces valeurs est le seuil d'incertitude de position tolérée.

Une deuxième probabilité est calculée à partir de la distance d_X entre ce point et son quatrième point le plus proche. La fonction r avec les valeurs a_{min} et a_{max} fixées respectivement à -500 et 500 est utilisée pour ramener cette valeur dans l'intervalle $[0 ; 1]$. Un point situé à plus de 500 mm d'un autre point est considéré comme isolé.

La probabilité finale P_X associée au point X est :

$$P_X = r(300 - \Delta z_X)r(500 - d_X) \quad (7.1)$$

Seuls les points pour lesquels cette probabilité est supérieure à un seuil empiriquement fixé à $0,4$ sont conservés pour l'estimation des caractéristiques géométriques

La seconde partie de la méthode consiste à déterminer la transformation permettant le passage d'un repère local associé à un système de stéréovision au repère de référence. Cette étape est réalisée de la façon suivante :

Le repère cartésien global WGS84 dans lequel sont exprimés les coordonnées GPS a pour origine le centre de la terre O et ses axes sont f , g et h . Un traitement est réalisé afin d'exprimer les coordonnées des points mesurés dans un repère local de centre P_0 et d'axes N , E et D . La figure 7.4 présente ces deux repères cartésiens.

La transformation qui permet le passage du repère WGS84 au repère local se fait par l'application de deux rotations (R_1 et R_2) suivies d'une translation déterminées de la façon suivante :

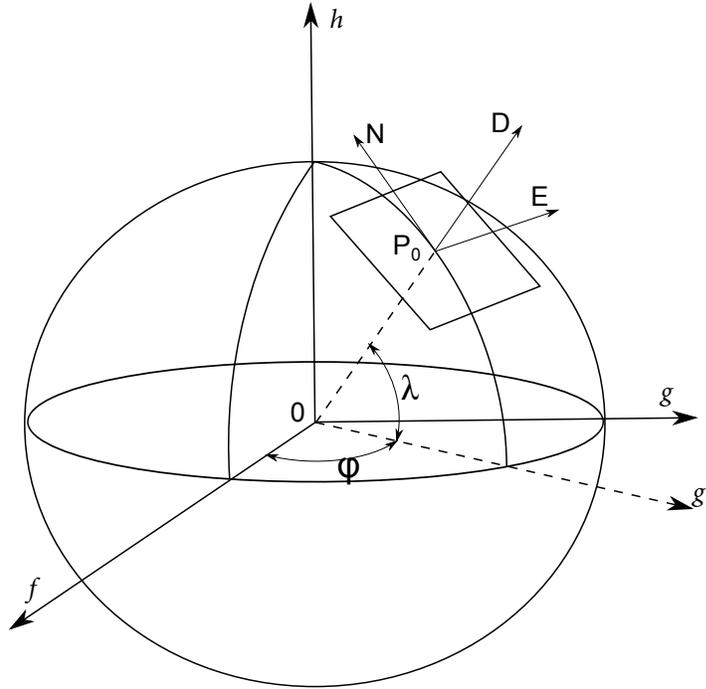


FIGURE 7.4 : Repères de référence. Repère global WGS84 (x, y, z) de centre 0 et repère local (N, E, D) de centre P_0 .

Soit ϕ et λ respectivement les angles de longitude et latitude du point P_0 . Ils sont fournis par le récepteur GPS (il est aussi possible de les calculer à partir des coordonnées WGS84). R_1 est la matrice correspondant à la rotation autour de l'axe z avec un angle ϕ . Cette rotation permet d'obtenir un nouveau repère (f', g', h') . R_2 est la matrice correspondant à la rotation autour de l'axe g' avec un angle λ . La matrice R_3 est définie comme la combinaison des matrices R_1 et R_2 :

$$\begin{aligned}
 R_3 &= R_1 R_2 \\
 &= \begin{pmatrix} \cos(\phi) & -\sin(\phi) & 0 \\ \sin(\phi) & \cos(\phi) & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} -\cos(\lambda) & 0 & -\sin(\lambda) \\ 0 & 1 & 0 \\ \sin(\lambda) & 0 & -\cos(\lambda) \end{pmatrix} \\
 &= \begin{pmatrix} -\cos(\lambda) \cos(\phi) & -\sin(\phi) & -\sin(\lambda) \cos(\phi) \\ -\cos(\lambda) \sin(\phi) & \cos(\phi) & -\sin(\lambda) \sin(\phi) \\ \sin(\lambda) & 0 & -\cos(\lambda) \end{pmatrix}
 \end{aligned}$$

La matrice homogène de passage $P_{WGS84 \rightarrow loc}$ entre le repère WGS84 et le repère local est la composée de la matrice R_3 et de la translation \mathbf{t}_{P_0} de vecteur $\overrightarrow{P_0 D}$. Elle s'écrit :

$$P_{WGS84 \rightarrow loc} = \begin{pmatrix} R_3 & \mathbf{t}_{P_0} \\ 0 & 1 \end{pmatrix}$$

La matrice de transformation, nommée $M_{i \rightarrow loc}$, qui permet le passage d'un repère local associé à un système de stéréovision au repère de référence (associé à l'autre repère de stéréovision) est la combinaison d'une matrice de translation et de trois matrices de rotation dont voici les détails de construction :

Soit N_i , E_i et D_i les coordonnées du système i dans le repère local, alors la matrice de translation T_i est définie par :

$$T_i = \begin{pmatrix} 1 & 0 & 0 & N_i \\ 0 & 1 & 0 & E_i \\ 0 & 0 & 1 & D_i \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Soit α_i , β_i et γ_i respectivement les angles de roulis, tangage et lacet de la caméra gauche du système i par rapport repère local i estimés par la carte de navigation. Soit R_{α_i} la matrice homogène de rotation autour de l'axe y_c , R_{β_i} la matrice homogène de rotation autour de l'axe x_c et R_{γ_i} la matrice homogène de rotation autour de l'axe z_c . Les matrices de rotation se présentent sous la forme :

$$R_{\alpha_i} = \begin{pmatrix} \cos(\alpha_i) & 0 & -\sin(\alpha_i) & 0 \\ 0 & 1 & 0 & 0 \\ \sin(\alpha_i) & 0 & \cos(\alpha_i) & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

$$R_{\beta_i} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos(\beta_i) & \sin(\beta_i) & 0 \\ 0 & -\sin(\beta_i) & \cos(\beta_i) & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

$$R_{\gamma_i} = \begin{pmatrix} \cos(\gamma_i) & \sin(\gamma_i) & 0 & 0 \\ -\sin(\gamma_i) & \cos(\gamma_i) & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

La matrice de transformation $M_{i \rightarrow loc}$ est définie par :

$$M_{i \rightarrow loc} = T_i R_{\alpha_i} R_{\beta_i} R_{\gamma_i} \quad (7.2)$$

La matrice $M_{i \rightarrow loc}$ est employée pour projeter les points 3D obtenus par les deux systèmes de stéréovision dans un repère commun.

Une méthode a été développée pour améliorer le recalage. Elle utilise la reprojection dans les images de feu obtenues par un dispositif de vision de points tridimensionnels issus de l'autre système de stéréovision. La performance du recalage est évaluée en fonction du nombre de points reprojétés qui n'appartiennent pas à la zone de pixels classés « feu » par la méthode de détection. Les cinq meilleures images stéréoscopiques prises au cours d'une propagation de feu qui génèrent un nombre important de points 3D et permettent l'obtention d'une forme 3D compatible avec les images sont considérées pour estimer la performance du recalage.

soit P_i un point 3D obtenu à partir du système de stéréovision i , $M_{i \rightarrow ref}$ la matrice de projection du repère du système i dans le repère de référence défini dans l'équation (7.2), $M_{ref \rightarrow j} = M_{j \rightarrow ref}^{-1}$ la matrice de projection du repère de référence dans le repère du système j , K_g^j la matrice de calibrage de la caméra gauche du système j . Le pixel \mathbf{x}_g , correspondant à la reprojection du point P_i dans l'image gauche du système j , est donné par la formule suivante :

$$\mathbf{x}_g = K_g^j M_{loc \rightarrow j} M_{i \rightarrow loc} P_i \quad (7.3)$$

La figure 7.5 présente un exemple de points reprojétés issus de deux systèmes de stéréovision placés respectivement à l'avant et à l'arrière d'un feu. Les points issus de la vue avant sont projetés sur l'image de détection de la vue arrière et *vice versa*. Il apparaît dans ces figures que de nombreux points sont à l'extérieur de la zone de feu et donc que le recalage initial n'est pas précis.

La troisième étape de cette méthode de calibrage est basée sur l'emploi d'une méthode d'optimisation par recherche de motifs (voir partie 3.3.1) qui permet de déterminer les angles α_i , β_i et γ_i minimisant le nombre de mauvais points reprojétés.

La figure 7.6 présente les points reprojétés obtenus après cette étape d'optimisation.

Cette étape se termine par la visualisation dans une interface graphique des points 3D d'une même scène projetés dans le repère de référence. Cette interface permet de valider ou de rechercher manuellement les valeurs des angles qui optimisent le recouvrement de points. Elle est présentée dans la figure 7.7.

La figure 7.10 présente un organigramme illustré des différentes étapes réalisées à partir de l'obtention des images stéréoscopiques multimodales jusqu'à l'obtention des points tridimen-

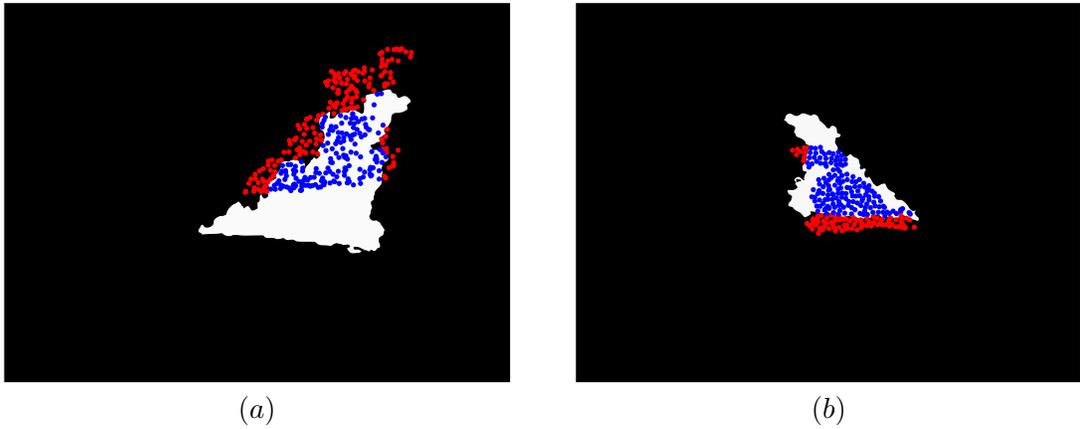


FIGURE 7.5 : Reprojection de points 3D. (a) Points issus de la vue arrière reprojétés sur l'image présentant les pixels de feu détectés à partir de la vue avant et (b) Points issus de la vue avant projetés sur l'image présentant les pixels de feu détectés à partir de la vue arrière. Les points rouges sont les points considérés comme de mauvaises reprojections.

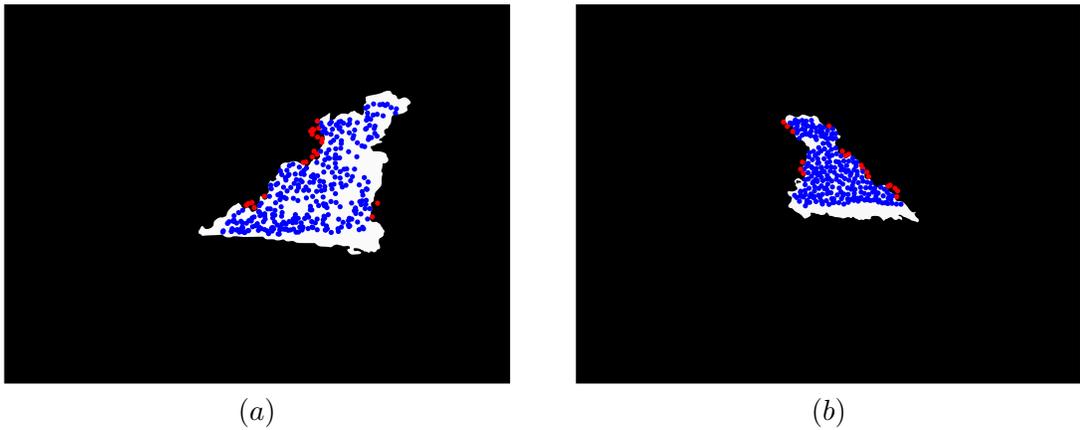


FIGURE 7.6 : Reprojection des points après optimisation. (a) Points issus de la vue arrière reprojétés sur l'image présentant les pixels de feu détectés à partir de la vue avant et (b) Points issus de la vue avant projetés sur l'image présentant les pixels de feu détectés à partir de la vue arrière. Les points rouges sont les points considérés comme des mauvaises reprojections.

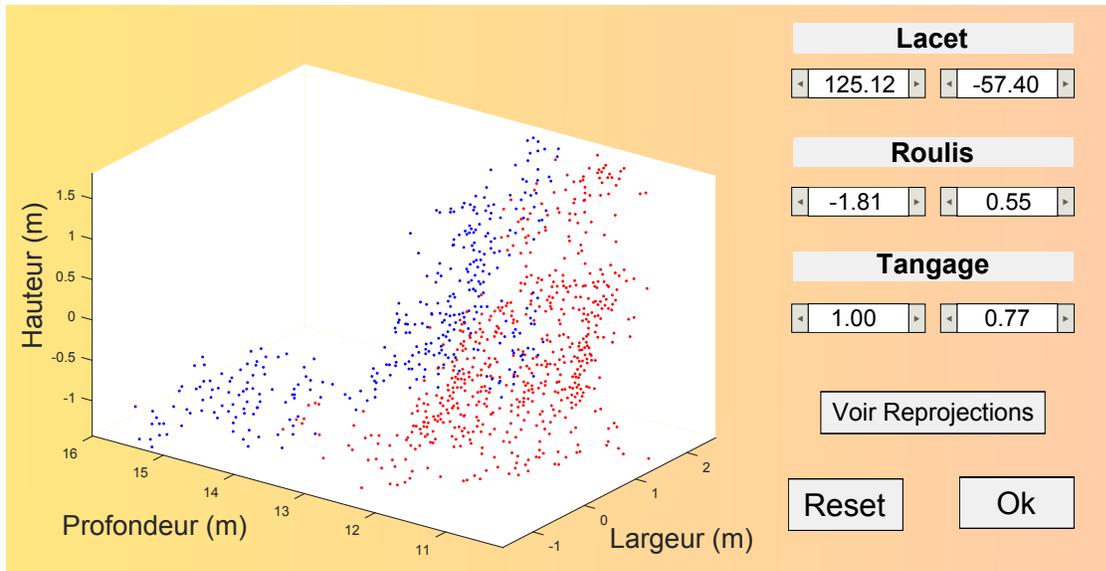


FIGURE 7.7 : Interface proposée pour la correction manuelle d'angles.

sionnels recalés.

7.3 Reconstruction tridimensionnelle de la forme de feu

La forme tridimensionnelle du feu est obtenue en appliquant à l'ensemble des points 3D recalés une triangulation de Delaunay. Cette procédure génère des tétraèdres dont les sommets sont les points qui se trouvent sur une sphère. Le rayon de cette sphère est un paramètre qui a été fixé. Les sommets des tétraèdres peuvent être éloignés d'une distance supérieure au rayon ce qui peut générer des formes tridimensionnelles différentes de l'objet à reconstruire comme ceci est illustré en figure 7.9 (a). Pour affiner la forme tridimensionnelle obtenue, une première procédure est réalisée de sorte à ne conserver que les tétraèdres dont le rayon de la sphère inscrite est inférieur à 60 cm. Dans une deuxième procédure, les centres des sphères inscrites dans les tétraèdres retenus sont ensuite reprojétés dans les images stéréoscopiques. Les tétraèdres dont le centre de la sphère inscrite projeté ne fait pas partie des pixels classés « feu » sont éliminés. La figure 7.9 (b) présente un exemple de raffinement.

Une option d'ombrage de Gouraud [182] peut être appliquée à forme tridimensionnelle obtenu par triangulation de Delaunay et raffinement. Elle génère une forme lissée dont un exemple est présentée dans la figure 7.10.

Aucune estimation de caractéristiques géométriques n'est faite à partir de la forme tridimensionnelle du feu mais elle constitue un bon indicateur visuel afin de déterminer en comparant la forme finale avec les images de départ si les points 3D ont été correctement positionnés.

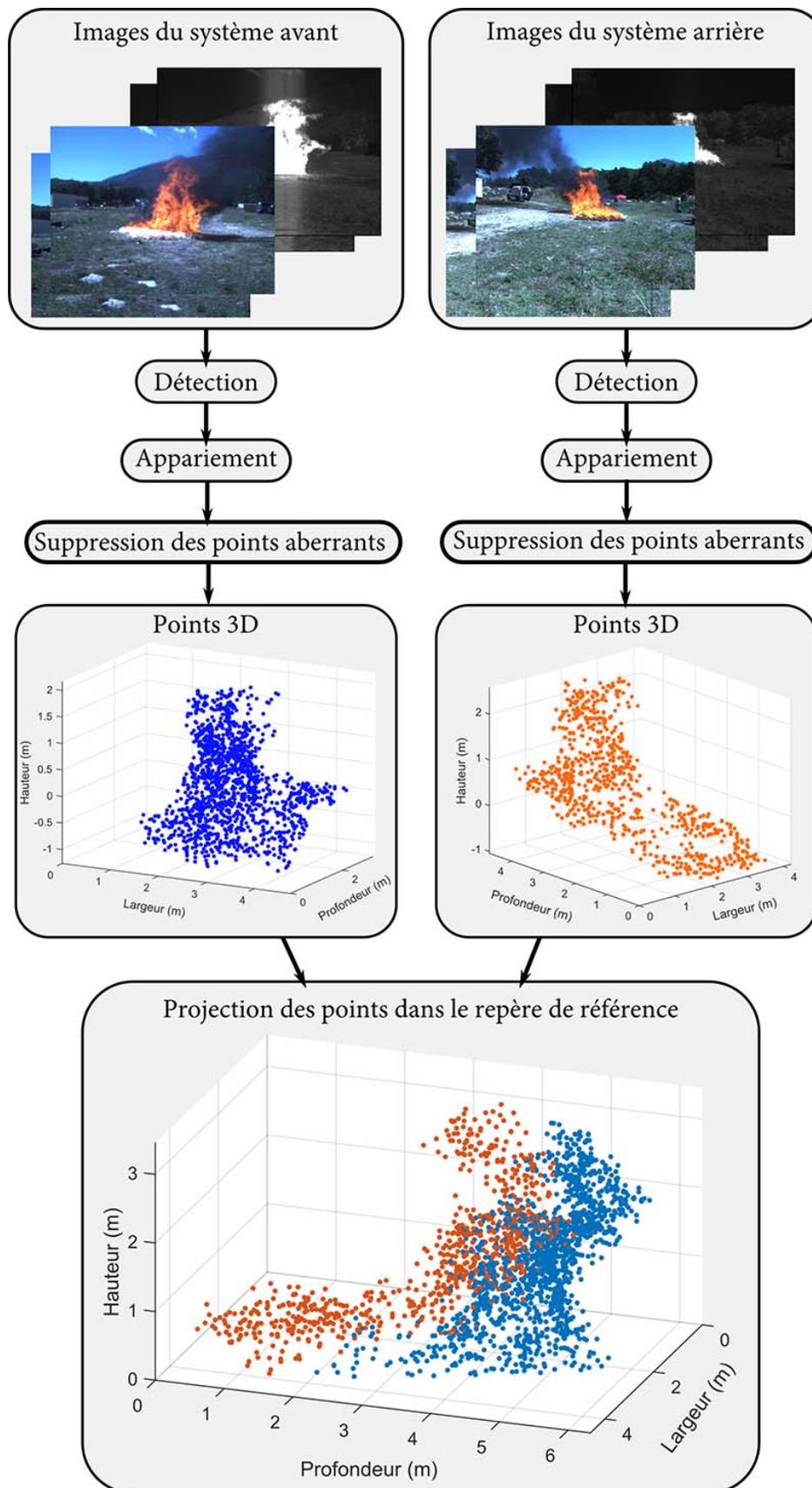


FIGURE 7.8 : Schéma des étapes de la reconstruction des ponts du feu.

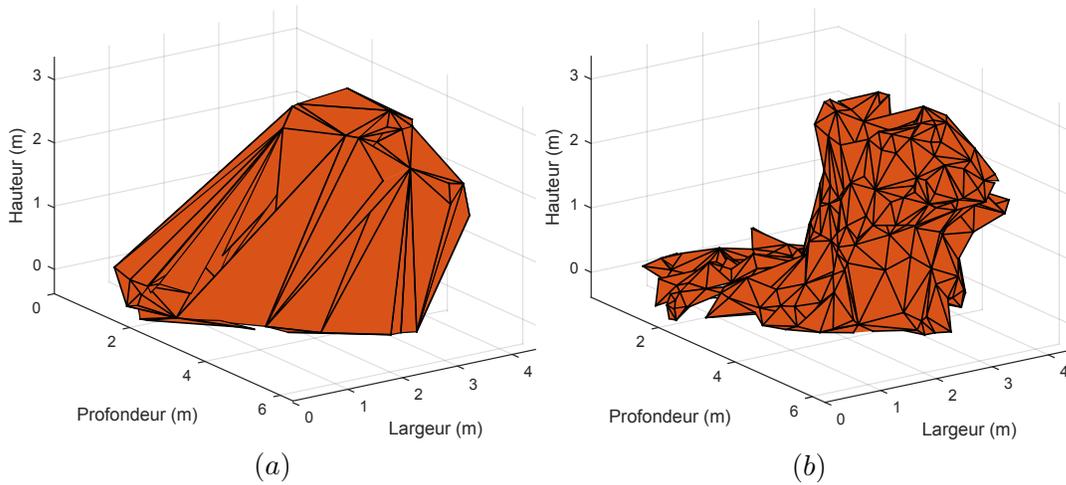


FIGURE 7.9 : Forme tridimensionnelle d'un feu obtenue par triangulation de Delaunay. (a) sans raffinement et (b) avec raffinement.

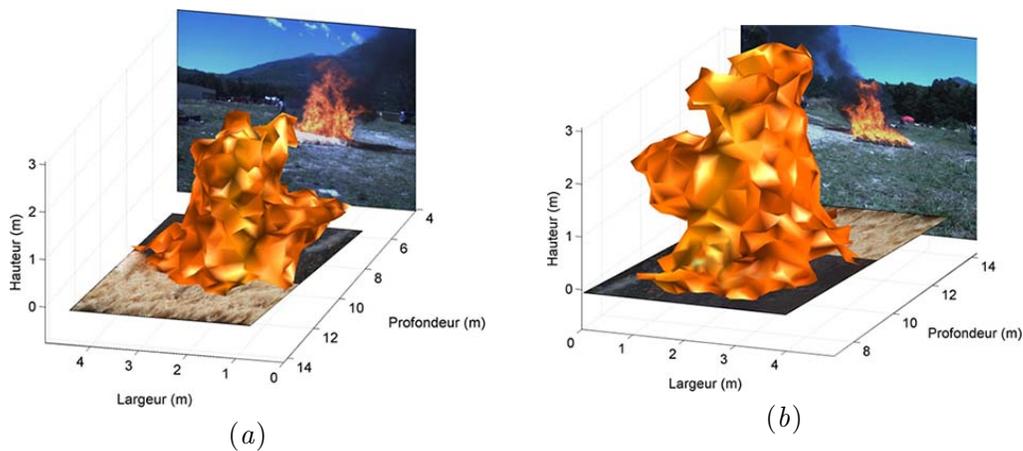


FIGURE 7.10 : Reconstruction complète du feu. (a) reconstruction tridimensionnelle finale vue de face et (b) vue de l'arrière.

7.4 Calcul de caractéristiques géométriques

A partir des points 3D (issus de systèmes de stéréovision positionnés de façon complémentaire sur le terrain), triés et recalés, il est possible d'estimer les caractéristiques géométriques de feu suivantes : la surface de végétal participant à la combustion, le périmètre, l'épaisseur et le volume d'un feu. Cette section détaille les méthodes permettant de calculer ces informations. Il est supposé dans la suite de cette partie que l'estimation du plan de base à partir des points 3D les plus bas est déjà réalisée.

7.4.1 Périmètre du feu

Le périmètre du feu est estimé à partir des points bas du front, c'est à dire dont la hauteur est inférieure à un seuil qui a été fixé à trente centimètres. Cette valeur permet de travailler avec un nombre suffisant de points tout en évitant de considérer ceux n'appartenant pas à la base du front. Les points sélectionnés sont projetés dans le plan (largeur - hauteur) et une triangulation de Delaunay est appliquée (avec un rayon de 30 cm) afin d'obtenir une surface fermée et d'en calculer le contour et l'aire. Les chercheurs travaillant sur la modélisation des feux de forêt considèrent que le végétal situé en dessous d'une flamme contribue à la combustion et approximent la surface par des formes simples [183]. La figure 7.11 présente le contour du feu présenté en figure 7.10. Le périmètre estimé pour ce feu est de 15 mètres.

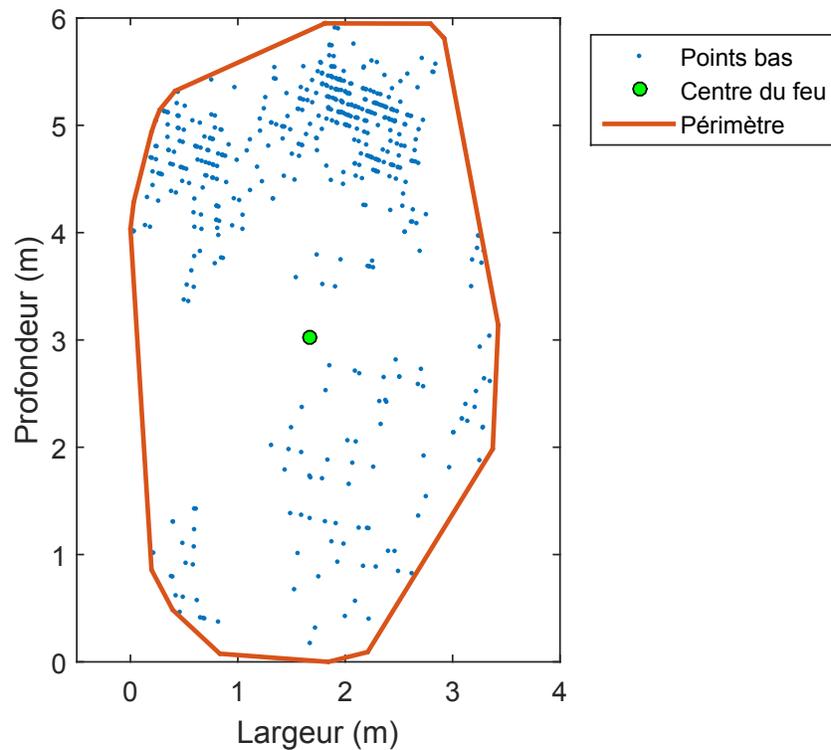


FIGURE 7.11 : Périmètre du feu.

7.4.2 Surface du végétal participant à la combustion

Comme pour le périmètre, la surface du végétal participant à la combustion est estimée à partir des points bas du front. Elle est approximée par l'aire A_s de la surface au sol délimitée par le périmètre et donnée par la formule suivante [184] :

$$A_s = \frac{1}{2} \sum_{i=1}^{n-1} (u_i v_{i+1} - u_{i+1} v_i) \quad (7.4)$$

Avec $(u_1, v_1) \dots (u_n, v_n)$ les coordonnées des points appartenant au contour du feu. Pour le feu montré en figure 7.11, l'aire de la surface au sol est de 17 mètres carrés.

7.4.3 Direction principale de propagation

Comme nous l'avons vu dans le chapitre 6, l'estimation de certaines caractéristiques nécessitent le calcul au préalable de la direction principale de propagation du feu et la transformation des points de sorte que le vecteur de direction principale soit colinéaire à leur axe de profondeur. Dans le cas d'images stéréoscopiques obtenues à partir d'un système positionné à l'avant du front, les lignes de front successives permettent de déterminer cet axe. Dans le cas d'images obtenues depuis plusieurs systèmes positionnés sur le terrain, un périmètre de feu est déterminé et il n'est pas possible de distinguer la ligne de front avant avec la ligne de front arrière à un instant t sans information complémentaire. Une méthode basée sur la détermination du centre de masse C_M des points de contour du feu de est utilisée. Les coordonnées u_{C_M} et v_{C_M} de ce point sont calculées par les formules suivantes [184] :

$$\begin{aligned} u_{C_M} &= \frac{1}{6A_s} \sum_{i=1}^{n-1} (u_i + u_{i+1})(u_i v_{i+1} - u_{i+1} v_i) \\ v_{C_M} &= \frac{1}{6A_s} \sum_{i=1}^{n-1} (v_i + v_{i+1})(u_i v_{i+1} - u_{i+1} v_i) \end{aligned} \quad (7.5)$$

La figure 7.11 montre un point de couleur verte qui est le centre de masse des points du contour de feu estimé par cette formule.

La direction principale est la moyenne de vecteurs de direction calculés en considérant deux à deux les centres de masse de points obtenus à des instants consécutifs sur un temps d'observation de cinq minutes avec un renouvellement d'informations par tranche de deux minutes. figure 7.12 présente la direction de propagation estimée par cette méthode. La propagation de ce feu ayant été rapide, la direction est calculée avec un intervalle de temps de 50 secondes.

7.4.4 Lignes de front

La ligne avant du front est estimée en interpolant avec une courbe de Bézier les points de la liste chaînée contenant le point de profondeur maximale et les deux points ayant une coordonnée de largeur extrême. La figure 7.13 présente les points du front avant et arrière détectés par cette méthode et les lignes de front associées.

7.4.5 Epaisseur

L'estimation de l'épaisseur du front de feu se fait en calculant la différence de profondeur entre le point de feu le plus avancé et celui le plus reculé sur les lignes de front avant et arrière. Pour le feu présenté en figure 7.10, l'épaisseur estimée est de 6 mètres.

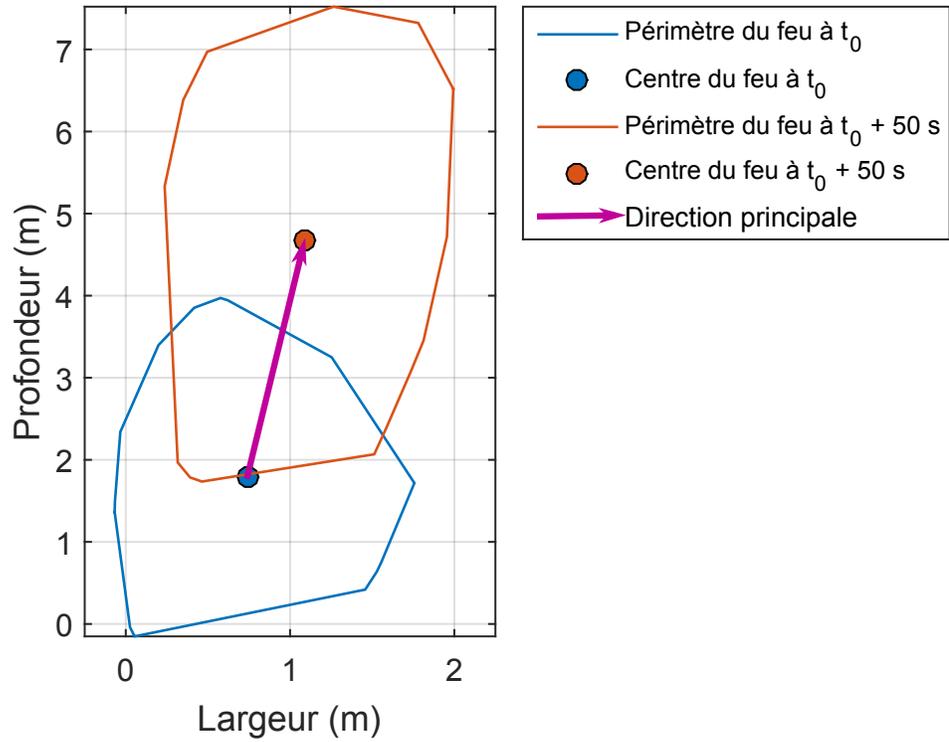


FIGURE 7.12 : Direction de propagation du feu.

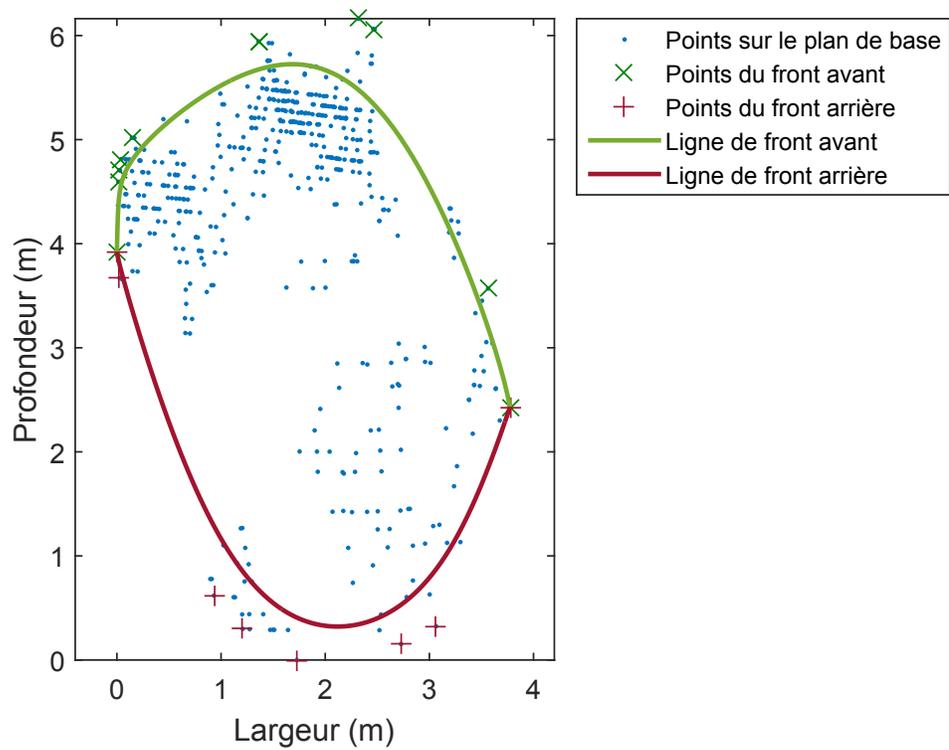


FIGURE 7.13 : Position du front de feu au sol et courbe de Bézier associé.

7.4.6 Volume

Le volume du feu est estimé à partir de la forme globale obtenue après avoir appliquée la triangulation de Delaunay sur les points 3D suivie de la procédure de raffinement de forme. Le volume global de la forme est la somme des volumes des tétraèdres élémentaires dont le volume est estimé de la façon suivante :

Soit a_1, a_2, a_3, a_4, a_5 et a_6 les six arrêtes d'un tétraèdre tel que celui illustré dans la figure 7.14.

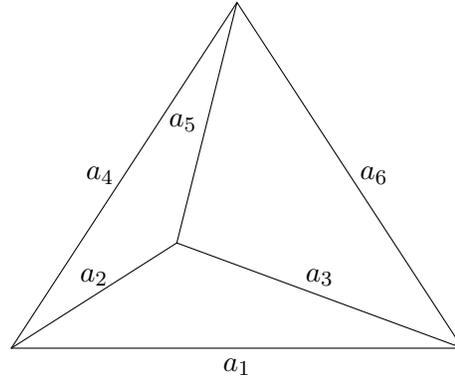


FIGURE 7.14 : Tétraèdre quelconque.

Le volume V_T de ce tétraèdre est donné par la formule d'Euler [185] :

$$V_T = \frac{1}{12} \sqrt{P - Q + R} \quad (7.6)$$

Avec :

$$\begin{aligned} P &= 4 a_1^2 a_2^2 a_3^2 \\ Q &= a_1^2 (a_2^2 + a_4^2 - a_5^2)^2 - a_2^2 (a_1^2 + a_4^2 - a_6^2)^2 - a_4^2 (a_1^2 + a_2^2 - a_3^2)^2 \\ R &= (a_2^2 + a_4^2 - a_5^2)(a_1^2 + a_4^2 - a_6^2)(a_1^2 + a_2^2 - a_3^2) \end{aligned}$$

Le volume de la forme tridimensionnelle présentée dans figure 7.9 a ainsi été estimée à 9 mètres cubes.

7.5 Incertitude de mesure

Il est difficile d'évaluer l'incertitude des mesures par vision de caractéristiques géométriques de feu obtenues car le feu est un phénomène non prévisible et non reproductible. L'incertitude de mesure de la méthode décrite dans cette thèse a été évaluée à partir d'images stéréoscopiques d'une voiture utilitaire sur laquelle des marques ont été collées (voir figure 7.15).



FIGURE 7.15 : Véhicule utilisé pour l'évaluation de l'incertitude des mesures.

Deux expériences de prise d'images stéréoscopiques ont été menées. L'une a été réalisée en utilisant un seul système de stéréovision et des images stéréoscopiques ont été acquises pour différentes distances entre le système de vision et la voiture. Les positions des caméras et du véhicule ont été relevées avec le récepteur GPS Leica Viva CS10 pour chaque configuration différente. L'autre expérience a été menée en maintenant la voiture à l'arrêt et en positionnant les systèmes de stéréovision autour de sorte à obtenir quatre vues complémentaires du véhicule. La figure 7.16 présente les deux configurations expérimentales.



FIGURE 7.16 : Configuration de prise d'images. (a) un seul système de vision et un véhicule en mouvement et (b) plusieurs systèmes de vision et un véhicule à l'arrêt.

La figure 7.17 montre la reconstruction tridimensionnelle du véhicule obtenue à partir des quatre paires d'images stéréoscopiques couleur. Il peut être observé que la forme globale de la voiture est compatible avec l'image présentée en figure 7.15. Un trou apparaît sur le haut de la forme reconstruite du fait du manque d'informations tridimensionnelles dans cette zone comme ceci est visible dans la figure 7.17(a). Il est donc impossible de calculer le volume total du véhicule à partir de cette reconstruction.

Le tableau 7.1 donne les différentes distances réelles (relevées par GPS) et estimées entre le véhicule et les caméras, la largeur et la hauteur du véhicule réelles (fournies par le constructeur) et estimées aux différentes distances et les erreurs de mesure.

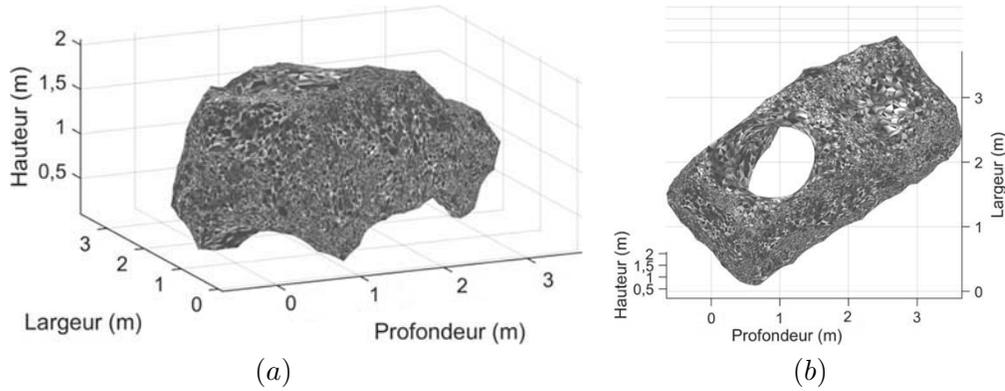


FIGURE 7.17 : Forme tridimensionnelle obtenue depuis les images de quatre systèmes de stéréovision positionnés autour d'une voiture. (a) Vue correspondante au système présenté dans la figure 7.16 et (b) vue de dessus.

TABLEAU 7.1 : Comparaison des mesures réelles et estimées.

Distance mesurée	4,8 m	8,1 m	11,4 m	14,7 m	18,0 m
Distance estimée	4,8 m	8,0 m	11,2 m	14,6 m	18,0 m
Erreur	0,0 m	0,1 m	0,2 m	0,1 m	0,0 m
Hauteur réelle	1,8 m				
Hauteur estimée	1,7 m	1,9 m	1,8 m	1,8 m	2,0 m
Erreur	0,1 m	0,1 m	0,0 m	0,0 m	0,2 m
Largeur réelle	1,81 m				
Largeur estimée	1,8 m	1,8 m	2,0 m	1,7 m	1,8 m
Erreur	0,0 m	0,0 m	0,2 m	0,1 m	0,0 m

Les données de ces deux tableaux montrent que les résultats obtenus sont de bonne qualité. L'incertitude maximum a été atteinte pour la troisième position.

Des images d'un objet de plus grande taille et situé à de grandes distances par rapport aux caméras aurait permis d'obtenir une analyse de performance plus complète mais ceci n'a pas pu être réalisé.

7.6 Conclusion

Dans ce chapitre, une méthode d'estimation de caractéristiques géométriques de feu a été présentée. Elle est basée sur le traitement de points tridimensionnels de feu obtenus depuis plusieurs systèmes de stéréovision et comprend trois étapes. Dans une première étape, une probabilité, dont la valeur dépend de la profondeur du point estimée et de la distance à ses plus proches voisins, est calculée pour tout point 3D et une sélection de points est faite en fonction de cette dernière. Dans une deuxième étape, les points 3D sont projetés dans un repère commun en utilisant des données relevées sur chaque système de vision. Une méthode d'optimisation a été présentée afin d'obtenir un meilleur recalage. Dans une troisième phase, les

caractéristiques géométriques sont estimées par traitement des points 3D. L'utilisation d'au moins deux systèmes de stéréovision placés de façon complémentaire sur le terrain permet la détermination de caractéristiques géométriques telle que la surface du végétal participant à la combustion, le périmètre du feu, son épaisseur et son volume. L'incertitude des mesures obtenues par le dispositif proposée a été évaluée à partir de deux expériences de prises d'images stéréoscopiques et dont l'objet de mesure était une voiture utilitaire. Les résultats obtenus ont montré que la méthode proposée permet l'estimation de caractéristiques géométriques de feu avec une incertitude compatible avec les exigences de la communauté scientifique travaillant sur la modélisation des feux de forêt.

Chapitre 8

Suivi d'une propagation de feu avec le dispositif de vision

8.1	Introduction	219
8.2	Suivi d'une propagation de feu réalisée sur un petit terrain plat	219
8.2.1	Description de l'expérience	219
8.2.2	Résultats obtenus et commentaires	220
8.3	Suivi de brûlages dirigés	220
8.4	Conclusion	226

8.1 Introduction

Ce chapitre concrétise en quelque sorte l'ensemble des éléments présentés dans ce rapport en présentant le suivi de propagation de feu dans des contextes différents. La section 8.2 présente le suivi d'un feu expérimental réalisé sur une plate-forme d'environ vingt-cinq mètres carrés. La section 8.3 décrit le suivi de brûlages dirigés réalisés sur des parcelles de terrain escarpées par des professionnels de la lutte contre les incendies.

8.2 Suivi d'une propagation de feu réalisée sur un petit terrain plat

8.2.1 Description de l'expérience

L'expérience décrite dans cette section a été réalisée sur le plateau technique de l'Unité d'Instruction et d'Intervention de la Sécurité Civile N°5 de Corte. Une surface de vingt-cinq mètres carrés a été recouverte de frisure de bois. Une ligne de feu a été initiée à l'extrémité de la parcelle de sorte que le feu, suivant la direction du vent, se propage sur toute la longueur de la surface de combustible. Deux systèmes de stéréovision contenant des caméras munies d'une focale de six millimètres ont été placés aux deux extrémités de la surface de propagation comme ceci est présenté dans la figure 8.1.



FIGURE 8.1 : Configuration de l'expérience. Les caméras gauches des systèmes de stéréovision avant et arrière sont respectivement entourés en bleu et rouge.

Des séquences d'images multimodales acquises avec une fréquence 1 Hz et un temps d'ouverture de $100 \mu s$ pour les images du spectre proche-infrarouge et de $500 \mu s$ pour le spectre visible ont été enregistrées.

8.2.2 Résultats obtenus et commentaires

La figure 8.2 présente les évolutions de la largeur, hauteur, épaisseur, longueur, inclinaison, vitesse du point maximum, surface et volume du front ainsi que de la surface de végétal participant à la combustion obtenues au cours de l'expérience décrite précédemment. La figure 8.3 montre l'évolution temporelle du périmètre du feu.

Il est malheureusement impossible de déterminer avec quelle précision ces caractéristiques ont été estimées. Certaines caractéristiques telles que la position du front avant et sa hauteur ont toutefois été comparées avec succès à des repères positionnés sur le terrain. Globalement, les données estimées sont compatibles avec ce qui pourrait être attendu pour ce type de feu.

8.3 Suivi de brûlages dirigés

Un brûlage dirigé est une opération d'aménagement et d'entretien de l'espace comprenant la réduction du combustible sur les ouvrages de prévention des incendies de forêts. Il est également une opération de gestion des peuplements forestiers, des pâturages, des landes et des friches. Il est réalisé par des professionnels de la lutte contre les incendies.

Il a été possible de suivre trois brûlages dirigés au cours de cette thèse. Dans ces expériences, un capteur de flux était positionné devant l'arrivée supposée du front. Les équipements de vision ont été installés sur le terrain afin de privilégier des mesures du front de feu lorsque

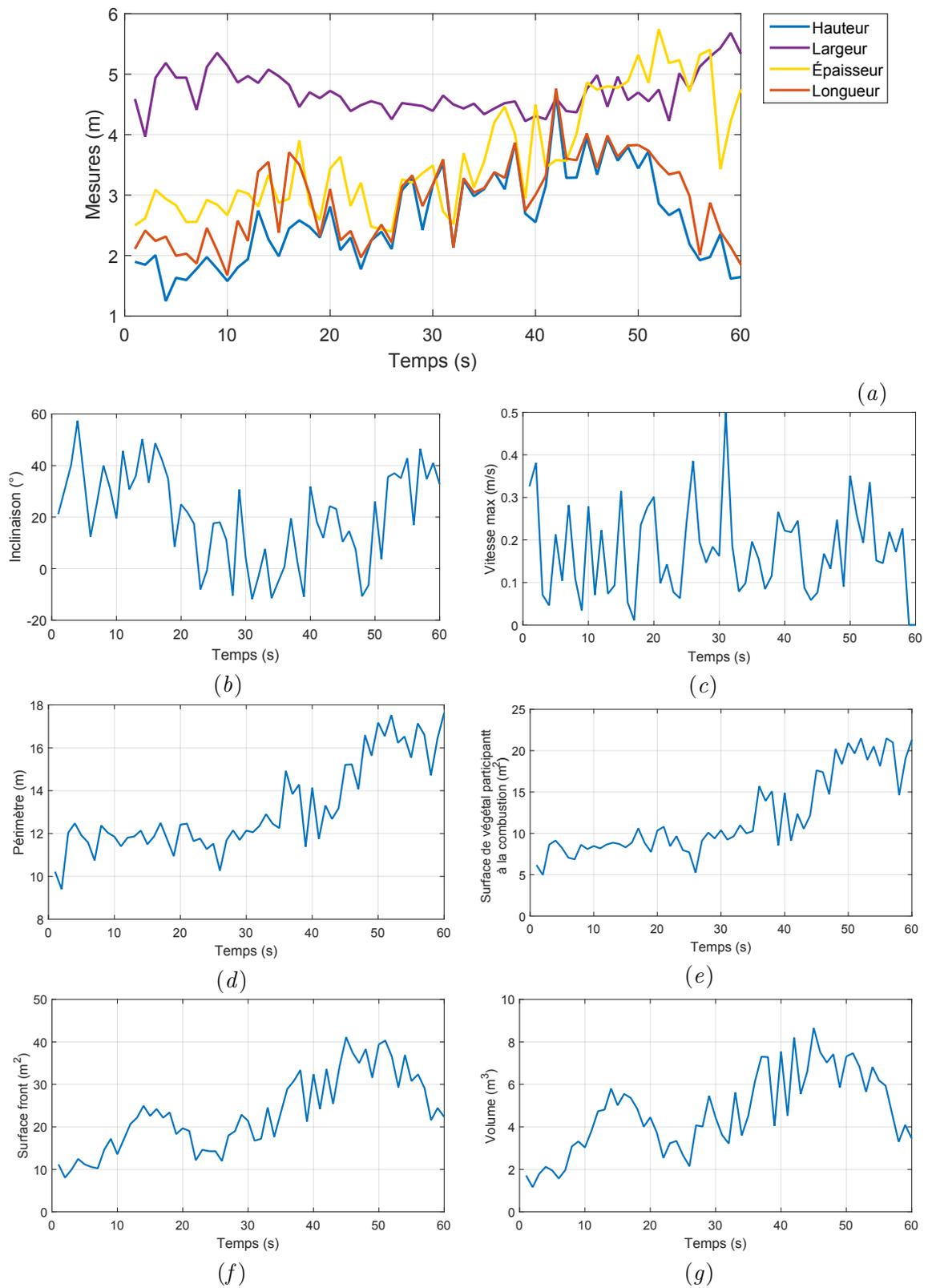


FIGURE 8.2 : Estimation des paramètres géométriques d'un feu au cours de sa propagation. (a) Hauteur, largeur, longueur et épaisseur, (b) inclinaison, (c) vitesse maximum d'un point de la ligne de front (d) périmètre, (e) surface de végétal participant à la combustion, (f) surface et (g) volume.

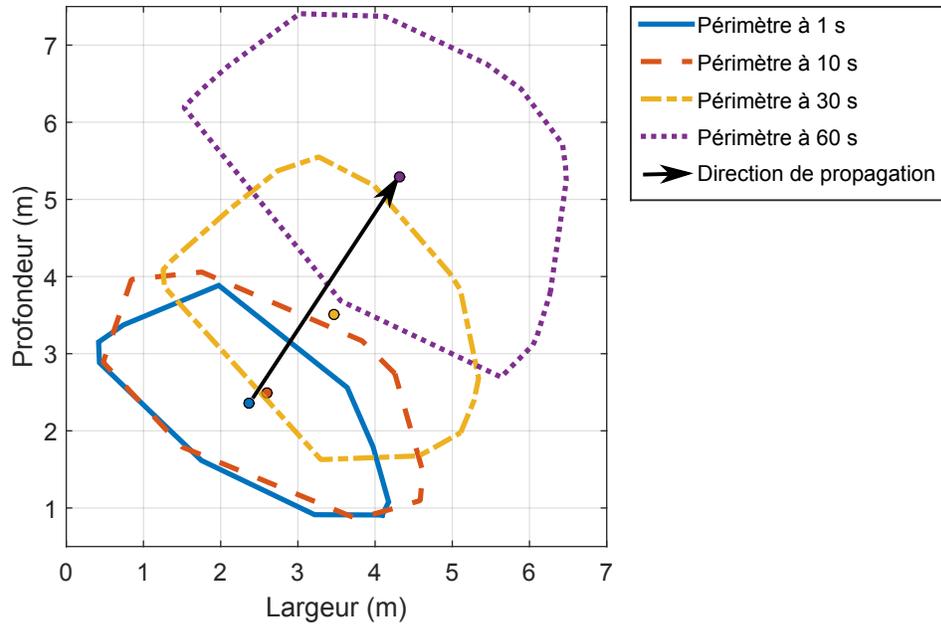


FIGURE 8.3 : Évolution temporelle du périmètre du feu.

ce dernier était proche du capteur (de sorte de pouvoir corrélérer les mesures obtenues par les différents dispositifs). Aussi, les dispositifs ont été disposés de sorte qu'il y ait une vue frontale et latérale du feu.

La figure 8.4 montre des informations relatives à un brûlage dirigé réalisé sur la commune de Corte (Haute-Corse). La figure 8.4 (a) et la figure 8.4 (b) présentent les photos des deux dispositifs de vision tels qu'ils ont été positionnés. La figure 8.4 (c) montre une image présentant le positionnement des systèmes de stéréovision par rapport à la zone de brûlage. La figure 8.4 (d) est une photo de la parcelle qui permet d'avoir une idée de la topologie du terrain. La figure (e) est une image où l'on peut discerner dans un cercle vert la position du capteur de flux, l'arrivée du front de flamme et la position relative du système de stéréovision.

Les principales difficultés dans le suivi par vision de brûlages dirigés sont dus principalement au fait que le feu ne démarre pas de manière identique sur la ligne de mise à feu, qu'il peut se transformer en un ensemble de plusieurs zones de feu et enfin que de la fumée dense noire peut être générée et masquer les zones de feu.

La figure 8.5 présente une des meilleures images multimodales qui a été acquise au cours de ce brûlage. Le feu est difficilement visible dans l'image couleur. Une zone de flamme principale apparaît dans l'image du proche-infrarouge. Les mesures obtenues sur cette image sont contenues dans le tableau 8.1.

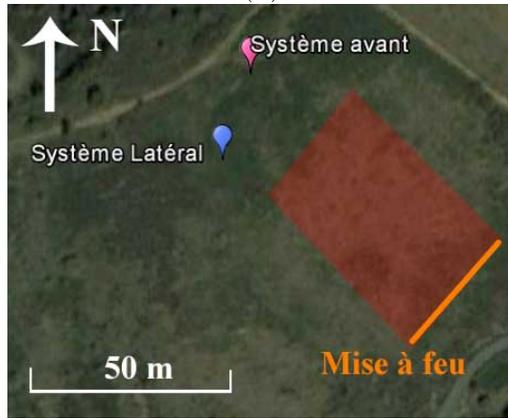
La figure 8.6 présente des images correspond à un brûlage dirigé qui a été fait sur la commune de Piedicorte (Haute-Corse). Les figure 8.6 (a) et (b) montrent les deux dispositifs de stéréovision



(a)



(b)



(c)



(d)



(e)

FIGURE 8.4 : Brûlage dirigé réalisé sur la commune de Corte. (a) Dispositif de stéréovision positionné à l'avant du front, (b) Dispositif de stéréovision positionné sur le côté (c) positions relatives des systèmes de vision par rapport à la zone de brûlage (d) mise à feu du brûlage par les forestiers et (e) positionnement du capteur de flux et de la caméra latérale.



FIGURE 8.5 : Image multimodale obtenue lors du premier brûlage dirigé. (a) image du spectre visible et (b) Image du spectre proche-infrarouge.

Distance feu-caméra	Largeur	Hauteur
35,4 m	2,0 m	1,5 m

TABLEAU 8.1 : Caractéristiques estimées du feu présenté dans la figure 8.5

et la topologie du terrain. Le terrain ne permettant pas de positionner un dispositif de vision à l'avant du front, les deux systèmes de stéréovision ont été placés dans une position latérale par rapport à la direction supposée du feu. La figure 8.6 (c) présente le placement relatif des systèmes de vision par rapport à la zone de brûlage.

Les difficultés rencontrées lors du suivi de ce brûlage ont été multiples. D'une part, l'un des deux systèmes de vision était situé trop loin du feu pour obtenir des mesures fiables. D'autre part, comme ceci est visible dans la figure 8.7, il n'y a jamais eu de front de feu établi.

Le troisième brûlage dirigé a été réalisé sur la commune de Solaro (corse du Sud). La figure 8.8 donne une vision des deux systèmes de stéréovision placés à l'avant et à l'arrière de la zone de propagation ainsi que leur position relative.

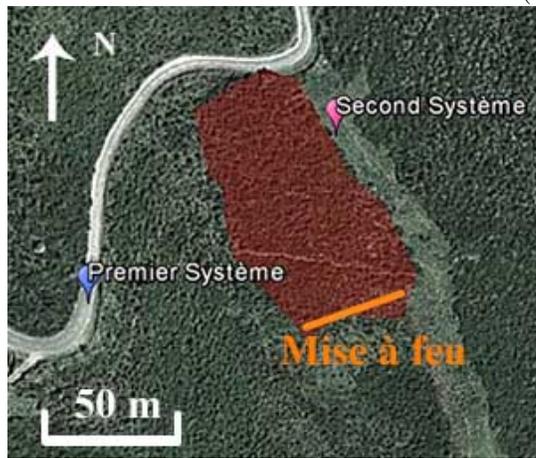
Au cours de ce brûlage dirigé, le feu n'a pas établi un régime de propagation et il a fallu le rallumer à plusieurs reprises. Ceci n'a pas permis de suivre l'évolution temporelle de ses caractéristiques. La figure 8.9 présente une image multimodale acquise alors que le feu atteignait la partie haute de la parcelle et à partir de laquelle les données présentées dans le tableau 8.2 ont été estimées.

Distance feu-caméra	Largeur	Hauteur
10,3 m	1,8 m	2,2 m

TABLEAU 8.2 : Caractéristiques estimées du feu présenté dans la figure 8.9



(a)



(b)

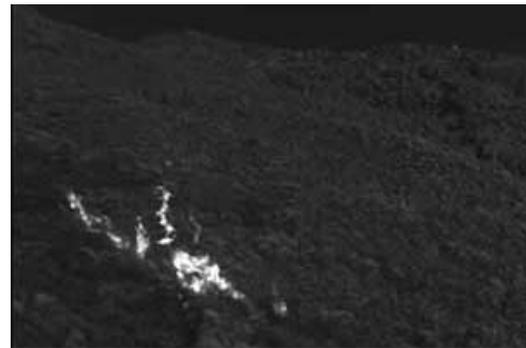


(c)

FIGURE 8.6 : Brûlage dirigé réalisé sur la commune de Piedicorte. (a) Photographie de la zone d'expérience prise depuis l'emplacement du premier système (b) positions relatives des systèmes de vision par rapport à la zone de brûlage et (c) photo du second système.



(a)



(b)

FIGURE 8.7 : Images des spectres visible et proche-infrarouge du second brûlage. (a) Image du spectre visible et (b) Image du spectre proche-infrarouge.

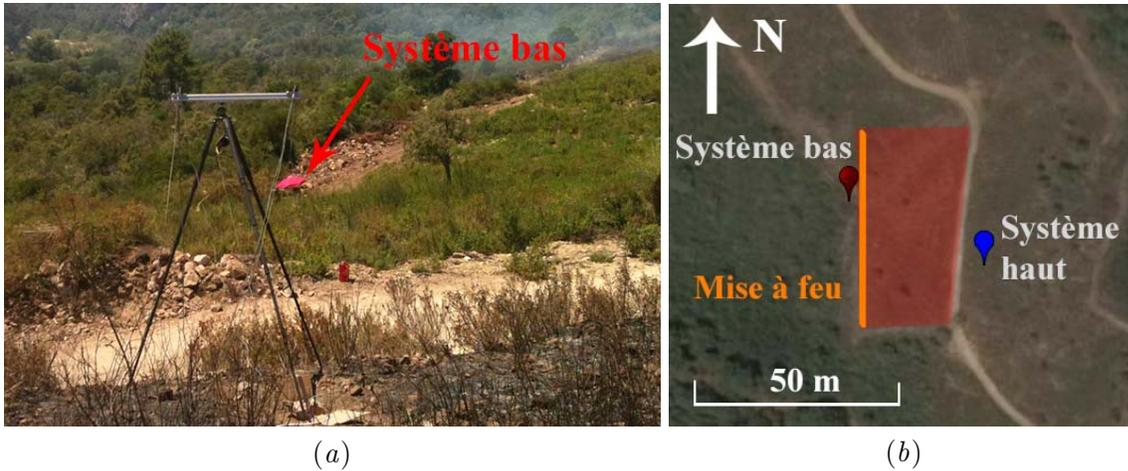


FIGURE 8.8 : Brûlage dirigé réalisé sur la commune de Solaro. (a) Photographie de la zone d'expérience prise depuis l'emplacement du système placé à l'avant de la parcelle (le second système se trouve au niveau du parasol fuchsia) et (b) positions relatives des systèmes de vision par rapport à la zone de brûlage.



FIGURE 8.9 : Image multimodale obtenue lors du troisième brûlage dirigé. (a) image du spectre visible et (b) Image du spectre proche-infrarouge.

8.4 Conclusion

Ce chapitre a présenté l'utilisation du dispositif de vision développée dans cette thèse et les résultats d'estimation de caractéristiques géométriques obtenus pour différentes propagations de feu. Il a été montré que le système est adapté et efficace pour le suivi de feu réalisé sur une plateforme. Son utilisation reste plus difficile pour le suivi de brûlages dirigés car, d'une part, la topologie du terrain ne permet pas toujours le positionnement optimum des systèmes de vision, et d'autre part, le nombre de systèmes de vision est trop petit pour suivre correctement un feu se propageant sur de grandes distances. Ce système n'a jamais été employé pour le suivi de vrais feux compte tenu du temps nécessaire à la mise en place du dispositif complet et de la vitesse de propagation de ce type de feu.

Conclusion et perspectives

L'estimation de caractéristiques géométriques de feux de végétation est une nécessité pour la communauté travaillant sur la problématique des feux de forêt. La dangerosité, les dimensions et les distances parcourues par les incendies rendent difficile la mesure de ses caractéristiques. Une des possibilités pour résoudre cette problématique est d'utiliser la vision. Comme ceci a été présenté au chapitre 1, peu de dispositifs de vision dédiés aux feux de forêt existent et ils nécessitent le positionnement de repères sur le terrain, ce qui ne peut pas être réalisé dans le cas de feux réels. Cette thèse propose d'estimer les caractéristiques géométriques de feux de végétation à partir d'un dispositif composé de plusieurs systèmes de stéréovision multimodale dont l'architecture est présentée au chapitre 2. Ce chapitre décrit l'étude relative aux émissions de longueur d'onde des feux et aux bandes spectrales propices à leur observation ainsi que l'argumentation associée au choix des paramètres du système de stéréovision choisi. Dans le processus de mesure par vision proposé, une des étapes clés est la détection des pixels de feu dans les images. Une étude de performance concernant onze algorithmes de détection de pixels travaillant sur des images du spectre visible pouvant être employés pour l'identification de pixels de feux de forêt et cinq nouveaux algorithmes est présentée dans le chapitre 3. Ce chapitre décrit notamment une base de données d'images contenant cinq cent images de feux de végétation dont les pixels ont été classés dans plus de onze catégories. La plus grande difficulté rencontrée par toutes ces méthodes de détection de pixels intervient lorsque de la fumée masque les zones de feu dans les images. Une solution est de travailler également avec des images du proche-infrarouge. Cinq nouvelles méthodes de détection de pixels de feu basées sur la fusion d'images du visible et du proche-infrarouge sont proposées dans le chapitre 4. Ce chapitre décrit une base de données contenant cent images multimodales utilisées pour évaluer les algorithmes proposés. La fusion d'images multimodales est également employée afin d'optimiser le nombre de points appariés entre les images stéréoscopiques comme ceci est décrit dans le chapitre 5. Les pixels appariés permettent d'obtenir des points tridimensionnels à partir desquels les caractéristiques géométriques des feux sont calculées. Les algorithmes permettant l'estimation de caractéristiques géométriques directement estimables en traitant les points tridimensionnelles obtenus depuis un seul système de stéréovision sont décrits au chapitre 6. Ils permettent la mesure de la position du front de feu, la vitesse de ses points, sa direction de propagation, sa hauteur, largeur, inclinaison, le profil de sa hauteur selon sa

largeur et sa surface. Le chapitre 6 décrit les méthodes qui permettent l'estimation des autres caractéristiques géométriques du feu que sont la position de son front avant et de son front arrière, son périmètre et son volume ainsi que la surface du végétal participant à la combustion. Ces méthodes nécessitent l'utilisation d'au moins deux systèmes de stéréovision positionnés de façon à obtenir des visions complémentaires du feu. Leur fiabilité a été évaluée en estimant l'incertitude de différentes mesures effectuées à partir d'images d'un véhicule utilitaire à l'arrêt et en mouvement. Les résultats de toutes les estimations des caractéristiques géométriques d'un feu réalisés sur une parcelle de cent mètres carrés sont présentés dans le chapitre 8. Ce chapitre présente également les limites du dispositif.

Perspectives

Le dispositif de vision et les méthodes proposées dans cette thèse ont prouvé être un moyen efficace pour estimer les caractéristiques géométriques d'un feu de végétation à l'échelle d'un terrain d'environ cent mètres carrés. Pour le suivi de feux de plus grande échelle et sur des terrains escarpés, des limites apparaissent et sont principalement liées au dispositif de vision. Le nombre de systèmes de stéréovision positionnés sur le terrain, leur distance au feu et la présence éventuelle de fumée dense noire (qui peut masquer des zones de feu dans les images proche-infrarouge) sont les principaux facteurs de cette limite. Plusieurs moyens permettraient d'améliorer le dispositif. Une première amélioration pourrait être apportée grâce à l'utilisation de caméras à zoom variable qui maintiendraient une taille d'objet dans les images comparable au cours du temps. Des recherches ont été initiées à l'Université de Corse en collaboration avec le Professeur Moulay Akhloufi sur ce sujet. Une deuxième amélioration consisterait à considérer des caméras thermiques de l'infrarouge à moyennes longueurs d'onde avec porte-filtre rotatif permettant la distinction de la base et du front de feu. Ces caméras sont toutefois d'un prix très élevé et l'achat de deux systèmes peut être une impossibilité. Enfin, les prix abordables de drone permettent désormais de considérer cet engin pour obtenir des images à courte distance d'un feu au cours de sa propagation. Des recherches ont déjà été développées à ce sujet [186]. Elles ont débuté cette année à l'Université de Corse et devraient se poursuivre en collaboration avec le Professeur Turgay Celik de l'Université de Witwatersrand (Afrique du Sud), notamment avec la réalisation de mon post-doc. Enfin des améliorations logicielles pourraient être apportées en améliorant la méthode de détection de pixels de feu avec notamment la prise en compte du mouvement [187] ou de la texture [46] du feu.

Publications et distinctions

Publications

Toulouse T., Rossi L., Celik T. et Akhloufi M. « Automatic fire pixel detection using image processing : a comparative analysis of rule-based and machine learning-based methods ». *Signal, Image and Video Processing*, 2015.

Toulouse T., Rossi L., Akhloufi M., Celik T. et Maldague X. « Benchmarking of wildland fire color segmentation algorithms ». *IET image processing*, doi : 10.1049/iet-ipr.2014.0935, 2015

Rossi L., Toulouse T., Cancellieri D., Rossi J.L., Morandini F., Akhloufi M. « Utilisation de la stéréovision visible et proche infrarouge pour la mesure de données expérimentales dans le cadre d'une recherche pluridisciplinaire sur les feux de forêt ». *Journal National de la Recherche en IUT*, 2014

Toulouse T., Rossi L., Cancellieri D., Rossi J.L., Morandini F., Akhloufi M. « Intégration de la vision dans une recherche pluridisciplinaire dédiée aux feux de forêt » Dans le Congrès National de la Recherche en IUT, Corte, France, 12-14 juin 2013

Distinctions

Prix d'excellence pour les résultats à l'examen prédoctoral en tant qu'étudiant au programme de doctorat en génie électrique de l'Université Laval (Mai 2015).

Prix de la Communication scientifique en « Sciences & Techniques et Santé » du Conseil Départemental de la Corse du Sud lors de la journée des doctorant du 25 juin 2015 pour le poster intitulé : *Estimation par stéréovision multimodale de caractéristiques géométriques d'un feu de végétation en propagation.*

Bibliographie

- [1] FAO. « Enlisting communities in wildfire prevention ». 2010. <http://www.fao.org/news/story/en/item/41230/icode/>, accédé le 13/08/2015.
- [2] EFFIS. « European Forest Fire Information System ». <http://forest.jrc.ec.europa.eu/effis/>. Accédé le 13/08/2015.
- [3] Serna J. « Rim fire started by a hunter's illegal fire, Forest Service says ». *Los Angeles Times*, 2013.
- [4] NSW Rural Fire Service. « Watch and Act? Linksvie Road Fire, Springwood (Blue Mountains) ». http://www.rfs.nsw.gov.au/dsp_content.cfm?cat_id=684. Accédé le 13/08/2015.
- [5] NSW Rural Fire Service. « Initial assessment of Winmalee and Springwood areas ». 2013. <http://www.rfs.nsw.gov.au/news-and-media/media-releases/initial-assessment-of-winmalee-and-springwood-areas>, accédé le 13/08/2015.
- [6] Fuglem P.L., Hirsch K.G. et Canadian Council of Forest Ministers. « Stratégie canadienne en matière de feux de forêt : synthèses de fond, analyses et perspectives. ». 2006.
- [7] « PROMÉTHÉE² : La banque de données sur les incendies de forêts en région Méditerranéenne en France ». <http://www.promethee.com/default/bilan-tableaux>, accédé le 13/08/2015.
- [8] Rigolot E. « Impact du changement climatique sur les feux de forêt ». 2008.
- [9] Wotton B.M., Nock C.A. et Flannigan M.D. « Forest fire occurrence and climate change in Canada ». *International Journal of Wildland Fire*, 19 (3) p. 253–271, 2010.
- [10] Balbi J.H., Rossi J.L., Marcelli T. et Santoni P.A. « A 3D physical real-time model of surface fires across fuel beds ». *Combustion Science and Technology*, 179 (12) p. 2511–2537, 2007.
- [11] Rossi J.L., Simeoni A., Moretti B. et Leroy-Cancellieri V. « An analytical model based on radiative heating for the determination of safety distances for wildland fires ». *Fire Safety Journal*, 46 (8) p. 520–527, 2011.

- [12] Balbi J.H., Chatelon F.J., Rossi J.L., Simeoni A., Viegas D.X. et Rossa C. « Modelling of Eruptive Fire Occurrence and Behaviour ». *Journal of Environmental Science and Engineering*, (3) p. 115–132, 2014.
- [13] McArthur A.G. « Weather and grassland fire behavior. ». *Forestry and Timber Bureau.*, Leaflet (100), 1966.
- [14] Albini F.A. « A Model for Fire Spread in Wildland Fuels by-Radiation ». *Combustion Science and Technology*, 42 (5-6) p. 229–258, 1985.
- [15] Grishin A.M. *Mathematical modeling of forest fires and new methods of fighting them.* Publishing house of the Tomsk state university, 1997.
- [16] Mendes-Lopes J.M., Ventura J.M. et Amaral J.M. « Flame characteristics, temperature-time curves, and rate of spread in fires propagating in a bed of Pinus pinaster needles ». *International Journal of Wildland Fire*, 12 (1) p. 67–84, 2003.
- [17] Pastor E., Zarate L., Planas E. et Arnaldos J. « Mathematical models and calculation systems for the study of wildland fire behaviour ». *Progress in Energy and Combustion Science*, 29 (2) p. 139–153, 2003.
- [18] Viegas D.X. « On the existence of a steady state regime for slope and wind driven fires ». *International Journal of Wildland Fire*, 13 (1) p. 101–117, 2004.
- [19] Pérez Y., Pastor E., Planas E., Plucinski M. et Gould J. « Computing forest fires aerial suppression effectiveness by IR monitoring ». *Fire Safety Journal*, 46 (1–2) p. 2–8, 2011.
- [20] Martínez-de Dios J.R., André J.C., Gonçalves J.C., Arrue B.C., Ollero A. et Viegas D.X. « Laboratory fire spread analysis using visual and infrared images ». *International Journal of Wildland Fire*, 15 (2) p. 179–186, 2006.
- [21] Rossi L., Molinier T., Akhloufi M., Tison Y. et Pieri A. « Estimating the surface and volume of laboratory-scale wildfire fuel using computer vision ». *Image Processing, IET*, 6 (8) p. 1031–1040, 2012.
- [22] Verstockt S., Van Hoecke S., Tilley N., Merci B., Sette B., Lambert P., Hollemeersch C.F.J. et Van de Walle R. « FireCube : a multi-view localization framework for 3D fire analysis ». *Fire Safety Journal*, 46 (5) p. 262–275, 2011.
- [23] Martínez-de Dios J.R., Merino L. et Ollero A. « Fire detection using autonomous aerial vehicles with infrared and visual cameras ». Dans *Proceedings of the 16th IFAC World Congress*, 2005.
- [24] Stipaničev D., Vuko T., Krstinić D., Štula M. et Bodrozić L. « Forest fire protection by advanced video detection system-Croatian experiences ». Dans *Third TIEMS Workshop-Improvement of Disaster Management System*, Trogir, 2006.

- [25] Verstockt S., Lambert P., Van de Walle R., Merci B. et Sette B. « State of the art in vision-based fire and smoke detection ». Dans 14th International Conference on Automatic Fire Detection, volume 2, p. 285–292. University of Duisburg-Essen. Department of Communication Systems, 2009.
- [26] Çetin A.E., Dimitropoulos K., Gouverneur B., Grammalidis N., Günay O., Habiboğlu Y.H., Töreyn B.U. et Verstockt S. « Video fire detection–Review ». *Digital Signal Processing*, 23 (6) p. 1827–1843, 2013.
- [27] Phillips III W., Shah M. et da Vitoria Lobo N. « Flame recognition in video ». *Pattern recognition Letters*, 23 (1) p. 319–327, 2002.
- [28] Celik T., Demirel H., Ozkaramanli H. et Uyguroğlu M. « Fire detection using statistical color model in video sequences ». *Journal of Visual Communication and Image Representation*, 18 (2) p. 176–185, 2007.
- [29] Töreyn B.U., Dedeoğlu Y., Güdükbay U. et Çetin A.E. « Computer vision based method for real-time fire and flame detection ». *Pattern Recognition Letters*, 27 (1) p. 49–58, 2006.
- [30] Collumeau J.F., Laurent H., Hafiane A. et Chetehouna K. « Fire scene segmentations for forest fire characterization : A comparative study ». Dans 18th IEEE International Conference on Image Processing (ICIP), p. 2973–2976. 2011.
- [31] Ko B.C., Cheong K.H. et Nam J.Y. « Fire detection based on vision sensor and support vector machines ». *Fire Safety Journal*, 44 (3) p. 322–329, 2009.
- [32] Borges P.V.K., Mayer J. et Izquierdo E. « Efficient visual fire detection applied for video retrieval ». Dans 16th European Signal Processing Conference, p. 1–5. IEEE, 2008.
- [33] Celik T. et Demirel H. « Fire detection in video sequences using a generic color model ». *Fire Safety Journal*, 44 (2) p. 147–158, 2009.
- [34] Wang D.C., Cui X., Park E., Jin C. et Kim H. « Adaptive flame detection using randomness testing and robust features ». *Fire Safety Journal*, 55 p. 116–125, 2013.
- [35] Rudz S., Chetehouna K., Hafiane A., Laurent H. et Séro-Guillaume O. « Investigation of a novel image segmentation method dedicated to forest fire applications ». *Measurement Science and Technology*, 24 (7) p. 075 403, 2013.
- [36] Marbach G., Loepfe M. et Brupbacher T. « An image processing technique for fire detection in video images ». *Fire safety journal*, 41 (4) p. 285–289, 2006.
- [37] Celik T. « Fast and Efficient Method for Fire Detection Using Image Processing ». *ETRI journal*, 32 (6) p. 881–890, 2010.
- [38] Horng W.B., Peng J.W. et Chen C.Y. « A new image-based real-time flame detection method using color analysis ». Dans Networking, Sensing and Control Proceedings, p. 100–105. IEEE, 2005.

- [39] Liu C.B. et Ahuja N. « Vision based fire detection ». Dans Proceedings of the 17th International Conference on Pattern Recognition,(ICPR), volume 4, p. 134–137. IEEE, 2004.
- [40] Chen J., He Y. et Wang J. « Multi-feature fusion based fast video flame detection ». *Building and Environment*, 45 (5) p. 1113–1122, 2010.
- [41] Chen T.H., Wu P.H. et Chiou Y.C. « An early fire-detection method based on image processing ». Dans International Conference on Image Processing, (ICIP), volume 3, p. 1707–1710. IEEE, 2004.
- [42] Rossi L., Molinier T., Akhloufi M., Tison Y. et Pieri A. « A 3D vision system for the measurement of the rate of spread and the height of fire fronts ». *Measurement Science and Technology*, 21 (10) p. 105 501, 2010.
- [43] Rossi L. et Akhloufi M. « Dynamic fire 3D modeling using a real-time stereovision system ». Dans Technological Developments in Education and Automation p. 33–38. Springer, 2010.
- [44] Ridler T. et Calvard S. « Picture thresholding using an iterative selection method ». *IEEE transactions on Systems, Man and Cybernetics*, 8 (8) p. 630–632, 1978.
- [45] Green D.M. et Swets J.A. *Signal detection theory and psychophysics*, volume 1. Wiley New York, 1966.
- [46] Chino D.Y., Avalhais L.P., Rodrigues Jr J.F. et Traina A.J. « BoWFire : Detection of Fire in Still Images by Integrating Pixel Color and Texture Analysis ». Dans 28th Conference on Graphics, Patterns, and Images (SIBGRAPI), August 26-29, Salvador, Brazil, 2015.
- [47] Russell S.J., Norvig P. et Davis E. *Artificial intelligence : a modern approach*. Prentice Hall series in artificial intelligence. Prentice Hall, Upper Saddle River, 3rd ed édition, 2010.
- [48] Gonzales R.C. et Woods R.E. *Digital Image Processing, 2-nd Edition*. Prentice Hall, 2002.
- [49] Quintiere J.G. *Principles of fire behavior*. Cengage Learning, 1998.
- [50] Chen T.H., Kao C.L. et Chang S.M. « An intelligent real-time fire-detection method based on video processing ». Dans Proceedings of 37th Annual International Carnahan Conference on Security Technology, p. 104–111. IEEE, 2003.
- [51] Westland S. et Ripamonti C. *Multispectral Imaging*. Wiley Online Library, 2004.
- [52] Vapnik V. *The nature of statistical learning theory*. springer, 1999.
- [53] Otsu N. « A threshold selection method from gray-level histograms ». *Automatica*, 11 (285-296) p. 23–27, 1975.

- [54] Gouverneur B., Verstockt S., Pauwels E., Han J., de Zeeuw P.M. et Vermeiren J. « Archeological treasures protection based on early forest wildfire multi-band imaging detection system ». Dans SPIE Security+ Defence, p. 85 410J–85 410J. International Society for Optics and Photonics, 2012.
- [55] Verstockt S., Vanoosthuyse A., Van Hoecke S., Lambert P. et Van de Walle R. « Multi-sensor fire detection by fusing visual and non-visual flame features ». Dans Image and Signal Processing volume 6134, p. 333–341. Springer, 2010.
- [56] Martínez-de Dios J. et Ollero A. « Wavelet applications to forest-fire monitoring and measurement ». Dans Proceedings of the 5th Biannual World Automation Congress, volume 13, p. 221–226. IEEE, 2002.
- [57] Torabnezhad M., Aghagolzadeh A. et HadiSeyedarabi H. « Visible and IR image fusion algorithm for short range smoke detection ». Dans First RSI/ISM International Conference on Robotics and Mechatronics (ICRoM), p. 38–42. IEEE, 2013.
- [58] Kittler J. et Illingworth J. « Minimum error thresholding ». *Pattern recognition*, 19 (1) p. 41–47, 1986.
- [59] Martínez de Dios J.R. *Detección y Monitorización en Entornos Exteriores no Estructurados empleando Imágenes de Infrarrojos. Aplicación a los Incendios Forestales*. thèse de doctorat, University of Seville, 2001.
- [60] Martínez-de Dios J.R., Merino L., Caballero F. et Ollero A. « Automatic forest-fire measuring using ground stations and unmanned aerial systems ». *Sensors*, 11 (6) p. 6328–6353, 2011.
- [61] Martínez-de Dios J. et Ollero A. « Automatic threshold selection for infrared images of fire ». Dans Proceedings of IV International Conference on Forest fire research and wild-land fire safetyt, Luso, Coimbra, Portugal, 18-23 November 2002., p. 129–139. Millpress Science Publishers, 2002.
- [62] Rockinger O. « Image sequence fusion using a shift-invariant wavelet transform ». Dans International Conference on Image Processing (ICIP), volume 3, p. 288–291. IEEE, 1997.
- [63] Zadeh L. « Fuzzy sets ». *Information and Control*, 8 (3) p. 338 – 353, 1965.
- [64] Kass M., Witkin A. et Terzopoulos D. « Snakes : Active contour models ». *International journal of computer vision*, 1 (4) p. 321–331, 1988.
- [65] Rousselle J.J. *Les contours actifs, une méthode de segmentation : application à l'imagerie médicale*. thèse de doctorat, Tours, 2003.
- [66] Xu C. et Prince J.L. « Generalized gradient vector flow external forces for active contours ». *Signal processing*, 71 (2) p. 131–139, 1998.

- [67] Cohen L.D. « On active contour models and balloons ». *CVGIP : Image understanding*, 53 (2) p. 211–218, 1991.
- [68] Amini A., Tehrani S., Weymouth T.E. *et al.* « Using dynamic programming for minimizing the energy of active contours in the presence of hard constraints ». Dans Second International Conference on Computer Vision, p. 95–99. IEEE, 1988.
- [69] Williams D.J. et Shah M. « A fast algorithm for active contours and curvature estimation ». *CVGIP : Image understanding*, 55 (1) p. 14–26, 1992.
- [70] Meléndez J., Aranda J.M., de Castro A.J. et López F. « Measurement of forest fire parameters with multi-spectral imaging in the medium infrared ». *Quantitative InfraRed Thermography Journal*, 3 (2) p. 183–199, 2006.
- [71] Richards J.A. *Remote sensing digital image analysis*. Springer, 2012.
- [72] Çetin A.E., Dimitropoulos K., Gouverneur B., Grammalidis N., Günay O., Habiboğlu Y.H., Töreyn B.U. et Verstockt S. « Video fire detection - Review ». *Digital Signal Processing*, 23 (6) p. 1827–1843, 2013.
- [73] Yu C., Mei Z. et Zhang X. « A real-time video fire flame and smoke detection algorithm ». *Procedia Engineering*, 62 p. 891–898, 2013.
- [74] Stauffer C. et Grimson W.E.L. « Learning patterns of activity using real-time tracking ». *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22 (8) p. 747–757, 2000.
- [75] Xuan Truong T. et Kim J.M. « Fire flame detection in video sequences using multi-stage pattern recognition techniques ». *Engineering Applications of Artificial Intelligence*, 25 (7) p. 1365–1372, 2012.
- [76] Chamberlin D. et Rose A. « The flicker of luminous flames ». *Proceedings of the Symposium on Combustion*, 1-2 p. 27–32, 1948.
- [77] Töreyn B.U., Cinbiş R.G., Dedeoğlu Y. et Çetin A.E. « Fire detection in infrared video using wavelet analysis ». *Optical Engineering*, 46 (6) p. 7204, 2007.
- [78] Hu M.K. « Visual pattern recognition by moment invariants ». *Information Theory, IRE Transactions on*, 8 (2) p. 179–187, 1962.
- [79] Tsetsos V., Sekkas O., Tsoublekas G., Hadjieythymiades S. et Zervas E. « A Forest Fire Detection System : The Meleager Approach ». Dans Informatics (PCI), 2012 16th Panhellenic Conference on, p. 320–326. IEEE, 2012.
- [80] Doretto G., Cremers D., Favaro P. et Soatto S. « Dynamic texture segmentation ». Dans Proceedings of Ninth IEEE International Conference on Computer Vision, p. 1236–1242. IEEE, 2003.

- [81] Chabrier S., Emile B., Rosenberger C. et Laurent H. « Unsupervised performance evaluation of image segmentation ». *EURASIP Journal on Applied Signal Processing*, 2006 p. 217–217, 2006.
- [82] Zhang H., Fritts J.E. et Goldman S.A. « Image segmentation evaluation : A survey of unsupervised methods ». *Computer Vision and Image Understanding*, 110 (2) p. 260–280, 2008.
- [83] Hafiane A., Chabrier S., Rosenberger C. et Laurent H. « A new supervised evaluation criterion for region based segmentation methods ». Dans *Advanced Concepts for Intelligent Vision Systems*, p. 439–448. Springer, 2007.
- [84] Unnikrishnan R., Pantofaru C. et Hebert M. « Toward objective evaluation of image segmentation algorithms ». *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 29 (6) p. 929–944, 2007.
- [85] Rudz S., Chetehouna K., Hafiane A., Sero-Guillaume O. et Laurent H. « On the evaluation of segmentation methods for wildland fire ». Dans *Advanced Concepts for Intelligent Vision Systems*, p. 12–23. Springer, 2009.
- [86] Santana P., Gomes P. et Barata J. « A vision-based system for early fire detection ». Dans *IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, p. 739–744. 2012.
- [87] Martin D., Fowlkes C., Tal D. et Malik J. « A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics ». Dans *Proceedings of Eighth IEEE International Conference on Computer Vision*, volume 2, p. 416–423. IEEE, 2001.
- [88] Polak M., Zhang H. et Pi M. « An evaluation metric for image segmentation of multiple objects ». *Image and Vision Computing*, 27 (8) p. 1223–1227, 2009.
- [89] Baldi P., Brunak S., Chauvin Y., Andersen C.A. et Nielsen H. « Assessing the accuracy of prediction algorithms for classification : an overview ». *Bioinformatics*, 16 (5) p. 412–424, 2000.
- [90] Van Rijsbergen C.J. *Information Retrieval ..* Butterworth-Heinemann, London, 2nd edition édition, 1979.
- [91] Kulkarni M. et Nicolls F. « Interactive image segmentation using graph cuts ». *Pattern Recognition Association of South Africa*, p. 99–104, 2009.
- [92] Alexander M.E. « Calculating and interpreting forest fire intensities ». *Canadian Journal of Botany*, 60 (4) p. 349–357, 1982.
- [93] Anderson W., Pastor E., Butler B., Catchpole E., Dupuy J.L., Fernandes P., Guijarro M., Mendes-Lopes J.M. et Ventura J. « Evaluating models to estimate flame characteristics

- for free-burning fires using laboratory and field data ». *Forest Ecology and Management*, 234 p. S77, 2006.
- [94] Ononye A., Vodacek A. et Saber E. « Automated extraction of fire line parameters from multispectral infrared images ». *Remote Sensing of Environment*, 108 (2) p. 179–188, 2007.
- [95] Paugam R., Wooster M.J. et Roberts G. « Use of Handheld Thermal Imager Data for Airborne Mapping of Fire Radiative Power and Energy and Flame Front Rate of Spread ». *IEEE Transactions on Geoscience and Remote Sensing*, PP (99) p. 1–15, 2012.
- [96] Merino L. et Ollero A. « Computer vision techniques for fire monitoring using aerial images ». Dans *IEEE 2002 28th Annual Conference of the Industrial Electronics Society*, volume 3, p. 2214–2218. IEEE, 2002.
- [97] Martínez-de Dios J., Arrue B.C., Ollero A., Merino L. et Gómez-Rodríguez F. « Computer vision techniques for forest fire perception ». *Image and vision computing*, 26 (4) p. 550–562, 2008.
- [98] Merino L., Martinez R. et Ollero A. « Cooperative perception for fire monitoring ». Dans *Proceedings World Automation Congress*, volume 15, p. 435–440. TSI Press, 2004.
- [99] Den Breejen E., Roos M., Schutte K., De Vries J.S. et Winkel H. « Infrared measurements of energy release and flame temperatures of forest fires ». Dans *Proceedings of 3rd International Conference on Forest Fire Research*, p. 2003–2012. 1998.
- [100] Rossi L., Akhloufi M., Pieri A., Rossi J.L. et Molinier T. « Image processing and vision for the study and the modeling of spreading fires ». Dans *3rd International Conference on Image Processing Theory, Tools and Applications (IPTA)*, p. 47–52. 2012.
- [101] Trucco E. et Verri A. *Introductory techniques for 3-D computer vision*, volume 93. Prentice Hall Englewood Cliffs, 1998.
- [102] Rossi L., Molinier T., Pieri A., Akhloufi M., Tison Y. et Bosseur F. « Measurement of the geometric characteristics of a fire front by stereovision techniques on field experiments ». *Measurement Science and Technology*, 22 (12) p. 125504, 2011.
- [103] Rossi L., Molinier T., Akhloufi M., Pieri A. et Tison Y. « Instrumentation of spreading fires : toward the development of a metrological system based on stereovision ». Dans *Proc. 7th Mediterranean Combustion Symposium*, p. 11. 2011.
- [104] Besl P.J. et McKay N.D. « Method for registration of 3-D shapes ». Dans *Robotics-DL tentative*, p. 586–606. International Society for Optics and Photonics, 1992.
- [105] Robinson J.M. « Fire from space : Global fire evaluation using infrared remote sensing ». *International Journal of Remote Sensing*, 12 (1) p. 3–24, 1991.

- [106] Le Maoult Y., Sentenac T., Orteu J.J. et Arcens J.P. « Fire detection : a new approach based on a low cost CCD camera in the near infrared ». *Process Safety and Environmental Protection*, 85 (3) p. 193–206, 2007.
- [107] Briz S. « Reduction of false alarm rate in automatic forest fire infrared surveillance systems ». *Remote Sensing of Environment*, 86 (1) p. 19–29, 2003.
- [108] Boulet P., Parent G., Collin A., Acem Z., Porterie B., Clerc J. et Kaiss A. « Spectral emission of flames from laboratory scale vegetation fires ». *International Journal of Wildland Fire*, 18 (7) p. 875–884, 2009.
- [109] Billaud Y., Boulet P., Pizzo Y., Parent G., Acem Z., Kaiss A., Collin A. et Porterie B. « Determination of Woody Fuel Flame Properties by Means of Emission Spectroscopy Using a Genetic Algorithm ». *Combustion Science and Technology*, 185 (4) p. 579–599, 2013.
- [110] MIVIM. « Electromagnetic Spectrum ». <http://mivim.gel.ulaval.ca/dynamique/index.php?idD=58&Lang1>.
- [111] Rodriguez F. et Moreno Robles A. « D-09-04.The infrared imager : its use for wildland fire monitoring ». Cahier de recherche, EUROFIRELAB, 2004.
- [112] Wooster M.J., Roberts G., Smith A.M., Johnston J., Freeborn P., Amici S. et Hudak A.T. « Thermal Remote Sensing of Active Vegetation Fires and Biomass Burning Events ». Dans *Thermal Infrared Remote Sensing* p. 347–390. Springer, 2013.
- [113] Li Y., Vodacek A. et Zhu Y. « An automatic statistical segmentation algorithm for extraction of fire and smoke regions ». *Remote sensing of environment*, 108 (2) p. 171–178, 2007.
- [114] « Imagingsource DFK 72BUC02 specifications ». http://www.theimagingsource.com/en_US/products/cameras/usb-ccd-color/dfk72buc02/, accédé le 02/08/2015.
- [115] « JAI AD-080GE specifications ». <http://www.jai.com/en/products/ad-080ge>, accédé le 02/08/2015.
- [116] « CEPIP Jade3MW specifications ». <http://alacron.com/clientuploads/directory/Cameras/CEDIP/jadealtair.pdf>, accédé le 02/08/2015.
- [117] « Jenoptik IR-TCM 284 specifications ». <http://www.jenoptik.com/us-ir-tcm-384-infrared-camera-series>, accédé le 02/08/2015.
- [118] Bedo M.V., Blanco G., Oliveira W.D., Cazzolato M.T., Costa A.F., Rodrigues Jr J.F., Traina A.J. et Traina Jr C. « Techniques for effective and efficient fire detection from social media images ». Dans *17th International Conference on Enterprise Information System*, p. 34–45. 2015.

- [119] Péteri R., Fazekas S. et Huiskes M.J. « DynTex : a Comprehensive Database of Dynamic Textures ». *Pattern Recognition Letters*, 2010. <http://projects.cwi.nl/dyntex/>.
- [120] Zhao J., Zhang Z., Han S., Qu C., Yuan Z. et Zhang D. « SVM based forest fire detection using static and dynamic features ». *Computer Science and Information Systems*, 8 (3) p. 821–841, 2011.
- [121] WorkWithColor.com. « Color Names ». <http://www.workwithcolor.com/color-names-01.htm>. Accédé le 02/08/2015.
- [122] Bradski G. « The OpenCV Library ». *Dr. Dobb's Journal of Software Tools*, 2000.
- [123] Hooke R. et Jeeves T.A. « “Direct Search” Solution of Numerical and Statistical Problems ». *Journal of the ACM (JACM)*, 8 (2) p. 212–229, 1961.
- [124] Liu J. et Tu Y. « Thresholding of images using an efficient c-mean clustering algorithm ». *Journal of electronics & information technology*, 14 (4) p. 424–427, 1992.
- [125] Mitchell T.M. « Machine learning. WCB ». 1997.
- [126] Palmann C. *Contribution au recalage d'images de modalités différentes à travers la mise en correspondance de nuages de points : Application à la télédétection*. thèse de doctorat, Aix Marseille 2, 2011.
- [127] Bilodeau G.A., Torabi A. et Morin F. « Visible and infrared image registration using trajectories and composite foreground images ». *Image and Vision Computing*, 29 (1) p. 41–50, 2011.
- [128] Verstockt S. *Multi-modal video analysis for early fire detection*. thèse de doctorat, Ghent University, 2011.
- [129] Hartley R. et Zisserman A. *Multiple view geometry in computer vision*. Cambridge university press, 2003.
- [130] Liggins II M., Hall D. et Llinas J. *Handbook of multisensor data fusion : theory and practice*. CRC press, 2008.
- [131] Solberg A.S., Jain A.K. et Taxt T. « Multisource classification of remotely sensed data : fusion of Landsat TM and SAR images ». *IEEE Transactions on Geoscience and Remote Sensing*, 32 (4) p. 768–778, 1994.
- [132] Ducourthial B., Cherfaoui V. et Denoeux T. « Self-stabilizing distributed data fusion ». Dans *Stabilization, Safety, and Security of Distributed Systems* p. 148–162. Springer, 2012.
- [133] Jimenez L., Morales-Morell A. et Creus A. « Classification of hyperdimensional data based on feature and decision fusion approaches using projection pursuit, majority voting, and neural networks ». *Geoscience and Remote Sensing, IEEE Transactions on*, 37 (3) p. 1360–1366, 1999.

- [134] Lee T., Richards J.A. et Swain P.H. « Probabilistic and evidential approaches for multisource data analysis ». *Geoscience and Remote Sensing, IEEE Transactions on*, GE-25 (3) p. 283–293, 1987.
- [135] Smets P. et Kennes R. « The transferable belief model ». *Artificial intelligence*, 66 (2) p. 191–234, 1994.
- [136] Ramasso E., Rombaut M. et Pellerin D. « Modèle des Croyances Transférables : Représentation des connaissances, Fusion d’informations, Décision ». 2007.
- [137] Dempster A.P. « A generalization of Bayesian inference ». *Journal of the Royal Statistical Society. Series B (Methodological)*, p. 205–247, 1968.
- [138] Shafer G. *et al.* *A mathematical theory of evidence*, volume 1. Princeton university press Princeton, 1976.
- [139] Smets P. « Belief functions ». *Non-standard logics for automated reasoning*, p. 253–286, 1988.
- [140] Delmotte F. et Smets P. « Target identification based on the transferable belief model interpretation of Dempster-Shafer model ». *IEEE Transactions on Systems, Man and Cybernetics, Part A : Systems and Humans*, 34 (4) p. 457–471, 2004.
- [141] Smets P. *Un modèle mathématico-statistique simulant le processus du diagnostic médical*. thèse de doctorat, Univ. Libre Bruxelles, Belgium, 1978.
- [142] Whitaker R.T. « A level-set approach to 3D reconstruction from range data ». *International Journal of Computer Vision*, 29 (3) p. 203–231, 1998.
- [143] Gales G. *Mise en correspondance de pixels pour la stéréovision binoculaire par propagation d’appariements de points d’intérêt et sondage de régions*. thèse de doctorat, Université Toulouse III-Paul Sabatier, 2011. Thèse de doctorat dirigée par Dalle, Patrice.
- [144] Galindo P.A. *Image matching for 3D reconstruction using complementary optical and geometric information*. thèse de doctorat, Université de Lorraine, 2015. Thèse de doctorat dirigée par Lévy, Bruno et Zayer, Rhaleb.
- [145] Pohl C. et Van Genderen J. « Review article multisensor image fusion in remote sensing : concepts, methods and applications ». *International journal of remote sensing*, 19 (5) p. 823–854, 1998.
- [146] Zeng J., Sayedelahl A., Gilmore T., Frazier P. et Chouikha M. « Assessing Image Fusion Methods for Unconstrained Outdoor Scenes ». Cahier de recherche, DTIC Document, 2006.
- [147] Mitchell H.B. *Image fusion : theories, techniques and applications*. Springer, 2010.

- [148] Akhloufi M. *Reconnaissance des visages par imagerie multispectrale, du visible à l'infrarouge lointain*. thèse de doctorat, Faculté des Sciences et Génie, Université Laval, 2013.
- [149] Sadjadi F. « Comparative image fusion analysais ». Dans IEEE Computer Society Conference on Computer Vision and Pattern Recognition-Workshops, p. 8–8. IEEE, 2005.
- [150] Chavez P., Sides S.C. et Anderson J.A. « Comparison of three different methods to merge multiresolution and multispectral data- Landsat TM and SPOT panchromatic ». *Photogrammetric Engineering and remote sensing*, 57 (3) p. 295–303, 1991.
- [151] Burt P.J. et Kolczynski R.J. « Enhanced image capture through fusion ». Dans Proceedings of Fourth International Conference on Computer Vision, 1993, p. 173–182. IEEE, 1993.
- [152] Burt P.J. et Adelson E.H. « The Laplacian pyramid as a compact image code ». *IEEE Transactions on Communications*, 31 (4) p. 532–540, 1983.
- [153] Mallat S.G. « A theory for multiresolution signal decomposition : the wavelet representation ». *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11 (7) p. 674–693, 1989.
- [154] Rosten E. et Drummond T. « Machine learning for high-speed corner detection ». Dans Computer Vision–ECCV p. 430–443. Springer, 2006.
- [155] Harris C. S.M. « A combined corner and edge detector ». *4th Alvey Vision Conf*, 1998.
- [156] Shi J. et Tomasi C. « Good features to track ». Dans IEEE Computer Society Conference on Computer Vision and Pattern Recognition, p. 593–600. IEEE, 1994.
- [157] Brown M., Szeliski R. et Winder S. « Multi-image matching using multi-scale oriented patches ». Dans IEEE Computer Society Conference on Computer Vision and Pattern Recognition, volume 1, p. 510–517. IEEE, 2005.
- [158] Noble J.A. « Finding corners ». *Image and Vision Computing*, 6 (2) p. 121–128, 1988.
- [159] Lowe D.G. « Distinctive image features from scale-invariant keypoints ». *International journal of computer vision*, 60 (2) p. 91–110, 2004.
- [160] Mikolajczyk K. et Schmid C. « Scale & affine invariant interest point detectors ». *International journal of computer vision*, 60 (1) p. 63–86, 2004.
- [161] Schaffalitzky F. et Zisserman A. « Multi-view matching for unordered image sets, or “How do I organize my holiday snaps?” ». Dans Computer Vision—ECCV 2002 p. 414–431. Springer, 2002.

- [162] Strecha C., Bronstein A.M., Bronstein M.M. et Fua P. « LDAHash : Improved matching with smaller descriptors ». *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 34 (1) p. 66–78, 2012.
- [163] Van De Weijer J. et Schmid C. « Coloring local feature extraction ». Dans *Computer Vision—ECCV 2006* p. 334–348. Springer, 2006.
- [164] Chambon S. et Crouzil A. « Similarity measures for image matching despite occlusions in stereo vision ». *Pattern Recognition*, 44 (9) p. 2063–2075, 2011.
- [165] Petrosian A.A. et Meyer F.G. *Wavelets in signal and image analysis : from theory to practice*, volume 19. Springer Science & Business Media, 2001.
- [166] Hill P.R., Canagarajah C.N. et Bull D.R. « Image fusion using complex wavelets. ». Dans *BMVC*, p. 1–10. Citeseer, 2002.
- [167] Donate A., Liu X. et Collins Jr E.G. « Efficient path-based stereo matching with subpixel accuracy ». *IEEE Transactions on Systems, Man and Cybernetics, Part B : Cybernetics*, 41 (1) p. 183–195, 2011.
- [168] Szeliski R. et Scharstein D. « Symmetric sub-pixel stereo matching ». Dans *Computer Vision—ECCV 2002* p. 525–540. Springer, 2002.
- [169] Otto G.P. et Chau T.K. « ‘Region-growing’ algorithm for matching of terrain images ». *Image and vision computing*, 7 (2) p. 83–94, 1989.
- [170] Lhuillier M. et Quan L. « Match propagation for image-based modeling and rendering ». *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24 (8) p. 1140–1146, 2002.
- [171] Tsai R. et Huang T.S. « Uniqueness and Estimation of Three-Dimensional Motion Parameters of Rigid Objects with Curved Surfaces ». *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 6 (1) p. 13–27, 1984.
- [172] Marquardt D.W. « An algorithm for least-squares estimation of nonlinear parameters ». *Journal of the Society for Industrial & Applied Mathematics*, 11 (2) p. 431–441, 1963.
- [173] Zhang Z. « A flexible new technique for camera calibration ». *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22 (11) p. 1330–1334, 2000.
- [174] Longuet-Higgins H.C. *Readings in Computer Vision : Issues, Problems, Principles, and Paradigms*, chapitre A Computer Algorithm for Reconstructing a Scene from Two Projections, p. 61–62. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 1987.
- [175] Nistér D. « An efficient solution to the five-point relative pose problem ». *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26 (6) p. 756–770, 2004.

- [176] Stefanovic P. « Relative orientation—a new approach ». *ITC Journal*, 3 p. 417–448, 1973.
- [177] Datta A., Kim J.S. et Kanade T. « Accurate camera calibration using iterative refinement of control points ». Dans *IEEE 12th International Conference on Computer Vision Workshops*, p. 1201–1208. IEEE, 2009.
- [178] Bouguet J.Y. « Camera Calibration Toolbox for Matlab ». http://www.vision.caltech.edu/bouguetj/calib_doc/. Accédé le 12/08/2015.
- [179] Rossi L., Toulouse T., Akhloufi M., Pieri A. et Tison Y. « Estimation of spreading fire geometrical characteristics using near infrared stereovision ». Dans *IS&T/SPIE Electronic Imaging*, p. 86 500A–86 500A. 2013.
- [180] Bézier P. *L'utilisation des courbes et surfaces en CAO*. Hermès, 1988.
- [181] Delaunay B. « Sur la sphere vide ». *Izv. Akad. Nauk SSSR, Otdelenie Matematicheskii i Estestvennyka Nauk*, 7 (793-800) p. 1–2, 1934.
- [182] Gouraud H. « Continuous shading of curved surfaces ». *Computers, IEEE Transactions on*, 100 (6) p. 623–629, 1971.
- [183] Moretti B. *Modélisation du comportement des feux de forêt pour des outils d'aide à la décision*. thèse de doctorat, Université de Corse Pasqual Paoli, 2015. Thèse de doctorat dirigée par Balbi J.H. et Rossi J.L.
- [184] Bourke P. « Calculating the area and centroid of a polygon ». 1988.
- [185] Dove K.L. et Sumner J.S. « Tetrahedra with integer edges and integer volume ». *Mathematics Magazine*, p. 104–111, 1992.
- [186] Yuan C., Zhang Y. et Liu Z. « A Survey on Technologies for Automatic Forest Fire Monitoring, Detection and Fighting Using UAVs and Remote Sensing Techniques ». *Canadian Journal of Forest Research*, (ja), 2015.
- [187] Varcheie P.D.Z., Sills-Lavoie M. et Bilodeau G.A. « A multiscale region-based motion detection and background subtraction algorithm ». *Sensors*, 10 (2) p. 1041–1061, 2010.