

République Algérienne Démocratique et Populaire
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique
Université Saad Dahleb – Blida1



Faculté des Sciences
Département d'Informatique

Domaine : Mathématiques Informatique
Filière : Informatique
Spécialité : Système Informatique et Réseaux

Mémoire de fin d'études en vue de l'obtention du
Diplôme de Master

Thème

Conception et implémentation d'un système de soustraction de fond pour les besoins en analyse vidéo.

Présenté par :

LAMARI Khadidja Houda
YAMIR Meriem

Soutenu le 30/09/ 2019 Devant le jury composé de

Mr. Cherif Zahar Sid-Ahmed Amine : Président
Mr. OUELD AISSA Ahmed : Examineur
Mr. ZAIR Mustapha : promoteur
Mme. SETITRA Insaf : Encadreur

2018/2019

Remerciements

Au terme de ce travail nous tenons à remercier tout d'abord ALLAH tout puissant de nous avoir donné le courage pour la réalisation de ce travail.

Nous exprimons notre grande reconnaissance et remerciements à notre encadreuse ; Mme. SITTRA Insaf pour son aide, ses conseils et de nous avoir orienté et mis à notre disposition tous les moyens nécessaires durant notre travail.

Nous exprimons également notre gratitude à notre promoteur Mr. ZAIR Mustapha ainsi qu'à l'ensemble de nos jurys pour avoir accepté d'analyser ce travail.

En fin, que nos chers parents et amis, trouvent ici l'expression de nos remerciements les plus sincères et les plus profonds en reconnaissance de leurs sacrifices, aides, soutiens et encouragements pour atteindre nos buts.

Dédicaces

A mes très chers parents qui n'ont jamais cessé de m'aider et m'encourager durant tout mon parcours et qui m'ont permis d'être aujourd'hui ce que je suis.

A mes chers frères et soeurs, mon fiancé Fethi et à toute ma grande famille sans exception.

A mon amie, mon binôme Meriem que je remercie d'avoir été patiente et collaborative.

A toutes mes amies, qu'elles m'excusent de ne pas pouvoir les citer au risque d'oublier quelqu'un, je vous dis merci pour tous les moments inoubliables qu'on a passé ensemble durant ces cinq années.

A tous ceux qui me sont chers

Lamari khadidja Houda.

Dédicaces

Je dédie ce modeste travail qui n'aurait jamais pu voir le jour sans les soutiens indéfectibles et sans limite de ma chère maman qui ne cesse de me donner avec amour nécessaire pour que je puisse arriver à ce que je suis aujourd'hui. Que dieux vous protège et que la réussite soit toujours à ma portée pour que je puisse vous combler de bonheur.

Je dédie aussi ce travail à :

*Ma très chère sœur Khadidja et son mari
Mohamed.*

Mes chers frères Hamza et Zoubir.

*Toutes mes amies particulièrement Imane,
Soumia.*

*Mes oncles et tantes ainsi que leurs épouses,
époux et enfants.*

Tous mes chers cousins et cousines.

Yamir Meriem

Résumé

La soustraction de fond est un moyen courant pour la détection d'objets en mouvement dans des séquences vidéo. C'est la première étape pour différents types d'applications dans le domaine de la vision par ordinateur. La soustraction de fond est liée à des problèmes naturels et matériels qui rendent les résultats de la détection moins performants et moins précis.

Détecter les objets de premier plan avec moins de fautes est l'un des objectifs majeurs à atteindre. C'est dans ce contexte que la division Systèmes d'Information et Systèmes Multimédia DSISM du Centre de Recherche en Information Scientifique et Technique CERIST consacre ses efforts en vue d'améliorer les résultats de la soustraction de fond pour des besoins de vision par ordinateur, principalement la reconnaissance et la classification d'objets en mouvement.

Pour cela notre travail dans le cadre de ce PFE se propose de concevoir et d'implémenter un système exploitant des algorithmes de soustraction de fond existants les plus performants, de les comparer à travers notre système, et de proposer des améliorations possibles.

Nous proposons ainsi une interface ergonomique pour la comparaison des algorithmes, combinons ces algorithmes à un extracteur de contour, exposons les résultats avant et après combinaisons, et analysons quantitativement et qualitativement les résultats obtenus. Les comparaisons quantitatives consistent à analyser les taux de précision, et les comparaisons qualitatives consistent à visualiser les masques de soustraction de fond obtenus, analyser leurs formes, et vérifier à quel point ces formes ressemblent aux objets originels en mouvement.

Mots clés : vision par ordinateur, soustraction de fond, détection des objets en mouvement, détection de contour.

Abstract

Background subtraction is a common means for detecting moving objects in video sequences. This is the first step for different types of applications in the field of computer vision. Background subtraction is linked to natural and material problems that make detection results less efficient and less accurate.

Detecting foreground objects with fewer mistakes is one of the major goals to achieve. It is in this context that the Information Systems and Multimedia Systems Division DSISM of the Center for Research in Scientific and Technical Information CERIST devotes its efforts to improve the results of the subtraction of background for computer vision needs, mainly the recognition and classification of moving objects.

For this reason, our work in the framework of this project graduation proposes to design and implement a system exploiting the most powerful existing subtraction algorithms, to compare them through our system, and to propose possible improvements.

We thus propose an ergonomic interface for the comparison of the algorithms, combine these algorithms with a contour extractor, expose the results before and after combinations, and analyze quantitatively and qualitatively the results obtained. Quantitative comparisons consist in analyzing accuracy rates, and qualitative comparisons consist in visualizing the subtraction masks obtained, analyzing their shapes, and verifying how these shapes resemble the original objects in motion.

Key words: computer vision, background subtraction, detecting moving objects, contour detection.

ملخص

يُعد طرح الخلفية وسيلة شائعة للكشف عن الكائنات المتحركة في تنبعات الفيديو. هذه هي الخطوة الأولى لأنواع مختلفة من التطبيقات في مجال رؤية الكمبيوتر. يرتبط طرح الخلفية بالمشكلات الطبيعية والمادية التي تجعل نتائج الكشف أقل كفاءة وأقل دقة.

يعد اكتشاف الكائنات الأمامية مع وجود عدد أقل من الأخطاء أحد الأهداف الرئيسية المراد تحقيقها. في هذا السياق، يكرس قسم نظم المعلومات وأنظمة الوسائط المتعددة التابع لمركز البحوث في المعلومات العلمية والتقنية جهوده لتحسين نتائج طرح الخلفية لاحتياجات رؤية الكمبيوتر، التعرف على الأجسام وتصنيف الأجسام المتحركة.

لهذا السبب، يقترح عملنا في إطار مشروع التخرج تصميم وتنفيذ نظام يستعمل أقوى خوارزميات الطرح الحالية، لمقارنتها، واقتراح التحسينات الممكنة.

لذلك، نقترح وجود واجهة مريحة لمقارنة الخوارزميات، والجمع بين نواتجها من دون استعمال تقنية تحديد الحدود وباستعمالها، وكشف النتائج قبل الدمج وبعده، وتحليلها من الناحية الكمية والنوعية. تتمثل المقارنات الكمية في تحليل معدلات الدقة، والمقارنات النوعية تتمثل في رؤية نواتج الطرح التي تم الحصول عليها، وتحليل أشكالها، والتحقق من كيف تشبه هذه الأشكال الكائنات الأصلية المتحركة.

الكلمات المفتاحية: رؤية الكمبيوتر، طرح الخلفية، الكشف عن الأجسام المتحركة، تقنية تحديد الحدود.

Table des matières

Introduction générale	1
Chapitre 1 :Vision par ordinateur et soustraction de fond	
Introduction.....	4
1 La vision par ordinateur	4
1.1 Définition	4
1.2 Etapes de la vision par ordinateur	5
1.3 Domaines d'application de la vision par ordinateur.....	7
2 La soustraction de fond	7
2.1 Définitions.....	7
2.2 Etapes de soustraction de fond.....	9
2.2.1 Initialisation de l'arrière-plan	9
2.2.2 Modélisation de fond	10
2.2.3 Maintenance de l'arrière-plan.....	10
2.2.4 Détection du premier plan.....	10
2.3 Difficultés liées à la soustraction de fond	11
2.3.1 Difficultés dues à la caméra.....	11
2.3.2 Difficultés dues à l'environnement.....	11
2.3.3 Difficultés dues aux objets de premier plan.....	13
2.4 Domaines d'applications de la soustraction de fond.....	14
2.4.1 Surveillance visuelle des activités humaines	14
2.4.2 Calcul visuel de coque	14
2.4.3 Interaction homme-machine (IHM).....	14
2.4.4 Reconnaissance des gestes des mains basée sur la vision	14
2.4.5 Observation visuelle des comportements d'animaux et insectes.....	15
2.4.6 Observation visuelle des environnements naturels	15
2.4.7 Codage vidéo basé sur le contenu.....	15

Conclusion	15
-------------------------	-----------

Chapitre 2 : Etat De L'art

Introduction.....	16
--------------------------	-----------

1. Approches de base de la soustraction de fond	16
--	-----------

1.1 Approches non- récursives :.....	16
--------------------------------------	----

1.1.1 Différence de trames	16
----------------------------------	----

1.1.2 Filtre Médian.....	19
--------------------------	----

1.2 Approches récursives	19
--------------------------------	----

1.2.1 Moyenne Gaussienne Mobile (Running Gaussian average).....	19
---	----

1.2.2 Mélange de Gaussiennes (Mixture of Gaussians MoG)	19
---	----

1.2.3 Filtre de Kalman	20
------------------------------	----

1.2.4 Filtre médian approximé (Approximated median filtering AMF).....	21
--	----

2 Approches récentes de soustraction de fond	21
---	-----------

2.1 LOBSTER	21
-------------------	----

2.2 SuBSENSE.....	23
-------------------	----

2.2.1 Modélisation au niveau des pixels et extraction des informations	23
--	----

2.2.2 Détection de changement via un consensus.....	24
---	----

2.2.3 Surveillance de l'arrière-plan et schéma de commentaires.....	25
---	----

2.3 Type-2 Fuzzy GMM-UV with MRF et Type-2 Fuzzy GMM-UM with MRF .	25
--	----

2.3.1 Maintenance du modèle	26
-----------------------------------	----

2.3.2 Détection de l'avant plan	26
---------------------------------------	----

2.3.3 Optimisation des champs d'étiquetage.....	27
---	----

2.3.4 Calcul de la fonction de vraisemblance (Likelihood Function).....	27
---	----

2.3.5 Probabilité postérieure	27
-------------------------------------	----

2.4 IMBS	28
----------------	----

2.4.1 Suppression de l'ombre	29
------------------------------------	----

2.4.2 Modèle de mise à jour.....	29
----------------------------------	----

2.4.3 Nouveautés de la méthode	30
--------------------------------------	----

2.5 MultiCue BGS.....	30
-----------------------	----

2.5.1 Phase initiale	30
----------------------------	----

2.5.2	Phase de modélisation de fond.....	30
2.5.3	Phase de la soustraction de fond	31
2.5.4	Phase de la maintenance du model	31
2.6	Détection d'avant plan par texture et MRF	32
2.7	Adaptatives SOM	33
➤	Sensibilité.....	34
➤	Taux d'apprentissage	34
➤	Seuil.....	34
➤	Bruit de Variance	34
2.8	Fuzzy Gaussian	34
2.8.1	Modélisation d'arrière-plan flou :.....	35
2.8.2	Détection de premier plan floue :.....	35
2.8.3	Maintenance d'arrière-plan flou	35
2.8.4	Post-traitement flou.....	35
2.9	Comparaison entre les algorithmes	35
Conclusion		37

Chapitre 3 : Conception

Introduction	39
1 Solution proposée	40
1.1 Phase 1 soustraction de fond basique	40
1.2 Phase 2 extraction de contour	40
1.3 Phase 3 soustraction de fond sur images de contour.....	41
1.4 Phase 4 combinaison de masques	41
1.5 Phase 5 calcul de métriques	41
1.6 Phase 6 Analyse quantitative et qualitative	41
2 Objectifs.....	41
3 Modélisation de l'application avec UML.....	42
3.1 Diagramme de cas d'utilisation.....	42
3.2 Diagramme de séquence.....	44
3.3 Diagramme de classe.....	48

Conclusion	49
Chapitre 4 : Implémentation et résultats	
Introduction.....	55
1 Environnement Matériel	55
2 Environnement Logiciel	55
2.1 Outil de modélisation UML	56
2.2 Environnement de développement et librairies.....	56
2.3 Langages de développement	58
2.4 Editeurs.....	58
2.5 Outils de travail collaboratif.....	59
❖ Git Hub.....	59
3 Interface de l’application	59
3.1 Interface d’accueil (soustraction de fond).....	59
3.2 Interface de contour.....	60
3.2 Interface d’addition	61
3.3 Interface de calcul	62
4 Évaluation des performances	63
5 Base d’images utilisée	64
5.1 CDnet 2012	64
5.2 CDnet 2014	65
5.3 Masque d’avant plan Vérité terrain.....	66
6 Expérimentation.....	67
6.1 Résultat d’évaluation des algorithmes de soustraction de fond pour CDnet 2014	68
6.2 Résultats d’évaluation des algorithmes de soustraction de fond après détection	
de contour	70
Conclusion	77
CONCLUSION GENERALE.....	78
RÉFÉRENCES BIBLIOGRAPHIQUE	
BIBLIOGRAPHIE	80

Liste des figures

Chapitre 1 : Vision par ordinateur et soustraction de fond

Figure 1- 1 : La vision par ordinateur [4].	5
Figure 1- 2 : Le processus de restauration [8].	6
Figure 1- 3 : Le modèle de fond à l'instant 600s de base d'images CDnet2014 catégorie baseline	8
Figure 1- 4 : L'instant 601s de la base d'images CDnet2014 catégorie baseline.	8
Figure 1- 5 : Le masque d'avant-plan à l'instant 601s de la base d'images CDnet 2014 catégorie baseline.	9
Figure 1- 6 : Diagramme représentant le processus de soustraction de fond [22].	11
Figure 1- 7: Le modèle de fond à l'instant 951s de la base d'images CDnet2014 catégorie camera.	12
Figure 1- 8 : La figure représentant l'instant 951s de la base d'images CDnet2014 catégorie cameraJitter.	12
Figure 1- 9 : L'avant -plan résultant par l'algorithme AGMM de la base d'images CDnet2014 catégorie cameraJitter.	13

Chapitre 2 : Etat de l'art

Figure 2- 1 : La figure représentant l'instant 629s de la base d'images CDnet2014 catégorie baseline.	18
Figure 2- 2 : La figure représentant l'instant 630s de la base d'images CDnet2014 catégorie baseline.	18
Figure 2- 3: le masque d'avant plan détecté par la méthode différence de trames de la catégorie baseline de la base d'images CDnet2014.	19
Figure 2- 4 : Les étapes de la méthode Texture et MRF [97].	33

Chapitre 3 : Conception

Figure 3- 1 : Diagramme de cas d'utilisation globale.....	43
Figure 3- 2 : Diagramme de cas d'utilisation « lancer la soustraction de fond »	44
Figure 3- 3 : Diagramme de séquence pour le cas « lancer la soustraction de fond ».....	45
Figure 3- 4 : Diagramme de classe.....	49

Chapitre 4 : Implémentation et résultat

Figure 4- 1: Interface d'accueil.	60
Figure 4- 2: interface de contour de fond.	61
Figure 4- 3: interface d'addition de deux images.....	62
Figure 4- 4 : interface de calcul des métriques.....	62
Figure 4- 5 : L'image représente l'instant 478s de CDnet 2014 catégorie baseline.....	66
Figure 4- 6 : Exemples d'images de chaque séquence vidéo de la base CDnet 2014. (a) mauvais temps BDW, (b) la ligne de base BSL, (c) instabilité de la caméra CJT, (d) fond dynamique DBG, (e) mouvement d'objet intermittent IOM, (f) faible taux de trames LFM, (g) vidéos de nuit NVD, (h) PTZ, (i) ombre SHD, (j) thermique THML, (k) turbulence TRBL.	67
Figure 4- 7 : Résultats qualitatifs de soustraction de fond obtenus à partir du subsense, IMBS, Fuzzy Gaussien, adaptive SOM, GMM uv with MRF, Texture, multicue BGS, LOBSTER, les images vidéo sont extraites de la catégorie mauvais temps BDW. (GT) est le masque de vérité terrain (ground truth).	70
Figure 4- 8 : Résultats qualitatifs de soustraction de fond de la méthode SUBSENSE, Multicue et Texture with MRF avant et après extraction du contour sur les séquences vidéos de CDnet 2014 pour la catégorie BDW.....	73

Figure 4- 9 : Résultats qualitatifs de soustraction de fond des algorithmes IMBS, LOBSTER, GMM uv with MRF et Fuzzy Gaussien, sur les séquences de vidéos de CDnet 2014 pour la catégorie BDW.....	73
Figure 4- 10 : Résultats qualitatifs de soustraction de fond des algorithmes GMM um with MRF et adaptive SOM sur les séquences vidéos de CDnet 2014 pour la catégorie BDW.....	74
Figure 4- 11 : Différence au niveau qualitatif entre la soustraction de fond normale et la combinaison entre le résultat de soustraction de fond sur les séquences vidéo originales de CDnet, sur les séquences après extraction du contour en utilisant le filtre de Sobel et combinaison des deux résultats. Méthodes LOBSTER et Subsense.....	75
Figure 4- 12 : Différence au niveau qualitatif entre la soustraction de fond normale et la combinaison entre le résultat de soustraction de fond sur les séquences vidéo originales de CDnet, sur les séquences après extraction du contour en utilisant le filtre de Sobel et combinaison des deux résultats. Méthodes IMBS et Texture.....	76
Figure 4- 13 : Différence au niveau qualitatif entre la soustraction de fond normale et la combinaison entre le résultat de soustraction de fond sur les séquences vidéo originales de CDnet, sur les séquences après extraction du contour en utilisant le filtre de Sobel et combinaison des deux résultats. Méthode IMBS.....	76

Liste des tableaux

Chapitre 2 : Etat de l'art

Tableau 2- 1 :Résultats approxumatives de l'usage de la mémoire et de la vitesse de processeur dans le cas ou les paramètres sont similaires[106].....	22
Tableau 2- 2 : Résultats des deux métriques F-mesure et la précision pour les deux méthodes Vibe et LOBSTER[106]......	23
Tableau 2- 3 : Comparaison entre les méthodes non-paramétrique en utilisant la base d'images CDnet 2012[84].....	25
Tableau 2- 4 :comparaison entre les neuf algorithmes de bgslibrary.....	37

Chapitre 4 : Implémentation et résultat

Tableau 4- 1 :Comparaison de la métrique FM (F Mesure) de différents algorithmes de soustraction de fond sur les vidéos originales des 11 catégories du CDnet2014.	69
Tableau 4- 2 :Évaluation de la méthode SUBSENSE sur les vidéos originales des 11 catégories du CDnet2014.	69
Tableau 4- 3 :Comparaison de la métrique FM de différents algorithmes de soustraction de fond sur les vidéos après l'extraction de contour des 11 catégories du CDnet2014.	71
Tableau 4- 4 :Évaluation de la méthode SUBSENSE sur les vidéos des 11 catégories du CDnet2014 après l'extraction du contour.	72
Tableau 4- 5 :Comparaison de métrique FM de combinaison de soustraction de fond et la détection de contour sur les 11 catégories de CDnet2014.	74

Introduction Générale

Introduction générale

La détection d'un objet dans une scène de caméra vidéo est un domaine de recherche relativement nouveau en informatique et, en raison de sa large applicabilité dans la vie réelle, cela a augmenté de plus en plus. La vidéosurveillance est une des principales raisons de l'intérêt croissant et de l'utilisation de la vidéo dans les systèmes de sécurité. La détection d'objet en mouvement dans un flux vidéo est une étape essentielle dans les applications de vidéosurveillance. De nombreux algorithmes de détection des objets en mouvement ont été utilisés et proposés ces dernières années. Il s'agit de soustraction de fond, de flot optique, de différence temporelle et de nombreux autres algorithmes pour détecter les objets en mouvement. De ceux-ci, l'algorithme le plus largement utilisé est la soustraction de fond.

Pour diverses applications de vision par ordinateur, la soustraction de fond est une méthode « rapide et efficace » pour localiser des objets en mouvement dans une vidéo filmée par une caméra statique. Dans cette perspective, la détection de mouvement est souvent la première étape d'un système de vision par ordinateur à plusieurs étapes tel que le suivi de voitures, la reconnaissance de plaques d'immatriculation, la reconnaissance de personnes, etc. Pour cette raison, il est généralement aussi rapide et simple que possible. Par conséquent, la plupart des méthodes de soustraction de fond indiquent un mouvement pour chaque pixel à l'instant t dont la couleur est significativement différente de celles du contexte. Cette solution a fait ses preuves lorsque la caméra est rigoureusement statique avec un fond fixe sans bruit.

Cependant, détecter un mouvement par soustraction de fond n'est pas une tâche simple. En effet, certaines vidéos avec un rapport signal sur bruit faible provoqué par une mauvaise qualité de la caméra, des artefacts de compression ou un environnement bruyant, sont susceptibles de générer de nombreux faux points positifs. Les faux positifs peuvent également être induits par des changements d'éclairages (graduels ou soudains), un fond animé (vagues sur l'eau, arbres secoués par le vent) ou instabilité de l'appareil photo pour en nommer quelques-uns. D'autre part, des faux négatifs peuvent également se produire lorsqu'un objet en mouvement est composé de couleurs similaires à celles de l'arrière-plan (effet de camouflage).

Avec de tels scénarios, une simple différence inter-images avec un seuil global se révèle être une solution faible. Pour faire face à ces défis, de nombreux modèles d'arrière-plan et mesures de distance liées à différents schémas d'optimisation ont été proposés dans la dernière décennie. Ces méthodes sont (du moins en théorie) plus robustes au bruit et aux

instabilités de l'arrière-plan que les approches de soustraction de fond de base. Mais sont-elles vraiment ? Et de combien sont-elles meilleures ?

Dans ce mémoire, nous comparons un ensemble de quelques-unes de ces méthodes de soustraction de fond sur différentes séquences vidéo réelles de la base d'images *CDnet 2014*. La première étape à suivre est d'abord étudier l'ensemble des algorithmes de soustraction de fond, puis de proposer une interface ergonomique et maintenable permettant un choix de ces méthodes, et enfin, explorer la possibilité d'intégrer les contours aux algorithmes de soustraction étudiés et comparer les résultats obtenus.

Le but de cette étude est donc de :

- Etudier des algorithmes de soustraction de fond récents.
- Etudier la possibilité d'améliorer les résultats obtenus des algorithmes de soustraction de fond étudiés en combinant la détection de contour.
- Proposer une application desktop permettant de :
 - Choisir un algorithme de soustraction de fond et une vidéo et lancer l'opération de soustraction de fond.
 - Sauvegarder le masque de soustraction de fond résultant.
 - Comparer le résultat obtenu de chaque méthode avec sa vérité terrain, calculer les métriques de précision et sauvegarder les résultats obtenus.
 - Extraire le contour de vidéos.
 - Appliquer les algorithmes de soustraction de fond aux images de contours et comparer les résultats obtenus.
 - Combiner les résultats obtenus des méthodes sur les images originelles et les images de contour et comparer les résultats quantitativement et qualitativement.
- Interpréter tous les résultats obtenus.

Ce présent rapport est composé de quatre chapitres :

Dans le premier chapitre nous présenterons les généralités de la vision par ordinateur ainsi que le sujet principal de ce travail : la soustraction de fond.

Le deuxième chapitre fera l'objet d'un l'Etat de l'art en décrivant clairement le fonctionnement des méthodes choisies pour comparaisons.

Dans le troisième chapitre nous présentons la conception du système de soustraction de fond en donnant la description de la conception exprimée à l'aide des diagrammes d'UML.

C'est dans le quatrième chapitre que nous citerons des différents outils de développement utilisés dans la réalisation de notre application ainsi que l'environnement de développement et quelques captures d'écrans qui permettront de présenter notre application ainsi que les résultats obtenus. Cela suivi de l'ensemble de nos expérimentations en utilisant notre système, et une critique des résultats obtenus.

Chapitre 1 : Vision par ordinateur et soustraction de fond

Introduction

Depuis quelques années, la vision par ordinateur témoigne d'un intérêt croissant, d'une part du fait des technologies meilleur marché et de plus en plus compétentes, et d'autre part des besoins en sécurité et télésurveillance.

La soustraction de fond est une étape cruciale dans de nombreux systèmes de vision par ordinateur, car elle est d'abord appliquée pour détecter des objets en mouvement dans un flux vidéo

La première partie de ce chapitre nous permettra d'étudier les notions de base de la vision par ordinateur ainsi que ses domaines d'applications. La deuxième partie sera consacrée à la soustraction de fond ; ses problèmes, ses étapes et ses domaines d'application.

1 La vision par ordinateur

1.1 Définition

La vision est un processus de traitement de l'information, elle utilise des stratégies bien définies pour atteindre ses buts, l'entrée d'un système de vision est constituée d'une séquence d'images, la sortie est une description de l'entrée en termes d'objets et de relations entre ces objets [1].

Il existe d'autres définitions de la vision par ordinateur :

Selon Richard Szeleski [2], la vision par ordinateur est de décrire le monde que nous voyons dans une ou plusieurs images par des propriétés, telles que la forme, l'éclairage et les distributions de couleur.

Wesley E. Snyder et Hairong Qi [3] ont défini la vision par un processus avec lequel une machine, généralement un ordinateur numérique, traite automatiquement une image et/ou signal et reconnaît le contenu de l'image.

Dans un système classique de vision par ordinateur, un support de capture tel qu'une caméra capture la scène, puis l'envoie à un terminal de traitement, souvent un ordinateur. De plus, un système de vision par ordinateur touche à plusieurs concepts, nous en citons dans ce qui suit quelques-uns.

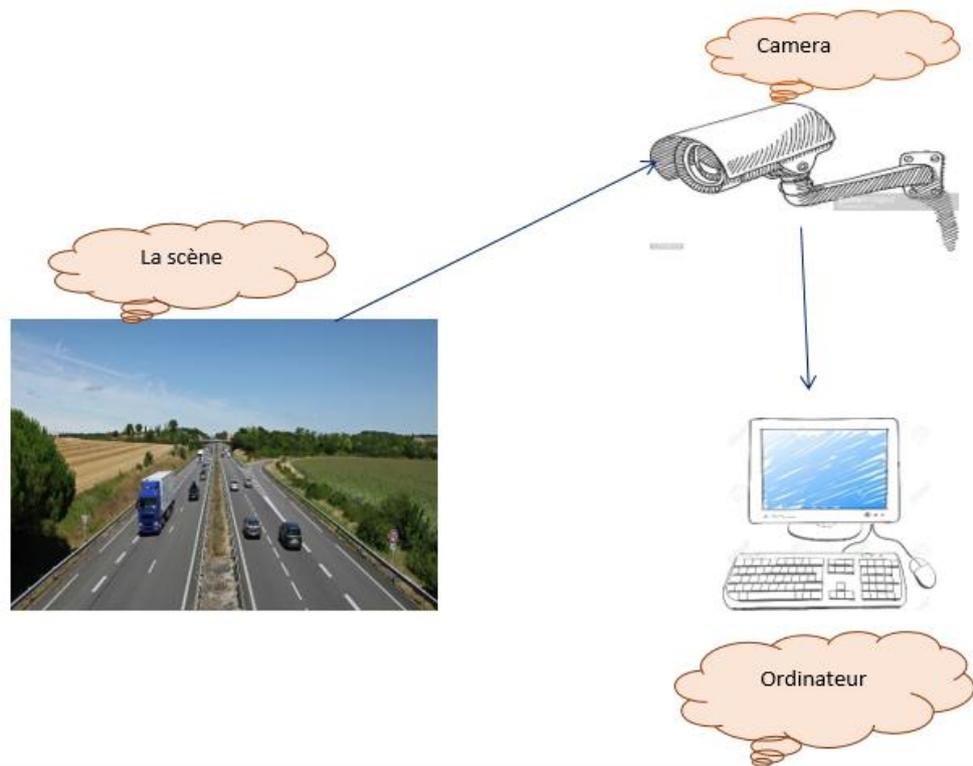


Figure 1- 1 : La vision par ordinateur [4].

1.2 Etapes de la vision par ordinateur

- a. **Le capteur de vision** : le capteur permet la perception du monde extérieur [4].
- b. **Le traitement d'images** : analyse, traite et interprète la scène capturée par le capteur de vision.
- c. **La compression d'images** : la compression d'images numériques est la technologie de réduction du débit de l'image. Celle-ci permet d'économiser de l'espace de stockage et de réduire le débit de transmission [5].
- d. **La restauration d'images** : la restauration d'image [6] est un domaine classique du traitement numérique de l'image. Le but de la restauration d'image [7] est d'éliminer ou de réduire numériquement les effets de dégradation tel que le flou et le bruit, sans augmenter la taille d'image. La restauration d'image comporte deux techniques ; les techniques du domaine spatial et les techniques du domaine fréquentiel [8]. La Figure 1-2 illustre le processus de restauration.

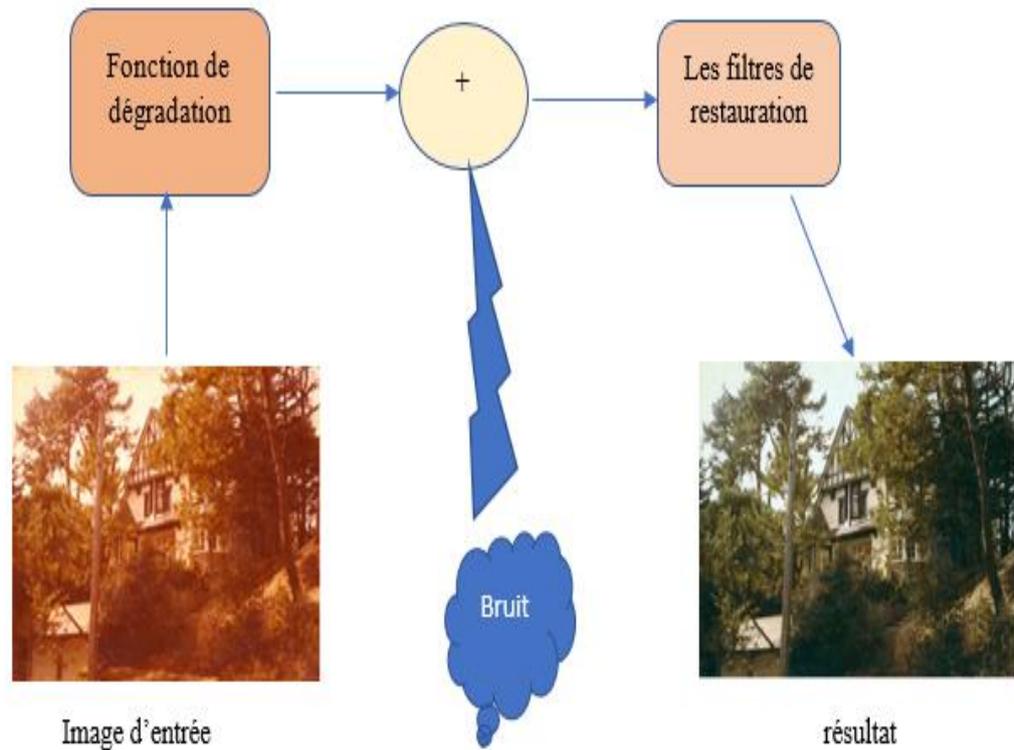


Figure 1- 2 : Le processus de restauration [8].

- e. **L'analyse d'image** : l'analyse d'images est un outil fondamental pour la reconnaissance, la différenciation et la quantification de divers types d'images [9], elle renvoie généralement des valeurs numériques et / ou des informations graphiques sur les caractéristiques de l'image. Le terme analyse d'image est utilisé lorsque le résultat est un nombre ou une décision, et non une image [10].
- f. **La segmentation d'image** : La segmentation d'image consiste à exprimer l'image sous forme de région connectée physiquement significative [11]. C'est une technique importante en vision par ordinateur, vise à partitionner une image en plusieurs images segments basées sur certaines propriétés telles que l'homogénéité de couleurs ou les motifs de texture similaires [12].
- g. **L'interprétation d'image** : c'est le passage de la description structurelle à la description sémantique en gardant certains objectifs (mesure de paramètres sur des

formes, description du contenu de la scène en matière de concepts non mathématiques) [4].

1.3 Domaines d'application de la vision par ordinateur

Il existe plusieurs domaines d'application de la vision par ordinateur, les deux principaux domaines du développement de la vision par l'ordinateur sont le domaine médical et le domaine militaire :

- **Militaire** : exemple (Détection de soldats ou de véhicules ennemis et guidage de missiles) [13].
- **Médicales** : exemple (la Cytologie, la Tomographie, l'Echographie) [4].
- **Industriel** : exemple (le contrôle de la qualité des détails ou des produits finis automatiquement inspecté afin de trouver des défauts,) [13].
- **Multimédia** : domaine multimédia : exemple (la compression et la décompression d'images, le stockage, l'amélioration de la qualité, etc.) [2].
- **Domaine d'agriculture** : exemple (enlever les aliments indésirables) [13].
- **Domaine de Surveillance** : exemple (la surveillance des intrus, analyse du trafic routier, et la surveillance des piscines pour les victimes de noyade) [2].

2 La soustraction de fond

2.1 Définitions

La méthode de soustraction de fond (ou soustraction d'arrière-plan) est considérée comme l'une des méthodes les plus fiables pour la détection d'objets en mouvement, elle est largement utilisée pour détecter le déplacement des objets dans des vidéos à partir d'une caméra statique [14]. C'est une étape de prétraitement importante dans de nombreux systèmes de vision par ordinateur modernes [15] tels que l'analyse vidéo, le suivi d'objets, la surveillance vidéo, le comptage d'objets et l'analyse du trafic. Selon [16] la soustraction de fond est la première étape pour détecter les objets en mouvement dans la scène.

Dans une scène, le processus de détection d'objets peut généralement être représenté avec un modèle appelé le modèle de fond (Figure 1-3), l'algorithme associé détecte les écarts par rapport au modèle de fond pour chaque trame entrante (Figure 1-4), résultant d'un masque d'avant plan (Figure 1-5).



Figure 1- 3 : Le modèle de fond à l'instant 600s de base d'images CDnet2014 catégorie baseline



Figure 1- 4 : L'instant 601s de la base d'images CDnet2014 catégorie baseline.

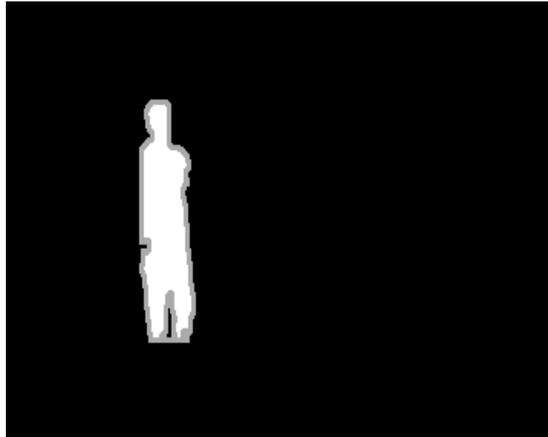


Figure 1- 5 : Le masque d'avant-plan à l'instant 601s de la base d'images CDnet 2014 catégorie baseline.

L'ancienne forme de cette méthode s'appelle parfois différentiation de trames (frame différenciation) [17]. Dans cette méthode, si la scène est immobile ou évoluant progressivement, la détection de premier plan peut être résolue [18]. Les méthodes de soustraction de fond trouvent des écarts par rapport à la représentation de la scène, ce qui pose des problèmes pour la détection précise des régions contenant des objets en mouvement.

Les problèmes les plus fréquents de la soustraction de fond sont les changements d'éclairage graduels ou soudains, le mouvement sans intérêt tel que les branches ou les vagues, le camouflage, la relocalisation de l'objet d'arrière-plan et l'initialisation avec des objets en mouvement [15].

Pour sa définition, selon [17] La soustraction de fond est une composante essentielle d'applications où l'objectif est de séparer le premier plan de l'arrière-plan. Selon [18] la soustraction de fond joue un rôle important dans l'analyse des images mouvement dans les systèmes de vidéosurveillance.

2.2 Etapes de soustraction de fond

Le processus de soustraction de fond est illustré dans la Figure1-6, il comprend les étapes suivantes :

2.2.1 Initialisation de l'arrière-plan

Cette étape également appelée la génération de fond, extraction de fond ou reconstruction de fond [19], consiste à calculer la première image de fond. Le module d'initialisation d'arrière-plan fournit la première image d'arrière-plan de N images.

2.2.2 Modélisation de fond

D'après [20] la modélisation de fond décrit le type de modèle utilisé pour représenter le fond et détermine principalement la capacité du modèle à traiter un ou plusieurs modèles d'arrière-plans.

Le moyen le plus simple de modéliser l'arrière-plan consiste à acquérir une image d'arrière-plan ne contenant aucun objet en mouvement.

Dans certains environnements, l'arrière-plan n'est pas disponible et peut toujours être modifié dans des situations critiques, telles que des modifications de l'éclairage ou des objets introduits supprimés de la scène.

Il existe 5 catégories de modélisation de l'arrière-plan :

- La modélisation de fond de base.
- La modélisation statistique de fond.
- La modélisation de fond via le clustering.
- La modélisation de fond via les réseaux neuronaux.
- L'estimation de fond.

2.2.3 Maintenance de l'arrière-plan

C'est une mise à jour de l'image d'arrière-plan en utilisant l'arrière-plan précédent, l'image actuelle et le masque de détection au premier plan pour que le modèle s'adapte aux changements qui apparaissent au fil du temps.

2.2.4 Détection du premier plan

La détection de premier plan est une étape clé dans le problème de soustraction d'arrière-plan [21].

Cette approche consiste à détecter des objets en mouvement provenant de caméras statiques via un processus de classification des pixels en avant-plan ou en arrière-plan.

Le processus de classification consiste à comparer l'image d'arrière-plan et l'image actuelle.

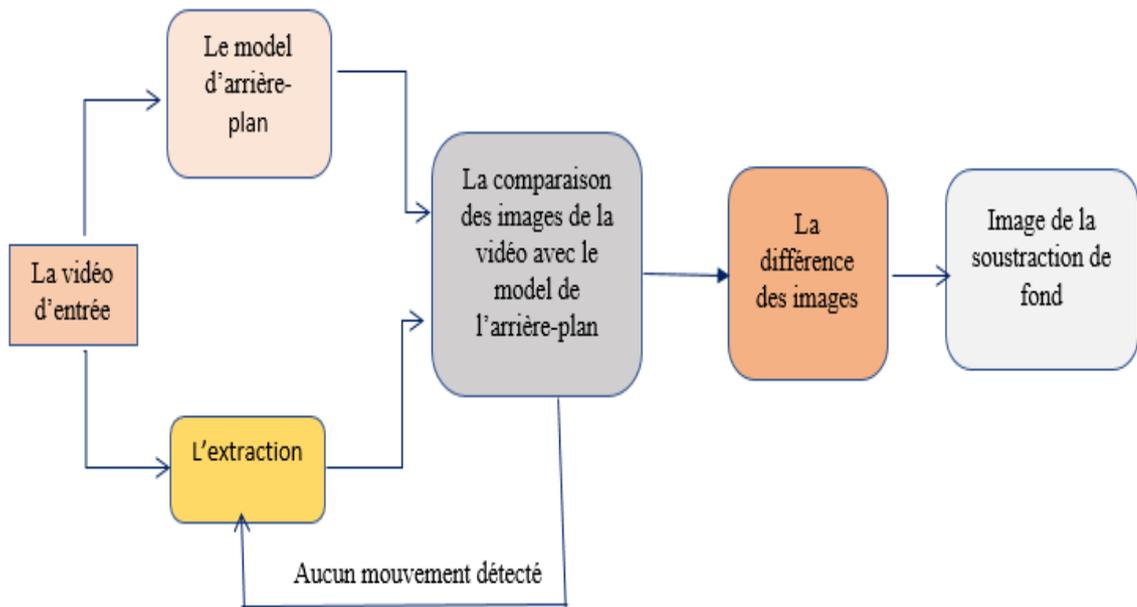


Figure 1- 6 : Diagramme représentant le processus de soustraction de fond [22].

2.3 Difficultés liées à la soustraction de fond

Les difficultés liées à la soustraction de fond sont énumérées dans ce qui suit.

2.3.1 Difficultés dues à la caméra

La mauvaise qualité de la vidéo, les artefacts de compression ainsi que la faible fréquence d'image ne permet pas d'étudier le déplacement des objets.

2.3.2 Difficultés dues à l'environnement

a. Changements soudains d'éclairage

L'arrière-plan s'adapte aux changements graduels de l'apparence de l'environnement. Par exemple, à l'extérieur, l'intensité lumineuse varie généralement pendant la journée. Ce type de changement se produit par exemple avec l'allumage / l'extinction soudaine d'une lumière dans un environnement intérieur [14]. Cela peut également se produire lors de scènes en extérieur (transition rapide de temps nuageux à lumineux), lumière du soleil. L'éclairage affecte fortement l'apparence de l'arrière-plan et provoque des détections faussement positives. Le modèle d'arrière-plan devrait en tenir compte [23].

b. Mouvement global de l'image

La charge du vent ou d'autres sources de mouvement appliqué sur la caméra, provoque un mouvement global dans les images. Ces causes résultent des fausses détections sur toute l'image dans un algorithme typique de soustraction d'arrière-plan (Figure 1-7). Le mouvement dans certaines parties de l'arrière-plan par exemple : des branches d'arbres se déplaçant avec le vent, ou eau ondulante. Ce type de mouvement d'arrière-plan provoque une variation des valeurs d'intensité de pixel de manière significative avec le temps [14].



Figure 1- 7: Le modèle de fond à l'instant 951s de la base d'images CDnet2014 catégorie camera



Figure 1- 8 : La figure représentant l'instant 951s de la base d'images CDnet2014 catégorie cameraJitter.



Figure 1- 9 : L'avant -plan résultant par l'algorithme AGMM de la base d'images CDnet2014 catégorie cameraJitter.

c. Camouflage

La plupart des méthodes de la soustraction de fond opèrent sur les intensités ou les couleurs des pixels, ce qui est un grand problème dans le domaine de surveillance. En effet, il existe de fortes chances que les objets de l'avant plan aient la même couleur que l'arrière-plan, ce qui provoque un faux classement. L'algorithme de soustraction de fond dans ce cas précis, considère les objets de premier plan comme objets d'arrière- plan [24].

d. Fond dynamique

Certaines parties du paysage peuvent contenir des mouvements (une fontaine, des mouvements de nuages, l'agitation des branches des arbres, les vagues d'eau, etc.), et doivent être considérés comme fond, selon leur pertinence.

Ce mouvement peut être périodique ou irrégulier (par exemple, les feux de circulation ou les arbres ondulants). La gestion de cette dynamique de fond est une tâche difficile [14].

e. Mauvais temps

La difficulté du mauvais temps survient dans le cas où, les vidéos souffrent des de conditions des intempéries hivernales telles que la neige tombante et la mauvaise visibilité. Ceci provoque une soustraction de fond difficile.

2.3.3 Difficultés dues aux objets de premier plan

Les objets en mouvement peuvent avoir des couleurs similaires à celles de la route et des ombres, ce qui fait que l'arrière-plan peut être faussement détecté en tant qu'objet ou vice-versa. Aussi, les véhicules peuvent s'arrêter occasionnellement aux intersections en raison de feux de circulation ou de signaux de contrôle. Ce genre d'arrêts transitoires augmente le

poids des non-fonds Gaussien et dégrade sérieusement la qualité d'estimation de fond d'un trafic des séquences d'images [25].

2.4 Domaines d'applications de la soustraction de fond

La soustraction de fond est utilisée dans plusieurs domaines, en voici quelques-uns :

2.4.1 Surveillance visuelle des activités humaines

L'objectif est d'identifier et de suivre des objets d'intérêts dans plusieurs environnements. Tels que les magasins pour connaître les habitudes des clients [26], la reconnaissance des personnes parmi d'autres et la biométrie en font partie [27], la sécurité dans les gares et les aéroports, où les bagages laissés sans surveillance peuvent être un objectif essentiel, surveillance des colis [28], la surveillance des scènes maritimes pour compter le nombre de navires en circulation dans un port [29].

L'analyse de trafic urbain afin de détecter des incidents tels que des véhicules immobilisés sur des autoroutes [30] [31], la détection de stationnement illégal, les places de stationnement libres, et congestion [32] [33]. La soustraction de fond est également utilisée dans le sport lorsque des décisions importantes doivent être prises rapidement, comme dans le football et le tennis avec "Hawk-Eye2 » [40] [41] [42].

2.4.2 Calcul visuel de coque

La coque visuelle est utilisée pour la synthèse d'images afin d'obtenir un modèle géométrique approximatif d'un objet, qu'il soit statique ou non [43].

2.4.3 Interaction homme-machine (IHM)

Plusieurs applications, nécessitent une interaction homme-machine, comme dans les arts [40], les jeux et les applications [44] [45] [46]. Dans le cas des jeux, le joueur peut observer sa propre image ou silhouette composée dans une scène virtuelle comme dans PlayStation Eye-Toy.

2.4.4 Reconnaissance des gestes des mains basées sur la vision

Cette application nécessite de détecter, suivre et reconnaître les gestes des mains pour plusieurs applications telles que l'interface homme-machine, les études de comportement, l'interprétation et l'apprentissage du langage des signes, la téléconférence, la formation à distance, la robotique, la sélection de jeux et la manipulation d'objets.

2.4.5 Observation visuelle des comportements d'animaux et insectes

L'interaction des animaux avec leurs environnement (plantes, etc.), comme dans le cas des oiseaux, afin de disposer d'informations sur l'impact du changement climatique sur l'écosystème, et étudier le comportement des animaux dans différentes conditions météorologiques, comme dans le cas de poissons en présence de typhon, de tempêtes ou de courants marins [34] [35] [36].

2.4.6 Observation visuelle des environnements naturels

L'objectif est de détecter des objets étrangers dans des environnements naturels tels que les forêts, les océans et les rivières afin de protéger la biodiversité de la faune et de la flore, exemple : des objets étrangers dans une rivière ou l'océan peuvent être des bouteilles flottantes [37] [38] [39].

2.4.7 Codage vidéo basé sur le contenu

Dans la téléconférence, les films numériques et les visiophones, seules les images clés sont transmises avec les objets en mouvement. Par exemple, la norme de communication multimédia MPEG-4[49] active la fonctionnalité basée sur le contenu en utilisant le VOP (Vidéo Object plane) comme élément de codage de base [50]. Le but de la substitution et du traitement de fond (aussi appelé coupe de fond et vidéo) est d'extraire le premier plan de la vidéo d'entrée, puis de le combiner avec un nouveau fond comme les vidéos ou les photos que l'on trouve sur les réseaux sociaux [51].

Conclusion

Dans ce chapitre, nous avons étudié la notion de vision par ordinateur ainsi que ses domaines d'application. Par la suite nous avons mis l'accent sur le processus de soustraction de fond et ses domaines d'application.

En utilisant ces notions de base, nous pouvons désormais étudier et comparer les travaux existants de soustraction de fond.

Chapitre 2 : Etat De L'art

Introduction

Le terme de soustraction de fond désigne un processus visant à séparer des objets d'avant-plan d'un motif relativement stationnaire [52]. De nombreux algorithmes de soustraction de fond et de détection de premier plan ont été proposés au cours des dernières années, chaque méthode est basée sur une méthode basique ou combine plusieurs méthodes pour améliorer son résultat. Dans ce travail nous étudions un ensemble de méthodes récentes de soustraction de fond proposés dans la littérature.

Ce chapitre est organisé en trois parties. Dans la première partie, nous étudions les différentes techniques de modélisation de l'arrière-plan. Dans la deuxième partie nous abordons le principe des algorithmes récents de soustraction de fond choisis et nous les comparerons. Nous inter-changeons le terme fond et arrière-plan tout au long de ce chapitre, où les deux termes ont le même sens.

1. Approches de base de la soustraction de fond

La soustraction de fond est un moyen courant de détecter et de localiser des objets en mouvement dans des séquences vidéo [53].

Différentes méthodes ont été proposées pour détecter le mouvement d'objet en utilisant différentes techniques de soustraction de fond au cours des dernières années [54].

Les approches basiques de la soustraction de fond sont divisées en deux catégories : les approches non-récurrentes et les approches récurrentes.

1.1 Approches non- récurrentes :

Les techniques récurrentes ne conservent pas de mémoire tampon pour l'estimation en arrière-plan [55]. Les techniques non-récurrentes stockent un tampon d'un certain nombre de trames de la séquence vidéo et estiment le modèle de fond basé sur la variation temporelle de chaque pixel dans le tampon [56].

1.1.1 Différence de trames

Cette technique est connue sous le nom de FD (Frame Difference) est considérée comme la méthode la plus simple parmi les techniques de base de la soustraction de fond pour détecter les objets en mouvement à partir d'une caméra fixe. Son but est de détecter les changements d'intensité dans les images vidéo [57]. Son principe consiste à comparer la

valeur absolue de la différence du pixel $I_t(x,y)$ à l'instant t et le pixel $I_{t+1}(x,y)$ à l'instant $t+1$ dans un niveau de gris, où x, y sont les coordonnées du pixel et I une trame de la vidéo. La valeur absolue est définie par l'équation suivante [58] :

$$I_{d(t,t+1)} = |I_{t+1} - I_t| \quad (1)$$

Après avoir créé l'image de différence (équation 1), celle-ci subit des opérations de post-extraction. Ces étapes sont les suivantes.

➤ **Transformation en niveau de gris**

La transformation de la valeur calculée par la valeur absolue du niveau RGB (image en couleur) au niveau de gris. L'image résultante de la différence absolue sera en couleurs. L'étape de transformation permet de passer cette image de l'espace de couleurs RVB (Rouge, Vert, Bleu) vers l'espace de niveau de gris (des dégradations de 0 ; le noir, à 255 ; le blanc).

➤ **Filtrage et binarisation**

Le filtrage est appliqué en utilisant la loi de Gausse afin d'éliminer les trous. L'étiquetage des composants connectés est également nécessaire pour acquérir le plus petit rectangle contenant le déplacement objets [58].

Les changements détectés dans l'image de différence absolue pour chaque pixel peuvent être causés par du bruit ou par un réel mouvement. Afin de distinguer si le changement trouvé à partir du calcul précédent est dû à un mouvement réel ou à un bruit, une binarisation est appliquée. Celle-ci met à 1 les pixels supérieurs à un certain seuil (objets en mouvement), et à 0 les pixels inférieurs au seuil (bruit et fond).

La méthode de différence de trames est illustrée par les figures : Figure 2-1, Figure 2-2 et Figure 2-3.



Figure 2- 1 : La figure représentant l'instant 629s de la base d'images CDnet2014 catégorie baseline.



Figure 2- 2 : La figure représentant l'instant 630s de la base d'images CDnet2014 catégorie baseline.

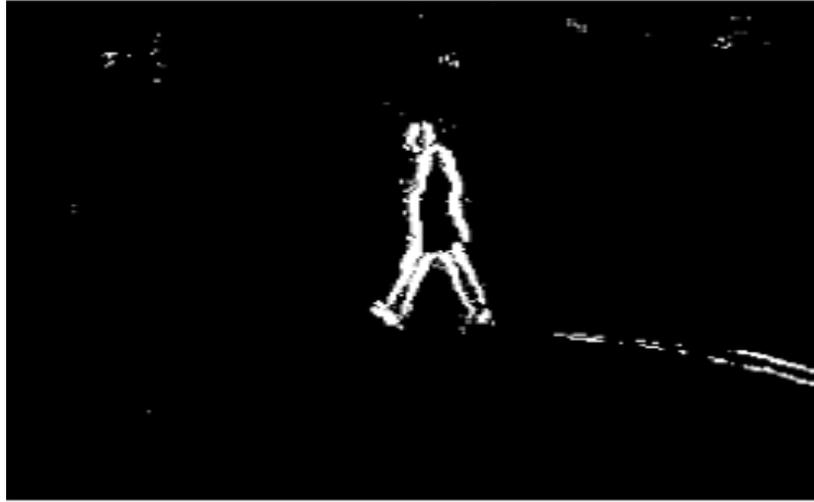


Figure 2- 3: le masque d'avant plan détecté par la méthode différence de trames de la catégorie baseline de la base d'images CDnet2014

Cette méthode a une vitesse de calcul minimale, ce qui fait que ce modèle est très adaptatif [48].

1.1.2 Filtre Médian

Selon [50][59][90], le filtre médian est basé sur l'estimation du modèle de l'arrière-plan selon une moyenne calculée pour chaque pixel en entrée. Le pixel est considéré comme un arrière-plan s'il dépasse la médiane du tampon de trames stockées.

1.2 Approches récursives

Les techniques récursives [54] ne stockent pas les images de la vidéo en mémoire tampon pour l'estimation de l'arrière-plan. Au lieu de cela, elles mettent à jour récursivement un modèle d'arrière-plan unique basé sur chaque image en entrée [86].

1.2.1 Moyenne Gaussienne Mobile (Running Gaussian average)

Cette approche est divisée en trois parties principales :

- La construction du modèle d'arrière-plan,
- L'exécution de la soustraction d'arrière-plan,
- La classification des scènes et la mise à jour du modèle de fond [62].

1.2.2 Mélange de Gaussiennes (Mixture of Gaussians MoG)

MoG suit plusieurs distributions gaussiennes simultanément, elle maintient une fonction de densité pour chaque pixel [60].

Chaque pixel a son propre ensemble de composants gaussiens et chaque gaussienne est capturée selon trois paramètres : Moyenne, poids et déviation [63].

D'après [64], le modèle MoG est un modèle puissant de soustraction, d'estimation et de prédiction.

la fonction de distribution est donnée par:

$$F(i_t = \mu) = \sum_{i=1}^k \omega_{i,t} \cdot \eta(\mu, \sigma) \quad (2)$$

Où η est la $i^{\text{ème}}$ composante gaussienne, μ est l'intensité moyenne, σ est l'écart type, $\omega_{i,t}$ est la partie des données représentées par la $i^{\text{ème}}$ composante.

Selon [65] MoG est plus adapté à l'environnement extérieur en raison de sa capacité à séparer et à supprimer les objets non stationnaires qui se classent comme des bruits tels que les feuilles en mouvement ou le ciel.

1.2.3 Filtre de Kalman

Le filtre de Kalman est largement utilisé pour le suivi dans divers domaines tel que la vision par ordinateur, l'économétrie, et la navigation [66].

Le Filtre de Kalman est également utilisé pour modéliser l'arrière-plan dynamique.

Le principe du filtrage de Kalman également implique une prédiction et une mise à jour, ce qui peut être utilisé dans la segmentation d'objets.

Le filtre de Kalman est caractérisé par deux principales équations : l'équation d'état et l'équation de mesure [67]

Le filtre de Kalman fonctionne en deux étapes :

- L'étape de prédiction.
- L'étape de correction.

L'équation (3) représente l'équation de l'état et l'équation (4) représente l'équation de mesure.

$$x_t = Ax_{t-1} + Bw_{t-1} \quad (3)$$

$$z_t = Cx_t + v_t \quad (4)$$

A est la matrice de transition d'état, B matrice externe de transition de contrôle, w représente le processus de bruit, C matrice de transition qui trace le processus état de la mesure, et v représente le bruit de mesure, $t-1$ représente le temps [68].

1.2.4 Filtre médian approximé (Approximated median filtering AMF)

Cette méthode consiste à prendre les N images récentes en mémoire tampon [88], l'arrière-plan est calculé à partir de la médiane des images choisies.

Si un pixel dans la trame actuelle a une valeur supérieure à la valeur correspondante du pixel d'arrière-plan, le pixel d'arrière-plan est incrémenté.

Contrairement au Filtrage médian temporel, cette approche ne nécessite pas de stocker des images dans une mémoire tampon et tente de mettre à jour le modèle de fond estimé en ligne. Il est donc extrêmement rapide et adapté aux applications en temps réel.

2 Approches récentes de soustraction de fond

2.1 LOBSTER

Est un algorithme de soustraction de fond basé sur l'analyse spatio-temporelle. L'algorithme LOBSTER [91] combine l'algorithme ViBe[92] (extracteur d'arrière-plan visuel) avec la fonctionnalité LBSP (modèles de similarité binaire local, Local Binary Similarity Process) pour obtenir d'excellentes performances de détection dans le scénario général.

La méthode LOBSTER est basée sur ViBe qui s'appuie sur la collecte et la maintenance des échantillons de fond. Une approche aléatoire est utilisée afin de déterminer si de nouveaux échantillons s'adaptent à son modèle en comptant leurs voisins les plus proches dans un seuil d'intersection.

ViBe met à jour le modèle d'arrière-plan de manière aléatoire et établit un modèle avec une précision rapide, élevée à une vitesse de traitement rapide également.

L'algorithme ViBe fournit la méthode du modèle d'arrière-plan mis à jour, mais élimine lentement la région fantôme.

LOBSTER utilise des modèles binaires locaux avec autosimilarité à travers les images, il est une adaptation de ViBe [81] et remplace la description du niveau d'intensité de pixel utilisée dans ViBe par la similitude spatio-temporelle des modèles binaires locaux.

Le modèle est mis à jour à l'aide de simples règles basées sur la distance entre les modèles de similitude. La première amélioration de cette méthode est au niveau de la modélisation de l'arrière-plan qui est basée sur l'observation en s'appuyant uniquement sur les comparaisons de LBSP. Certains pixels peuvent être faussement étiquetés comme arrière-plan dû à ce problème. LOBSTER utilise une vérification et attribution d'étiquette en

s'appuyant sur des comparaisons directes d'intensité de couleur afin de réduire le taux de faux négatifs de la méthode finale LBSP.

La version LBSP amélioré qui était utilisée dans LOBSTER est défini par la formule suivante :

$$LBSP(x, y) = \sum_{p=0}^{P-1} d_{T_\tau}(i_{x,y,p}, i_{x,y}) \cdot 2^p \quad (5)$$

$$\text{Avec } d_{T_\tau}(i, i_c) = \begin{cases} 1 & \text{if } |i_{x,y,p} - i_{x,y}| \leq T_\tau \cdot i_{x,y,p} \\ 0 & \text{autrement} \end{cases} \quad (6)$$

Où $i_{x,y,p}$, correspond à l'intensité du pième voisin de (x, y) sur un motif prédéfini, i_c , Tr serait la nouvelle valeur de seuil relative

LOBSTER utilise la méthode L1 pour effectuer les calculs de similarité pour des raisons de performance et de la faible complexité. Toutefois, la version améliorée de ViBe utilise d'autres méthodes pour le calcul de la similarité.

Afin de diminuer les calculs inutiles et les couts résultant du descripteur LBSP, une nouvelle stratégie est utilisée dans LOBSTER connue sous le nom de « stratégie par canal »

Dans le Tableau 2-1 LOBSTER est comparé avec son ancêtre ViBe au niveau de l'utilisation de la mémoire et la vitesse des configurations proposées à l'original de la méthode ViBe.

D'après ces résultats, clair il est que LOBSTER semble beaucoup plus lent et volumineux que son homologue, même avec l'amélioration de la détection de changement par canal.

Méthode	Mémoire (bytes par RGB par échantillon)	320x240 relatif FPS moyen
LOBSTER (C_{Rel})	72	0.122
LOBSTER (A_{Rel})	72	0.147
LOBSTER (C_{Abs})	72	0.151
LOBSTER (A_{Abs})	72	0.181
Vibe	24	1.000

Tableau 2- 1 : Résultats approximatifs de l'usage de la mémoire et de la vitesse de processeur dans le cas où les paramètres sont similaires [106].

Méthode	La précision	F-Mesure
LOBSTER(A_{Rel})	0.80	0.75
LOBSTER(A_{Abs})	0.77	0.75
ViBe	0.74	0.67

Tableau 2- 2 : Résultats des deux métriques F-mesure et la précision pour les deux méthodes Vibe et LOBSTER[106].

Le Tableau 2-2 montre les résultats globaux des deux configurations (A_{Rel} , A_{Abs}), les résultats montrent que les deux configurations de l'algorithme LOBSTTER avec ses deux configurations marquent des meilleurs résultats par rapport à son ancêtre Vibe.

2.2 SuBSENSE

St-Charles et al. [93] ont proposé une approche fondée sur l'adaptation et l'intégration des modèles de similarité binaire locale (LBSP). Cette approche utilise un mécanisme de retour adaptatif pour la mise à jour du fond caractéristique lorsque la dynamique de fond change avec le temps.

Tout d'abord, les informations spatio-temporelles sont extraites de chaque pixel en utilisant les valeurs RVB et les descripteurs LBSP. Ensuite, un modèle de consensus comme ViBe [92] est utilisé pour la classification des entités binaires. Les étapes de fonctionnement de la méthode sont détaillées dans ce qui suit.

2.2.1 Modélisation au niveau des pixels et extraction des informations

Les représentations au niveau des pixels en utilisant non seulement leurs valeurs RVB, mais aussi le modèle de similarité binaire local (LBSP) identifié dans les équations (7) et (8) fonctionnant dans le domaine spatio-temporel , ou i_x est la référence centrale et i_p correspond à l'intensité du piém voisin de x , τ_d est le seuil de similarité interne.

$$LBSP(x) = \sum_{p=0}^{P-1} d(i_p, i_x) \cdot 2^p \quad (7)$$

$$\text{Avec } d(i_p, i_x) \begin{cases} 1 & \text{if } |i_p - i_x| \leq T_d \\ 0 & \text{autrement} \end{cases} \quad (8)$$

Cette approche améliore les chances de détecter les objets camouflés lorsque leur texture diffère de celle de l'arrière-plan. L'approche peut même tolérer des changements d'éclairage lorsque les couleurs locales et les intensités varient également dans le temps. Dans le cas des « transitions lumineuses vers sombre » (qui sont moins susceptibles d'être des ombres, et plus susceptibles d'être des changements pertinents), le seuil de similarité restera faible,

évitant ainsi d'éventuels classifications faux-fond. Dans ce cas, il suffit de remplacer (7) et (8) par l'équation suivante, où Tr est le nouveau seuil interne relatif (lié à $[0,1]$).

$$\begin{cases} 1 & \text{if } |i_p - i_x| \leq Tr \cdot i_x \\ 0 & \text{autrement} \end{cases} \quad (9)$$

2.2.2 Détection de changement via un consensus

Le type d'approche de consensus est similaire à celle de ViBe [79], il détermine si une observation donnée doit être considéré comme étant au premier plan ou à l'arrière-plan en fonction de sa similitude avec les échantillons récemment observés.

SuBSENSE utilise une version simplifiée (13) pour la classification des pixels en arrière-plan et en avant plan

$$S_t(x) = \begin{cases} 1 & \text{if } \{ \text{dist}(I_t(x), B_n(x)) < R \forall n \neq_{min} \} \\ 0 & \text{otherwise} \end{cases} \quad (10)$$

Où S_t est la carte de segmentation en sortie, $dist$ renvoie la distance entre l'observation actuelle et un échantillon de fond donné, R est le seuil de distance maximum et $\#_{min}$ est le nombre minimum de correspondances requis pour une classification de fond.

Dans ce contexte, un R de petite valeur signifie que le modèle doit être très précis pour réussir à classer les pixels en arrière-plan.

La mise à jour des modèles de pixels est faite en utilisant un conservateur stochastique, l'approche est faite en deux étapes. Tout d'abord, à chaque fois qu'un pixel x est classé comme arrière-plan en utilisant (13), un élément choisi au hasard de l'échantillon de $B(x)$ à une probabilité $1/T$ d'être remplacé par l'observation à $I_t(x)$ où T est un " facteur de sous-échantillonnage temporel ".

Ensuite, l'un des voisins de $B(x)$ a également une probabilité de $1/T$ de voir l'un de ses échantillons remplacés par cette même observation.

Ce nouveau paramètre contrôle la vitesse d'adaptation du modèle d'arrière-plan ; les petites valeurs mènent à des probabilités de mise à jour élevées (et donc à l'évolution rapide du modèle) et inversement.

Le fait que les échantillons soient remplacés aléatoirement au lieu que ce soit en fonction de la date de leur dernière modification, assure qu'un solide historique des représentations d'arrière-plan à court et à long terme peut être conservé dans les modèles de pixels. De même, de nouveaux échantillons ne peuvent être insérés que lorsqu'un pixel local est reconnu

en fond, cette approche évite les avant-plans statiques et évite d'assimiler trop rapidement les objets au fond.

La deuxième étape de mise à jour, "la diffusion", permet aux régions qui sont homogènes avec le fond de se rejoindre plus vite. Les artefacts fantômes, qui sont généralement définies comme des régions de fond faussement classées dues à la suppression d'un objet de la scène observée, peuvent être éliminés rapidement car ils partagent de nombreuses similitudes avec d'autres parties de l'arrière-plan.

De plus, cette même étape de 'diffusion' améliore la cohérence spatiale du modèle d'arrière-plan au point où un mouvement de caméra limité peut être toléré.

2.2.3 Surveillance de l'arrière-plan et schéma de commentaires

Dans cette étape, la dynamique de fond basée sur la comparaison entre les modèles de pixels et les observations locales doit être mesurée d'une façon continue. Cela garantit une réponse optimale en temps dans toutes les régions de la trame observée. Ensuite, les niveaux de bruit de segmentation locale doivent être mesurés en fonction de la détection de pixels clignotants, cela permet aux régions d'arrière-plan dynamiques de se distinguer des régions statiques où les objets de premier plan peuvent être présents. Deux nouveaux contrôleurs dynamiques pour R et T sont alors introduits, qui reposent tous les deux sur des "indicateurs" locaux émanant de la surveillance de la dynamique de fond et de la segmentation.

L'utilisation d'un filtre médian à l'opération de post-traitement élimine tout bruit

Méthode	$FM_{overall}$	$FM_{baseline}$	$FM_{camerajet}$	FM_{shadow}	FM_{therma}	$FM_{dyn.back}$
SuBSENSE	0.826	0.950	0.815	0.899	0.817	0.818
ViBe	0.668	0.870	0.600	0.803	0.742	0.665
ViBe+	0.722	0.871	0.754	0.815	0.665	0.720

Tableau 2- 3 : Tableau 2- 3 : Comparaison entre les méthodes non-paramétrique en utilisant la base d'images CDnet 2012[84].

2.3 Type-2 Fuzzy GMM-UV with MRF et Type-2 Fuzzy GMM-UM with MRF

Type-2 Fuzzy GMM-UV et Type-2 Fuzzy GMM-UM modélisent le fond comme un modèle de mixture gaussienne GMM (Gaussian Mixture Model) avec des moyennes incertaines (UM) ou variances incertaines (UV) spécifiées par la logique floue (Fuzzy).

Zhao et al [94] ont proposé une nouvelle méthode de modélisation de fond basée sur un modèle de mélange gaussien "Fuzzy" de type 2 (T2-FGMM) et chaînes markoviennes aléatoires (Markov Random Field MRF), pour la détection de mouvement dans une scène dynamique. L'arrière-plan est rendu adaptatif par les mises à jour des fonctions 'Fuzzy' pour la moyenne ou la variance afin de régler le problème causé par une scène multimodale¹.

Des contraintes spatio-temporelles sont introduites dans la T2-FGMM par un cadre bayésien. Les résultats montrent que cette approche fonctionne mieux que le modèle de mélange gaussien (GMM) en dynamique typique.

Les étapes de la méthode sont détaillées par la suite.

Premièrement, la sortie de T2-FGMM est considérée comme le champ de marquage initial de la MRF. Ensuite, l'énergie locale du domaine de l'étiquetage est élégamment combinée avec celui de l'observation par un cadre bayésien.

L'algorithme ICM (Conditionnel Modes), est utilisé en raison de son efficacité de calcul, afin d'obtenir la probabilité a posteriori maximale MAP (Maximum A Posteriori Probability).

2.3.1 Maintenance du modèle

La maintenance du modèle est faite comme le GMM original [4] avec le taux d'apprentissage α .

2.3.2 Détection de l'avant plan

Tout d'abord, les Gaussiennes sont ordonnées en ordre croissant. Puis une valeur f_c définie comme la partie maximale de x appartenant à l'avant plan. Les Gaussiennes dont le poids total dépasse f_c seront considérées comme une distribution d'arrière-plan. Un test de correspondance est fait par la distance entre deux limites de l'intervalle log-vraisemblance de la MF principale pour chaque pixel. Le classement sera fait par un seuil constant déterminé expérimentalement. Si la condition précédente est remplie, alors le pixel correspond à la gaussienne. Si aucune correspondance n'est trouvée pour toutes les Gaussiennes K , alors le pixel est étiqueté comme avant plan.

Dans le cas où une correspondance est trouvée, si cela se produit dans les meilleurs Gaussiennes B , alors le pixel est étiqueté en tant que fond, sinon le pixel sera toujours

¹ Une scène multimodale signifie qu'elle contient une ou plusieurs situations telles qu'un arbre agitant, des moniteurs vacillants, les ondulations de l'eau, neige ou pluie.

étiqueté en tant qu'avant plan. Après l'appariement de tous les pixels, un champ d'étiquetage initial est obtenu.

2.3.3 Optimisation des champs d'étiquetage

Le champ d'étiquetage initial n'est généralement pas optimal et peut éventuellement être amélioré avec les contraintes spatio-temporelles.

Sur la base de cette idée, la méthode utilise le modèle MRF pour optimiser le champ d'étiquetage.

Pour le système de voisinage spatial, supposons que les voisins du pixel i se composent de 8 pixels les plus proches, avec la définition spécifique suivante :

$$N_i = \{j \mid 0 < \text{dist}(\text{pixel}_i, \text{pixel}_j) < 2\} \quad (11)$$

ou $\text{dist}(\text{pixel}_i, \text{pixel}_j)$ est la distance entre i et j

Par conséquent, l'énergie a priori totale est calculée comme suit :

$$U(l) = U_e(l) + U_s(l) = \sum_i U_e(i) + \sum_i U_s(i) \quad (12)$$

Où U_e est l'énergie spatiale du pixel_i , U_s est l'énergie temporelle, $U_e(l)$ est « the prior energy » de U_e , $U_s(l)$ est « the prior energy » de U_s , $U(l)$ est l'énergie « the total prior energy »

La distribution antérieure $P(l)$ est une distribution de Gibbs:

$$P(l) = \frac{1}{Z} \times \exp\left(-\frac{U(l)}{T}\right) \quad (13)$$

Où Z est une constante positive pour la normalisation et T désigne la constante de température.

2.3.4 Calcul de la fonction de vraisemblance (Likelihood Function)

La distribution du premier plan est également considérée comme une gaussienne. Elle a la même variance avec le fond mais une moyenne différente, ce qui garantit que le premier plan soit totalement différent de l'arrière-plan.

2.3.5 Probabilité postérieure

Le travail suivant consiste à maximiser la probabilité postérieure et à obtenir un meilleur champ d'étiquetage :

$$l^* = \arg \max P(x|l)P(l) \quad (14)$$

Où $P(l)$ est la distribution de Gibbs et $P(x|l)$ est la fonction de vraisemblance, l^* est le meilleur champ d'étiquetage.

La méthode ICM est utilisée pour optimiser l^* . Bien que l'ICM ne puisse garantir la solution maximum globale, est très efficace et donne des résultats satisfaisants selon les auteurs.

2.4 IMBS

C'est une approche de soustraction de fond par pixel [95], non récursive et non prédictive qui a été conçue spécialement pour traiter le fond de l'eau.

IMBS a trois composantes :

➤ La première composante

La première composante est un algorithme de clustering en ligne. Les valeurs RVB observées à un pixel sont représentées par des histogrammes de taille de bin variable. Cela permet de modéliser des données non gaussiennes, et les modèles d'intensité irrégulière.

➤ La deuxième composante

C'est est une compréhension au niveau régional du fond pour la mise à jour du modèle de fond. Les régions au premier plan persistant un certain nombre d'images sont inclus comme arrière-plan dans le modèle de fond mis à jour.

➤ La troisième composante

C'est un module de suppression de bruit qui permet de filtrer les fausses détections en raison des ombres, des reflets et des sillages du bateau.

Cette méthode est la seule traitant spécialement le problème de vision par ordinateur dans le transport maritime en particulier. Le fonctionnement de la méthode est détaillé comme suit.

La méthode est appelée soustraction indépendante de fond multimodal (Independent Multimodal Background subtraction IMBS). L'algorithme et a été conçu pour effectuer une analyse de fond rapide et efficace. Le modèle d'arrière-plan est calculé au moyen d'une analyse statistique en ligne pour les pixels d'un ensemble L de N images afin d'atteindre une vitesse de calcul élevée.

Selon une période d'échantillonnage P , la trame actuelle I est ajoutée à L , devenant ainsi un échantillon de fond S_n , $1 \leq n \leq N$.

Soit $I(t)$ la trame d'entrée de taille $W \times H$ à l'instant t , et $F(t)$ le masque de premier plan correspondant.

Le modèle d'arrière-plan B est une matrice de H lignes et W colonnes.

Chaque élément $B(i, j)$ de la matrice est un ensemble des tuples (r, g, b, d) , où r, g, b sont des valeurs RVB et $d \in [1, N]$ est le nombre de pixels $S_n(i, j)$ associés à ces valeurs r, g, b .

La modélisation de chaque pixel sous forme de tuple présente l'avantage de capturer les dépendances statistiques entre les canaux RVB, au lieu de considérer chaque canal séparément.

La procédure d'enregistrement du fond (Register Background) crée le modèle d'arrière-plan, tandis que la procédure récupération de l'avant plan (Get Foreground) calcule l'image binaire de premier plan F représentant la sortie du processus.

Le mécanisme de création de modèle en ligne permet d'éviter le stockage des images appartenant à L , qu'est le principal inconvénient des techniques soustractions de fond non-récurrentes.

2.4.1 Suppression de l'ombre

Pour faire face aux pixels de premier plan classés par erreur qui peuvent déformer la forme de l'objet détecté, un module de suppression des ombres est nécessaire.

IMBS adopte une stratégie qui est une légère modification de la méthode à base de VHS proposée par Cucchiara et al. Dans [107]. Le module de suppression des ombres est essentiel pour le succès de l'algorithme.

Après la suppression de l'ombre un processus de remplissage est utilisé pour remplir les trous et petites lacunes.

2.4.2 Modèle de mise à jour

Etant donné l'échantillon d'arrière-plan S_n et le masque d'avant-plan actuel F , si $F(i, j) = 1$ et que $S_n(i, j)$ est associé à un tuple T dans le modèle d'arrière-plan en cours de développement, alors T est étiqueté 'tuple au premier plan'.

Lors du calcul du premier plan, si $I(i, j)$ est associé à un tuple au premier plan, il est classé comme un point potentiel de premier plan.

Une telle solution permet d'identifier des régions de la scène représentant des objets de premier plan non mobiles.

La décision d'inclure ou non les points potentiels du premier plan faisant partie de l'arrière-plan sont pris sur la base d'une carte de persistance.

Si un pixel est classé consécutivement comme premier plan potentiel pendant une période plus longue qu'une valeur prédéfinie (par exemple, $R / 3$), il devient alors partie intégrante du modèle de fond.

En outre, le processus d'étiquetage fournit des informations supplémentaires aux utilisateurs de modules de niveau supérieur (par exemple, un module de suivi visuel) aidant à réduire les observations de fantômes.

Cette solution permet de résoudre les problèmes de la mise à jour sélective et aveugle proposée par (Elgammal et al. [108]).

2.4.3 Nouveautés de la méthode

Les principales nouveautés de la méthode sont :

- 1) un algorithme de classification en ligne permettant de saisir la nature multimodale de l'arrière-plan sans conserver de mémoire tampon avec les images précédentes,
- 2) un mécanisme de mise à jour du modèle pouvant détecter les changements dans la géométrie de fond.

2.5 MultiCue BGS

Cette méthode utilise plusieurs signaux [96] tels que la texture, la couleur et l'apparence de la région. Le modèle de texture avec les descripteurs de modèle binaires locaux adaptatifs à la scène sont utilisés pour détecter les régions de premier plan initiales.

Les statistiques de valeur de couleurs sont intégrées avec des descripteurs de texture pour affiner les résultats des segmentations de premier plan à base de texture.

Enfin, la distance entre la carte de contour (edgemap) et le premier plan segmenté en régions est utilisé pour affiner le premier plan. Les phases de la méthode sont décrites dans ce qui suit.

2.5.1 Phase initiale

Cette phase consiste à appliquer le filtre Gaussien (smoothing) sur les cadres d'images pour filtrer le bruit de caméra et les réponses inutilement hautes fréquences, puis la conversion de son vecteur de couleur RVB en un vecteur HSV couleur-hexagone.

2.5.2 Phase de modélisation de fond

Les informations sur la scène sont regroupées en approximant les statistiques de modèle de fond. Une version améliorée du schéma codebook permet de mettre à jour les modèles de texture de fond T-BGM (texture- background model) et les modèles de couleur de fond C-BGM (color- background model).

T-BGM se concentre sur la formation de plusieurs seuils qui stockent des informations de scène complexes à générer des caractéristiques de texture plus cohérentes et correctes. Ces caractéristiques de texture sont basées sur le nouvel opérateur de texture proposé qui est basé sur la distribution des textures observées.

2.5.3 Phase de la soustraction de fond

Cette dernière se divise en deux processus importants :

❖ Processus au niveau des pixels :

Dans ce dernier, plusieurs premiers plans sont détectés sur la base du T-BGM construit, et pour cette cause C-BGM est utilisé pour corriger les erreurs dans les résultats de la détection initiale. Plus concrètement, pour chaque position de pixel pour laquelle il n'y a pas assez d'informations sur la texture dans le cadre de test et dans le fond formé, le C-BGM est utilisé pour affiner sortie faussement négative provoquée par des schémas basés sur une texture qui ne peuvent pas détecter les changements dans de telles régions, l'extraction des deux descripteurs de couleurs et de texture se fait par un nouveau opérateur connu sous le nom SALBP(scene adaptive local binary pattern)(modèle binaire local adaptatif à la scène).

❖ Processus au niveau des régions :

Ensuite, des opérations morphologiques simples sont effectuées pour éliminer les alarmes de faux positifs de points isolés. Dans les processus au niveau régional, après la suppression des petits blobs inutiles, une vérification de blob candidat basée sur l'apparence est effectuée pour toutes les régions candidates restantes, ce qui donne les résultats de détection finaux.

La phase de maintenance comme l'étape de la modélisation de fond, les informations sur la scène sont regroupées en approximer les statistiques de modèle de base.

2.5.4 Phase de la maintenance du model

Cette phase utilise trois procédures basées sur le schéma de codebook pour toutes les régions d'image sauf les pixels dans les boîtes de sélection de premier plan :

- Mise à jour du codebook (Codebook update).
- Absorption de la région fantôme (Ghost region absorption).
- Suppression des codeword dépassés (Outdated codeword removal).

2.6 Détection d'avant plan par texture et MRF

La Détection d'avant plan par texture et MRF TBMOD (Texture-Based Foreground Detection with MRF) a adapté certaines parties du modèle MoG pour mettre à jour les histogrammes LBP des pixels [97]. La mise à jour du modèle TBMOD a été reprise avec des modifications qui accélèrent la phase initiale de la construction du modèle et optimisent le temps de calcul.

L'étape suivante de l'algorithme incorpore l'image dans le modèle et recommence avec une nouvelle image, si seules les mises à jour du modèle sont nécessaires. Sinon, un graphique est construit sur les suppositions initiales des histogrammes LBP et coupé pour trouver une approximation optimale des régions de premier plan.

Les principales améliorations par rapport aux autres algorithmes sont comme suit.

- Pour reconnaître les changements d'éclairage dans la scène, l'image de la caméra est convertie en espace colorimétrique Luv (ou Lab) et le canal L est extrait et normalisé dans la plage [0, 255] au lieu d'améliorer les anciennes méthodes.
- L'utilisation d'un nouvel opérateur LBP : Le nouvel opérateur est calculé dans un voisinage 5x 5 par conséquent, le paramètre LBP_{new} n'est pas défini avec une largeur de deux pixels sur les bordures de l'image.
- Le modèle d'échecs peut optimiser le temps de calcul du processus de mise à jour. Au lieu de tous les pixels, chaque pixel pair dans les lignes paires et chaque pixel impair dans les colonnes impaires sont calculées.
- Les pixels restants sont étiquetés au premier plan si au moins deux pixels dans leur voisinage à 4 connexions sont des pixels au premier plan, sinon il s'agit d'un pixel d'arrière-plan. Le temps de calcul est réduit de moitié environ et la précision est toujours maintenue au niveau des pixels.
- La mise à jour des histogrammes LBP_{new} est similaire au modèle MoG, optimisé avec une fonction MRF dans SFBS (smooth foreground-background segmentation) . Cette MRF contenait des étiquettes d'arrière-plan et de premier plan et son minimum global était équivalent au résultat de la coupe minimale d'un graphique.

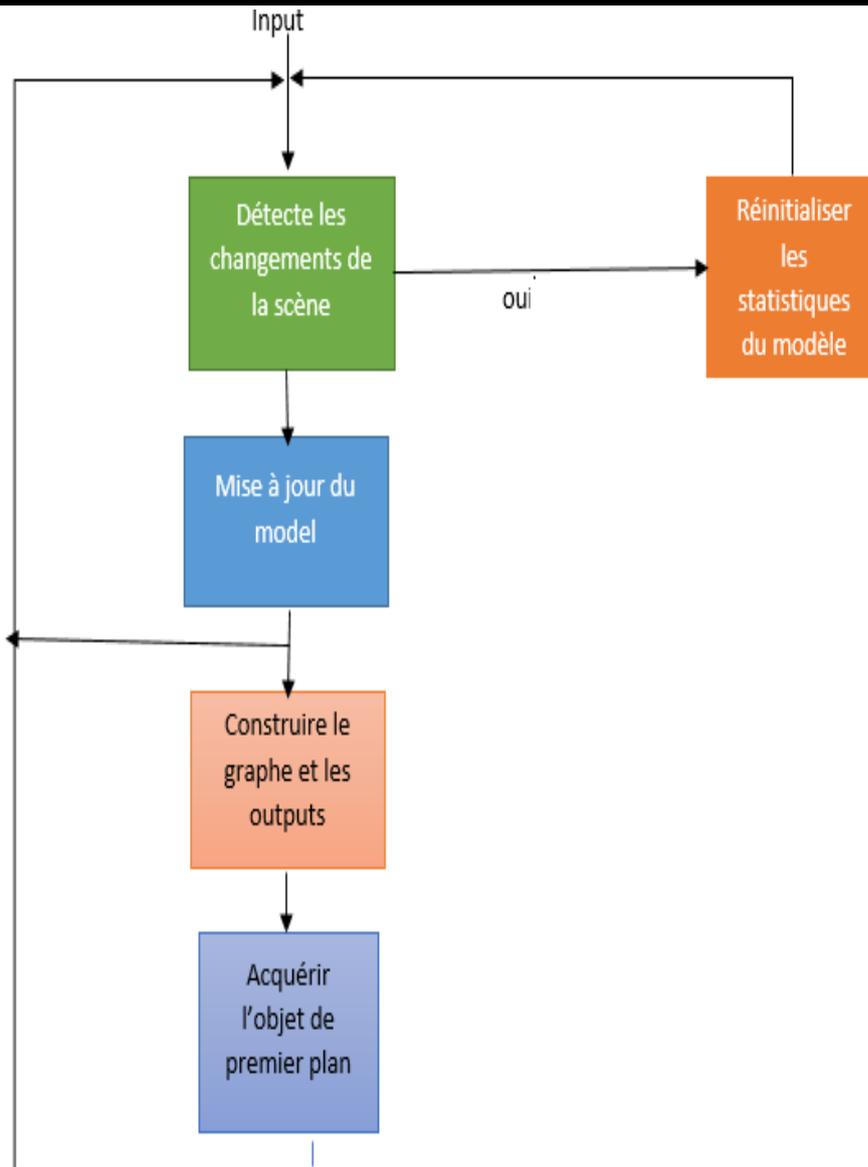


Figure 2- 4 : Les étapes de la méthode Texture et MRF [97].

2.7 Adaptatives SOM

La soustraction de fond basée sur les cartes auto organisatrices adaptatives SOM (Adaptive self organizing maps) est une méthode inspirée biologiquement et basée sur des mécanismes d'attention visuelle. L'objectif est d'obtenir les objets qui retiennent l'attention de l'utilisateur conformément aux ensembles de fonctionnalités prédéfinies, y compris le niveau de gris, le mouvement et les caractéristiques de forme.

L'approche également appelé algorithme de soustraction d'arrière-plan auto organisant (SOB), modélise de manière adaptative l'arrière-plan à l'aide d'un réseau de neurones concurrentiel similaire à la carte d'auto-organisation (Ko) de Kohonen [98]. Pour chaque

pixel, une carte neuronale composée de 3x3 vecteurs de poids est définie. Les pixels source entrants sont mappés sur le vecteur de pondération le plus proche selon une métrique de distance euclidienne, et les vecteurs de pondération situés dans son voisinage sont mis à jour. L'ensemble de vecteurs de pondération sert de modèle d'arrière-plan utilisé pour la soustraction d'arrière-plan afin d'identifier les pixels de premier plan.

Le modèle peut gérer des scènes contenant des arrière-plans multimodaux avec des objets en mouvement et des changements d'éclairage graduels. Il utilise une procédure de mise à jour sélective qui empêche l'inclusion d'objets de premier plan fixes dans l'arrière-plan. Les paramètres de la méthode sont énumérés dans ce qui suit.

➤ **Sensibilité**

Détermine la sensibilité aux changements d'arrière-plan, Les valeurs faibles améliorent la détection des objets dans la scène, mais rendent également le modèle plus sensible au bruit.

➤ **Taux d'apprentissage**

Le taux auquel le modèle s'adapte aux modifications de l'image vidéo. Les valeurs basses correspondent à un modèle s'adaptant lentement. Des valeurs élevées permettent au modèle de s'adapter rapidement aux changements de scène.

➤ **Seuil**

Définit le seuil pour la règle de classification floue, les valeurs faibles ont tendance à inclure plus de pixels dans les objets de premier plan détectés.

➤ **Bruit de Variance**

Définit la valeur minimale de la variance pour le modèle gaussien, des valeurs plus élevées sont recommandées pour les vidéos avec des images bruitées.

2.8 Fuzzy Gaussian

Une version modifiée du modèle gaussien [99] qui utilise une règle de classification floue et effectue des moyennes cumulatives en ligne floues pour la mise à jour sélective de la moyenne et de la matrice de covariance. La mise à jour sélective floue de l'arrière-plan permet une meilleure segmentation des objets au premier plan fixes par rapport au modèle gaussien simple. Les paramètres de la méthode sont identiques à la méthode Adaptive SOM.

Les étapes de cette méthode sont détaillées comme suit :

2.8.1 Modélisation d'arrière-plan flou :

De nombreuses approches récentes se fondent sur un modèle d'arrière-plan multimodal. Le modèle habituellement utilisé est le mélange gaussien.

Les paramètres sont souvent initialisés à l'aide d'une séquence d'apprentissage, qui contient des données insuffisantes ou bruitées. Donc, les paramètres ne sont pas bien déterminés.

Le principal défi consiste à modéliser le fond multimodal. L'algorithme habituellement utilisé est le modèle de mélange gaussien. Les paramètres sont déterminés à l'aide d'une séquence d'apprentissage contenant des données insuffisantes ou bruitées. Donc, les paramètres ne sont pas bien déterminés. Dans ce contexte, les modèles de mélange gaussien flou de type 2 sont utilisés pour modéliser les incertitudes lorsque des arrière-plans dynamiques se produisent.

2.8.2 Détection de premier plan flou :

Le modèle d'arrière-plan peut être unimodal tel que la moyenne mobile ou multimodal tel que la modélisation d'arrière-plan avec mesure de confiance. Une autre approche consiste à agréger différentes caractéristiques telles que les caractéristiques de couleur et de texture en utilisant l'intégrale de Sugeno ou l'intégrale de Choquet. La détection de premier plan flou est plus robuste aux changements d'éclairage et aux ombres que la détection de premier plan nette.

2.8.3 Maintenance d'arrière-plan flou

L'idée est de mettre à jour l'arrière-plan après l'appartenance du pixel à l'arrière-plan ou au premier plan de la classe. Cette appartenance provient de la détection de premier plan flou. Cette maintenance d'arrière-plan adaptative floue permet de gérer de manière robuste les changements d'éclairage et les ombres.

2.8.4 Post-traitement flou

Un moteur d'inférence floue peut être appliqué sur les masques de premier plan courant et antérieurs, pour bien détacher l'objet en mouvement de l'arrière-plan dynamique. MANJULA et SIVABALAKRISHNAN[107] ont proposé une méthode de ce type, qui utilise la caractéristique de couleur et un filtrage médian .

2.9 Comparaison entre les algorithmes

Le Tableau 2-4 compare les différentes méthodes étudiées selon différents critères.

Nom de la méthode	Année	Méthode de modélisation	Type de la méthode	Type de modélisation	Méthode de modélisation de l'arrière-plan	Précision (x100)	F-mesure ² (x100)
Lobster	2014	Récente	Caractéristiques multiples	Non récurrente	Non-paramétrique (Modèles statistiques avancés).	0.6867	0.6230
subscene	2014	Récente	Caractéristiques multiples	Non récurrente	Modèle statistique non paramétrique basé sur l'échantillon.	0.7668	0.7509
T2FMRF_UV	2012	Traditionnelle	Type Fuzzy-2	Récurrente	Gaussien (modèle statistique)	0.888	0.763
T2FMRF_UV	2012	Traditionnelle	Type Fuzzy-2	Récurrente	Gaussien (modèle statistique)	0.670	0.743
multicueBGS	2012	Récente	Caractéristiques multiples	Récurrente	Approches basées sur les pixels et approche basées sur les régions.	0.2602	0.4025
IMBS	2012	Récente	Non-paramétrique	Récurrente	Modèles statistiques avancés (non paramétriques).	0.3703	0.4946
Texture based foreground with mrf	2011	Récente	Caractéristique multiples	Non récurrente	Approche basé sur les régions.	0.35	0.40
Adaptative Som	2008	Traditionnelle	Reseau de neurons	Récurrente	Méthodes basées sur l'apprentissage automatique.	0.907	0.867
Fuzzy gaussien	2008	Traditionnelle	Fuzzy gaussien	Récurrente	Gaussien (modèle statistique)	0.704	0.808

Tableau 2- 3: comparaison entre les neuf algorithmes de bgslibrary.

² Les tests ont été faits sur la base d'images CDnet 2014.

Conclusion

Dans ce chapitre nous avons illustré les techniques de bases de la soustraction de fond, avec un ensemble d'algorithmes récents. Nous avons comparé ces derniers selon des critères qui nous ont semblé intéressants.

En utilisant ces connaissances, nous pouvons à présent modéliser notre solution proposée pour comparer ces algorithmes quantitativement et qualitativement et proposer une amélioration.

Chapitre 3 : Conception

Introduction

À ce jour, de nombreux algorithmes de détection de mouvement ont été développés, ils fonctionnent bien dans certains types de vidéos, mais la plupart sont sensibles aux changements soudains d'éclairage, aux conditions environnementales (nuit, pluie, neige, turbulence de l'air), aux mouvements de fond et/ou de la caméra, aux ombres, et aux effets de camouflage (similarité photométrique de l'objet et de l'arrière-plan). À l'heure actuelle, aucun algorithme ne semble capable de traiter simultanément tous les défis clés associés aux vidéos réelles (non synthétiques).

Dans ce chapitre nous proposons notre solution qui permette de choisir entre un ensemble d'algorithmes récents de soustraction de fond, et leur rajoute la possibilité d'être exécutés en combinaison avec l'extraction de contour.

1 Solution proposée

Afin d'améliorer les résultats de la soustraction de fond nous proposons une nouvelle technique qui consiste à appliquer la méthode d'extraction de contour sur les images de vidéos au préalable à la soustraction de fond proprement dite. Nous démontrons ainsi l'efficacité d'utiliser la détection de contour sur les trames de vidéos. La proposition est faite en 6 étapes essentielles comme suit.

1.1 Phase 1 soustraction de fond basique

Appliquer une Soustraction de fond sur l'ensemble de vidéos dont nous disposons en utilisant l'ensemble d'algorithmes récents de soustraction de fond étudiés dans le chapitre de l'état de l'art.

1.2 Phase 2 extraction de contour

Extraire le contour de chaque trame de chaque vidéo dont nous disposons. Nous choisissons le filtre de Sobel durant cette étude.

Le filtre de Sobel est un opérateur utilisé en traitement d'image pour la détection de contours. Il s'agit d'un des opérateurs les plus simples qui donne toutefois des résultats corrects. Pour faire simple, l'opérateur calcule le gradient de l'intensité de chaque pixel. Ceci indique la direction de la plus forte variation du clair au sombre, ainsi que le taux de changement dans cette direction. On connaît alors les points de changement soudain de luminosité, correspondant probablement à des bords, ainsi que l'orientation de ces bords.

En termes mathématiques, le gradient d'une fonction de deux variables (ici l'intensité en fonction des coordonnées de l'image) est un vecteur de dimension 2 dont les coordonnées sont les dérivées selon les directions horizontale et verticale. En chaque point, le gradient pointe dans la direction du plus fort changement d'intensité, et sa longueur représente le taux de variation dans cette direction. Le gradient dans une zone d'intensité constante est donc nul. Au niveau d'un contour, le gradient traverse le contour, des intensités les plus sombres aux intensités les plus claires [W1].

1.3 Phase 3 soustraction de fond sur images de contour

Appliquer une Soustraction de fond sur l'ensemble de vidéos dont nous disposons avec contour extrait en utilisant l'ensemble d'algorithmes récents de soustraction de fond étudiés dans le chapitre de l'état de l'art.

1.4 Phase 4 combinaison de masques

Combiner les résultats obtenus dans la phase 1 et la phase 3. Ceci est fait en appliquant une addition des masques issus des deux phases.

L'addition matricielle est une opération mathématique qui consiste à produire une matrice qui est le résultat de l'addition de deux matrices de même type [W2].

1.5 Phase 5 calcul de métriques

Calculer les métriques d'analyse de précision de la soustraction de fond pour les résultats de la phase 1, 3 et 4.

1.6 Phase 6 Analyse quantitative et qualitative

Comparer les résultats quantitatifs et qualitatifs des algorithmes étudiés sur les cas cités ci-haut.

2 Objectifs

Notre objectif principal est le développement d'une application logicielle sous Windows permettant la génération de l'avant plan et du fond à partir d'une séquence d'images ou une vidéo. En utilisant cette application, les opérations des phases citées ci haut peuvent être réalisées. De ce fait l'application doit répondre aux objectifs suivants.

- Premier Objectif : Lecture d'une vidéo ou d'une séquence d'images.

-
- Deuxième objectif : Choix d'une méthode de soustraction de fond.
 - Troisième objectif : Génération de l'avant plan et de l'arrière-plan de la méthode et des paramètres choisis précédemment.
 - Quatrième objectif : Application d'une méthode d'extraction de contour sur une séquence vidéo choisi par l'utilisateur.
 - Cinquième objectif : Combinaisons de résultats de soustraction de fond (addition de deux séquences vidéo).
 - Sixième objectif : Analyse de l'amélioration quantitative et qualitative des résultats de la soustraction de fond sur les algorithmes étudiés.

3 Modélisation de l'application avec UML

Dans ce qui suit nous allons présenter le fonctionnement de notre système en utilisant les diagrammes suivants du langage de modélisation UML :

- Diagramme de cas d'utilisation.
- Diagramme de classes.
- Diagramme de séquence.

3.1 Diagramme de cas d'utilisation

Un cas d'utilisation permet de décrire l'interaction entre l'acteur et le système à développer. Plus précisément un cas d'utilisation décrit une séquence d'actions réalisées par le système qui produit un résultat observable pour un acteur [W3].

Il existe Trois types de relations qui sont pris en charge par la norme UML et sont graphiquement représentées par des types particuliers de ces relations :

- Les inclusions.
- Les extensions.
- Les généralisations.

La seule relation entre les acteurs est la généralisation.

La Figure 3-1 présente les différentes fonctionnalités de notre système décrites par un diagramme de cas d'utilisation global. Les principales fonctionnalités du système sont présentées dans le diagramme.

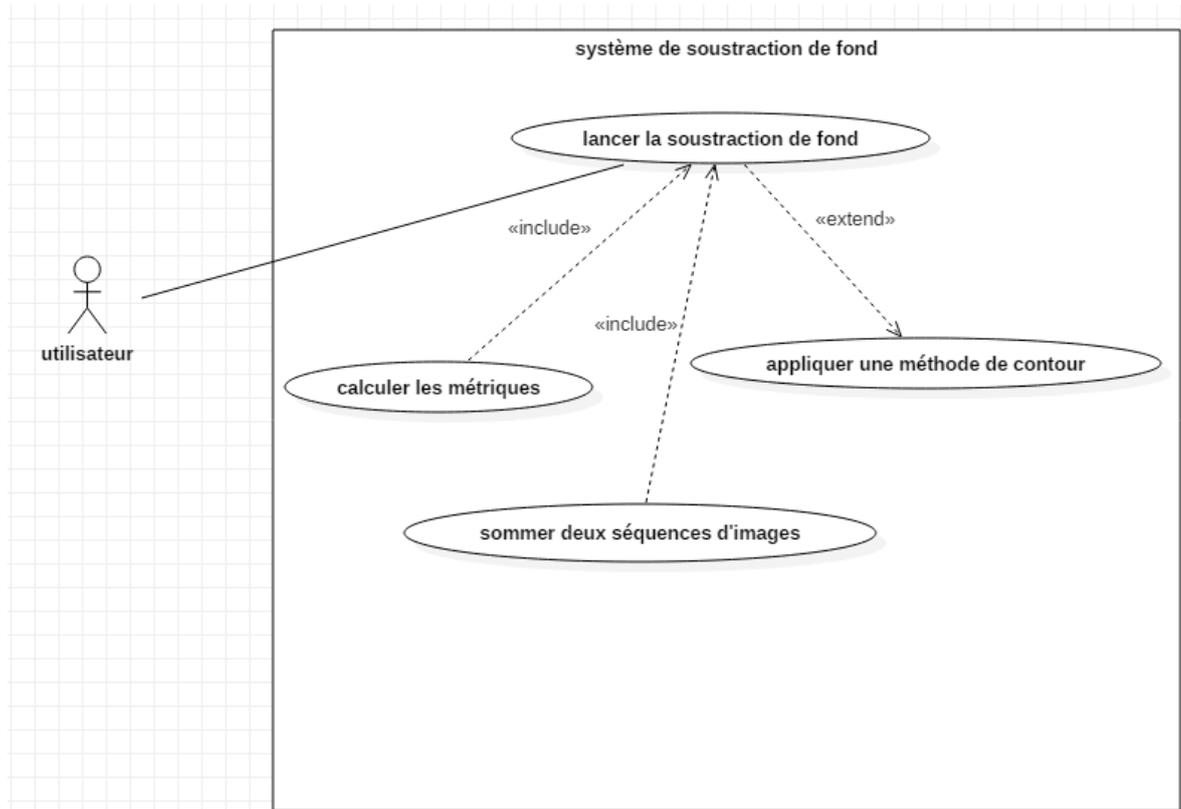


Figure 3- 1 : Diagramme de cas d'utilisation globale.

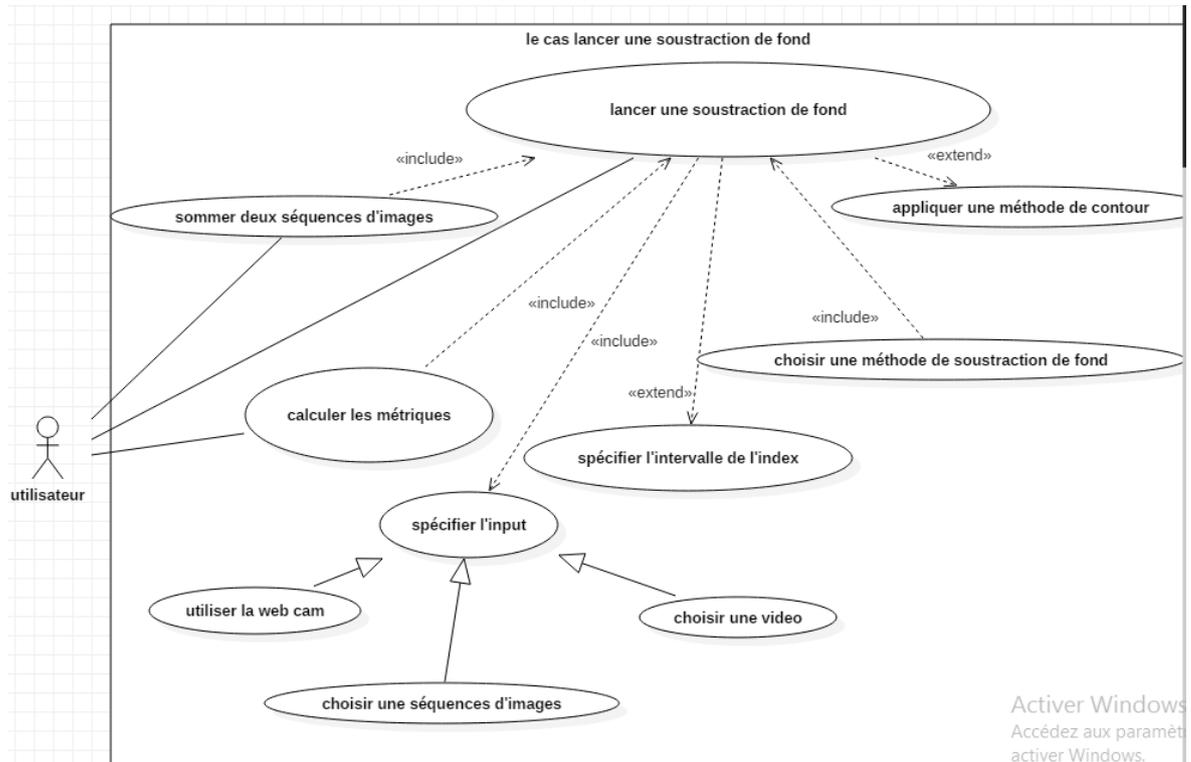
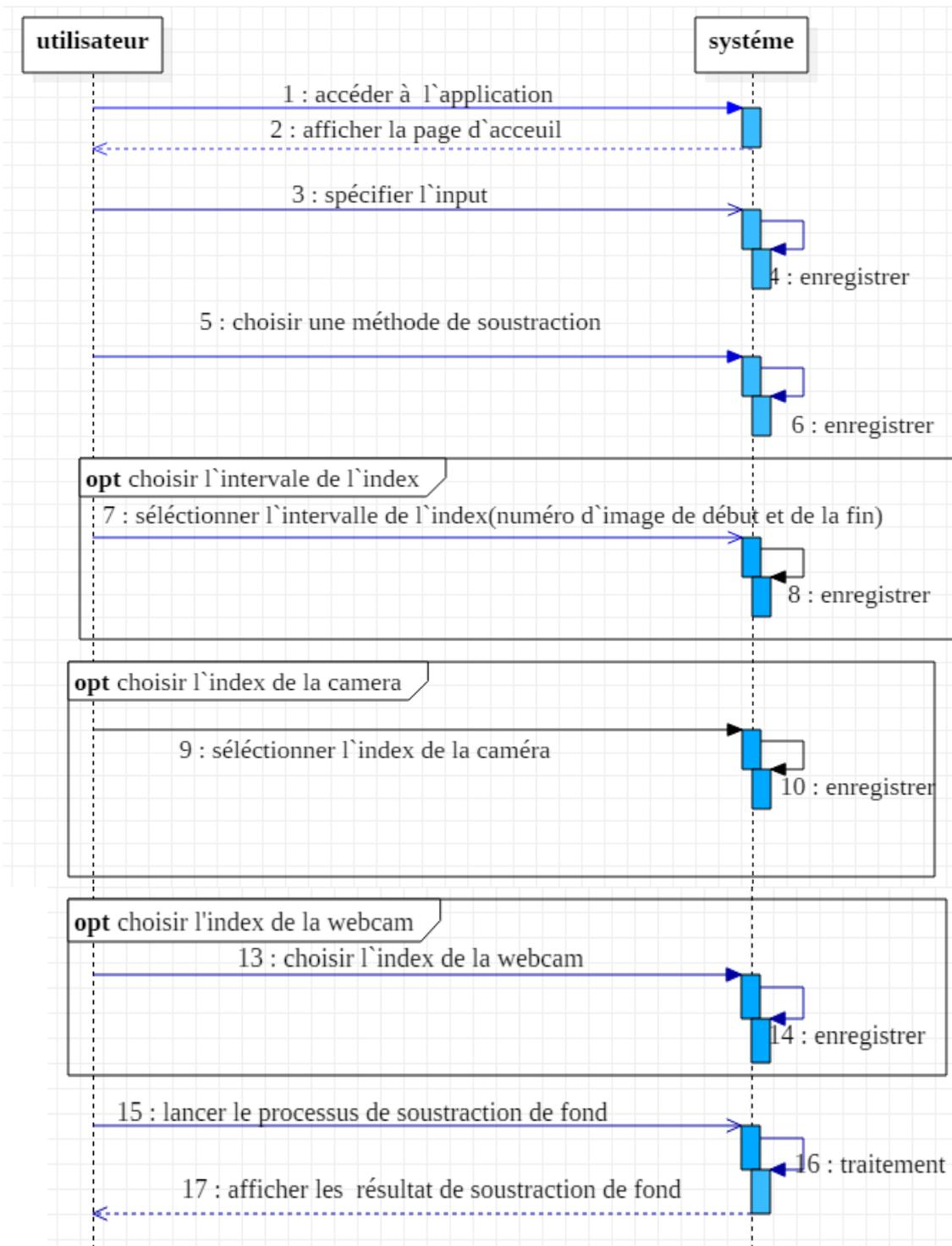


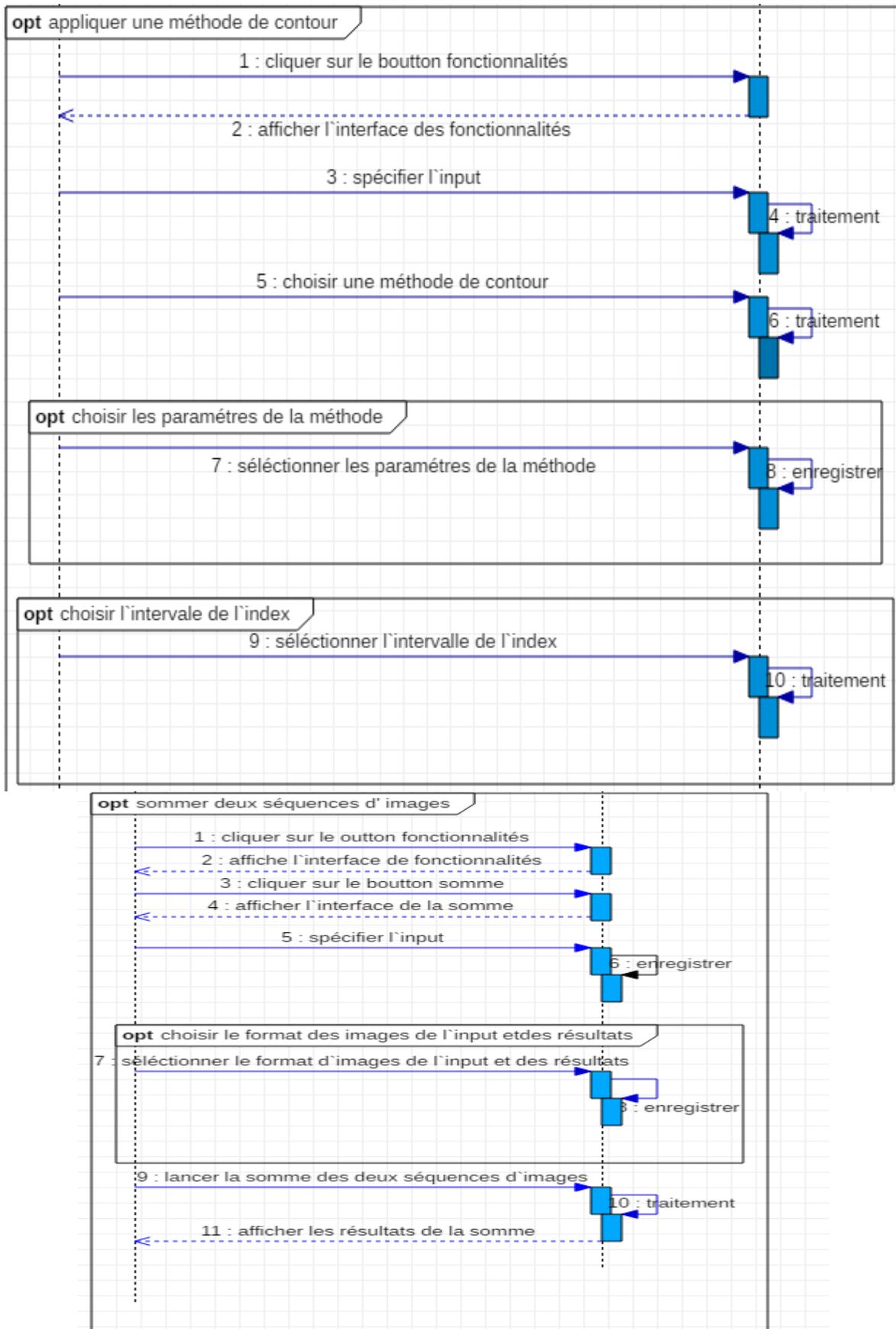
Figure 3-2 : Diagramme de cas d'utilisation « lancer la soustraction de fond ».

3.2 Diagramme de séquence

Le diagramme de séquence permet de montrer les interactions d'objets dans le cadre d'un scénario d'un Diagramme des cas d'utilisation. Dans un souci de simplification, l'acteur principal est représenté à gauche du diagramme, et les acteurs secondaires éventuels à droite du système. La ligne de vie de l'objet représente une durée de son interaction avec les autres objets du diagramme. Le but étant de décrire comment se déroulent les actions entre les acteurs ou objets [W4].

Le diagramme de séquence pour lancer une soustraction de fond est représenté à la Figure 3 -3.





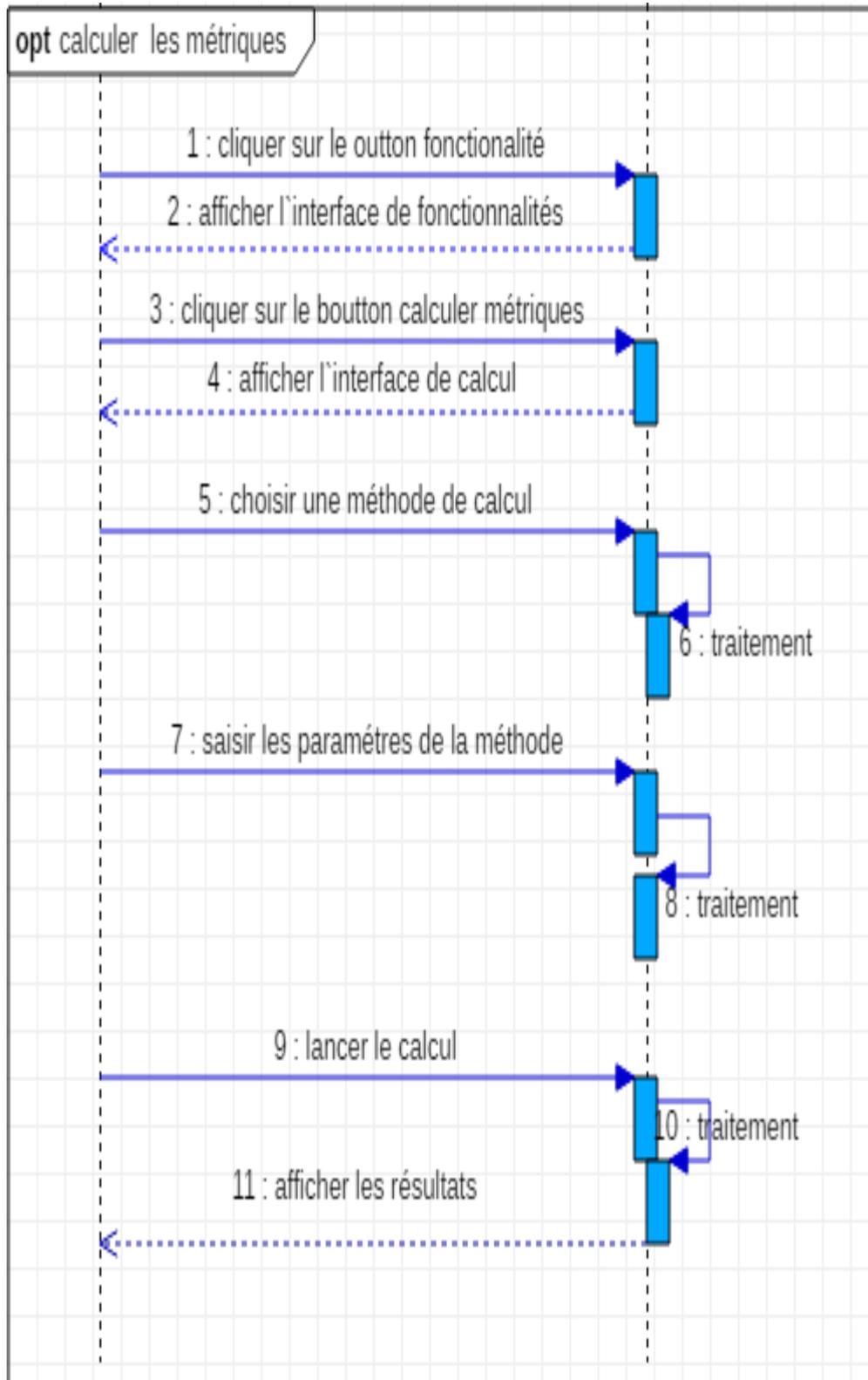


Figure 3-3 : Diagramme de séquence pour le cas « lancer la soustraction de fond ».

3.3 Diagramme de classe

Le diagramme de classes est un schéma utilisé en génie logiciel pour présenter les classes et les interfaces des systèmes ainsi que les différentes relations entre celles-ci. Ce diagramme fait partie de la partie statique d'UML car il fait abstraction des aspects temporels et dynamiques [W5].

Classiquement il existe deux principaux types de relation entre les objets :

- Les associations, bien connues des vieux modèles entité/association utilisés dans la conception des bases de données depuis les années 1970.
- Les sous-types, particulièrement utilisés en conception orientée objets, puisqu'ils s'expriment très bien à l'aide de l'héritage en programmation.

Après l'étape de définition et d'analyse des besoins, nous avons proposé un schéma pour résoudre la problématique en prenant en considération les besoins définis précédemment. Nous décrivons ainsi dans la Figure 3 -4 l'architecture du système et les relations entre ses différents composants à l'aide du diagramme de classes.

Les principales classes utilisées dans notre conception sont les suivantes :

- Classe Utilisateur.
- Classe Algorithme.
- Classe Input.

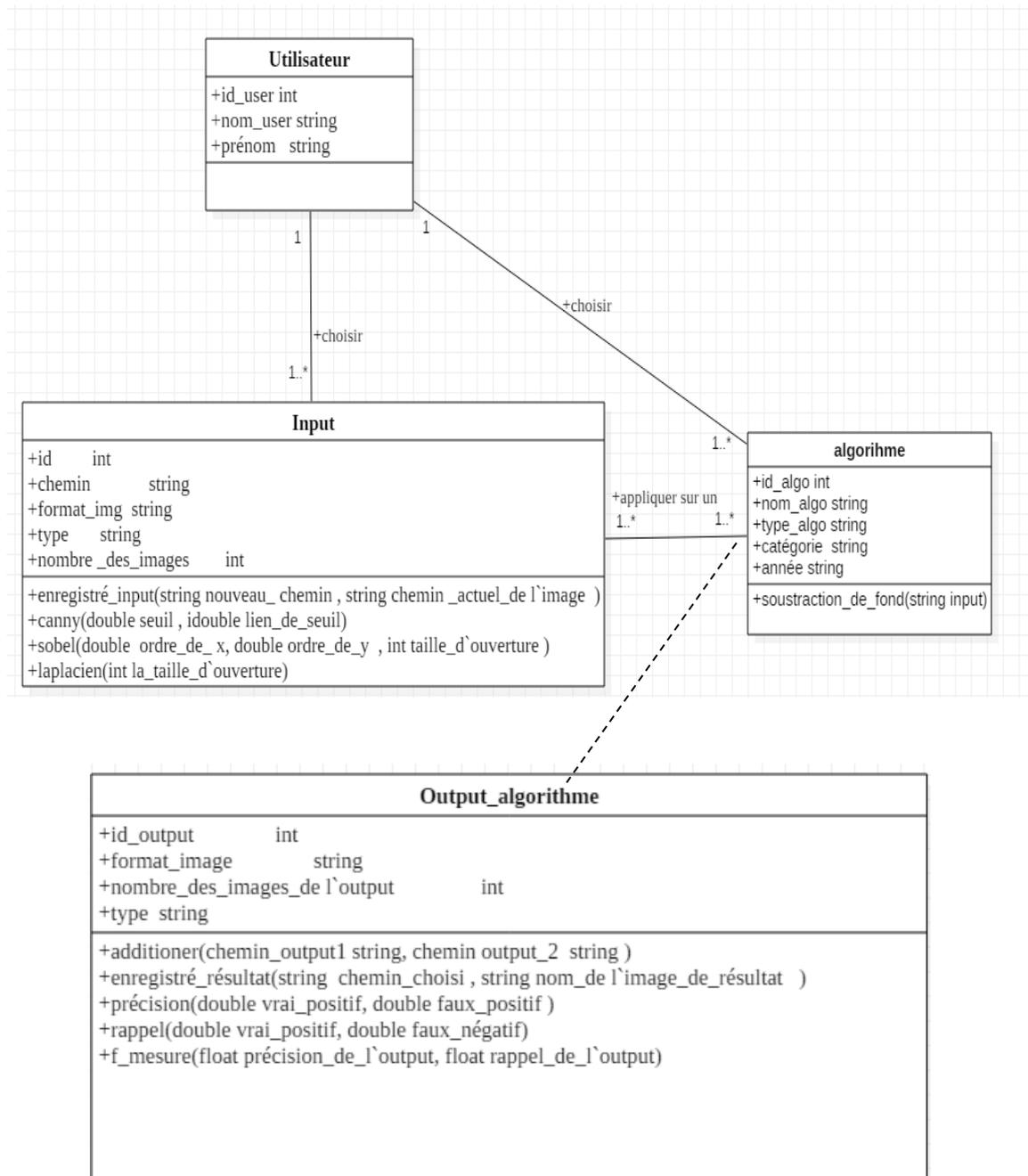


Figure 3-4 : Diagramme de classe.

Conclusion

Au terme de ce chapitre, nous avons donné une vue sur l'amélioration qu'on voudrait apporter aux algorithmes actuels de soustraction de fond, nous avons également présenté notre system à l'aide du langage de modélisation UML.

Nous avons décrit les différentes fonctionnalités à l'aide des diagrammes disponibles en UML tels que le diagramme de classes pour montrer la vue structurelle du système, les diagrammes de cas d'utilisation et les diagrammes de séquence. Nous pouvons maintenant entamer la phase d'implémentation informatique des différentes fonctionnalités de notre système.

Chapitre 4 : Implémentation et résultats

Introduction

La soustraction de fond est l'une des tâches de bas niveau les plus fréquemment rencontrées dans les domaines de la vision par ordinateur et du traitement vidéo. Cependant, il n'existe pas de solution universelle permettant de traiter avec succès tous les problèmes. Y compris les mauvaises conditions d'éclairage, leurs changements soudains et les mouvements parasites de fond. Dans le chapitre précédent, nous avons pensé à une solution à ce problème.

Par conséquent, après avoir conçu notre solution, nous entamons l'étape de la réalisation de l'application. Nous présentons d'abord l'environnement de développement, l'architecture de déploiement, les choix techniques que nous avons adoptés pour sa réalisation et la base d'images que nous avons utilisé. Enfin, nous présentons les résultats obtenus et les critiquons.

1 Environnement Matériel

Pc 1

- ✓ Processeur : Intel ® Pentium ® CPU2020M @2.40GHz 2.40 GHz
- ✓ RAM : 2 Go
- ✓ Type du système : système d'exploitation 32 bit

Pc 2

- ✓ Processeur : Intel ® Pentium ® CPU2020M @2.40GHz 2.40 GHz
- ✓ RAM : 2 Go
- ✓ Type du système : système d'exploitation 32bit

2 Environnement Logiciel

Pour mettre en place notre solution, nous nous sommes basés sur des technologies fiables, performantes et extensibles, afin de pouvoir satisfaire au mieux les exigences en termes de qualité et de performance.

2.1 Outil de modélisation UML

❖ STAR UML



STAR UML est un logiciel de modélisation UML, qui a été "cédé comme open source" par son éditeur, à la fin de son exploitation commerciale, qui continue sous une licence modifiée de GNU GPL. Aujourd'hui la version StarUML V3 n'existe qu'en licence propriétaire. StarUML gère la plupart des diagrammes spécifiés dans la norme UML 2.0. StarUML est écrit en Delphi1, et dépend de composants Delphi propriétaires (non open-Source) [W6].

2.2 Environnement de développement et librairies

❖ QT Creator



Qt Creator est un environnement de développement intégré multi plate-forme faisant partie du Framework Qt. Il est donc orienté pour la programmation en C++. Il intègre directement dans l'interface un débogueur, un outil de création d'interfaces graphiques, des outils pour la publication de code sur Git et Mercurial ainsi que la documentation Qt. L'éditeur de texte intégré permet l'auto complétion ainsi que la coloration syntaxique. Qt Creator utilise sous Linux le compilateur gcc. Il peut utiliser MinGW ou le compilateur de Visual Studio sous Windows [W7].

❖ Opencv



OpenCV (pour Open Computer Vision) est une bibliothèque graphique libre, initialement développée par Intel, spécialisée dans le traitement d'images en temps réel. La société de robotique Willow Garage et la société ItSeez se sont succédé au support de cette bibliothèque. Depuis 2016 et le rachat de ItSeez par Intel, le support est de nouveau assuré par Intel. Cette bibliothèque est distribuée sous licence BSD. NVidia a annoncé en septembre 2010 qu'il développerait des fonctions utilisant CUDA pour Open CV [W8].

❖ Emgu cv



Emgu CV est une bibliothèque multi-plateforme de traitement d'images. Elle est étroitement liée à OpenCV car Emgu CV est un wrapper .NET pour OpenCV. Nous pouvons dire qu'Emgu CV est OpenCV dans

.NET. Le wrapper étonnant permet aux fonctions OpenCV d'être appelées à partir de langages de programmation .NET. C #, VB, IronPython et VC ++ sont quelques-uns des langages pris en charge. Emgu CV peut être compilé en Mono et fonctionne sous Linux, Windows, Mac OS X et sur des plates-formes mobiles populaires telles que les appareils Android, iPhone, iPod Touch et iPad [W9].

❖ BGSLibrary

BGSLibrary est une bibliothèque de soustraction de fond qui fournit un Framework C++ et des outils pour effectuer une soustraction de fond. Elle est publiée pour la première fois en mars 2012 par Andrews Sobral. La bibliothèque propose actuellement 32 algorithmes de soustraction de fond. Le code source est disponible sous licence GNU GPL v3 et la bibliothèque est gratuite pour une utilisation non commerciale, open source et indépendante de la plate-forme. La bibliothèque BGSLibrary fournit également une interface graphique basée sur Java permettant aux utilisateurs de configurer la source vidéo d'entrée, la région d'intérêt et les paramètres de chaque algorithme. Afin d'utiliser la bibliothèque, Open CV doit être installé au préalable [102].

❖ Cmake



CMake est un système de construction logicielle multiplateforme. Il permet de vérifier les prérequis nécessaires à la construction, et de déterminer les dépendances entre les différents composants d'un projet, afin de planifier une construction ordonnée et adaptée à la plateforme. La construction du projet est ensuite déléguée à un logiciel spécialisé dans l'ordonnancement de tâches et spécifique à la plateforme, Make, Ninja ou Microsoft Visual Studio. CMake reprend le concept de configuration initié par autoconf, mais s'en distingue par son caractère multiplateforme, le rendant adapté à la construction des logiciels destinés à fonctionner sous Linux et Windows. Le nom « CMake » est l'abréviation de « cross Platform make ». Malgré l'utilisation de « make » dans son nom, CMake est une application séparée et de plus haut niveau que l'outil make [W10]

2.3 Langages de développement

❖ Langage C++



C++ est un langage de programmation compilé permettant la programmation sous de multiples paradigmes (programmation procédurale, orientée objet ou générique). Ses bonnes performances, et sa compatibilité avec le C en font un des langages de programmation les plus utilisés dans les applications où la performance est critique. Créé initialement par Bjarne Stroustrup dans les années 1980, le langage C++ est aujourd'hui normalisé par l'ISO. Sa première normalisation date de 1998 (ISO/CEI 14882:1998), ensuite amendée par l'erratum technique de 2003 (ISO/CEI 14882:2003). Une importante mise à jour a été ratifiée et publiée par l'ISO en septembre 2011 sous le nom de ISO/IEC 14882:2011, ou C++11. Depuis, des mises à jour sont publiées régulièrement : en 2014 (ISO/CEI 14882:2014, ou C++14) puis en 2017 (ISO/CEI 14882:2017, ou C++17) [W11].

❖ Langage C#



Est un langage de programmation orientée objet, commercialisé par Microsoft depuis 2002 et destiné à développer sur la plateforme Microsoft .NET. Il est dérivé du C++ et très proche du Java dont il reprend la syntaxe générale ainsi que les concepts, y ajoutant des notions telles que la surcharge des opérateurs, les indexeurs et les délégués. Il est utilisé notamment pour développer des applications web sur la plateforme ASP.NET [W12].

2.4 Editeurs

❖ Visuel studio



Microsoft Visual Studio est une suite de logiciels de développement pour Windows et mac OS conçue par Microsoft. La dernière version s'appelle Visual Studio 2019. Visual Studio est un ensemble complet d'outils de développement permettant de générer des applications web ASP.NET, des services web XML, des applications bureautiques et des applications mobiles. Visual Basic, Visual C++, Visual C# utilisent tous le même environnement de développement intégré (IDE), qui leur permet de partager des outils et facilite la création de solutions faisant appel à plusieurs langages. Par ailleurs, ces langages permettent de mieux tirer parti des fonctionnalités du Framework .NET, qui fournit un accès à des technologies clés simplifiant le développement

d'applications web ASP et de services web XML grâce à Visual Web Developer. Les logiciels de développement de visual Studio sont conçus par Microsoft pour Windows et mac OS. [W13].

2.5 Outils de travail collaboratif

❖ Git Hub



Git Hub est une plateforme open source de gestion de versions et collaboration destinée aux développeurs de logiciels. Livrée en tant que logiciel à la demande, la solution GitHub a été lancée en 2008. Elle repose sur Git, un système de gestion de code open source créé par Linus Torvald dans le but d'accélérer le développement logiciel [W14].

3 Interface de l'application

3.1 Interface d'accueil (soustraction de fond)

Dès que l'utilisateur spécifie l'input, la méthode de soustraction de fond, et les options de paramètres de l'input, le processus de la soustraction commence.

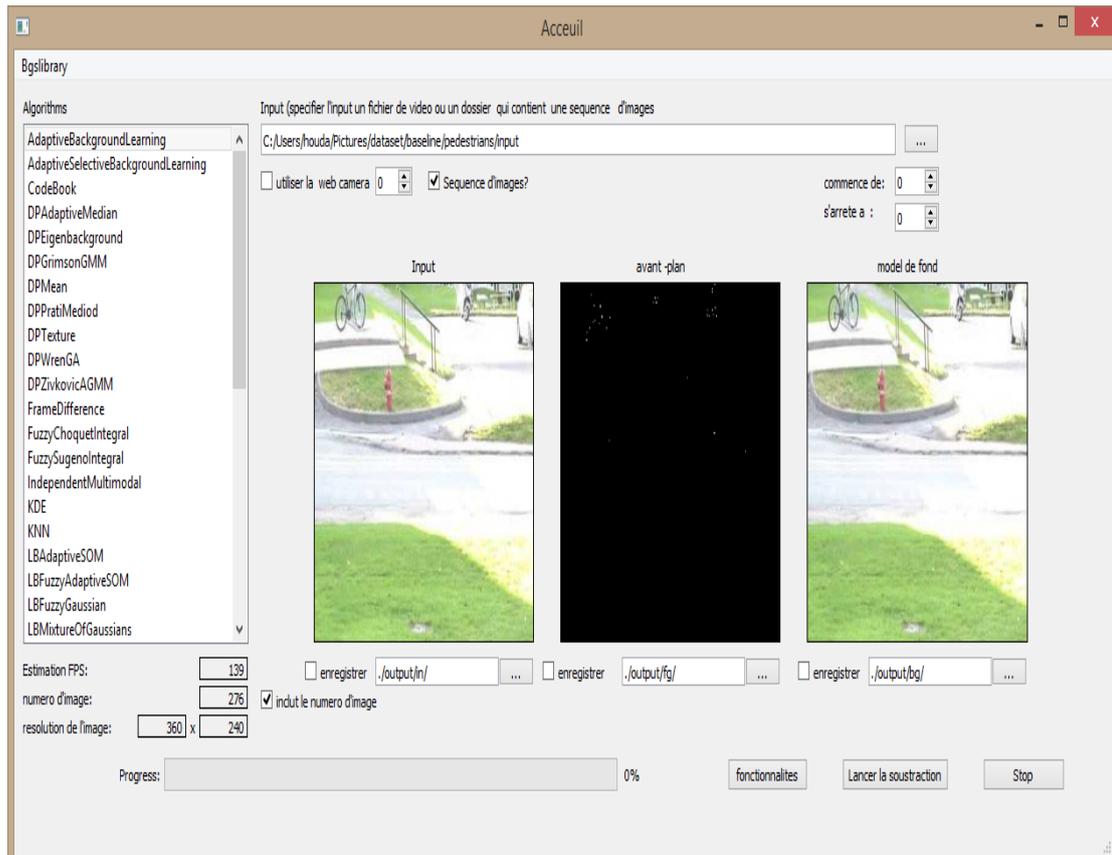


Figure 4- 1:Interface d'accueil.

3.2 Interface de contour

Dès que l'utilisateur spécifie l'input, la méthode d'extraction de contour, et les options, le processus de filtrage commence.

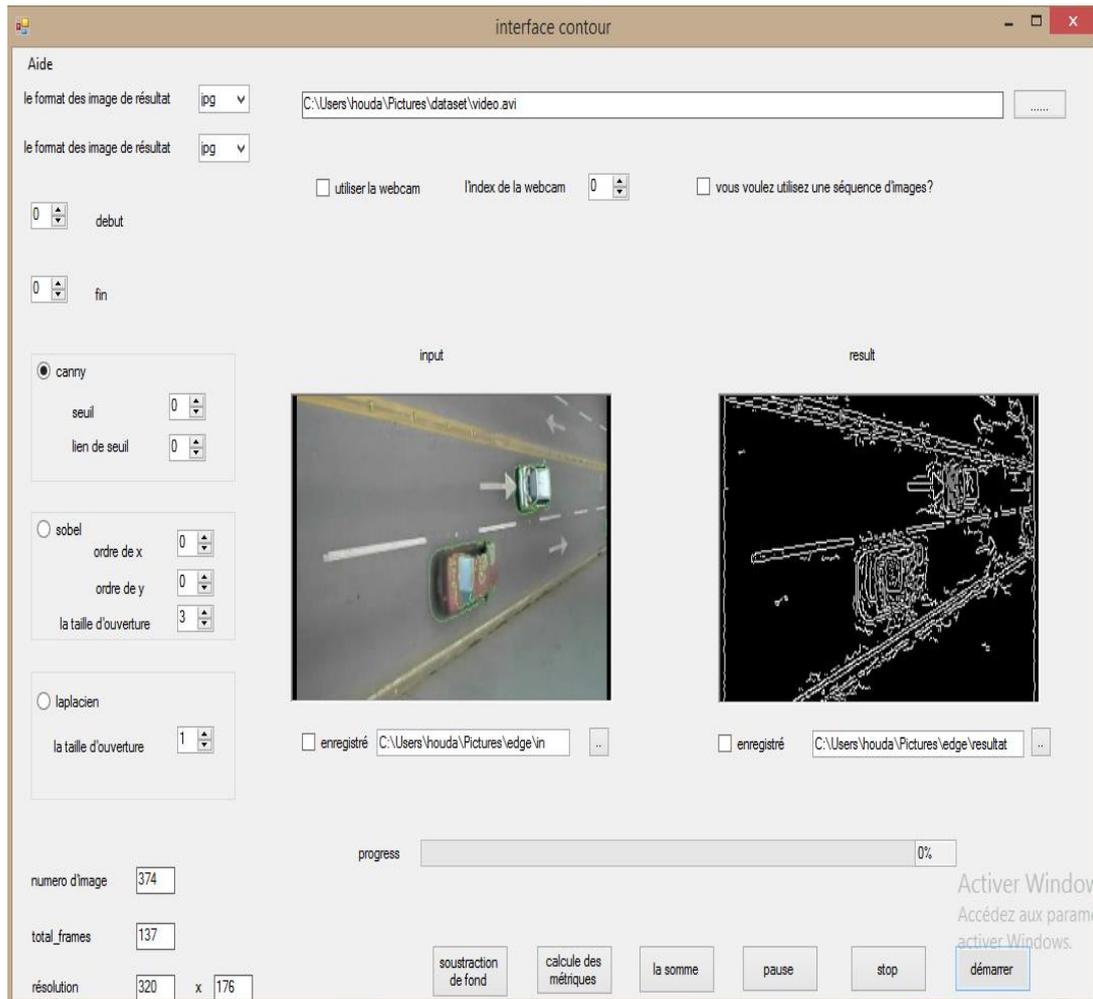


Figure 4- 2: interface de contour de fond.

3.2 Interface d'addition

Dès que l'utilisateur spécifie l'input, et les options de paramètres de l'input et de résultat, le processus de l'addition se lance.

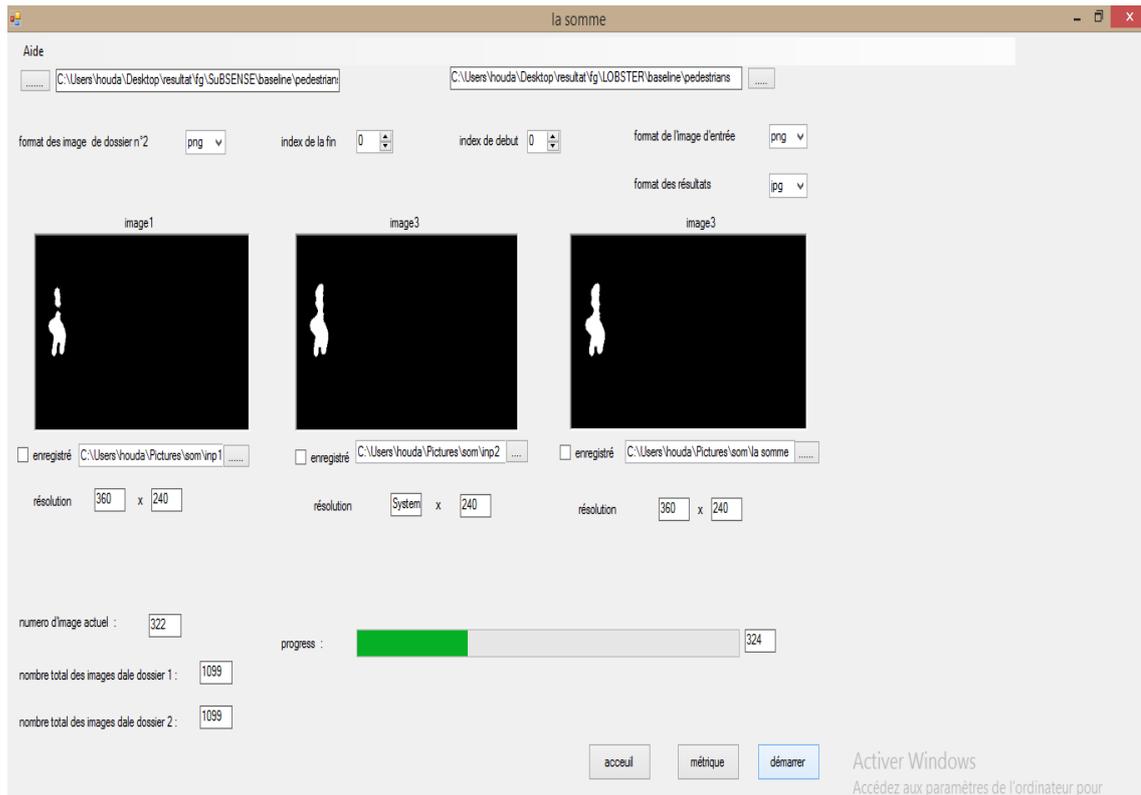


Figure 4- 3: interface d'addition de deux images.

3.3 Interface de calcul

Dès que l'utilisateur spécifie la méthode de calcul, saisir ses paramètres, le processus de calcul se lance.



Figure 4- 4 : interface de calcul des métriques.

4 Évaluation des performances

Afin de calculer la précision et la performance des algorithmes de soustraction de fond choisis, des mesures permettent de calculer toutes les métriques de performance ont été explorées. Ces métriques, utilisent les séquences vidéo sur lesquelles nous voudrions appliquer une soustraction de fond, ainsi que et les masques d'avant-plan dits vérité terrain (ground truth) fournis par l'utilisateur. Nous évaluons la performance en comparant la soustraction de fond résultant des algorithmes à la vérité terrain pour obtenir les mesures de performance (TP, FP, TN, FN, Re, Precision, Sp, FPR, FNR, PWC, FPR-S) définies dans ce qui suit.

Selon [103] :

- True Positive (TP) : les pixels de premier plan dans la segmentation en sortie qui sont également des pixels de premier plan dans la segmentation de vérité terrain.
- Faux positif (FP) : les pixels de premier plan dans la segmentation en sortie qui ne sont pas des pixels de premier plan dans la segmentation de vérité terrain.
- True Negative (TN) : les pixels d'arrière-plan dans la segmentation de sortie qui sont également pixels d'arrière-plan dans la segmentation de la vérité terrain.
- False Negative (FN) : les pixels d'arrière-plan de la segmentation en sortie qui ne sont pas des pixels d'arrière-plan dans la segmentation de la vérité terrain.

À l'aide des mesures calculées précédemment nous calculons les différentes métriques suivantes [104] :

$$\text{Re (rappel)} = \frac{tp}{tp+fn} . \quad (15)$$

$$\text{Précision} = \frac{tp}{tp+fp} . \quad (16)$$

$$\text{Sp (spécificité)} = \frac{tn}{tn+fp} . \quad (17)$$

$$\text{FPR (taux de faux positifs)} = \frac{fp}{fp+tn} . \quad (18)$$

$$\text{FNR (faux taux négatif)} = \frac{fn}{tp+fn} . \quad (19)$$

$$\text{PWC (pourcentage de mauvaises classifications)} = \frac{100x(fn+fp)}{tp+fp+fn+tn} . \quad (20)$$

$$\text{Mesure F : } \frac{(2 \times \text{précision} \times \text{rappel})}{(\text{précision} + \text{rappel})} \quad (21)$$

FPR-S : Taux de faux positifs moyens dans les zones d'ombre dure.

5 Base d'images utilisée

Nous avons étudié deux bases d'images pour la validation des algorithmes de soustraction de fond, soit CDnet 2012 et CDnet 2014³. Les deux fournissent un ensemble diversifié de vidéos réalistes, capturées par des caméras. Ils ont été sélectionnés pour couvrir un large éventail de problèmes de détection et sont représentatifs des données visuelles intérieures et extérieures. Les types de vidéos capturées sont dans des scénarios de surveillance, d'environnement intelligent et de base de données vidéo. Dans ce travail nous en avons gardé la base CDnet 2014 car elle englobe les vidéos de CDnet 2012 et y rajoute d'autres nouveaux challenges.

5.1 CDnet 2012

Cette base d'images comprend les défis suivants : arrière-plan dynamique, instabilité de la caméra, mouvement d'objet intermittent, ombres et signatures thermiques. Elle couvre 6 catégories pour inclure divers défis de détection de changement et de mouvement [100] :

Catégorie Baseline : représente un mélange de défis légers typiques des 4 catégories suivantes. Certaines vidéos ont un mouvement de fond subtil, d'autres ont des ombres isolées, certaines ont un objet abandonné et d'autres des piétons qui s'arrêtent quelques instants puis s'éloignent. Le traitement de ces vidéos est assez facile, mais pas trivial.

Catégorie dynamic background : La catégorie comprend les scènes avec un fort mouvement de fond (parasite) : Bateaux sur une eau scintillante, voitures passant à côté d'une fontaine, ou piétons, camions passant devant un arbre secoué par le vent.

Catégorie Camera Jitter : contient des vidéos intérieures et extérieures capturées par des caméras instables (par exemple vibrantes). La magnitude du mouvement varie d'une vidéo à l'autre.

Catégorie Intermittent Object Motion : inclut des vidéos avec des scénarios connus pour causer des artefacts "fantômes" dans le mouvement détecté, c'est-à-dire que les objets

³ <http://changedetection.net/>

bougent, puis s'arrêtent pendant un court instant, puis recommencent à bouger. Certaines vidéos incluent des objets immobiles qui commencent soudainement à bouger, par exemple un véhicule garé qui s'éloigne. Cette catégorie est destinée à tester l'adaptation de divers algorithmes aux changements d'arrière-plan.

Catégorie Shadow : comprend les vidéos intérieures et extérieures qui présentent des ombres fortes et légères. Certaines ombres sont assez étroites tandis que d'autres occupent la majeure partie de la scène. De plus, certaines ombres sont projetées par des objets en mouvement alors que d'autres sont projetées par des arbres et des bâtiments.

Catégorie Thermal : comprend les vidéos capturées par des caméras infrarouges lointaines. Ces vidéos contiennent des artefacts thermiques typiques tels que des tampons thermiques (par exemple, des points lumineux laissés sur un siège, une personne qui se lève et s'en va), une réflexion de la chaleur sur les sols et les fenêtres et des effets de camouflage lorsqu'un objet en mouvement a la même température que l'environnement des régions.

5.2 CDnet 2014

Cette base d'images inclut toutes les vidéos de CDnet 2012, ainsi que des vidéos supplémentaires présentant les difficultés suivantes : conditions météorologiques difficiles, faible débit, la nuit, capture PTZ (Pan Tilt Zoom) et turbulence de l'air.

Chaque jeu de données est accompagné d'une segmentation précise du terrain et d'une annotation des zones de changement de mouvement pour chaque image vidéo couvrant 5 nouvelles catégories [101] :

Catégorie Bad Weather : comprend des vidéos en extérieur enregistrées dans des conditions météorologiques hivernales difficiles : tempête de neige, neige au sol, brouillard.

Catégorie Low Frame-Rate : contient des vidéos capturées à des nombres d'images par seconde compris entre 0,17 et 1 image par seconde.

Catégorie Night : comprend les vidéos capturées la nuit (conditions de lumière difficiles) de, principalement, la circulation automobile.

Catégorie PTZ (Pan Tilt Zoom) : contient les séquences vidéo capturées par les caméras pan-tilt-zoom en mode panoramique continu lent, mode panoramique intermittent, PTZ en mode de surveillance à 2 positions ou zoom avant / arrière.

Catégorie Air Turbulence : inclut les vidéos en plein air montrant la turbulence de l'air causée par la montée de la chaleur.

5.3 Masque d'avant plan Vérité terrain

Les jeux de données en ligne actuels ont été conçus principalement pour tester le suivi de la scène et la compréhension des algorithmes, et donc la vérité terrain est fournie sous la forme de boîtes englobantes. [100] Bien que cela puisse être utilisé pour valider les méthodes de détection de changement, une validation précise nécessite une vérité de terrain à la résolution en pixels.

Pour permettre une comparaison et un classement quantitatifs précis de divers algorithmes, toutes les vidéos incluses dans CDnet 2014 sont fournies avec une segmentation précise du terrain et une annotation des zones de changement et/ou mouvement pour chaque image vidéo. Dans chaque cadre, les régions « réellement changeantes » ont été soigneusement localisées à l'aide d'opérateurs humains. Ces régions rencontrent les contraintes suivantes :

- Elles ne font pas partie de l'arrière-plan (c'est-à-dire que les arbres, les vagues, les drapeaux, etc. sont exclus) ce sont des personnes, des animaux ou des objets fabriqués par l'homme, par exemple voitures, camions, bateaux, trains.
- Un objet en mouvement qui s'arrête soudainement, exemple une voiture un piéton devant un réverbère etc., doit être détecté pendant un court instant avant de se fondre dans l'arrière-plan. La réflexion de la lumière, la turbulence de l'air causée par la chaleur et les halos de projecteurs ne sont pas considérés comme des objets en mouvement, même lorsqu'ils sont générés par des objets en mouvement.

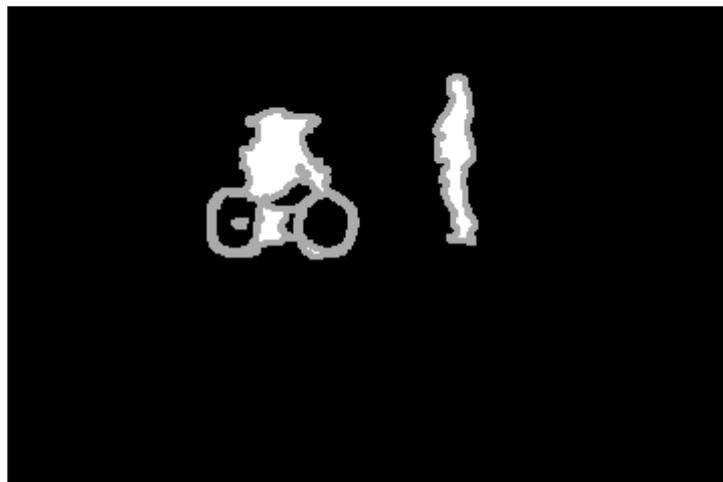


Figure 4- 5 : L'image représente l'instant 478s de CDnet 2014 catégorie baseline.

Les ombres dures ont été étiquetées manuellement afin de permettre la comparaison des algorithmes en fonction de leur robustesse aux ombres.

6 Expérimentation

Les expériences ont été menées sur les vidéos de CDnet 2014 composé de 53 vidéos représentant une grande variété de défis. La base d'images est regroupée en 11 catégories citées précédemment : mauvais temps (Bad Weather BDW), ligne de base (Baseline BSL), fond dynamique (Dynamic background DBG), mouvement d'objet intermittent (intermittence object motion IOM), faible taux de trames (Low framerate LFM), vidéos de nuit (Night videos NVD), ombre (Shadow SHD), Pan Tilt Zoom (PTZ), Turbulence (Turbulence TRBL), Thermique (Thermal THML), Instabilité de la caméra (Camera Jitter CJT). Un exemple d'image de chaque catégorie est représenté à la Figure 4-6.

Nous nous intéressons dans nos expérimentations tout particulièrement à la métrique FM, car la plupart des algorithmes de soustraction d'arrière-plan présentent généralement des valeurs FM plus élevées que les algorithmes de soustraction de fond performants. Cela est dû à la combinaison de plusieurs mesures d'évaluation pour le calcul de la FM. Ainsi, la performance globale d'un algorithme de soustraction de fond est hautement couplée avec sa performance FM.

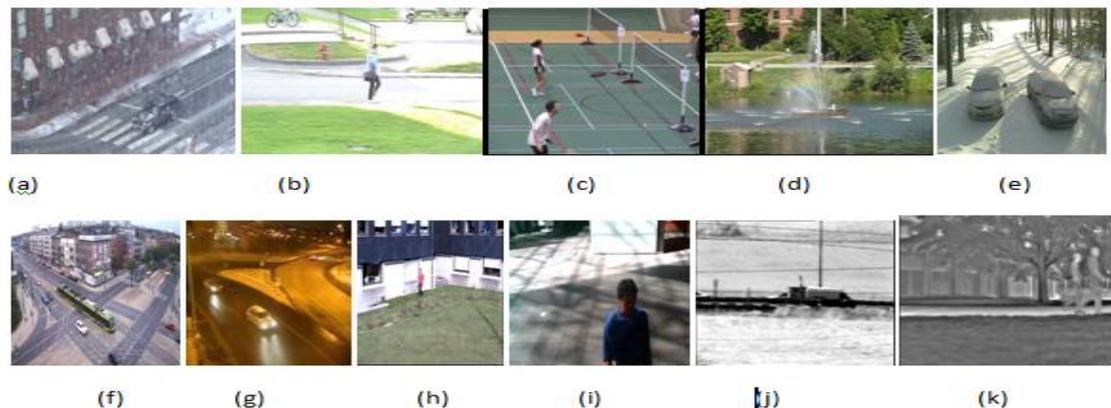


Figure 4- 6 : Exemples d'images de chaque séquence vidéo de la base CDnet 2014. (a) mauvais temps BDW, (b) la ligne de base BSL, (c) instabilité de la caméra CJT, (d) fond dynamique DBG, (e) mouvement d'objet intermittent IOM, (f) faible taux de trames LFM,

(g) vidéos de nuit NVD, (h) PTZ, (i) ombre SHD, (j) thermique THML, (k) turbulence TRBL.

Afin de mesurer la qualité des algorithmes de soustraction de fond les plus récents, nous calculons les différentes métriques d'évaluation pour ces méthodes. De plus, nous comparons les F Mesure (FM) de toutes les méthodes de soustractions de fond.

6.1 Résultat d'évaluation des algorithmes de soustraction de fond pour CDnet 2014

Subsense donne des valeurs élevées de la mesure FM dans 6 catégories parmi les 11 catégories (voir Tableau 4-1). Pour les catégories telles que la catégorie BSL, la catégorie PTZ et la catégorie LFM, IMBS est la méthode la plus précise. Pour la catégorie IOM, Adaptive SOM est la meilleure méthode. Enfin la meilleure méthode pour la catégorie THML est la méthode texture based MRF.

L'algorithme Subsense utilise des informations spatio-temporelles basées sur la couleur et la texture pour détecter les changements. Cela permet aux objets de premier plan camouflés d'être détectés plus facilement tout en ignorant la plupart des variations d'éclairage. En outre, au lieu d'utiliser des constantes définies à l'échelle de l'ensemble du cadre pour dicter la sensibilité et la vitesse d'adaptation du modèle, une boucle de rétroaction au niveau des pixels pour ajuster dynamiquement les paramètres internes de la méthode est utilisée sans intervention de l'utilisateur [93].

Subsense peut être utilisé efficacement dans des scénarios de surveillance complexes présentant de nombreux défis simultanément.

La sortie des algorithmes est une image binaire (masque de soustraction de fond) qui contient les informations de classification des trames de la vidéo actuelle. Les objets de premier plan sont marqués comme pixels blancs et les pixels noirs représentent les pixels appartenant au modèle d'arrière-plan (voir figure 4-7).

Méthode / catégorie	LOBSTER	SUBSENSE	UV	UM	TEXTURE	IMBS	MULTICUE	ADAPTI VE SOM	FUZZY GAUSSIAN	MAX FM
FM BDW	53,05	80,24	45,4	0	50,58	65,69	57,43	27,48	12,69	80,24
FM BSL	91,84	95,45	42,34	0	65,39	95,49	72,01	69,41	27,45	95,49
FM DBG	82,39	94,37	5,97	0	4,18	85,32	47,86	9,4	2,09	94,37
FM IOM	27,49	36,63	14,07	0	63,3	45,79	27,09	89,27	42,29	89,27
FM LFM	43,64	87,54	36,46	0	67,7	91,03	72,86	79,87	12,17	91,03
FM NVD	34,49	45,07	26,56	0	43,32	41,81	18,56	34,14	18,56	45,07
FM SHAD	89,4	89,83	16,8	0	81,62	85,35	84,99	85,04	28,04	89,83
FM PTZ	0,98	4,98	1,15	0	0,66	46,92	0,82	0,51	0,43	46,92
FM TRBL	86,16	88,12	9,91	0	19,15	79,65	80,82	48,76	18,56	88,12
FM TRML	40,24	52,93	52,73	0	86,96	72,04	83,42	68,74	42,29	86,96
FM CJT	72,84	96,49	20,67	0	9,48	89,85	77,54	51,32	18,08	96,49

Tableau 4- 1 : Comparaison de la métrique FM (F Mesure) de différents algorithmes de soustraction de fond sur les vidéos originales des 11 catégories du CDnet2014.

Métrique / catégorie	TP	FP	FN	Rappel	Précision	F mesure
FM BDW	4399074	550317	1615789	73,13	88,88	80,24
FM BSL	629241	18363	41621	93,79	97,16	95,45
FM DBG	246598	8762	20631	92,27	96,56	94,37
FM IOM	505144	1561405	186226	73,06	24,44	36,63
FM LFM	298554	27920	57017	83,96	91,44	87,54
FM NVD	1326681	2560917	671844	66,38	34,12	45,07
FM SHAD	4801922	983011	103434	97,89	83	89,83
FM PTZ	45250	1719362	5477	89,2	2,56	4,98
FM TRBL	3118312	64165	776274	80,06	97,98	88,12
FM TRML	277131	54871	438006	38,75	83,47	52,93
FM CJT	4030178	224696	68129	98,33	94,71	96,49

Tableau 4- 2 :Évaluation de la méthode SUBSENSE sur les vidéos originales des 11 catégories du CDnet2014.

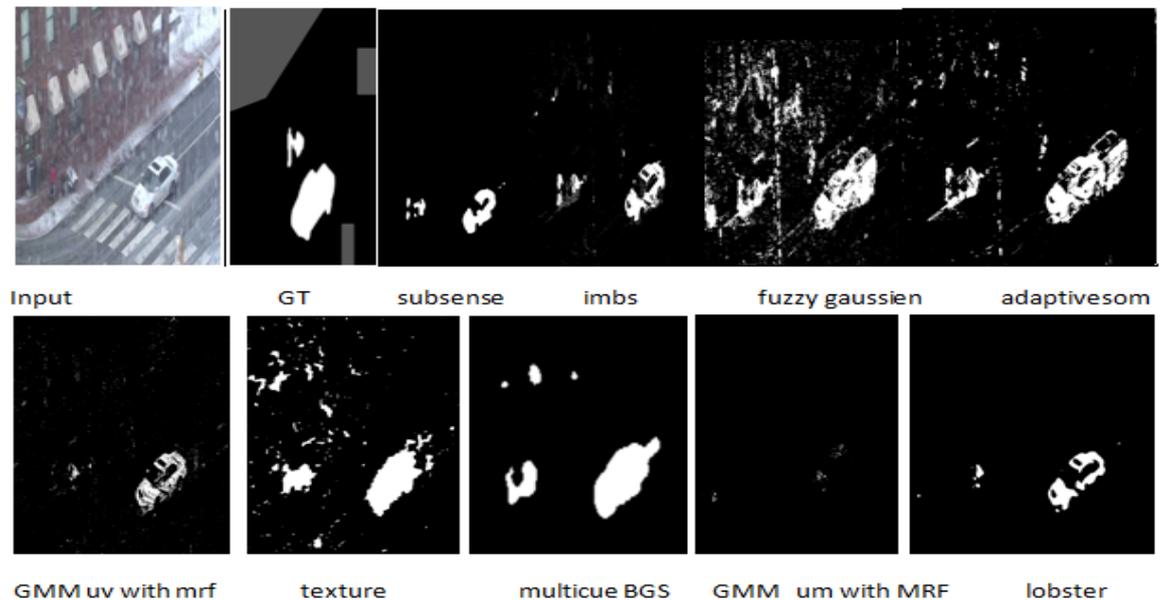


Figure 4- 7 : Résultats qualitatifs de soustraction de fond obtenus à partir du subsense, IMBS, Fuzzy Gaussien, adaptive SOM, GMM uv with MRF, Texture, multicue BGS, LOBSTER, les images vidéo sont extraites de la catégorie mauvais temps BDW. (GT) est le masque de vérité terrain (ground truth).

6.2 Résultats d'évaluation des algorithmes de soustraction de fond après détection de contour

Sobel est un algorithme connu utilisé pour la détection de contour. L'opérateur Sobel, parfois appelé opérateur Sobel–Feldman ou filtre Sobel, est utilisé dans le traitement des images et la vision par ordinateur, en particulier dans les algorithmes de détection de bord où il crée une image accentuant les bords.

Techniquement, il s'agit d'un opérateur de différenciation discret, calculant une approximation du gradient de la fonction d'intensité d'image. À chaque point de l'image, le résultat de l'opérateur Sobel–Feldman est le vecteur de gradient correspondant à la norme de ce vecteur. L'opérateur Sobel–Feldman est basé sur la convolution de l'image avec un filtre de petite taille, séparable et entier, dans les directions horizontales et verticales. Il est donc relativement peu coûteux en termes de calculs. En revanche, l'approximation de gradient qu'il produit est relativement grossière, en particulier pour les variations à haute fréquence de l'image [105].

Subsense dans ce cas-là possède la plus grande valeur de FM dans 7 catégories parmi les 11 catégories avant et après la détection de contour (voir Tableau4-3 et Tableau 4-4). Pour la catégorie intermittent Object Motion et la catégorie Shadow, texture réalise une bonne soustraction et donne les plus haut FM. Pour les catégories turbulence et thermal la méthode Multicue est la meilleure.

Méthode / catégorie	LOBSTER	SUBSENSE	UV	UM	TEXTURE	IMBS	MULTICUE	ADAPTIVE SOM	FUZZY GAUSSIAN	MAX FM
FM BDW	43,25	72,52	7,38	0	20,11	40,77	51,1	16,77	5,56	72,52
FM BSL	78,61	81,23	0	0	64,44	69,29	53,97	18,59	6,42	81,23
FM DBG	47,99	87,36	1,89	0	3,74	55,48	25,54	12,54	12,554	87,36
FM IOM	29,06	22,34	9,41	0	61,43	42,3	15,98	5,83	25,48	61,43
FM LFM	77,94	85,14	33,02	0	63,12	67,14	69,49	47,64	10,05	85,14
FM NVD	46,07	61,38	30,38	0	41,14	44,27	54,98	38,25	6,85	61,38
FM SHAD	53,87	57,09	9,41	0	78,59	44,32	66,42	40,27	0,24	78,59
FM PTZ	2,04	2,08	1,36	0	0,64	1,64	0,57	0,68	0,51	2,08
FM TRBL	40,09	57,33	7,38	0	61,43	38,18	66,45	19,68	7,37	66,45
FM TRML	58,59	66,14	1,36	0	73,8	43,95	76,15	36,76	7,76	76,15
FM CJT	33,92	59,56	14,96	0	8,99	34,44	58,4	25,84	25,86	59,56

Tableau 4- 3 : Comparaison de la métrique FM de différents algorithmes de soustraction de fond sur les vidéos après l'extraction de contour des 11 catégories du CDnet2014.

Métrique / catégorie	TP	FP	FN	Rappel	Précision	F mesure
FM BDW	3710254	505941	2304609	61,68	88	72,52
FM BSL	656406	288819	14456	97,84	69,44	81,23
FM DBG	215817	11030	51412	80,76	95,13	87,36
FM IOM	414046	2600104	277324	59,88	13,73	22,34
FM LFM	275638	16209	79933	77,51	94,44	85,14
FM NVD	1371629	1098980	626896	68,63	55,51	61,38
FM SHAD	2165571	514717	2739785	44,14	80,79	57,09
FM PTZ	47438	4446905	3289	93,51	1,05	2,08
FM TRBL	1623832	146360	2270754	41,69	91,73	57,33
FM TRML	376130	46067	339007	52,59	89,08	66,14
FM CJT	1946049	489470	2152258	47,48	79,9	59,56

Tableau 4- 4 :Évaluation de la méthode SUBSENSE sur les vidéos des 11 catégories du CDnet2014 après l'extraction du contour.

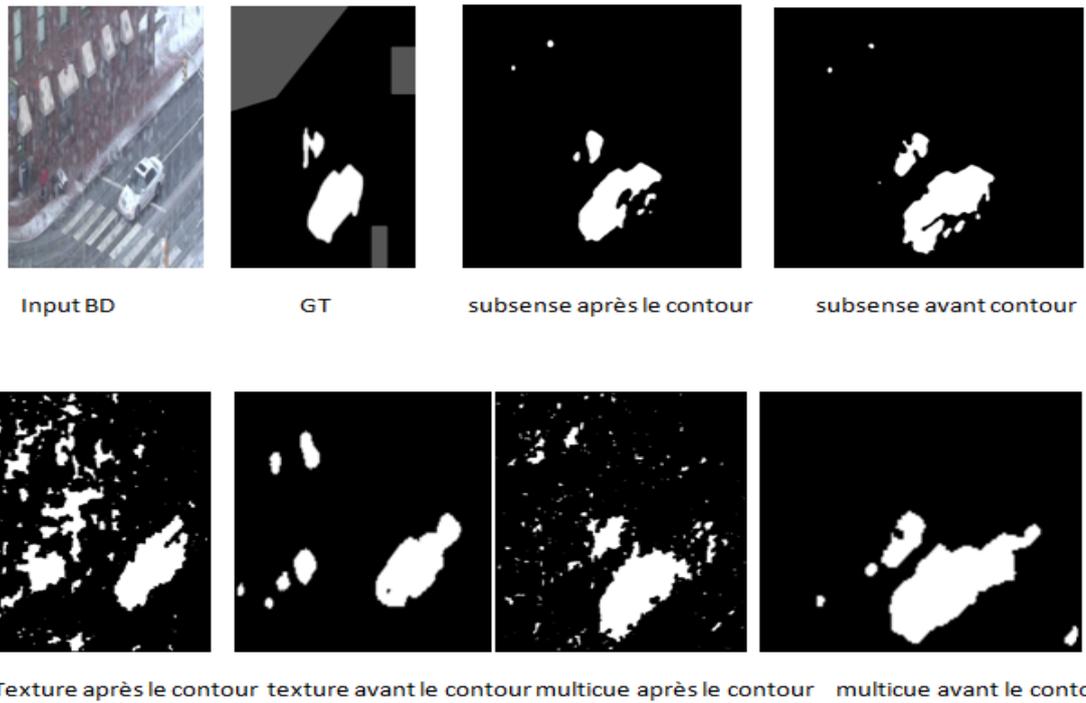


Figure 4- 8 : Résultats qualitatifs de soustraction de fond de la méthode SUBSENSE, Multicue et Texture with MRF avant et après extraction du contour sur les séquences vidéos de CDnet 2014 pour la catégorie BDW.

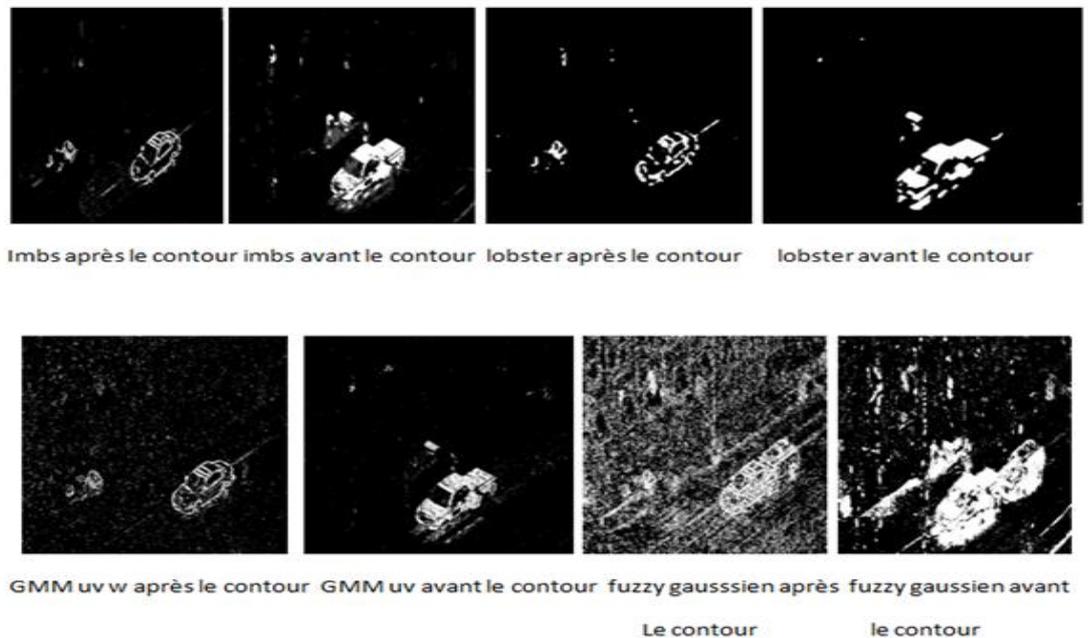


Figure 4- 9 : Résultats qualitatifs de soustraction de fond des algorithmes IMBS, LOBSTER, GMM uv with MRF et Fuzzy Gaussien, sur les séquences de vidéos de CDnet 2014 pour la catégorie BDW.

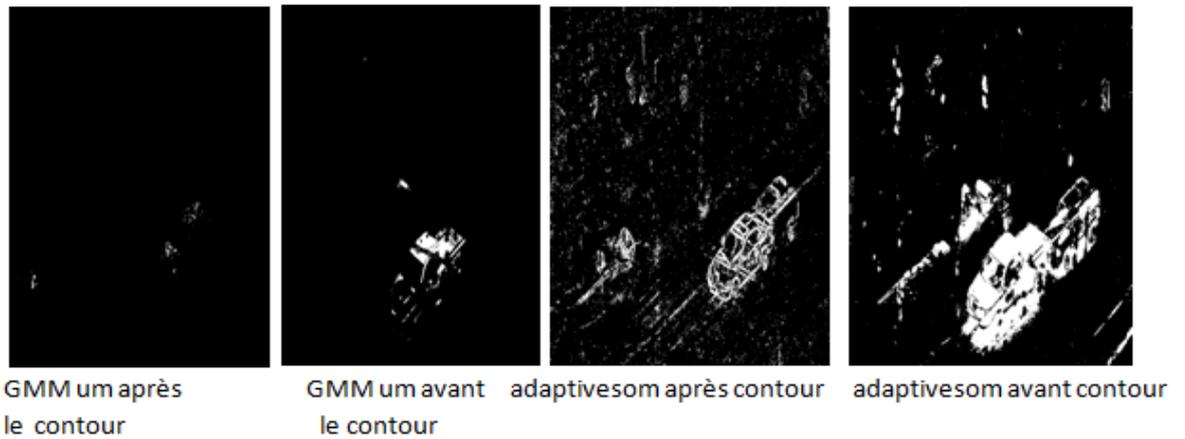


Figure 4- 10 : Résultats qualitatifs de soustraction de fond des algorithmes GMM um with MRF et adaptive SOM sur les séquences vidéos de CDnet 2014 pour la catégorie BDW.

LOBSTER réussi avec une meilleure performance dans la catégorie Baseline. Dans les catégories Badweather, Dynamic Background, Shadaw, Turbulence ainsi que la catégorie thermal, subsense donne des meilleurs résultats donc elle est la méthode idéale. La méthode Texture based foreground with mrf donne un bon résultat en intermittent Object Motion, lowFramerate, Thermal et en camera Jitter. IMBS effectue une bonne soustraction de fond car cette dernière possède une meilleure précision (Tableau 4-5).

Méthode / catégorie	LOBSTER	SUBSENSE	UV	UM	TEXTURE	IMBS	MULTICUE	ADAPTIVE SOM	FUZZY GAUSSIA N	MAX FM
FM BDW	65,26	82,23	53,05	0	19,09	72,84	47,23	20,84	6,47	82,23
FM BSL	86,72	81,48	31,49	0	56,85	84,74	52,66	23,21	7,06	86,72
FM DBG	12,54	94,37	5,97	0	4,18	85,32	47,86	2,93	0,8	94,37
FM IOM	25,95	22,83	11,14	0	59,22	38,09	15,5	6,75	3,14	59,22
FM LFM	43,64	89,89	38,67	0	56,59	90,02	63,85	61,88	9,24	90,02
FM NVD	35,52	47,5	25,14	0	37,81	43,15	54,23	34,82	16,32	54,23
FM SHAD	52,64	88,77	16,8	0	25,32	84,93	81,15	70,57	20,75	88,77
FM PTZ	0,93	1,52	1,24	0	0,62	1,42	0,56	0,48	0,42	1,52
FM TRBL	85,17	87,78	8,65	0	7,21	79,83	79,47	27,87	7,6	87,78
FM TRML	65,37	69,69	29,36	0	71,97	78,54	78,57	53,06	9,6	78,57
FM CJT	72,84	81,25	20,67	0	9,48	89,85	77,54	51,32	18,08	89,85

Tableau 4- 5 : Comparaison de métrique FM de combinaison de soustraction de fond et la détection de contour sur les 11 catégories de CDnet2014.

Les résultats expérimentaux ont prouvé que la combinaison des résultats obtenus par les algorithmes de soustraction de fond sur les séquences de vidéo de CDnet 2014 sans filtre et les résultats avec filtre sobel avec la même base d'images donne une soustraction de fond plus précise au niveau qualitative et quantitative pour toutes les méthodes testées (voir Figure 4-11, Figure 4-12 et Figure 4.13).

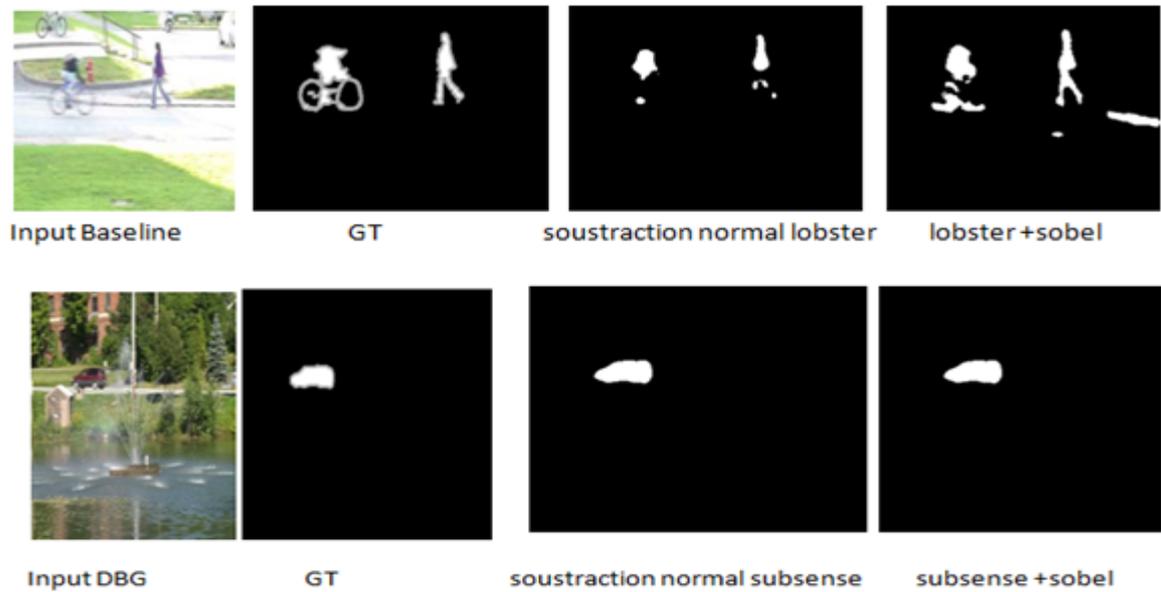


Figure 4- 11 : Différence au niveau qualitatif entre la soustraction de fond normale et la combinaison entre le résultat de soustraction de fond sur les séquences vidéo originales de CDnet, sur les séquences après extraction du contour en utilisant le filtre de Sobel et combinaison des deux résultats. Méthodes LOBSTER et Subsense.

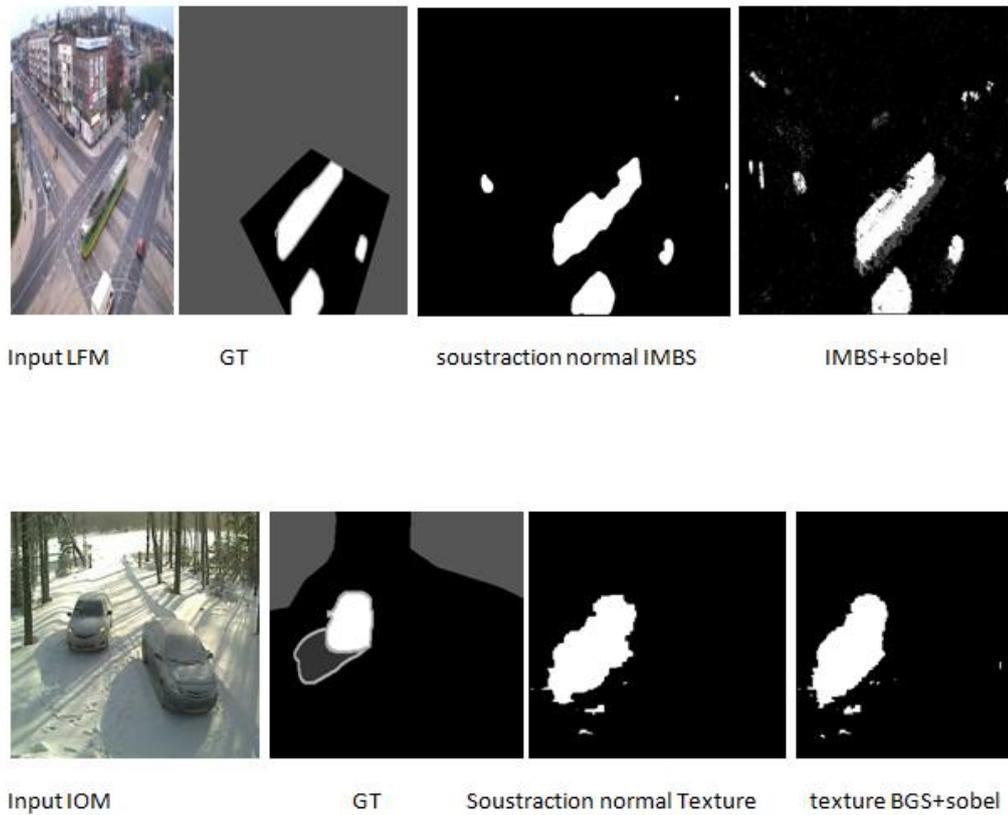


Figure 4- 12 : Différence au niveau qualitatif entre la soustraction de fond normale et la combinaison entre le résultat de soustraction de fond sur les séquences vidéo originales de CDnet, sur les séquences après extraction du contour en utilisant le filtre de Sobel et combinaison des deux résultats. Méthodes IMBS et Texture.

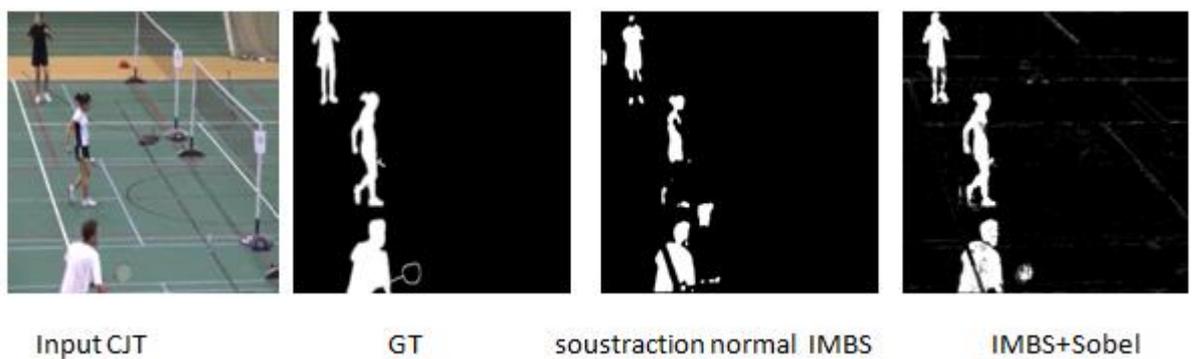


Figure 4- 13 : Différence au niveau qualitatif entre la soustraction de fond normale et la combinaison entre le résultat de soustraction de fond sur les séquences vidéo originales de CDnet, sur les séquences après extraction du contour en utilisant le filtre de Sobel et combinaison des deux résultats. Méthode IMBS.

Conclusion

Dans ce chapitre nous avons présenté l'environnement de développement, l'architecture de déploiement et tous les choix techniques que nous avons adoptés pour la réalisation de l'application.

Ensuite nous avons comparé les résultats quantitatifs et qualitatifs des récents algorithmes de soustraction de fond avant et après l'utilisation de la méthode d'extraction du contour. Nous avons également comparé ces résultats aux résultats obtenus en combinant les deux techniques. Les résultats quantitatifs sont calculés à l'aide des métriques de performance.

Nous avons conclu que la méthode subsense est la meilleure méthode de soustraction de fond parmi les méthodes considérées dans les trois cas testé (sans contour et avec contour ainsi qu'après la combinaison des deux techniques). Nous avons aussi conclu que la combinaison des masques d'avant-plan de soustraction de fond résultant des méthodes en utilisant la base d'images CDnet 2014 sans filtre Sobel et avec filtre donne des résultats plus précis au niveau qualitatif et quantitatif et que l'utilisation de la méthode de contour Sobel sans la combinaison donne des résultats moins précis. En effet, l'analyse qualitative a montré que les objets reflètent mieux leurs classes respectives et leur forme peut être utilisée pour d'autres opérations telles que le suivi, la reconnaissance ou la classification.

CONCLUSION GENERALE

Conclusion générale

Tout au long de la préparation de notre projet de fin d'études, nous avons pu mettre en pratique les connaissances acquises durant nos études universitaires et cela dans le but de réaliser une application de détection d'objet en mouvement sur une scène vidéo.

Dans ce travail nous nous sommes intéressés à la conception et la réalisation d'une application de soustraction de fond pour des vidéos issues de cameras fixes.

Ce projet de fin d'étude consistait en une étude comparative entre les différentes méthodes de soustraction de fond, les appliquer sur des séquences de vidéos, inclure la détection de contour, puis combiner les masques obtenus avec un filtre de détection de contour et sans filtre pour enfin calculer les métriques et analyser la précision des approches choisis.

Lors de la réalisation de ce projet, une étude bibliographique a été faite sur la vision par ordinateur, la soustraction de fond et ses problèmes, ainsi que les méthodes de soustraction de fond. Cette étude a été complétée par une étude conceptuelle de l'application et une implémentation des solutions proposées aux différentes questions posées durant la réalisation de cette approche. Enfin, des tests ont été réalisés en utilisant notre système, et nous avons conclu les points suivants.

- La méthode Subsense est la meilleure méthode d'après les résultats obtenus dans un ensemble considérable de cas.
- La combinaison de la technique d'extraction de contour avec une méthode de soustraction de fond améliore les résultats quantitatifs et qualitatifs pour toutes les méthodes testées.
- Chacun des algorithmes étudiés est plus efficace que d'autres dans des cas précis.
- Il n'existe pas une méthode de soustraction idéale parmi les méthodes testées qui puisse gérer tous les problèmes à la fois.

Comme perspective à notre travail, nous souhaitons pour les prochains travaux de recherche dans cet axe :

- Réaliser un algorithme de soustraction de fond qui englobe une nouvelle technique de contour.
- Faire une étude comparative entre d'autres méthodes de détection de contour et étudier les performances obtenues (filtre laplacien, extraction de contour basée sur les arbres de décision, etc.).

- Proposer une méthode ou un système qui analyse les séquences de vidéos et à partir de cette analyse, définit le type de brouillage au niveau des images et propose l'algorithme de soustraction de fond le plus adapté.

RÉFÉRENCES
BIBLIOGRAPHIQUE

BIBLIOGRAPHIE

- [1] HORAUD, R., & MONGA, O. VISION PAR ORDINATEUR: OUTILS FONDAMENTAUX, (1995). EDITIONS HERMÈS, PARIS.
- [2] SZELISKI, R. (2010). COMPUTER VISION: ALGORITHMS AND APPLICATIONS. SPRINGER SCIENCE & BUSINESS MEDIA.
- [3] SNYDER, W. E., & QI, H. (2017). FUNDAMENTALS OF COMPUTER VISION. CAMBRIDGE UNIVERSITY PRESS.
- [4] AMAZOUZ, S., & MEHDI, T. (2016). ESTIMATION DU MOUVEMENT D'UNE CAMERA DANS UNE SEQUENCE D'IMAGES (MASTER)
- [5] KOU, W. (1995). DIGITAL IMAGE COMPRESSION: ALGORITHMS AND STANDARDS (VOL. 333). SPRINGER SCIENCE & BUSINESS MEDIA.
- [6] HISTACE, A. (2012). IMAGE RESTORATION-RECENT ADVANCES AND APPLICATIONS.
- [7] GUNTURK, B. K., & LI, X. (EDS.). (2012). IMAGE RESTORATION: FUNDAMENTALS AND ADVANCES. CRC PRESS.
- [8] SENTHILSEVI, A., & SUKUMAR, R. (2014). A SURVEY ON IMAGE RESTORATION TECHNIQUES. INTERNATIONAL JOURNAL OF EMERGING ENGINEERING RESEARCH AND TECHNOLOGY, 2(8), 123-128.
- [9] PARK, B., & LU, R. (EDS.). (2015). HYPERSPECTRAL IMAGING TECHNOLOGY IN FOOD AND AGRICULTURE. NEW YORK, NY, USA: SPRINGER.
- [10] MENDOZA, F., & LU, R. (2015). BASICS OF IMAGE ANALYSIS. IN HYPERSPECTRAL IMAGING TECHNOLOGY IN FOOD AND AGRICULTURE (PP. 9-56). SPRINGER, NEW YORK, NY.
- [11] MING, L. (2010, AUGUST). IMAGE SEGMENTATION ALGORITHM RESEARCH AND IMPROVEMENT. IN 2010 3RD INTERNATIONAL CONFERENCE ON ADVANCED COMPUTER THEORY AND ENGINEERING (ICACTE) (VOL. 5, PP. V5-211). IEEE.
- [12] SYU, J. H., WANG, S. J., & WANG, L. C. (2017). HIERARCHICAL IMAGE SEGMENTATION BASED ON ITERATIVE CONTRACTION AND MERGING. IEEE TRANSACTIONS ON IMAGE PROCESSING, 26(5), 2246-2260.
- [13] LOPES, D. (2012). IMPORTANT CONCEPTS IN SIGNAL PROCESSING, IMAGE PROCESSING AND DATA COMPRESSION. WORLD TECHNOLOGIES.

-
- [14] SHAIKH, S. H., SAEED, K., & CHAKI, N. (2014). MOVING OBJECT DETECTION USING BACKGROUND SUBTRACTION. IN MOVING OBJECT DETECTION USING BACKGROUND SUBTRACTION (PP. 15-23). SPRINGER, CHAM.
- [15] GREFF, K., BRANDÃO, A., KRAUB, S., STRICKER, D., & CLUA, E. (2012, FEBRUARY). A COMPARISON BETWEEN BACKGROUND SUBTRACTION ALGORITHMS USING A CONSUMER DEPTH CAMERA. IN VISAPP (1) (PP. 431-436).
- [16] KOMAGAL, E., & VINODHINI, A. (2012, FEBRUARY). REAL TIME BACKGROUND SUBTRACTION TECHNIQUES FOR DETECTION OF MOVING OBJECTS IN VIDEO SURVEILLANCE SYSTEM. IN 2012 INTERNATIONAL CONFERENCE ON COMPUTING, COMMUNICATION AND APPLICATIONS(PP. 1-5). IEEE.
- [17] RAMESH, V. (2003, OCTOBER). BACKGROUND MODELING AND SUBTRACTION OF DYNAMIC SCENES. IN PROCEEDINGS NINTH IEEE INTERNATIONAL CONFERENCE ON COMPUTER VISION (PP. 1305-1312). IEEE.
- [18] BRUTZER, S., HÖFERLIN, B., & HEIDEMANN, G. (2011, JUNE). EVALUATION OF BACKGROUND SUBTRACTION TECHNIQUES FOR VIDEO SURVEILLANCE. IN CVPR 2011 (PP. 1937-1944). IEEE.
- [19] BOUWMANS, T., & GARCIA-GARCIA, B. (2019). BACKGROUND SUBTRACTION IN REAL APPLICATIONS: CHALLENGES, CURRENT MODELS AND FUTURE DIRECTIONS. ARXIV PREPRINT ARXIV:1901.03577]
- [20] BOUWMANS, T. (2014). TRADITIONAL AND RECENT APPROACHES IN BACKGROUND MODELING FOR FOREGROUND DETECTION: AN OVERVIEW. COMPUTER SCIENCE REVIEW, 11, 31-66.
- [21] EL BAF, F., BOUWMANS, T., & VACHON, B. (2008, MAY). FOREGROUND DETECTION USING THE CHOQUET INTEGRAL. IN 2008 NINTH INTERNATIONAL WORKSHOP ON IMAGE ANALYSIS FOR MULTIMEDIA INTERACTIVE SERVICES(PP. 187-190). IEEE.
- [22] MOULAY, A. A., & AMINE, A. ANOVEL BACKGROUND SUBTRACTION ALGORITHM FOR PERSON TRACKING BASED ON K-NN
- [23] VOSTERS, L. P., SHAN, C., & GRITTI, T. (2010, AUGUST). BACKGROUND SUBTRACTION UNDER SUDDEN ILLUMINATION CHANGES. IN 2010 7TH IEEE INTERNATIONAL CONFERENCE ON ADVANCED VIDEO AND SIGNAL BASED SURVEILLANCE (PP. 384-391). IEEE.

-
- [24] YAO, G., LEI, T., ZHONG, J., JIANG, P., & JIA, W. (2017). COMPARATIVE EVALUATION OF BACKGROUND SUBTRACTION ALGORITHMS IN REMOTE SCENE VIDEOS CAPTURED BY MWIR SENSORS. *SENSORS*, 17(9), 1945.
- [25] BOUWMANS, T., & GARCIA-GARCIA, B. (2019). BACKGROUND SUBTRACTION IN REAL APPLICATIONS: CHALLENGES, CURRENT MODELS AND FUTURE DIRECTIONS. *ARXIV PREPRINT ARXIV:1901.03577*.
- [26] LEYKIN, A., & TUCERYAN, M. (2005). A VISION SYSTEM FOR AUTOMATED CUSTOMER TRACKING FOR MARKETING ANALYSIS: LOW LEVEL FEATURE EXTRACTION. IN *HUMAN ACTIVITY RECOGNITION AND MODELLING WORKSHOP (VOL. 3, P. 6)*.
- [27] MOLINA-GIRALDO, S., INSUASTI-CEBALLOS, H. D., ARROYAVE, C. E., MONTOYA, J. F., LOPEZ-VILLA, J. S., ALVAREZ-MEZA, A., & CASTELLANOS-DOMINGUEZ, G. (2015, SEPTEMBER). PEOPLE DETECTION IN VIDEO STREAMS USING BACKGROUND SUBTRACTION AND SPATIAL-BASED SCENE MODELING. IN *2015 20TH SYMPOSIUM ON SIGNAL PROCESSING, IMAGES AND COMPUTER VISION (STSIVA) (PP. 1-6)*. IEEE
- [28] NOCETI, N., ZINI, L., & ODONE, F. (2018). A MULTI-CAMERA SYSTEM FOR DAMAGE AND TAMPERING DETECTION IN A POSTAL SECURITY FRAMEWORK. *EURASIP JOURNAL ON IMAGE AND VIDEO PROCESSING*, 2018(1), 11.
- [29] ZHANG, D., O'CONNOR, E., MCGUINNESS, K., O'CONNOR, N. E., REGAN, F., & SMEATON, A. (2012, NOVEMBER). A VISUAL SENSING PLATFORM FOR CREATING A SMARTER MULTI-MODAL MARINE MONITORING NETWORK. IN *PROCEEDINGS OF THE 1ST ACM INTERNATIONAL WORKSHOP ON MULTIMEDIA ANALYSIS FOR ECOLOGICAL DATA (PP. 53-56)*. ACM.
- [30] MONTEIRO, G., MARCOS, J., RIBEIRO, M., & BATISTA, J. (2008, OCTOBER). ROBUST SEGMENTATION FOR OUTDOOR TRAFFIC SURVEILLANCE. IN *2008 15TH IEEE INTERNATIONAL CONFERENCE ON IMAGE PROCESSING (PP. 2652-2655)*. IEEE.
- [31] MONTEIRO, G., MARCOS, J., & BATISTA, J. (2008). STOPPED VEHICLE DETECTION SYSTEM FOR OUTDOOR TRAFFIC SURVEILLANCE. *RECPAD 2008*.
- [32] MUNIRUZZAMAN, S., HAQUE, N., RAHMAN, F., SIAM, M., MUSABBIR, R., HADIUZZAMAN, M., & HOSSAIN, S. (2016). DETERMINISTIC ALGORITHM FOR TRAFFIC DETECTION IN FREE-FLOW AND CONGESTION USING VIDEO SENSOR. *JOURNAL OF BUILT ENVIRONMENT, TECHNOLOGY AND ENGINEERING*.

-
- [33] LIN, R., CAO, X., XU, Y., WU, C., & QIAO, H. (2009, JUNE). AIRBORNE MOVING VEHICLE DETECTION FOR VIDEO SURVEILLANCE OF URBAN TRAFFIC. IN 2009 IEEE INTELLIGENT VEHICLES SYMPOSIUM (PP. 203-208). IEEE.
- [34] REZAEI, B., & OSTADABBAS, S. (2018). MOVING OBJECT DETECTION THROUGH ROBUST MATRIX COMPLETION AUGMENTED WITH OBJECTNESS. IEEE JOURNAL OF SELECTED TOPICS IN SIGNAL PROCESSING, 12(6), 1313-1323.
- [35] ASHRAPHIJUO, M., AGGARWAL, V., & WANG, X. (2017). ON DETERMINISTIC SAMPLING PATTERNS FOR ROBUST LOW-RANK MATRIX COMPLETION. IEEE SIGNAL PROCESSING LETTERS, 25(3), 343-347.
- [36] KNAUER¹, U., HIMMELSBACH¹, M., WINKLER¹, F., ZAUTKE, F., BIENEFELD, K., & MEFFERT¹, B. (2005). APPLICATION OF AN ADAPTIVE BACKGROUND MODEL FOR MONITORING HONEYBEES.
- [37] ZHONG, J. (2003, OCTOBER). SEGMENTING FOREGROUND OBJECTS FROM A DYNAMIC TEXTURED BACKGROUND VIA A ROBUST KALMAN FILTER. IN PROCEEDINGS NINTH IEEE INTERNATIONAL CONFERENCE ON COMPUTER VISION (PP. 44-50). IEEE.
- [38] BORGHGRAEF, A., BARNICH, O., LAPIERRE, F., VAN DROOGENBROECK, M., PHILIPS, W., & ACHEROY, M. (2010). AN EVALUATION OF PIXEL-BASED METHODS FOR THE DETECTION OF FLOATING OBJECTS ON THE SEA SURFACE. EURASIP JOURNAL ON ADVANCES IN SIGNAL PROCESSING, 2010(1), 978451.
- [39] ALI, I., MILLE, J., & TOUGNE, L. (2014). ADDING A RIGID MOTION MODEL TO FOREGROUND DETECTION: APPLICATION TO MOVING OBJECT DETECTION IN RIVERS. PATTERN ANALYSIS AND APPLICATIONS, 17(3), 567-585.
- [40] BATISTA, J., PEIXOTO, P., FERNANDES, C., & RIBEIRO, M. (2006, SEPTEMBER). A DUAL-STAGE ROBUST VEHICLE DETECTION AND TRACKING FOR REAL-TIME TRAFFIC MONITORING. IN 2006 IEEE INTELLIGENT TRANSPORTATION SYSTEMS CONFERENCE (PP. 528-535). IEEE
- [41] JOHN, F., HIPINY, I., UJIR, H., & SUNAR, M. S. (2018). ASSESSING PERFORMANCE OF AEROBIC ROUTINES USING BACKGROUND SUBTRACTION AND INTERSECTED IMAGE REGION. ARXIV PREPRINT ARXIV:1810.01564.
- [42] TAMÁS, B. (2016). DETECTING AND ANALYZING ROWING MOTION IN VIDEOS. IN BME SCIENTIFIC STUDENT CONF (PP. 1-29).

-
- [43] LAURENTINI, A. (1994). THE VISUAL HULL CONCEPT FOR SILHOUETTE-BASED IMAGE UNDERSTANDING. *IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE*, 16(2), 150-162.
- [44] LEVIN, G. (2006). COMPUTER VISION FOR ARTISTS AND DESIGNERS: PEDAGOGIC TOOLS AND TECHNIQUES FOR NOVICE PROGRAMMERS. *AI & SOCIETY*, 20(4), 462-482.
- [45] BOUWMANS, T. (2011). RECENT ADVANCED STATISTICAL BACKGROUND MODELING FOR FOREGROUND DETECTION-A SYSTEMATIC SURVEY. *RECENT PATENTS ON COMPUTER SCIENCE*, 4(3), 147-176.
- [46] EL BAF, F., BOUWMANS, T., & VACHON, B. (2007, JUNE). COMPARISON OF BACKGROUND SUBTRACTION METHODS FOR A MULTIMEDIA APPLICATION. IN 2007 14TH INTERNATIONAL WORKSHOP ON SYSTEMS, SIGNALS AND IMAGE PROCESSING AND 6TH EURASIP CONFERENCE FOCUSED ON SPEECH AND IMAGE PROCESSING, MULTIMEDIA COMMUNICATIONS AND SERVICES (PP. 385-388). IEEE.
- [49] EBRAHIMI, T. (1997). MPEG-4 VIDEO VERIFICATION MODEL: A VIDEO ENCODING/DECODING ALGORITHM BASED ON CONTENT REPRESENTATION. *SIGNAL PROCESSING: IMAGE COMMUNICATION*, 9(4), 367-384.
- [50] CHIEN, S. Y., MA, S. Y., & CHEN, L. G. (2002). EFFICIENT MOVING OBJECT SEGMENTATION ALGORITHM USING BACKGROUND REGISTRATION TECHNIQUE. *IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY*, 12(7), 577-586.
- [51] HUANG, H., FANG, X., YE, Y., ZHANG, S., & ROSIN, P. L. (2017). PRACTICAL AUTOMATIC BACKGROUND SUBSTITUTION FOR LIVE VIDEO. *COMPUTATIONAL VISUAL MEDIA*, 3(3), 273-284.
- [52] KUMAR, A. N., & SURESHKUMAR, C. (2013, FEBRUARY). BACKGROUND SUBTRACTION BASED ON THRESHOLD DETECTION USING MODIFIED K-MEANS ALGORITHM. IN 2013 INTERNATIONAL CONFERENCE ON PATTERN RECOGNITION, INFORMATICS AND MOBILE ENGINEERING (PP. 378-382). IEEE.
- [53] YAO, G., LEI, T., ZHONG, J., JIANG, P., & JIA, W. (2017). COMPARATIVE EVALUATION OF BACKGROUND SUBTRACTION ALGORITHMS IN REMOTE SCENE VIDEOS CAPTURED BY MWIR SENSORS. *SENSORS*, 17(9), 1945.
- [54] BHARTI, T. T. (2013). BACKGROUND SUBTRACTION TECHNIQUE REVIEW. *INTERNATIONAL JOURNAL OF INNOVATIVE TECHNOLOGY AND EXPLORING ENGINEERING (IJITEE)*, 2(3), 166-168.

-
- [55] THU, N. T. H. (2018). RESEARCH ON COORDINATE ROTATION DIGITAL COMPUTER HARDWARE ARCHITECTURES AND APPLICATIONS.
- [56] SEHAIRI, K., CHOUIREB, F., & MEUNIER, J. (2017). COMPARATIVE STUDY OF MOTION DETECTION METHODS FOR VIDEO SURVEILLANCE SYSTEMS. JOURNAL OF ELECTRONIC IMAGING, 26(2), 023025.
- [57] RAO, C. S., & DARWIN, P. (2012). FRAME DIFFERENCE AND KALMAN FILTER TECHNIQUES FOR DETECTION OF MOVING VEHICLES IN VIDEO SURVEILLANCE. INT. J. ENG. RES. APPL. IJERA, 2(6), 1168-1170.
- [58] ZHAN, C., DUAN, X., XU, S., SONG, Z., & LUO, M. (2007, AUGUST). AN IMPROVED MOVING OBJECT DETECTION ALGORITHM BASED ON FRAME DIFFERENCE AND EDGE DETECTION. IN FOURTH INTERNATIONAL CONFERENCE ON IMAGE AND GRAPHICS (ICIG 2007) (PP. 519-523). IEEE.
- [59] DAS, D., & SAHARIA, D. (2014). IMPLEMENTATION AND PERFORMANCE EVALUATION OF BACKGROUND SUBTRACTION ALGORITHMS. ARXIV PREPRINT ARXIV:1405.1815.
- [60] VARWANI, H., CHOITHWANI, H., SAHATIYA, K., GANGAN, S., GYANCHANDANI, T., & MANE, D. (2013). UNDERSTANDING VARIOUS TECHNIQUES FOR BACKGROUND SUBTRACTION AND IMPLEMENTATION OF SHADOW DETECTION. INTERNATIONAL JOURNAL OF COMPUTER TECHNOLOGY AND APPLICATIONS, 4(5), 822.
- [61] SRINIVASAN, K., PORKUMARAN, K., & SAINARAYANAN, G. (2009, AUGUST). IMPROVED BACKGROUND SUBTRACTION TECHNIQUES FOR SECURITY IN VIDEO APPLICATIONS. IN 2009 3RD INTERNATIONAL CONFERENCE ON ANTI-COUNTERFEITING, SECURITY, AND IDENTIFICATION IN COMMUNICATION (PP. 114-117). IEEE.
- [62] MARTIN, C. (2014). BACKGROUND SUBTRACTION USING RUNNING GAUSSIAN AVERAGE: A COLOR CHANNEL COMPARISON. IN SEMINAR AUS BILDVERARBEITUNG UND MUSTERERKENNUNG.
- [63] TABKHI, H., BUSHEY, R., & SCHIRNER, G. (2013, NOVEMBER). ALGORITHM AND ARCHITECTURE CO-DESIGN OF MIXTURE OF GAUSSIAN (MOG) BACKGROUND SUBTRACTION FOR EMBEDDED VISION. IN 2013 ASILOMAR CONFERENCE ON SIGNALS, SYSTEMS AND COMPUTERS (PP. 1815-1820). IEEE.
- [64] VEGA-HERNÁNDEZ, D., HERRERA-NAVARRO, A. M., & JIMÉNEZ-HERNÁNDEZ, H. (2012, NOVEMBER). BACKGROUND SUBTRACTION MODEL BASED ON ADAPTABLE MOG. IN 2012 IEEE NINTH ELECTRONICS, ROBOTICS AND AUTOMOTIVE MECHANICS CONFERENCE(PP. 54-59). IEEE

-
- [65] MOHAMED, S. S., TAHIR, N. M., & ADNAN, R. (2010, MAY). BACKGROUND MODELLING AND BACKGROUND SUBTRACTION PERFORMANCE FOR OBJECT DETECTION. IN 2010 6TH INTERNATIONAL COLLOQUIUM ON SIGNAL PROCESSING & ITS APPLICATIONS (PP. 1-6). IEEE.
- [66] TAH, A., ROY, S., DAS, P., & MITRA, A. (2017). MOVING OBJECT DETECTION AND SEGMENTATION USING BACKGROUND SUBTRACTION BY KALMAN FILTER. INDIAN J. OF SCIENCE AND TECH, 10.
- [67] AHMAD, K. A., SAAD, Z., ABDULLAH, N., HUSSAIN, Z., & NOOR, M. M. (2011, MARCH). MOVING VEHICLE SEGMENTATION IN A DYNAMIC BACKGROUND USING SELF-ADAPTIVE KALMAN BACKGROUND METHOD. IN 2011 IEEE 7TH INTERNATIONAL COLLOQUIUM ON SIGNAL PROCESSING AND ITS APPLICATIONS (PP. 439-442). IEEE.
- [68] SEKAR, G., & DEEPIKA, M. (2015). COMPLEX BACKGROUND SUBTRACTION USING KALMAN FILTER. INTERNATIONAL JOURNAL OF ENGINEERING RESEARCH AND APPLICATIONS, 5(3), 15-20.
- [69] CHEN, LC, A. HERMANS, G. PAPANDREOU, F. SCHROFF, P. WANG ET H., H. (2018). MASKLAB: SEGMENTATION D'INSTANCE EN AFFINANT LA DETECTION D'OBJET A L'AIDE DE CARACTERISTIQUES SEMANTIQUES ET DE DIRECTION. DANS ACTES DE LA CONFERENCE DE L'IEEE SUR LA VISION PAR ORDINATEUR ET LA RECONNAISSANCE DE FORMES (PP. 4013-4022).
- [70] LIM, L. A., & KELES, H. Y. (2018). LEARNING MULTI-SCALE FEATURES FOR FOREGROUND SEGMENTATION. ARXIV PREPRINT ARXIV:1808.01477.
- [71] LIM, L. A., & KELES, H. Y. (2018). FOREGROUND SEGMENTATION USING A TRIPLET CONVOLUTIONAL NEURAL NETWORK FOR MULTISCALE FEATURE ENCODING. ARXIV PREPRINT ARXIV:1801.02225.
- [72] WANG, K., GOU, C., & WANG, F. Y. (2018). M^4 CD: A ROBUST CHANGE DETECTION METHOD FOR INTELLIGENT VISUAL SURVEILLANCE. IEEE ACCESS, 6, 15505-15520.
- [73] BRAHAM, M., PIÉRARD, S., & VAN DROOGENBROECK, M. (2017, SEPTEMBER). SEMANTIC BACKGROUND SUBTRACTION. IN 2017 IEEE INTERNATIONAL CONFERENCE ON IMAGE PROCESSING (ICIP) (PP. 4552-4556). IEEE.
- [74] SEHAIRI, K., CHOUIREB, F., & MEUNIER, J. (2017). COMPARATIVE STUDY OF MOTION DETECTION METHODS FOR VIDEO SURVEILLANCE SYSTEMS. JOURNAL OF ELECTRONIC IMAGING, 26(2), 023025.

-
- [75] MARTINS, I., CARVALHO, P., CORTE-REAL, L., & ALBA-CASTRO, J. L. (2017, JUNE). BMOG: BOOSTED GAUSSIAN MIXTURE MODEL WITH CONTROLLED COMPLEXITY. IN IBERIAN CONFERENCE ON PATTERN RECOGNITION AND IMAGE ANALYSIS (PP. 50-57). SPRINGER, CHAM.
- [76] BABAEI, M., DINH, D. T., & RIGOLL, G. (2018). A DEEP CONVOLUTIONAL NEURAL NETWORK FOR VIDEO SEQUENCE BACKGROUND SUBTRACTION. *PATTERN RECOGNITION*, 76, 635-649.
- [77] JIANG, S., & LU, X. (2018). WESAMBE: A WEIGHT-SAMPLE-BASED METHOD FOR BACKGROUND SUBTRACTION. *IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY*, 28(9), 2105-2115.
- [78] VARGHESE, A., & SREELEKHA, G. (2017). SAMPLE-BASED INTEGRATED BACKGROUND SUBTRACTION AND SHADOW DETECTION. *IPSIJ TRANSACTIONS ON COMPUTER VISION AND APPLICATIONS*, 9(1), 25.
- [79] BARNICH, O., & VAN DROOGENBROECK, M. (2011). ViBE: A UNIVERSAL BACKGROUND SUBTRACTION ALGORITHM FOR VIDEO SEQUENCES. *IEEE TRANSACTIONS ON IMAGE PROCESSING*, 20(6), 1709-1724.
- [80] VAN DROOGENBROECK, M., & PAQUOT, O. (2012, JUNE). BACKGROUND SUBTRACTION: EXPERIMENTS AND IMPROVEMENTS FOR ViBE. IN 2012 IEEE COMPUTER SOCIETY CONFERENCE ON COMPUTER VISION AND PATTERN RECOGNITION WORKSHOPS (PP. 32-37). IEEE.
- [81] VAN DROOGENBROECK, M., & BARNICH, O. (2014). ViBE: A DISRUPTIVE METHOD FOR BACKGROUND SUBTRACTION. *BACKGROUND MODELING AND FOREGROUND DETECTION FOR VIDEO SURVEILLANCE*, 7-1.
- [82] MADDALENA, L., & PETROSINO, A. (2019). SELF-ORGANIZING BACKGROUND SUBTRACTION USING COLOR AND DEPTH DATA. *MULTIMEDIA TOOLS AND APPLICATIONS*, 78(9), 11927-11948.
- [83] MADDALENA, L., & PETROSINO, A. (2010). A FUZZY SPATIAL COHERENCE-BASED APPROACH TO BACKGROUND/FOREGROUND SEPARATION FOR MOVING OBJECT DETECTION. *NEURAL COMPUTING AND APPLICATIONS*, 19(2), 179-186.
- [84] ST-CHARLES, P. L., BILODEAU, G. A., & BERGEVIN, R. (2015). SUBSENSE: A UNIVERSAL CHANGE DETECTION METHOD WITH LOCAL ADAPTIVE SENSITIVITY. *IEEE TRANSACTIONS ON IMAGE PROCESSING*, 24(1), 359-373.

-
- [85] WANG, R., BUNYAK, F., SEETHARAMAN, G., & PALANIAPPAN, K. (2014). STATIC AND MOVING OBJECT DETECTION USING FLUX TENSOR WITH SPLIT GAUSSIAN MODELS. IN PROCEEDINGS OF THE IEEE CONFERENCE ON COMPUTER VISION AND PATTERN RECOGNITION WORKSHOPS (PP. 414-418).
- [86] CHACON-MURGUIA, M. I., GONZALEZ-DUARTE, S., & VEGA-PINEDA, J. (2009, JUNE). SIMPLIFIED SOM-NEURAL MODEL FOR VIDEO SEGMENTATION OF MOVING OBJECTS. IN INT. JOINT CONF. ON NEURAL NETWORKS (PP. 474-480).
- [87] PARKS, D. H., & FELS, S. S. (2008, SEPTEMBER). EVALUATION OF BACKGROUND SUBTRACTION ALGORITHMS WITH POST-PROCESSING. IN 2008 IEEE FIFTH INTERNATIONAL CONFERENCE ON ADVANCED VIDEO AND SIGNAL BASED SURVEILLANCE (PP. 192-199). IEEE.
- [88] RAO, G. M., & SATYANARAYANA, C. (2014). OBJECT TRACKING SYSTEM USING APPROXIMATE MEDIAN FILTER, KALMAN FILTER AND DYNAMIC TEMPLATE MATCHING. INTERNATIONAL JOURNAL OF INTELLIGENT SYSTEMS AND APPLICATIONS, 6(5), 83.
- [89] LUMENTUT, J. S., & GUNAWAN, F. E. (2015). EVALUATION OF RECURSIVE BACKGROUND SUBTRACTION ALGORITHMS FOR REAL-TIME PASSENGER COUNTING AT BUS RAPID TRANSIT SYSTEM. PROCEA COMPUTER SCIENCE, 59, 445-453.
- [90] MOHAMAD, A. A., & OSMAN, M. (2013, AUGUST). ADAPTIVE MEDIAN FILTER BACKGROUND SUBTRACTIONS TECHNIQUE USING FUZZY LOGIC. IN 2013 INTERNATIONAL CONFERENCE ON COMPUTING, ELECTRICAL AND ELECTRONIC ENGINEERING (ICCEEE)(PP. 115-120). IEEE.
- [91] ST-CHARLES, P. L., & BILODEAU, G. A. (2014, MARCH). IMPROVING BACKGROUND SUBTRACTION USING LOCAL BINARY SIMILARITY PATTERNS. IN IEEE WINTER CONFERENCE ON APPLICATIONS OF COMPUTER VISION (PP. 509-515). IEEE.
- [92] BARNICH, O., & VAN DROOGENBROECK, M. (2010). ViBE: A UNIVERSAL BACKGROUND SUBTRACTION ALGORITHM FOR VIDEO SEQUENCES. IEEE TRANSACTIONS ON IMAGE PROCESSING, 20(6), 1709-1724.
- [93] ST-CHARLES, P. L., BILODEAU, G. A., & BERGEVIN, R. (2014). FLEXIBLE BACKGROUND SUBTRACTION WITH SELF-BALANCED LOCAL SENSITIVITY. IN PROCEEDINGS OF THE IEEE CONFERENCE ON COMPUTER VISION AND PATTERN RECOGNITION WORKSHOPS (PP. 408-413).

-
- [94]NOH, S., & JEON, M. (2012, NOVEMBER). A NEW FRAMEWORK FOR BACKGROUND SUBTRACTION USING MULTIPLE CUES. IN ASIAN CONFERENCE ON COMPUTER VISION (PP. 493-506). SPRINGER, BERLIN, HEIDELBERG.
- [95]ZHAO, Z., BOUWMANS, T., ZHANG, X., & FANG, Y. (2012, DECEMBER). A FUZZY BACKGROUND MODELING APPROACH FOR MOTION DETECTION IN DYNAMIC BACKGROUNDS. IN INTERNATIONAL CONFERENCE ON MULTIMEDIA AND SIGNAL PROCESSING (PP. 177-185). SPRINGER, BERLIN, HEIDELBERG.
- [96]BLOISI, D., & IOCCHI, L. (2012, SEPTEMBER). INDEPENDENT MULTIMODAL BACKGROUND SUBTRACTION. IN COMPIMAGE (PP. 39-44).
- [97]KERTÉSZ, C., & OY, V. (2011). TEXTURE-BASED FOREGROUND DETECTION.
- [98]MADDALENA, L., & PETROSINO, A. (2008). A SELF-ORGANIZING APPROACH TO BACKGROUND SUBTRACTION FOR VISUAL SURVEILLANCE APPLICATIONS. IEEE TRANSACTIONS ON IMAGE PROCESSING, 17(7), 1168-1177.
- [99]SIGARI, M. H., MOZAYANI, N., & POURREZA, H. (2008). FUZZY RUNNING AVERAGE AND FUZZY BACKGROUND SUBTRACTION: CONCEPTS AND APPLICATION. INTERNATIONAL JOURNAL OF COMPUTER SCIENCE AND NETWORK SECURITY, 8(2), 138-143.
- [100]. YOO, S., ET KIM, C. (2013). SOUSTRATION DE FOND EN UTILISANT LE CODAGE DES CARACTERISTIQUES HYBRIDES DANS LA STRUCTURE DES SACS DE CARACTERISTIQUES. PATTERN RECOGNITION LETTERS , 34 (16), 2086-2093.
- [101]WANG, Y., JODOIN, P.-M., PORIKLI, F., KONRAD, J., BENEZETH, Y., & ISHWAR, P. (2014). CDNET 2014: AN EXPANDED CHANGE DETECTION BENCHMARK DATASET. 2014 IEEE CONFERENCE ON COMPUTER VISION AND PATTERN RECOGNITION WORKSHOPS.
- [102] SOBRAL, A. (2013, JUNE). BGS LIBRARY: AN OPENCV C++ BACKGROUND SUBTRACTION LIBRARY. IN IX WORKSHOP DE VISAO COMPUTACIONAL (VOL. 2, NO. 6, P. 7).
- [103]GOYETTE, N., JODOIN, P. M., PORIKLI, F., KONRAD, J., & ISHWAR, P. (2012, JUNE). CHANGEDETECTION. NET: A NEW CHANGE DETECTION BENCHMARK DATASET. IN 2012 IEEE COMPUTER SOCIETY CONFERENCE ON COMPUTER VISION AND PATTERN RECOGNITION WORKSHOPS (PP. 1-8). IEEE
- [104] BABAEE, M., DINH, D. T., & RIGOLL, G. (2018). A DEEP CONVOLUTIONAL NEURAL NETWORK FOR VIDEO SEQUENCE BACKGROUND SUBTRACTION. PATTERN RECOGNITION, 76, 635-649.

-
- [105] SOBEL, I. (2014). HISTORY AND DEFINITION OF THE SOBEL OPERATOR. RETRIEVED FROM THE WORLD WIDE WEB, 1505.
- [106] ST-CHARLES, P. L., & BILODEAU, G. A. (2014, MARCH). IMPROVING BACKGROUND SUBTRACTION USING LOCAL BINARY SIMILARITY PATTERNS. IN IEEE WINTER CONFERENCE ON APPLICATIONS OF COMPUTER VISION (PP. 509-515). IEEE.
- [107] PRATI, A., MIKIC, I., TRIVEDI, M. M., & CUCCHIARA, R. (2003). DETECTING MOVING SHADOWS: ALGORITHMS AND EVALUATION. IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE, 25(7), 918-923.
- [108] ELGAMMAL, A., HARWOOD, D., & DAVIS, L. (2000, JUNE). NON-PARAMETRIC MODEL FOR BACKGROUND SUBTRACTION. IN EUROPEAN CONFERENCE ON COMPUTER VISION (PP. 751-767). SPRINGER, BERLIN, HEIDELBERG.

WEBOGRAPHIE

- [W1][HTTPS://FR.WIKIPEDIA.ORG/WIKI/FILTRE_DE_SOBEL](https://fr.wikipedia.org/wiki/Filtre_de_Sobel) (DEFINITION DU FILTRE SOBEL) [CONSULTE LE 17/07/2019].
- [W2][HTTPS://FR.WIKIPEDIA.ORG/WIKI/ADDITION_MATRICIELLE](https://fr.wikipedia.org/wiki/Addition_matricielle)(DEFINITION DE L'ADDITION MATRICIELLE) [CONSULTE 17/07/2019]

-
- [W3] [HTTPS://FR.WIKIPEDIA.ORG/WIKI/DIAGRAMME_DES_CAS_D'UTILISATION](https://fr.wikipedia.org/wiki/Diagramme_des_cas_d'utilisation)(DEFINITION DIAGRAMME CAS D'UTILISATION) [CONSULTE LE 17/07/2019]
- [W4] [HTTPS://FR.WIKIPEDIA.ORG/WIKI/DIAGRAMME_DE_CLASSES](https://fr.wikipedia.org/wiki/Diagramme_de_classes) (DEFINITION DE DIAGRAMME CAS D'UTILISATION) [CONSULTE LE 17/07/2019].
- [W5] [HTTPS://FR.WIKIPEDIA.ORG/WIKI/DIAGRAMME_DE_S%C3%A9QUENCE](https://fr.wikipedia.org/wiki/Diagramme_de_s%C3%A9quence) (DEFINITION DE DIAGRAMME DE SEQUENCE) [CONSULTE LE 17/07/2019].
- [W6] [HTTPS://FR.WIKIPEDIA.ORG/WIKI/STARUML](https://fr.wikipedia.org/wiki/Staruml)(DEFINITION DE STAR UML) [CONSULTE LE 24/07/2019]
- [W7] [HTTPS://FR.WIKIPEDIA.ORG/WIKI/QT](https://fr.wikipedia.org/wiki/Qt) (DEFINITION QT) [CONSULTE LE 24/07/2019]
- [W8] [HTTPS://FR.WIKIPEDIA.ORG/WIKI/OPENCV](https://fr.wikipedia.org/wiki/OpenCV) (DEFINITION DE OPENCV) [CONSULTE LE 24/07/2019]
- [W9] [HTTP://WWW.EMGU.COM/WIKI/INDEX.PHP/MAIN_PAGE](http://www.emgu.com/wiki/index.php/Main_Page)(DEFINITION DE EMGU CV) [CONSULTE 24/07/2019]
- [W10] [HTTPS://FR.WIKIPEDIA.ORG/WIKI/CMAKE](https://fr.wikipedia.org/wiki/Cmake) (DEFINITION DE CMAKE) [CONSULTE LE 24/07/2019]
- [W11] [HTTPS://FR.WIKIPEDIA.ORG/WIKI/C%2B%2B](https://fr.wikipedia.org/wiki/C%2B%2B) (DEFINITION DE C++) [CONSULTE LE 24/07/2019]
- [12] [HTTPS://FR.WIKIPEDIA.ORG/WIKI/C_SHARP](https://fr.wikipedia.org/wiki/C_sharp)(DEFINITION DEC#) [CONSULTE LE 24/07/2019]
- [W13] [HTTPS://FR.WIKIPEDIA.ORG/WIKI/MICROSOFT_VISUAL_STUDIO](https://fr.wikipedia.org/wiki/Microsoft_Visual_Studio) (DEFINITION DE VISUEL STUDIO) [CONSULTE LE 24/07/2019]
- [W14] [HTTPS://FR.WIKIPEDIA.ORG/WIKI/GITHUB](https://fr.wikipedia.org/wiki/Github) (DEFINITION DE GITHUB) [CONSULTE LE 24/07/2019]

