

الجمهورية الجزائرية
الجمهورية الجزائرية الديمقراطية الشعبية
République Algérienne Démocratique et Populaire
وزارة التعليم العالي و البحث العلمي
Ministère de l'enseignement Supérieur et de la recherche scientifique



Université Mohamed Khider Biskra
Faculté des Sciences et de la Technologie
Département de Génie Electrique
Filière : Electronique
Option : Réseaux et télécommunication

Mémoire de Fin d'Etudes
En vue de l'obtention du diplôme:

MASTER

Thème

*Suivi d'objets avec Mean shift en utilisant BWH
et CBWH*

Présenté par :

LAMARI Iliasse

Avis favorable de l'encadreur :

MEDOUAKH Saadia

Avis favorable du Président du Jury

BAARIR Zine-Eddine

Cachet et signature

الجمهورية الجزائرية الديمقراطية الشعبية
République Algérienne Démocratique et Populaire
وزارة التعليم العالي والبحث العلمي
Ministère de l'enseignement Supérieur et de la Recherche scientifique



Université Mohamed Khider Biskra
Faculté des Sciences et de la Technologie
Département de Génie Electrique
Filière : Electronique
Option : Réseaux et Télécommunication

Thème :
Suivi d'objets avec Mean shift en utilisant BWH
et CBWH

Proposé par : MEDOUAKH Saadia

Dirigé par : MEDOUAKH Saadia

RESUMES (Français et Arabe)

Résumé

Le suivi d'objets est une tâche importante dans la vision par ordinateur. De nombreux algorithmes ont été proposés pour résoudre les divers problèmes posés par les bruits, les encombrements et les occultations, etc. Parmi les différents algorithmes de suivi, le tracker Mean shift est l'un des algorithmes de suivi les plus efficaces pour les applications en temps réel, en raison de sa simplicité et de sa robustesse. Dans ce travail, nous nous intéressons au suivi d'objet avec l'algorithme Mean Shift en utilisant l'histogramme pondéré corrigé en fonction du fond CBWH (Corrected Background-Weighted Histogram) pour réduire les interférences du fond dans la localisation de la cible. L'algorithme CBWH a corrigé la formule BWH (Background-Weighted Histogram) en transformant uniquement le modèle cible mais pas le modèle candidat cible. Cet algorithme peut effectivement réduire l'interférence de l'arrière-plan dans la localisation de la cible et peut conduire à une convergence plus rapide et à une localisation plus précise que la représentation cible habituelle dans le tracker Mean shift. Les algorithmes Mean shift, BWH et CBWH sont appliqués à base de données OTB qui contient la vérité terrain et leur comparaison mutuelle par rapport à leurs performances respectives concernant des points de vue qualitatifs (visuel) et quantitatifs (erreur de localisation de centre et taux de recouvrement) a été effectuée.

Mots clés: suivi d'objet, mean shift, histogramme pondéré, BWH, CBWH.

المخلص:

تتبع الأشياء مهمة مهمة في رؤية الكمبيوتر. لقد تم اقتراح العديد من الخوارزميات لحل المشكلات المختلفة التي تسببها الضوضاء والازدحام والسحابات وما إلى ذلك. من بين خوارزميات التتبع المختلفة، يعد متتبع Mean shift واحدًا من أكثر خوارزميات التتبع فعالية للتطبيقات في الوقت الفعلي نظرًا لبساطته وقوته. في هذا العمل، نحن مهتمون بتتبع الأشياء مع خوارزمية Mean Shift باستخدام المدرج التكراري للخلفية المصححة (CBWH) لتقليل تداخل الخلفية في الموقع المستهدف. قامت خوارزمية CBWH بتصحيح صيغة BWH (رسم بياني مرجح مع الخلفية) من خلال تحويل النموذج المستهدف فقط ولكن ليس نموذج المرشح المستهدف. يمكن لهذه الخوارزمية أن تقلل بشكل فعال من تداخل الخلفية في الموقع المستهدف ويمكن أن تؤدي إلى تقارب أسرع وتوطين أكثر دقة من تمثيل الهدف المعتاد في متتبع Mean shift. يتم تطبيق خوارزميات Mean shift و BWH و CBWH على قاعدة بيانات OTB التي تحتوي على القيم الحقيقية ومقارنتها المتبادلة فيما يتعلق بأداء كل منها في وجهات النظر النوعية (المرئية) والكمية (خطأ موقع المركز ونسبة التداخل).

كلمات مفتاحية: تتبع الأشياء، رسم بياني مرجح، CBWH، BWH.

Abstract :

Object tracking is an important task in computer vision. Many algorithms have been proposed to solve the various problems posed by noise, congestion and occultations, etc. Among the different tracking algorithms, the Mean shift tracker is one of the most effective tracking algorithms for real-time applications because of its simplicity and robustness. In this work, we are interested in object tracking with the Mean Shift algorithm using the Corrected Background-Weighted Histogram (CBWH) corrected background histogram to reduce background interference in target location. The CBWH algorithm corrected the BWH (Background-Weighted Histogram) formula by transforming only the target model but not the target candidate model. This algorithm can effectively reduce background interference in target location and can lead to faster convergence and more accurate localization than the usual target representation in the Mean shift tracker. The Mean shift, BWH and CBWH algorithms are applied to the OTB database which contains the terrain truth and their mutual comparison with respect to their respective performances regarding qualitative (visual) and quantitative (center location error and recovery) has been done.

Key words: object tracking, mean shift, weighted histogram, BWH, CBWH.



Université Mohamed Khider de Biskra
Faculté des Sciences et de la Technologie
Département de Génie Electrique

MÉMOIRE DE MASTER

Sciences et Technologies
Télécommunications
Réseaux et télécommunications

Réf. :

Présenté et soutenu par :
LAMARI Iliasse Charef eddine

Le : lundi 8 juillet 2019

Suivi d'objets par Mean Shift en utilisant BWH et CBWH

Jury :

M.	BAARIR Zine-Eddine	Pr	Université de Biskra	Président
M.	TOUMI Abida	MCA	Université de Biskra	Examineur
M.	MEDOUAKH Saadia	MAA	Université de Biskra	Encadreur

Dédicace :

Avec grande reconnaissance je dédie ce travail qui n'aura jamais pu voir le jour sans les soutiens indéfectibles et sans limite de mes chers parents qui ne cessent de me donner le nécessaire pour que je puisse arriver à ce que je suis aujourd'hui. Que dieux vous protège

Je dédie aussi ce travail à :

- A ma sœur pour son support.*
- A mon encadreur pour ses efforts et conseils.*
- A tous mes professeurs et enseignants.*
- A tous les membres de ma famille et mes proches pour leur soutien.*
- A tous mes amis et mes connaissances pour leurs encouragements.*
- A toutes personnes ayant contribué à ce travail de près ou de loin sans exception.*

Remerciement :

Je remercie tout d'abord « Allah » de m'avoir donné le courage, la force et la volonté d'entamer et finir ce travail dans de bonnes conditions.

Je tiens à remercier chaleureusement mon encadreur Dr. MEDOUAKH Saadia d'avoir encadré ce travail pour son soutien et ses conseils pour la réalisation de celui-ci. Je remercie également sincèrement l'ensemble des mes professeurs qui m'ont permis grâce aux enseignements dispensés tout au long de mes études universitaires d'arriver jusque-là aujourd'hui. Je remercie également les membres du jury Prof. BAARIR Zine-Eddine et Dr. TOUMI Abida d'avoir accepté d'évaluer ce travail.

Enfin, un chaleureux remercie va naturellement vers mes chers parents, toute ma famille et mes amis qui ont su m'apporter le soutien nécessaire à l'accomplissement de mes efforts et de mon travail.

ILLASSE.

Liste des tableaux :

Chapitre 3 : Simulation, Résultats expérimentaux et discussion

Tableau 3.1 : Séquences d'images utilisées.....	41
Tableau 3.2 : Les moyennes des résultats obtenus de l'erreur de localisation du centre et du taux de recouvrement. .	53

Liste des figures :

Chapitre 1 : Généralités sur le suivi d'objets

Figure 1.1 : Exemple de suivi d'objets5
Figure 1.2 : Exemples d'applications de suivi d'objet 6
Figure 1.3 : Exemples de représentation de la forme d'objet. (a) : un point, (b) : points d'intérêt, (c) : rectangle, (d) : ellipse, (e) : multiple blocs, (f) : squelette, (g)-(h) contour, (i) : Silhouette 8
Figure 1.4 : Les défis de suivi 9
Figure 1.5 : Taxonomie des méthodes de suivi d'objets 14

Chapitre 2 : Suivi d'objet robuste avec Mean shift via CBWH

Figure 2.1 : Description intuitive de la convergence de la procédure Mean shift. 22
Figure 2.2: l'itération du Processus Mean Shift 23
Figure 2.3 : la similarité de la méthode par Mean shift 25
Figure 2.4 : Exemple de construction d'un histogramme pondéré 26
Figure 2.5 : La similarité de la méthode BWH..... 27
Figure 2.6 : La similarité de la méthode CBWH 31

Chapitre 3 : Simulation, Résultats expérimentaux et discussion

Figure 3.1 : La structure de l'algorithme de suivi par Mean shift via CBWH 37
Figure 3.2 : Exemple de la transformation d'espace couleur RGB vers HSV 38
Figure 3.3: Exemple sur l'histogramme 2D CBWH de composante S, V de l'espace HSV .. 39
Figure 3.4 : Exemple sur l'image de poids 40
Figure 3.5 : Métriques de performance : (a) Erreur de localisation du centre, (b) Taux de recouvrement. 42
Figure 3.6 : Résultats de suivi sur la séquence Bolt avec Mean shift (rectangle bleu), BWH (rectangle noir) et CBWH (rectangle rouge). 44

Liste des figures

Figure 3.7: Comparaison quantitative entre MS, BWH et CBWH la séquence Bolt : (a) Taux de recouvrement. (b) Erreur de localisation du centre:	44
Figure 3.8 : Résultats de suivi sur la séquence Bird2 avec Mean shift (rectangle bleu), BWH (rectangle noir) et CBWH (rectangle rouge)	45
Figure 3.9 : Comparaison quantitative entre MS, BWH et CBWH pour la séquence Bird2 : (a) Taux de recouvrement. (b) Erreur de localisation du centre.	46
Figure 3.10 : Résultats de suivi sur la séquence Human8 avec Mean shift (rectangle bleu), BWH (rectangle noir) et CBWH (rectangle rouge).	46
Figure 3.11 : Comparaison quantitative entre MS, BWH et CBWH pour la séquence Human8 : (a) Taux de recouvrement. (b) Erreur de localisation du centre.	47
Figure 3.12: Résultats de suivi sur la séquence Tiger1 avec Mean shift (rectangle bleu), BWH (rectangle noir) et CBWH (rectangle rouge)	48
Figure 3.13 : Comparaison quantitative entre MS, BWH et CBWH pour la séquence Tiger1 : (a) Taux de recouvrement. (b) Erreur de localisation du centre.	48
Figure 3.14 : Résultats de suivi sur la séquence David3 avec Mean shift (rectangle bleu), BWH (rectangle noir) et CBWH (rectangle rouge).	49
Figure 3.15 : Comparaison quantitative entre MS, BWH et CBWH pour la séquence David3 : (a) Taux de recouvrement. (b) Erreur de localisation du centre.	50
Figure 3.16 : Résultats de suivi sur la séquence BlurOwl avec Mean shift (rectangle bleu), BWH (rectangle noir) et CBWH (rectangle rouge).	51
Figure 3.17 : Comparaison quantitative entre MS, BWH et CBWH pour la séquence BlurOwl: (a) Taux de recouvrement. (b) Erreur de localisation du centre.	52

Liste des abréviations

q : Distribution du modèle.

$P(y)$: Distribution du candidat.

o : Le fond.

ρ : Coefficient de Bhattacharyya.

$d(y)$: Distance de Bhattacharyya.

$K(x)$: Noyau de densité.

δ : Fonction de Kronecker.

Dist : Distance.

ELC : Erreur de Localisation du Centre.

TR : Taux de recouvrement

MS : Mean shift.

BWH: Background Weighted Histogram.

CBWH: Corrected Background Weighted Histogram.

RESUMES (Français et Arabe et Anglais)

Résumé :

Le suivi d'objets est une tâche importante dans la vision par ordinateur. De nombreux algorithmes ont été proposés pour résoudre les divers problèmes posés par les bruits, les encombrements et les occultations, etc. Parmi les différents algorithmes de suivi, le tracker Mean shift est l'un des algorithmes de suivi les plus efficaces pour les applications en temps réel, en raison de sa simplicité et de sa robustesse. Dans ce travail, nous nous intéressons au suivi d'objet avec l'algorithme Mean Shift en utilisant l'histogramme pondéré corrigé en fonction du fond CBWH (Corrected Background-Weighted Histogram) pour réduire les interférences du fond dans la localisation de la cible. L'algorithme CBWH a corrigé la formule BWH (Background-Weighted Histogram) en transformant uniquement le modèle cible mais pas le modèle candidat cible. Cet algorithme peut effectivement réduire l'interférence de l'arrière-plan dans la localisation de la cible et peut conduire à une convergence plus rapide et à une localisation plus précise que la représentation cible habituelle dans le tracker Mean shift. Les algorithmes Mean shift, BWH et CBWH sont appliqués à base de données OTB qui contient la vérité terrain et leur comparaison mutuelle par rapport à leurs performances respectives concernant des points de vue qualitatifs (visuel) et quantitatifs (erreur de localisation de centre et taux de recouvrement) a été effectuée.

Mots clés : suivi d'objet, mean shift, histogramme pondéré, BWH, CBWH.

المخلص:

تتبع الأشياء مهمة مهمة في رؤية الكمبيوتر. لقد تم اقتراح العديد من الخوارزميات لحل المشكلات المختلفة التي تسببها الضوضاء والازدحام والسحابات وما إلى ذلك. من بين خوارزميات التتبع المختلفة، يعد متتبع Mean shift واحدًا من أكثر خوارزميات التتبع فعالية للتطبيقات في الوقت الفعلي نظرًا لبساطته وقوته. في هذا العمل، نحن مهتمون بتتبع الأشياء مع خوارزمية Mean Shift باستخدام المدرج التكراري للخلفية المصححة (CBWH) لتقليل تداخل الخلفية في الموقع المستهدف. قامت خوارزمية CBWH بتصحيح صيغة BWH (رسم بياني مرجح مع الخلفية) من خلال تحويل النموذج المستهدف فقط ولكن ليس نموذج المرشح المستهدف. يمكن لهذه الخوارزمية أن تقلل بشكل فعال من تداخل الخلفية في الموقع المستهدف ويمكن أن تؤدي إلى تقارب أسرع وتوطين أكثر دقة من تمثيل الهدف المعتاد في متتبع Mean shift. يتم تطبيق خوارزميات Mean shift و BWH و CBWH على قاعدة بيانات OTB التي تحتوي على القيم الحقيقية ومقارنتها المتبادلة فيما يتعلق بأداء كل منها في وجهات النظر النوعية (المرئية) والكمية (خطأ موقع المركز ونسبة التداخل).

كلمات مفتاحية: تتبع الأشياء، رسم بياني مرجح، BWH، CBWH.

Abstract :

Object tracking is an important task in computer vision. Many algorithms have been proposed to solve the various problems posed by noise, congestion and occultations, etc. Among the different tracking algorithms, the Mean shift tracker is one of the most effective tracking algorithms for real-time applications because of its simplicity and robustness. In this work, we are interested in object tracking with the Mean Shift algorithm using the Corrected Background-Weighted Histogram (CBWH) corrected background histogram to reduce background interference in target location. The CBWH algorithm corrected the BWH (Background-Weighted Histogram) formula by transforming only the target model but not the target candidate model. This algorithm can effectively reduce background interference in target location and can lead to faster convergence and more accurate localization than the usual target representation in the Mean shift tracker. The Mean shift, BWH and CBWH algorithms are applied to the OTB database which contains the terrain truth and their mutual comparison with respect to their respective performances regarding qualitative (visual) and quantitative (center location error and recovery) has been done.

Key words: object tracking, mean shift, weighted histogram, BWH, CBWH.

Sommaire

Introduction générale.....	1
----------------------------	---

Chapitre 1 : Généralités sur le suivi d'objets

1.1 Introduction	5
1.2 Le suivi d'objet.....	5
1.2.1 L'objectif du suivi d'objet.....	7
1.2.1.1 Le principe du suivi d'objet.....	7
1.2.1.2 La représentation de l'apparence de l'objet.....	7
1.2.2 Les défis du suivi.....	8
1.3 Modélisation de l'apparence de l'objet	9
1.4 Les caractéristiques visuelles	10
1.5 Localisation de l'objet	12
1.5.1 Méthode d'optimisation.....	12
1.6 Etat de l'art du suivi d'objet	13
1.6.1 Suivi par détection.....	14
1.6.1.1 Méthode générative	14
1.6.1.2 Méthode discriminative	15
1.6.2 Suivi par correspondance	16
1.6.3 Suivi par filtre de corrélation.....	17
1.6.4 Suivi par deep learning	18
1.7 Conclusion.....	19

Chapitre 2 : Suivi d'objet robuste par Mean shift via CBWH

2.1 Introduction	21
2.2 Etat de l'art sur l'algorithme Mean shift	21
2.3 Procédure Mean shift.....	22
2.3.1 Histogramme pondéré.....	25
2.4 Suivi par Mean shift via BWH	27
2.4.1 Algorithme de l'histogramme pondéré de fond (BWH) :	28
2.4.2 L'équivalence de la représentation de BWH à la représentation habituelle	29
2.5 Suivi par mean shift via CBWH	30
2.5.1 Mise à jour du modèle de fond dans CBWH.....	32
2.6 Conclusion.....	34

Chapitre 3 : Simulation, Résultats expérimentaux et discussions

3.1 Introduction	36
3.2 Structure de l’algorithme	36
3.2.1 Détection d’objet	38
3.2.2 Transformation RGB vers HSV	38
3.2.3 Histogramme CBWH.....	39
3.2.3.1 Poids et le nouveau centre	40
3.2.4. Environnement de travail	40
3.2.4.1 Environnement matériel	40
3.2.4.2 Langage de programmation	40
3.2.4.3 Bases de donnés.....	41
3.2.5 Mesure des performances d’un système de suivi	42
3.2.5.1 Taux de recouvrement	42
3.2.5.2 Erreur de localisation du centre	42
3.2.6 Résultats Expérimentaux	43
3.2.6.1 Comparaison entre Mean shift BWH et CBWH	43
3.2.6.1.1 La séquence Bolt	43
3.2.6.1.2 La séquence Bird2	45
3.2.6.1.3 La séquence Human8.....	46
3.2.6.1.4 La séquence Tiger1	47
3.2.6.1.5 La séquence David3	49
3.2.6.1.6 La séquence BlurOwl.....	50
3.2.7 Comparaison entre les moyennes d’erreur de localisation du centre et taux de recouvrement pour Mean Shift et BWH et CBWH.....	52
3.3 Conclusion	53
Conclusion générale	55
Bibliographie	58

Introduction générale

Introduction générale :

Le suivi visuel est une fonction élémentaire d'analyse des signaux visuels, présente dans le système visuel humain, et des primates supérieurs. Mais cette fonction est également essentielle pour la conception des algorithmes de vision artificielle dès lors que l'on analyse des séquences vidéo, car elle assure une cohérence et une stabilité spatio-temporelle des entités d'intérêts présents dans la scène, nécessaires à leur interprétation ou à une interaction. Le suivi visuel d'objet est une fonction élémentaire de suivi d'un seul objet, tandis que le «suivi d'objets multi-objets» se rapporte à un problème d'association de données temporelle, d'une image à la suivante, pour reconstituer les pistes des objets présents dans la scène. Ces données sont par exemple des détections issues d'un radar.

Le suivi d'objet a un rôle indispensable, nécessaire à un grand nombre d'applications d'interprétation et d'exploitation de vidéos. C'est une brique de base utilisée par d'autres fonctionnalités telles que pistage multi-objets. Selon [1] le suivi, dans sa forme la plus simple, est l'estimation de la trajectoire d'un objet en mouvement dans le plan image et la détermination de la position d'un objet cible dans chaque image d'une séquence vidéo pris par une caméra fixe ou embarquée, avec comme seules informations sa position et ses caractéristiques images à l'instant initial.

L'un des problèmes que le suivi d'objet rencontre est le phénomène de dérive qui est une mauvaise estimation de la position de l'objet et qui mène des fois à sa perte définitivement ce qui affecte la robustesse du suivi. Ce phénomène est dû à plusieurs perturbations : Occlusion, Déformation de l'objet, mouvement rapide et flou du mouvement, changement d'échelle et d'éclairage.

Plusieurs algorithmes ont été proposés [2] [3] [4] [5] certaines méthodes sont basées sur l'histogramme de couleur de l'objet, d'autres ont exploité l'information donnée par le contour de l'objet. Certaines méthodes ont utilisé le mouvement d'objets dans une séquence d'images, à savoir la technique classique du flot optique. Une récente étude de [6] a analysé plus finement les performances des trackers et a obtenu comme conclusion : les trackers se spécialisent soit en précision de localisation, soit en robustesse ; aucun tracker ne surpasse tous les autres sur tous les critères de performance (précision, robustesse) et pour toutes les perturbations visuelles.

Mean shift est l'un des algorithmes de suivi proposé, connu pour sa simplicité, sa rapidité et sa robustesse. Cet algorithme utilise le modèle d'apparence qui se base sur la densité de

couleur pour représenter l'objet cible. Mean shift suit la cible en trouvant le modèle de distribution le plus similaire dans une séquence de trames avec son modèle d'échantillon par recherche itérative et se repose sur l'histogramme pondéré de couleur afin de modéliser l'objet cible [5] [7]. L'un des majeurs défis qui affecte la robustesse du mean shift est la ressemblance entre le fond et l'objet cible. En conséquent une suggestion de combinaison de plusieurs caractéristiques visuelles a été proposée pour résoudre ce problème [1].

Dans ce mémoire, nous nous intéressons l'étude du tracker mean shift via CBWH (Corrected Background-Weighted Histogram) qui utilise les informations de fond pour réduire les interférences du fond dans la localisation de la cible où ce tracker peut fonctionner de manière plus robuste, même si le modèle cible contient beaucoup d'informations de fond. Cet algorithme se base sur les caractéristiques de couleurs du fond qui sont associés avec les caractéristiques de couleur HSV pour créer un histogramme pondéré CBWH. Une étude comparative est également faite entre les trackers Mean shift, BWH et CBWH.

Ce mémoire est divisé en 3 chapitres de la façon suivante :

- **Le chapitre 1** : tout d'abord nous découvrons le principe du suivi d'objet et les défis que les trackers sont susceptibles de rencontrer, ensuite un état de l'art sur le suivi d'objet afin de montrer la diversité des approches développées et leurs classifications.
- **Le chapitre 2** : est consacré l'état de l'art du tracker mean shift, le principe de l'algorithme mean shift et les méthodes BWH et CBWH et les améliorations qui apportent.
- **Le chapitre 3** : conclut ce mémoire en présentant la structure de l'algorithme CBWH et l'ensemble des résultats expérimentaux obtenus pour le suivi d'objets dans une séquence d'images de base de données OTB accompagnés de leurs discussions.

Nous terminerons ce mémoire par une conclusion générale.

CHAPITRE 1

Généralités sur le suivi d'objets

1 Introduction :

Les progrès de la technologie informatique ont entraîné de nombreux changements dans la manière de procéder, notamment dans la surveillance ou le contrôle d'accès. Ces dernières années, le traitement des images a tenté de transformer la caméra en œil humain et d'apporter une contribution importante aux systèmes de contrôle d'accès par le suivi d'objet qui a deux catégories : le suivi d'objet unique et le suivi d'objet multiple. Le suivi d'objet estime en permanence la trajectoire d'un objet dans l'image et localiser la position de l'objet. La plupart des techniques de suivi d'objet impliquent la création d'un modèle cible pour la représentation d'objet, l'extraction de propriétés d'objet et la différenciation d'objet cible à partir de candidats non cibles.

Dans ce chapitre, nous découvrons d'abord le principe du suivi d'objets et le défis qui se posent. Ensuite, on va trouver la représentation de la forme et de l'apparence de l'objet à l'aide de plusieurs caractéristiques, et on va présenter les méthodes de la localisation de l'objet. Enfin, nous présentons l'état de l'art des méthodes de suivi d'objets, qui sont classées en des plusieurs catégories : Suivi par détection, Suivi par correspondance, Suivi par filtre de corrélation, ... etc.

2. Le suivi d'objet :

Le suivi d'objet dans des séquences d'images est un thème de recherche très actif en vision par ordinateur. Le suivi correspond à l'estimation de la localisation de l'objet dans chacune des images d'une séquence vidéo, la caméra et/ou l'objet pouvant être simultanément en mouvement (figure.1.1). Le procédé de localisation se fonde sur la reconnaissance de l'objet d'intérêt à partir d'un ensemble de caractéristiques visuelles telles que la couleur, la forme, la vitesse, ...etc. [8]



Figure 1.1 : Exemple de suivi d'objet

De nombreuses méthodes de suivi d'objets ont été proposées dans la littérature et la différence entre ces méthodes réside en partie dans le choix de la représentation et de la forme des objets, des caractéristiques (composantes) de l'image utilisées, de la nature du mouvement estimé, etc. Ce choix dépend de l'application ainsi que de la vidéo traitée. [1]

Le suivi visuel d'objets a retenu une attention considérable de la communauté de la vision par ordinateur, durant les deux dernières décennies. Cet intérêt accru s'explique par les avancées importantes réalisées dans la modélisation de contenu visuel et par la large gamme d'applications utilisant les algorithmes de suivi : [9]

- Vidéosurveillance (détection, suivi, reconnaissance du comportement de personnes, d'intrus)
- Robotique (suivi d'obstacles pendant une phase d'évitement)
- Gestion et l'analyse du trafic (le suivi de voitures)
- Militaire (suivi de cibles et guidage de missiles)
- Imageries medical.

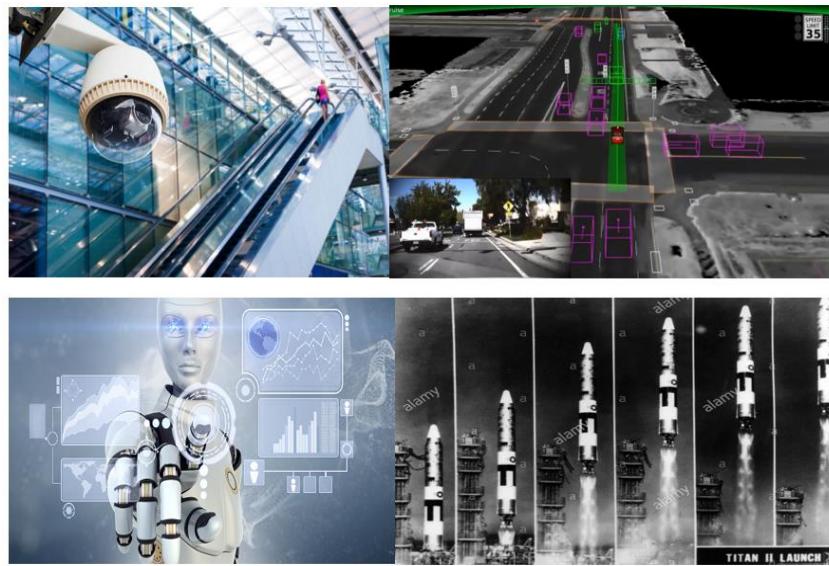


Figure 1.2 : Exemples d'applications de suivi d'objet

2.1 L'objectif du suivi :

2.1.1 Principe du suivi :

Le suivi, dans sa forme la plus simple, est l'estimation de la trajectoire d'un objet en mouvement dans le plan image [1], où un objet est un ensemble de pixels dépeignant la même représentation sémantique (ex : une voiture, un visage, un personnage...). Schématiquement, le principe consiste à extraire de l'objet des caractéristiques que l'on va s'efforcer de retrouver à chaque image de la séquence vidéo [10] En d'autres termes, le suivi détermine la position d'un objet cible dans chaque image d'une séquence vidéo. Toujours selon [1], toutes les méthodes de suivi comportent deux couches techniques : la première permet de détecter l'ensemble des candidats potentiels (objets similaires à la cible suivie) dans chaque image de la séquence et la seconde effectue la mise en correspondance d'une image à l'autre d'un de ces candidats avec la cible afin de maintenir la cohérence du suivi au fil du temps. [1][11]

Le but d'une méthode de suivi d'objet est d'estimer, dans chaque image de la séquence, les paramètres de l'objet ou des objets présents dans le champ de vision de la caméra. Les paramètres à estimer peuvent être divers, mais comprennent principalement une composante géométrique, indiquant la position dans l'image du centre de l'objet.

2.1.2 Le représentation de la forme d'un objet :

Les objets peuvent être représentés de nombreuses façons et le choix de la représentation d'un objet dépend du domaine d'application, Les représentations basées sur la forme d'un objet ont nombreuses : un ensemble de points, une forme géométrique, un contour, une silhouette, un modèle 2D ou 3D,... [12]

- **Points :**

L'objet peut être représenté par un point, Il s'agit d'une représentation simple de la localisation 2D (ou3D) de l'objet. Cette représentation se généralise à un ensemble de points aux quells peuvent être associés des descripteurs locaux de couleur, de texture ou de mouvement.

- **Formes géométriques:**

L'objet est représenté par une forme géométrique, par exemple un rectangle ou une ellipse, permettant une description de la dimension de l'objet, Cette représentation se généralise par des modèles articulés, composé d'un ensemble de formes géométriques 2D ou 3D particulièrement utilisé dans la modélisation du corps humain. [13][14].

- **Contour:**

La représentation d'un objet par son contour permet une description plus complète de sa forme. Un contour peut être vu comme étant un ensemble de points ordonnés généralement estimé à l'aide d'une analyse du gradient, d'intensité au voisinage d'un pixel. La région interne du contour est appelée la silhouette de l'objet et peut être utilisée conjointement à l'information de contour pour le suivi d'objets [15].

- **Squelette:**

Le squelette d'un objet peut être extrait pour caractériser la forme d'un objet ou d'une forme géométrique. Ce modèle est utilisé en tant que descripteur de forme pour la reconnaissance d'objets. Les auteurs utilisent le squelette des formes obtenues par un module de détection de mouvement afin d'estimer la position de la tête et d'effectuer une tâche de comptage [1].

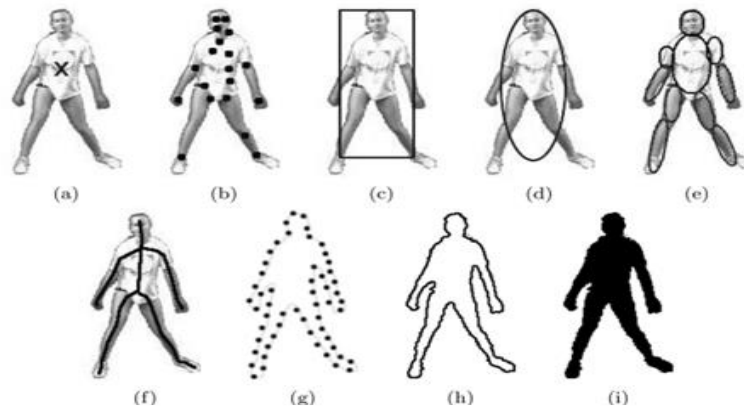


Figure 1.3 : Exemples de représentation de la forme d'objet. (a) : un point, (b) : points d'intérêt, (c) : rectangle, (d) : ellipse, (e) : multiple blocs, (f) : squelette, (g)-(h) contour, (i) : Silhouette [1]

2.2 Les défis du suivi :

Le suivi d'objets est une problématique difficile qui se pose dans un grand nombre d'applications de traitement d'images. La difficulté du suivi d'objets dépend de plusieurs facteurs relatifs aux données ou à l'application : [16]

- Changement d'illumination.
- Changement d'échelle.
- Occultations partielles ou totales.
- Déformation de l'objet.
- Objet en mouvement rapide.

- Présence du bruit dans les images.
- Mouvement de caméra.
- Objet de petite taille.

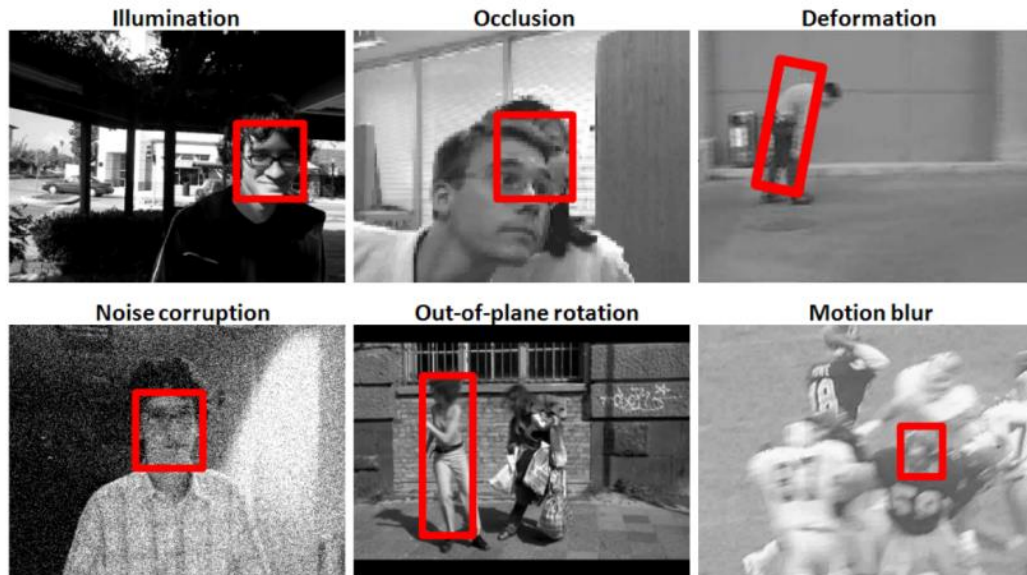


Figure1.4: Les défis de suivi [17]

3 Modélisation de l'apparence de l'objet :

L'apparence de l'objet est la principale information exploitée par les trackers, sa modélisation est donc une étape importante pour réussir à suivre l'objet correctement. Il existe une diversité de représentations d'un objet, chacune étant adaptée à un type de difficulté rencontrée en suivi (nature de l'objet, nature de la scène, qualité d'image, etc...)[18].

- **Templat:**

Cette représentation considère directement le signal lumineux dans les images. Dans le cadre du suivi d'objet, les méthodes basées sur un template effectuent directement la mise en correspondance 2D sur une partie de l'image sans passer par une phase d'extraction de caractéristiques. La recherche des paramètres de la transformation se fait généralement en optimisant un critère de corrélation [19].

- **Densité de probabilité d'apparence :**

L'apparence d'un objet peut être modélisée à l'aide de la répartition des valeurs des couleurs qu'il contient (ou de toute autre caractéristique) sous forme de densité de probabilité. La fonction de densité de probabilité peut être estimée par un estimateur à noyau (Kernel Density Estimator, KDE) [20], représentée sous la forme d'un histogramme, ou encore sous une forme paramétrique à l'aide d'une gaussienne ou d'un mélange de gaussiennes [21].

- **Points d'intérêts:**

Une des techniques répandues est le suivi de points d'intérêt extraits de l'objet : points de Harris [22], SIFT [23], SURF [24], etc. Cette représentation par points hérite des techniques de recalage de la même façon que l'approche par *template*. Les points d'intérêt sont invariants à certaines transformations géométriques ou radiométriques, ce qui rend leur suivi plus robuste.

4 Les caractéristiques visuelles :

Le terme "caractéristiques d'objet" englobe toutes les données utilisées pour caractériser et discriminer l'objet d'intérêt du reste de l'image, y compris d'autres objets présents dans la séquence vidéo. Les caractéristiques idéales pour le suivi d'objet est un invariant de l'objet, ou du moins robuste à tout type de transformation, tout changement d'illumination, toute dégradation de la qualité d'image (vidéo). [25]

- **Couleur:**

La couleur est sans doute la primitive la plus utilisée pour décrire un objet. Elle est directement accessible depuis le signal lumineux (image) et fournit une description intuitive de l'apparence d'un objet. La couleur prédominante d'un objet peut être utilisée directement en recherchant dans l'image les pixels de mêmes valeurs ou, de façon plus générale, rechercher l'objet dans l'image à l'aide de sa distribution couleur. L'information de couleur d'une image peut être représentée de différentes manières suivant l'espace colorimétrique utilisé. Il existe de très nombreux espaces de couleurs possédant chacune leurs caractéristiques propres mais partageant le même principe de base. Chaque espace crée une décomposition, selon ses propres paramètres, de toutes les couleurs en un ensemble de composantes dites primaires. [11]. La couleur est un premier descripteur visuel simple capable de distinguer plusieurs objets dans une scène [26]

- **Texture:**

A l'instar de la couleur, la texture est une composante clé de la perception humaine, ce qui en fait une caractéristique essentielle à considérer dans les algorithmes de détection ou de reconnaissance. Bien que chacun soit capable de reconnaître une texture, il est très difficile d'en donner une définition [11]. Comparé à la couleur, la texture nécessite une étape de traitement pour générer les descripteurs [1]. Les caractéristiques de texture sont moins sensibles aux changements d'éclairage par rapport à la couleur.

Les méthodes de modélisation de texture peuvent être classées dans trois catégories : Les modèles statistiques mesurent la distribution spatiale des valeurs des pixels (histogrammes) Les méthodes structurelles, la texture est représenté par une répétition d'éléments structurels et la texture est modélisée comme étant un arrangement spatial de ces éléments les modèles fondés sur les filtres, ils consistent à appliquer un ensemble de filtres à l'image afin d'en étudier la réponse. On retrouve les filtres dans le domaine spatial (filtres de Sobel, de Canny, Robert,), ceux dans le domaine fréquentiel (filtre de Fourier

- **Gradient :**

Cette information est extraite à partir de l'analyse spatiale de l'intensité lumineuse de l'image. Une propriété importante du gradient est sa sensibilité plus faible aux changements de luminosité comparée aux caractéristiques couleurs. Les contours issus du gradient sont exploités dans de nombreuses approches de suivi d'objet. Le gradient permet de définir des points caractéristiques dans les objets. L'algorithme CONDENSATION [27] (Conditionnal density propagation). Les descripteurs SIFT [23] (Scale Invariant Feature Transform) Les histogrammes d'orientations de gradient (HOG) [28].

- **Deep features :**

Récemment, l'utilisation des caractéristiques calculées par des réseaux profonds (en anglais Deep Features) est devenue très populaire. A cause de leur grande capacité de représentation des objets, les réseaux profonds issus de l'apprentissage profond (en anglais Deep learning) sont utilisés comme fournisseurs de caractéristiques pour le suivi d'objet. Le réseau est tout d'abord pré-entraîné sur une grande base d'images, par exemple ImageNet (plusieurs millions d'images), puis utilisé en tant que boîte noire pour fournir les cartes de caractéristiques des images passées en entrée du réseau. Ces caractéristiques alimentent un classifieur SVM (en anglais Support Vector Machine) en ligne (dernière couche du réseau remplacée par un SVM) dont l'objectif est d'apprendre les caractéristiques spécifiques à la cible permettant de discriminer la cible du fond. Cependant, l'information spatiale de

localisation de la cible n'est pas conservée par les deep features à cause des opérations de pooling (groupements de neurones) qui ont lieu dans différentes couches du réseau. Par conséquent, pour avoir accès à la localisation spatiale de la cible, l'idée est de générer des cartes de saillance spécifiques à la cible en rétro-propageant les caractéristiques spécifiques à la cible classées par SVM dans le réseau jusqu'à la première couche. [18]

5 Localisation de l'objet

Différentes approches ont été utilisées pour estimer la position ou l'état de l'objet dans chaque image, et l'estimation de la meilleure localisation de l'objet soulève le problème du coût de calcul nécessaire pour le réaliser.

5.1 Méthodes d'optimisation

Le problème de suivi est formulé sous la forme d'une optimisation de fonction différentiable par rapport à la variable position et où l'estimation de position de l'objet est résolue par une descente de gradient, comme c'est le cas de l'algorithme Meanshift ou du flot optique [18]. Le Meanshift [29] est une procédure itérative de recherche du maximum d'une densité de probabilité où la position estimée de l'objet est déplacée vers un centre de gravité local jusqu'à convergence. Le flot optique [30] estime le déplacement 2D d'un pixel d'une image à l'autre en imposant une contrainte de conservation d'intensité de celui-ci. Cette estimation peut être rendue plus robuste éventuellement par des contraintes supplémentaires d'intensité ou géométriques de déplacement [31] [32].

- ***Filter Kalman:***

Une autre stratégie de recherche est d'utiliser un modèle de mouvement pour restreindre l'espace de recherche de l'objet, notamment un filtre de Kalman [33]. Ce filtre a été proposé par Kalman au début des années 1960. Il est basé sur un principe récursif. En fait, la position courante de l'objet cible est calculée à l'aide de l'estimation de la dernière position de l'objet. Le filtre de Kalman se fait sur deux étapes : prédiction de l'état courant et mise à jour. La phase de prédiction permet de prédire la position courante de l'objet cible en se basant seulement sur son état précédent. Tandis que l'étape de mise à jour consiste à corriger l'estimation de l'état à l'instant courant afin d'obtenir plus de précision. [34]

- **Le filtrage particulaire :**

Le filtrage particulaire (Ou méthodes de Monte-Carlo séquentielles) [35], [36] est représenté la solution de base de la majorité des méthodes d'estimation dites Monte Carlo, il est une solution aux limitations du filtre de Kalman [18]. La méthode du Filtre de Particules consiste à estimer la localisation de l'objet cible en fonction de particules pondérées. Initialement, des particules sont générées aléatoirement autour de la dernière position de l'objet cible. Ces particules doivent être comparées (en utilisant un modèle d'apparence comme l'histogramme de couleur) avec le modèle de l'objet cible. Ainsi, la particule la plus similaire peut être choisie comme le nouvel état de l'objet suivi. [34] Le filtrage particulaire est largement utilisé en suivi mono-objet et multi-objets. Son coût de calcul dépend du nombre de particules utilisées. Il est adapté aux trajectoires complexes des objets et aux occultations, par exemple deux piétons se croisant. [18].

6 Etat de l'art du suivi d'objet

Il existe de nombreux des méthodes dans la littérature traitant du sujet de suivi d'objets. Plusieurs revues de l'état de l'art ont été publiées et rendent compte des approches de suivi proposées [18]. Yilmaz et al [1] proposent une classification des méthodes de suivi en fonction de la représentation d'objet utilisée (correspondance de points, modèles géométriques et modèles d'évolution de contours). Yang et al [14] réalisent un état de l'art sur les différents composants d'un tracker : descripteurs de caractéristiques, méthodes d'apprentissage en ligne, prise en compte de l'information de contexte et méthodes d'échantillonnage de Monte-Carlo. Wu *et al* [37] évaluent quantitativement 29 trackers de l'état de l'art sur une base de 50 vidéos. Li et al [17] fournissent une description très détaillée de tous les modèles d'apparence existant en suivi et discutent de leur composition (représentation visuelle et modélisation statistique de l'apparence). Dans cette section, nous prenons la classification des algorithmes de suivi qui a présenté dans la figure suivante.

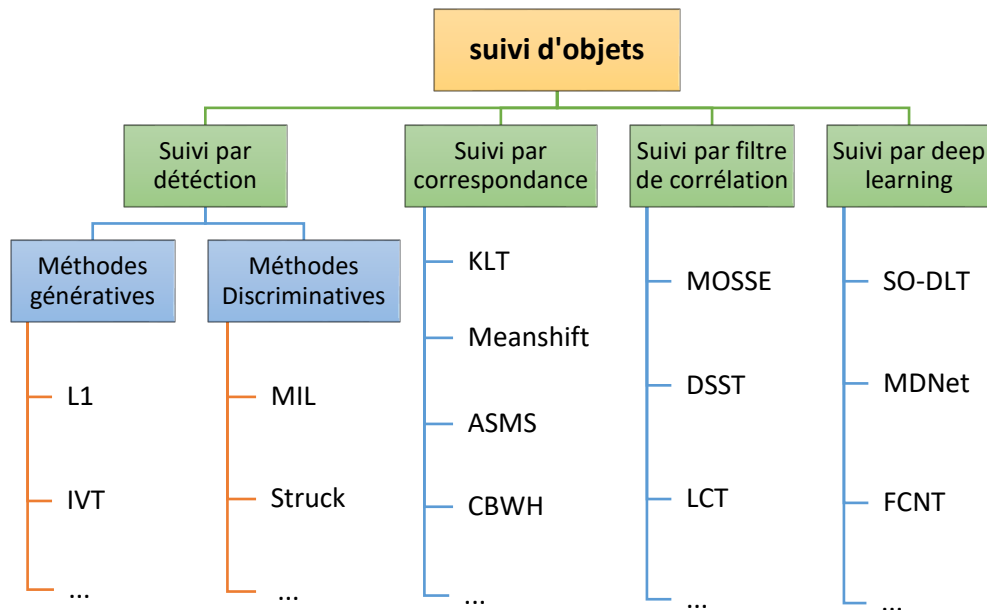


Figure 1.5 : Taxonomie des méthodes de suivi d'objets

6.1 Suivi par détection :

6.1.1 Méthode générative :

Dans la méthode générative, l'objet cible est représenté dans un espace de caractéristiques puis une recherche est effectuée afin de trouver le meilleur score de correspondance avec les objets candidats. En général, cette catégorie de méthode ne nécessite pas un grand ensemble de données pour l'apprentissage [34].

- **Tracker L1 :**

Est une méthode de suivi robuste proposée par Mei et Ling [38]. Cette méthode traite le suivi d'objets comme un problème d'approximation parcimonieuse et d'introduire le modèle trivial pour approcher le bruit et l'occultation. Pendant le suivi, les candidats cibles sont représentés sous la forme d'une combinaison linéaire parcimonieuse d'ensemble des modèles incluant des modèles de cible qui sont obtenus à partir de trames précédentes et des modèles triviaux. Le tracker L1 nécessite des ressources de calcul élevé en raison de nombreux calculs de minimisation L1.

- **Tracker IVT :**

Est une méthode de suivi proposée par Ross et al [39], utilise un modèle sous-espace incrémental pour décrire l'objet cible afin d'adapter les changements d'apparence. Elle réalise l'apprentissage incrémental d'un sous-espace de représentation de l'objet (PCA) et adapte le modèle en intégrant la nouvelle apparence de l'objet avec un facteur d'oubli sur les apparences passées de l'objet. Cette méthode est peu robuste en particulier lorsque la localisation de l'objet est imprécise.

6.1.2 Méthode Discriminative :

Le suivi sera traité comme un problème de classification binaire (classe de l'objet cible et classe pour les autres objets). En général, cette catégorie de méthode nécessite un grand ensemble de données pour l'apprentissage. [34].

- **SVM :**

Avidan [40] a proposé une approche discriminative, où le classifieur SVM (Support Vector Machines) et un tracker basé sur le flot optique sont combinés pour suivre des véhicules sur de longues séquences vidéo. Le suivi par cet approche est effectué en maximisant le score de classification du SVM sur des régions de l'image afin d'estimer la position de l'objet.

- **MIL :**

Est une méthode de suivi basée sur l'apprentissage en ligne à instance multiple qui permet au classificateur de sélectionner un nombre d'échantillons positifs potentiels en fonction de son état actuel. Le tracker MIL traite les échantillons d'entraînement comme des sacs. Un sac est considéré comme positif s'il contient au moins une instance positive, sinon le sac est mis à négatif. Ce tracker conserve suffisamment d'échantillons d'entraînement et tolère le bruit d'étiquetage lors de la mise à jour de son modèle. [14]

- **Struck :**

[Hare et al. [41] ont proposé l'algorithme Struck pour le suivi adaptatif d'objets qui se base sur la prédiction de sortie structurée. Cet algorithme est basé sur l'utilisation d'un noyau sur la sortie structuré de machine à vecteurs de support (SVM), qui est apprise en ligne pour fournir un suivi adaptatif. Struck intègre les étapes de la localisation cible et la mise à jour du modèle dans un cadre d'apprentissage de sortie structuré unifié. Parce que, la séparation entre les deux étapes introduite des erreurs supplémentaires de l'étape d'étiquetage à l'étape de mise à jour du modèle, car l'échantillon choisi par le classificateur peut ne pas correspondre à la meilleure estimation de l'emplacement de l'objet.

6.2 Suivi par correspondance :

Correspondant à la représentation d'un objet cible entre deux images consécutives est une façon naturelle d'estimer son mouvement et de le suivre au fil du temps. Il s'agissait d'une approche de suivi dominante à ses débuts, en raison de ses performances relativement bonnes, de sa structure simple et de ses faibles besoins en calcul [42].

- **KLТ :**

[Lucas and Kanade, [43] ont proposé le tracker KLT qu'est la méthode la plus efficace basée sur la mise en correspondance de Template. Le tracker KLT trouve des correspondances affines transformées entre deux trames successives au moyen de dérivés spatio-temporels. La nouvelle localisation de la cible est déterminée en mettant en correspondance sa position dans la trame précédente à la localisation dans la trame actuelle en utilisant la transformation affine estimée [42].

- **Meanshift :**

La méthode de mean shift est un estimateur non paramétrique du gradient de densité basé sur l'utilisation noyau. Elle a été proposée en 1975 par Fukunaga et al, et a été utilisée pour la première fois en 1997 dans le cadre de la segmentation d'images. Elle a ensuite été adoptée par Comaniciu [2], pour le suivi en temps réel d'objets déformables dans une séquence d'image. Cette méthode utilise un histogramme pondéré calculé sur une région géométrique (rectangle, ellipse) pour représenter l'objet. Le système de suivi Mean Shift maximise la similarité d'apparence itérativement en comparant les histogrammes de l'objet modèle et une fenêtre autour de la position hypothèse d'objet candidat.

- **ASMS:**

Vojir et al [44] ont proposé l'algorithme ASMS, version améliorée de MS. Il dispose d'un modèle statique construit sur l'apparence initiale de l'objet. ASMS propose deux améliorations par rapport au tracker Meanshift classique :

1. prendre en compte l'apparence du fond dans la modélisation de l'apparence de la cible pour aider à mieux la discriminer du fond.
2. Estimer l'échelle de la cible par une vérification aller-retour.

- **CBWH:**

Ning et al [45] ont proposé un histogramme pondéré d'arrière-plan correct CBWH afin de réduire réellement les interférences des caractéristiques d'arrière-plan saillantes. CBWH [45] et BWH [2] sont les méthodes améliorées basées sur la méthode mean shift. L'algorithme CBWH est une correction de BWH en transformant uniquement le modèle cible mais pas le modèle candidat cible. CBWH peut effectivement réduire l'interférence de l'arrière-plan dans la localisation de la cible et peut conduire à une convergence plus rapide et à une localisation plus précise que la représentation cible habituelle dans le suivi Mean Shift. Même si la cible n'est pas bien initialisée, l'algorithme proposé peut toujours suivre l'objet de manière robuste, ce qui est difficile à atteindre avec la représentation de cible conventionnelle. Le tracker CBWH est détaillé dans le chapitre 2.

6.3 Suivi par filtre de corrélation :

Récemment, les méthodes de suivi discriminatives basées sur le filtre de corrélation [46] [47] ont été prouvées pour être capable d'atteindre une vitesse assez élevée et des performances de suivi robustes. Classiquement, les filtres de corrélation sont conçus pour produire des pics de corrélation pour chaque cible intéressée dans la scène tout en donnant de faibles réponses à l'arrière-plan, qui sont généralement utilisés pour détecter les modèles attendus [42].

MOSSE :

Bolme et al. [48] ont proposé le filtre de la somme de sortie minimale de l'erreur au carré (MOSSE) pour le suivi visuel sur des images en niveaux de gris. Le suivi basé sur le filtre MOSSE est efficace du point de vue du calcul avec une vitesse atteignant plusieurs centaines de trames par seconde, et robuste aux variations d'illumination, d'échelle, de pose et de déformations non rigides. [42].

- **DSST :**

Danelljan et al. [49] ont proposé le tracker DSST (Discriminative Scale Space Tracker) qui modélise la cible par un filtre de corrélation discriminant espace-échelle basé sur des caractéristiques PCA-HOG extraites de la cible. La particularité du DSST par rapport à d'autres filtres de corrélation est la capacité à s'adapter aux variations de taille de la cible par une estimation précise de l'échelle [18].

- **LCT :**

Ma et al [50] ont proposé LCT en étendant le tracker populaire basé sur le filtre de corrélation au problème de suivi à long terme. LCT décompose la tâche de suivi à long terme dans l'estimation des paramètres de conversion et d'échelle des objets cibles, conjointement avec un composant re-détection est ajouté dans le système de suivi.

6.4 Suivi par deep learning :

Au cours des dernières années les approches fondées sur le Deep Learning, notamment le réseau de neurones convolutionnels (CNN), ont obtenu un grand succès empirique et ont dominé de nombreux problèmes de vision par ordinateur, tels que la classification d'objets Krizhevsky et al. [51], la détection Girshick et al. [52] et la vérification des visages Taigman et al. [53]. L'application du deep learning au problème de suivi d'objet visuel nécessite un effort supplémentaire, en raison du manque de données de formation appropriées, et de l'objet cible évoluant au cours de la séquence vidéo.

- **SO-DLT :**

Wang et al, [54] ont présenté le tracker SO-DLT (structured output deep learning tracking), permettant de pré-former et d'ajuster CNN en ligne, avec une sortie structurée, afin de distinguer la cible et l'arrière-plan. Tout d'abord, un CNN de sortie structuré hors ligne est formé à l'aide de données ImageNet afin d'apprendre les caractéristiques d'objet générique pour la classification des régions d'objet et non d'objet. Ensuite, deux CNN adoptés à partir du réseau CNN hors ligne sont configurés en ligne pour le suivi. Le résultat de suivi final est déterminé par CNN avec la plus grande confiance. [42].

- **MDNet :**

Nam et Han, [55] proposent le réseau MDNet (*Multi-Domain*) basé sur l'apprentissage de représentations multi-domaines pour le suivi d'objet (un domaine correspond à une séquence vidéo). Le réseau est pré-entraîné sur une large base de vidéos de suivi, les dernières couches (*fully-connected*) et des couches spécifiques au domaine sont entraînées en ligne afin de s'adapter au nouveau domaine. [18].

- **FCNT :**

[Wang et al [56] a utilisé les représentations apprises dans les couches convolutionnelles intermédiaires, au lieu d'utiliser les réponses de la dernière couche CNN en tant qu'entité de type boîte noire dans le tracker FCNT (fully convolutional network based tracker), comprenant deux couches de convolution de niveaux différents. Ce tracker est conçu pour un suivi robuste. [42].

7 Conclusion:

Le suivi d'objets en mouvement est parmi les problèmes les plus étudiés ces dernières années et il est à la base de tout système de vision. Dans ce chapitre, nous avons vu le principe de suivi d'objet, la représentation de l'objet en fonction de leurs formes et de leurs apparences, les caractéristiques visuelles et les différentes méthodes de la localisation d'objets. Ensuite nous avons présenté l'état de l'art des méthodes de suivi d'objets et la différence entre ces méthodes réside en partie dans le choix de la représentation et de la forme de l'objet et des caractéristiques de l'image utilisées. nous avons présenté plusieurs catégories de ces méthodes : suivi par détection qui se divise en méthode générative et en méthode discriminative, suivi par correspondance, suivi par filtre de corrélation et suivi par deep learning qu'a obtenu un grand succès empirique. La difficulté du suivi d'objets dépend de plusieurs facteurs relatifs aux données ou à l'application : Mouvements complexes, Occultations partielles ou totales, Nécessité d'un suivi en temps réel.

CHAPITRE 2

**Suivi d'objet robuste avec
l'algorithme mean shift via CBWH**

1 Introduction :

Le suivi des objets est une tâche importante dans la vision par ordinateur. De nombreux algorithmes ont été proposés pour résoudre les divers problèmes posés par les bruits, les encombrements et les occultations dans le modèle d'apparence de la cible à suivre. Mean shift est un algorithme de suivi d'objet robuste et efficace pour suivre des objets non rigides et en temps réel, en raison de sa simplicité et de son efficacité. Bien qu'il soit robuste à l'occultation partielle, la rotation, le mouvement de fond et la déformation non-rigide de la cible, il est très sensible aux changements d'échelles, aux occultations importantes, et il peut échouer en présence d'un autre objet de couleurs similaires, ou de fond de couleurs similaires ou dans le cas de grands déplacements, parce qu'il est basé sur l'histogramme de couleur pour représenter le modèle de l'objet cible.

Dans ce chapitre, nous allons voir l'état de l'art sur l'algorithme mean shift. Puis l'étude de suivi d'objet via le tracker mean shift en utilisant la méthode CBWH pour réduire les interférences de fond dans la localisation de la cible où mean shift peut fonctionner de manière plus robuste, même si le modèle cible contient beaucoup d'informations de fond.

2 Etat de l'art sur l'algorithme Mean shift :

L'algorithme mean shift était présenté en 1975 par Fukunaga et Hostetler [57]. Mean Shift est une procédure non paramétrique qui recherche la probabilité de densité maximale par convergence récursive. Mean Shift a été étendu à l'application de la vision par ordinateur et de la technique de classification [58] Cheng [59] a utilisé le premier algorithme mean shift en 1995 comme solution au problème de recherche de mode [60]. Il s'agit d'un algorithme de clustering des données par l'estimation de leur densité. Dans une fenêtre de recherche de taille constante on trouve le centroïde (position moyenne des données) et la fenêtre qui est centrée sur le centroïde. Cette procédure est répétée jusqu'à la convergence, et le point de convergence est considéré comme le centre de cluster pour les données visitées (voir la figure 2.1). Plus tard, Comaniciu [2] [61] a proposé le suivi d'objets basé sur le noyau en utilisant un histogramme de couleur pondéré. Cependant, il est à noter que l'algorithme mean shift présente certains inconvénients. Premièrement, l'algorithme ne parvient pas à suivre les informations spatiales de couleur. Deuxièmement, les mesures de similarité ne sont pas très discriminantes. Troisièmement, mean shift est enclin à choisir des minima locaux plutôt que des minima globaux, ce qui entraîne des erreurs de classification [62]. L'algorithme Cam Shift (Continuously Adaptive Mean Shift) a été introduit comme solution en ce sens qu'il

peut gérer des distributions de probabilité dynamiques en ajustant la taille de la zone de cible [63]. Cependant, le suivi de mean shift est utilisé afin de simplifier le calcul et sa faiblesse est compensée par le filtrage de Kalman [64] ou par le filtre à particules pour gérer le problème d'occultation [65] [66] Ning et al. [67] ont proposé un algorithme de suivi adapté à l'échelle et à l'orientation (SOAMST) afin de résoudre le problème de l'estimation des changements d'échelle et d'orientation de l'objet dans le cadre de suivi par Mean shift. Ning et al. [67] ont appliqué un histogramme conjoint de couleur-texture pour représenter le modèle de la cible, et les informations de texture sont extraites par un motif binaire local LBP. Ning et al. [68] ont proposé l'algorithme CBWH (Corrected Background-Weighted Histogram) qui utilise le mécanisme de mise à jour de fond, dans le cadre de l'algorithme Mean shift. Cet algorithme peut efficacement réduire les interférences de fond dans la localisation de l'objet cible.

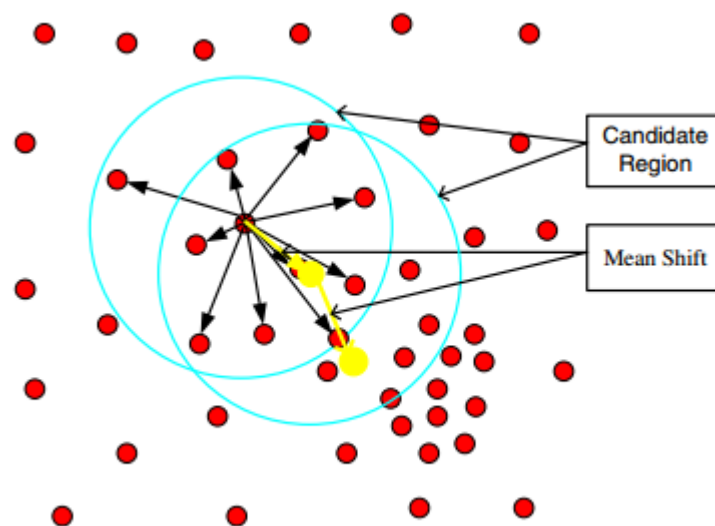


Figure 2.1 : Description intuitive de la convergence de la procédure Mean shift.

3 Procédure Mean Shift :

L'algorithme Mean Shift est un estimateur de gradient de densité non paramétrique, basé sur une procédure itérative de recherche de maximum local qui est basée sur une montée de gradient (figure 2.2).

L'approche Mean Shift [69] utilise un histogramme de couleur pondéré calculé sur une région rectangulaire pour représenter l'objet. Le système de suivi Mean Shift maximise la similarité d'apparence itérativement en comparant les histogrammes de l'objet modèle, et une

fenêtre autour de la position hypothèse d'objet candidat. La similarité entre deux histogrammes est définie en termes de coefficient de Bhattacharyya. À chaque itération, le vecteur Mean Shift est calculé tel que la similarité entre les histogrammes est augmentée. Ce processus est répété jusqu'à ce que la convergence soit réalisée, qu cinq à six itérations [50].

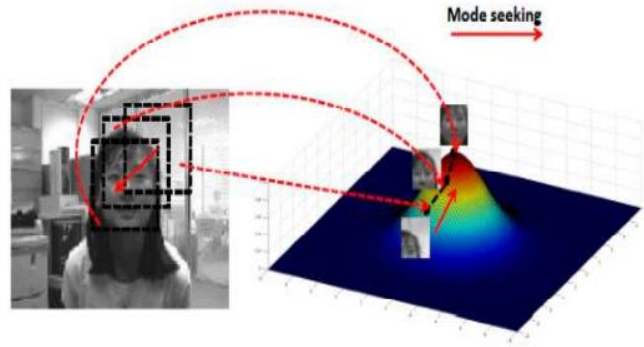


Figure 2.2 : l'itération du Processus Mean Shift

Mean shift représente le modèle de la cible par une région rectangulaire de taille (h_x, h_y) . On note $\{x_i^*\}_{i=1 \dots n}$ l'ensemble des coordonnées des n pixels du modèle cible, centré à 0. Et normalisé les demi-rayons du rectangle h_x et h_y . La loi de probabilité des couleurs $u \in \{1 \dots m\}$ est calculée en utilisant une fonction (noyau isotrope) de profil convexe $k(x)$, $k : [0; \infty[\rightarrow \mathbb{R}$, monotone et décroissante, attribuant un poids plus faible aux coordonnées éloignées du centre de la cible modèle. Le rayon de cette fonction de profil est égal à 1. La fonction $b : \mathbb{R}^2 \rightarrow \{1 \dots m\}$ qui associe à chaque pixel de coordonnées x_i^* l'indice de sa couleur $b(x_i^*)$ dans le m -histogramme.

Le rayon de cette fonction de profil est égal à 1. L'histogramme de la cible modèle pour un bins u est donné par :

$$\hat{q}_u = C \sum_{i=1}^n k(\|x_i^*\|^2) \delta[b(x_i^*) - u] \quad (2.1)$$

Où δ est la fonction de Kronecker et C constante de normalisation tel que : $\sum_{u=1}^m \hat{q}_u = 1$

Le constante C est donné par :

$$C = \frac{1}{\sum_{i=1}^m k(\|x_i^*\|^2)}$$

Dans l'image courante, l'objet candidat est à l'intérieur de la même rectangle et centré sur position y . En utilisant le même profil du noyau $k(x)$, la loi de probabilité de la cible candidats pour un bins u est donnée par :

$$\hat{p}_u(y) = C \sum_{i=1}^m k(\|y - x_i\|^2) \delta[b(x_i) - u] \quad (2.2)$$

Où C est un facteur de normalisation. Telle que : $\sum_{u=1}^m \hat{p}_u = 1$

La fonction de la similarité définit la similitude entre la cible modèle et les candidats. Le coefficient de Bhattacharyya est utilisé comme mesure de similarité entre deux histogrammes \hat{q} et $\hat{p}(y)$ Le coefficient de Bhattacharyya est défini par :

$$\rho(\hat{p}(y), \hat{q}) = \sum_{u=1}^m \sqrt{\hat{p}_u(y) \hat{q}_u} \quad (2.3)$$

La distance de Bhattacharyya est définie par:

$$d(y) = \sum_{u=1}^m \sqrt{1 - \rho[\hat{p}_u(y), \hat{q}_u]} \quad (2.4)$$

La localisation y la plus probable de l'objet dans l'image courante est obtenue en minimisant la distance ce qui est équivalent à maximiser le coefficient de Bhattacharyya $\rho(y)$ Pour l'image courante, il utilise le centre du rectangle (y_0) à l'image précédente comme initialisation. Il calcule la probabilité de la cible candidat centré sur $\hat{p}(y_0)$ dans l'image courante: $\{\hat{p}_u(y_0)\}_{u=1\dots m}$ et en utilisant l'approximation linéaire de Taylor du $\rho(y)$.

$$\rho(\hat{p}_u(y), \hat{q}_u) \approx \frac{1}{2} \sum_{u=1}^m \sqrt{\hat{p}_u(y) \hat{q}_u} + \frac{C}{2} \sum_{u=1}^{n_h} w_i k\left(\left\|\frac{y - x_i}{h}\right\|^2\right) \quad (2.5)$$

Où

$$w_i = \sum_{u=1}^m \sqrt{\frac{\hat{q}_u}{\hat{p}_u(y)}} \delta[b(x_i) - u] \quad (2.6)$$

Le premier terme de (2.5) est indépendante de y , le deuxième terme de (2.5) doit être maximisée. Le déplacement de vecteur Mean Shift de y_0 à y_1 peuvent être calculées de manière itérative par le second terme, et la nouvelle position de la cible candidat est calculée comme suit:

$$y_1 = \frac{\sum_{i=1}^n x_i w_i g(\|y_0 - x_i\|^2)}{\sum_{i=1}^n w_i g(\|y_0 - x_i\|^2)} \quad (2.7)$$

Où $g(x) = -k'(x) = 1$ et $k(x)$ est un noyau Epanechnikov.

Les itérations Mean Shift peuvent être utilisées pour maximiser le coefficient de Bhattacharyya comme fonction de y dans le voisinage d'une position donnée (figure 2.3).

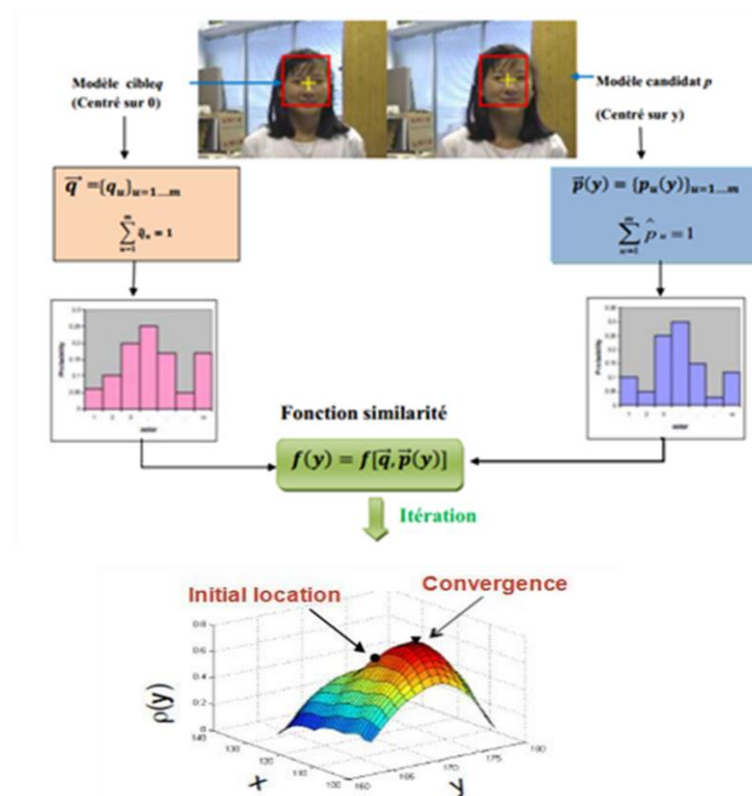


Figure 2.3 : La similarité de méthode Mean shift

3.1 Histogramme pondéré :

La région d'intérêt peut dans certains cas englober l'objet d'intérêt et des éléments du fond. Pour limiter l'influence du fond et privilégier l'information pertinente, un histogramme pondéré peut être utilisé pour calculer l'histogramme cible. La pondération de l'histogramme donne plus d'importance aux pixels proches du centre et inversement, c.-à-d. plus les pixels sont loin du centre de l'objet, plus le poids pris en compte dans l'histogramme final est faible[1].

L'exemple suivant (figure 2.4) montre comment calcule l'histogramme pondéré par un Noyau gaussienne d'une image en niveau de gris, avec la quantification de l'intensité en 4 niveaux.

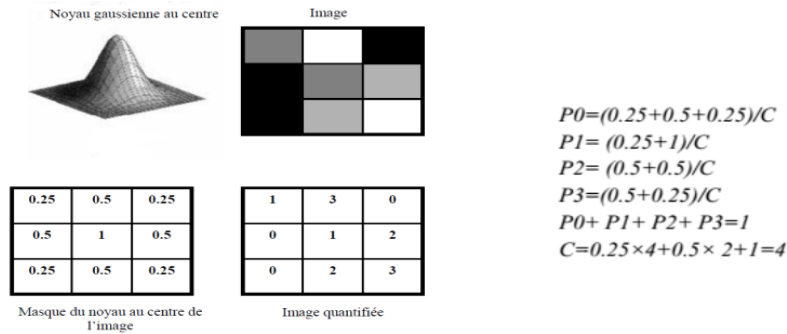


Figure 2.4 : Exemple de construction d'un histogramme pondéré [6]

La maximisation peut être effectuée efficacement en utilisant les itérations Mean Shift, à l'aide de l'algorithme suivant [2] [8] [70]

Algorithme de suivi par mean shift :

1. Calculer la distribution $p(y_0) = \langle p_u(y_0) \rangle_{u=1 \dots m}$
2. Calculer le coefficient de Battasharyya : $\sum_{u=1}^m \sqrt{\hat{p}_u(y_1) \hat{q}_u}$
3. Calculer les poids $\{ \omega_i \}_{i=0 \dots n}$ $w_i = \sum_{u=1}^m \sqrt{\frac{\hat{q}_u}{\hat{p}_u(y)}} \delta[b(x_i) - u]$
4. A partir du vecteur Mean shift calculer la nouvelle position de l'objet :

$$y_1 = \frac{\sum_{i=1}^n x_i w_i g(\|y_0 - x_i\|^2)}{\sum_{i=1}^n w_i g(\|y_0 - x_i\|^2)}$$

5. Mettre à jour $p(y_1) = \{p_u(y_1)\}_{u=1 \dots n}$ puis évaluer $\rho(p(y_1), q) = \sum_{u=1}^m \sqrt{\hat{p}_u(y_1) \hat{q}_u}$
6. Tant que $\rho(p(y_1), q) < \rho(p(y_0), q)$ faire $y \leftarrow \frac{y_0 + y_1}{2}$
7. Si $|y_1 - y_0| < \epsilon$ Stop. Sinon $y_0 \leftarrow y_1$ et retourner à l'étape 1.

4. Suivi par Mean shift via BWH :

Dans le suivi d'objets, les informations de fond sont souvent incluses dans la région cible détectée. Si la corrélation entre la cible et le fond est élevée (lorsque certaines des caractéristiques cibles sont présentes au fond), la précision de localisation de l'objet sera diminuée. Comaniciu et al [2], ont proposé l'histogramme pondéré de fond BWH (*Background-Weighted Histogram*), pour réduire les interférences de fond dans la représentation de la cible. La stratégie de BWH consiste à dériver une représentation simple des caractéristiques de fond et à l'utiliser pour sélectionner les caractéristiques de fond saillantes du modèle cible et du modèle candidat cible (voir la figure 2.5)

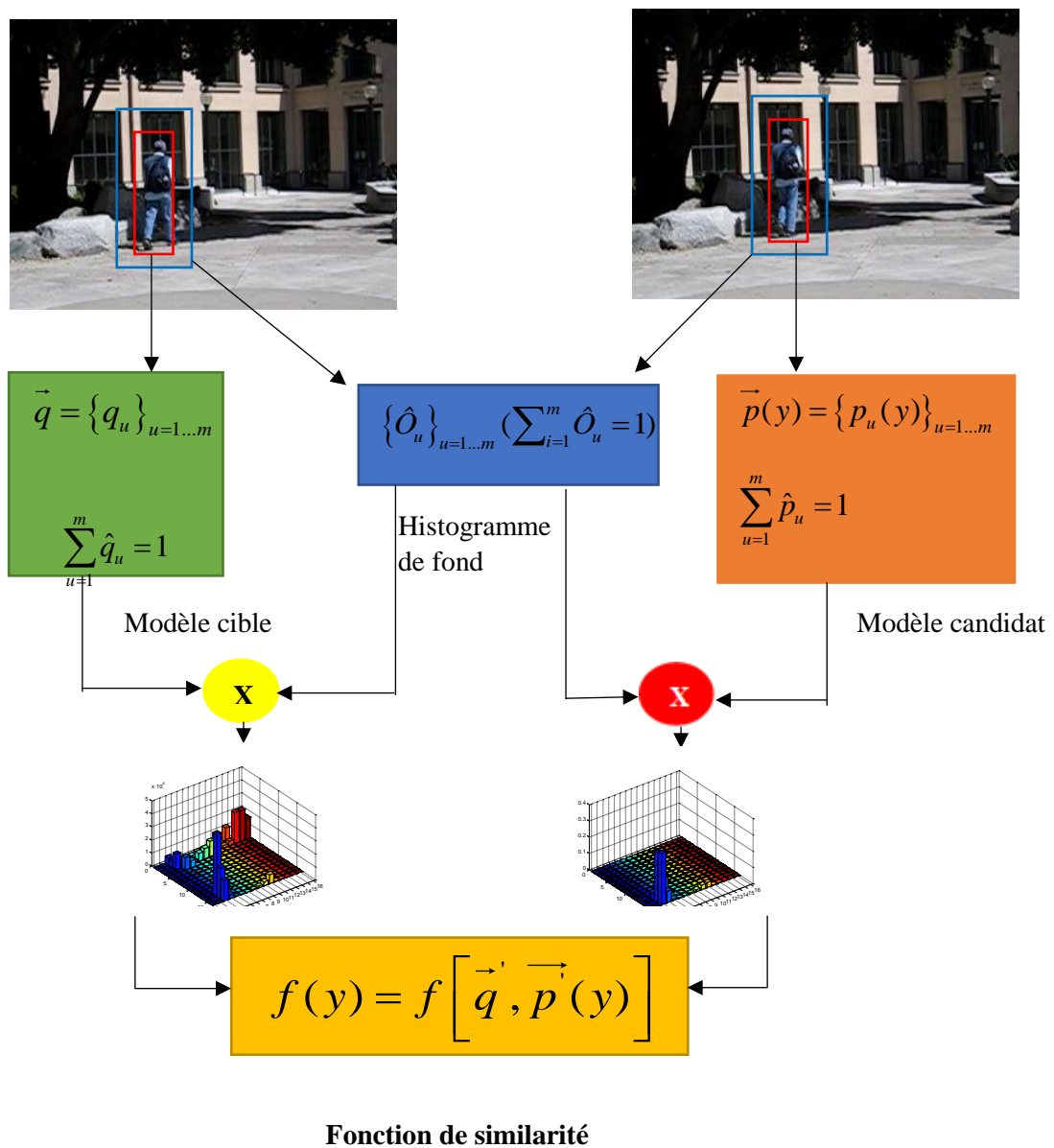


Figure 2.5 : la similarité de la méthode BWH.

4.1 Algorithme de l'histogramme pondéré de fond (BWH) :

Dans [2], le fond est représenté par : $\{\hat{O}_u\}_{u=1\dots m}$ (with $\sum_{i=1}^m \hat{O}_u = 1$) et il est calculé par la zone autour de la cible. La région de fond est trois fois la taille de la cible comme suggéré dans [2]. Noter par la valeur \hat{O}_u^* minimale non nulle dans $\{\hat{O}_u\}_{u=1\dots m}$. Les coefficients :

$$\{v_u = \min(\hat{O}^*/\hat{O}_u, 1)\}_{u=1\dots m} \quad (2.8)$$

Sont utilisés pour définir une transformation entre les représentations du modèle cible et du modèle candidat cible. La transformation réduit le poids de ces caractéristiques avec un faible v_u , c'est-à-dire les caractéristiques saillantes au fond. Ensuite, la nouvelle cible modèle est :

$$\hat{q}'_u = C' V_u \sum_{i=1}^n k(\|x_i^*\|^2) \delta[b(x_i^*) - u] \quad (2.9)$$

Avec la constante de normalisation :

$$C' = \frac{1}{\sum_{i=1}^n k(\|x_i^*\|^2) \sum_{u=1}^m V_u \delta[b(x_i^*) - u]}$$

La nouvelle cible candidat est :

$$p'_u(y) = C'_h V_u \sum_{i=1}^{n_h} k\left(\frac{\|y - x_i\|^2}{h}\right) \delta[b(x_i) - u] \quad (2.10)$$

Où

$$C'_h = \frac{1}{\sum_{i=1}^{n_h} k\left(\left\|\frac{(y - x_i)}{h}\right\|^2\right) \sum_{u=1}^m V_u \delta[b(x_i) - u]}$$

La transformation BWH ci-dessus vise à réduire les effets des caractéristiques d'arrière-plan importantes dans la région candidate cible sur la localisation cible. Cependant, BWH ne peut pas atteindre cet objectif car il est équivalent à la représentation cible habituelle dans l'algorithme mean shift. C'est ce que nous allons prouver ça dans la section suivante.

4.2 L'équivalence de la représentation de BWH à la représentation habituelle :

Par la formule d'itération de mean shift (2.7), dans la région candidate cible, les poids de points (équation (2.6)) déterminent la convergence de l'algorithme de suivi. Ce n'est que lorsque le poids des caractéristiques saillants du fond est réduit que la pertinence des informations de fond pour la localisation de la cible peuvent être réduite. On analyse les changements de poids à l'aide de la transformation BWH. On indique w'_i le poids du point x_i calculé par le BWH dans la région candidate cible.

Il peut être dérivé par Eq. (2.8) que :

$$w'_i = \sum_{u=1}^m \sqrt{\frac{\hat{q}'_u}{\hat{p}'_u(y)}} \delta[b(x_i) - u] \quad (2.11)$$

Soit u_i l'indice de bin dans l'espace des caractéristiques qui correspond au point x_i dans la région candidate On a :

$$\delta[b(x_i) - u_i]$$

Alors l'équation (2.11) peut être simplifié comme :

$$w'_i = \sqrt{\frac{\hat{q}'_{u_i}}{\hat{p}'_{u_i}(y)}} \quad (2.12)$$

Remplacez les équations (2.9) et (2.10) par l'équation. (2.12), on trouve :

$$w'_i = \frac{C' V_{u'} \sum_{j=1}^n k(\|x_j^* - u'\|^2) \delta[b(x_j^*) - u']}{C'_h V_{u'} \sum_{j=1}^{n_h} k\left(\left\|\frac{y - x_j}{h}\right\|^2\right) \delta[b(x_j) - u']}$$

En supprimant le facteur commun $V_{u'}$ du numérateur et du dénominateur et en substituant les facteurs de normalisation C et C_h dans l'équation ci-dessus, nous avons :

$$w'_i = \frac{CC_h \frac{C' \sum_{i=1}^n k(\|x_i^* - u'\|^2) \delta[b(x_i^*) - u']}{CC_h}}{CC_h \frac{C'_h \sum_{i=1}^{n_h} k\left(\left\|\frac{y - x_i}{h}\right\|^2\right) \delta[b(x_i) - u']}{CC_h}} = \sqrt{\frac{C' C_h}{CC_h}} \sqrt{\frac{\hat{q}_{u'}}{\hat{p}_{u'}}} = \sqrt{\frac{C' C_h}{CC_h}} w_i \quad (2.13)$$

Où w_i calculé par l'équation (2.8) est le poids du point i dans la région candidate cible en utilisant la représentation habituelle du modèle cible et du modèle candidat cible.

L'équation (2.19) suggère que w'_i est proportionnel à w_i . De plus, en combinant l'itération mean shift (Eq. (2.9)), nous avons :

$$y_1 = \frac{\sum_{i=1}^{n_h} x_i g_i w'_i}{\sum_{i=1}^{n_h} g_i w'_i} = \frac{\sum_{i=1}^{n_h} x_i g_i w_i \sqrt{C' C_h / CC_h}}{\sum_{i=1}^{n_h} g_i w_i \sqrt{C' C_h / CC_h}} = \frac{\sum_{i=1}^{n_h} x_i g_i w_i}{\sum_{i=1}^{n_h} g_i w_i} \quad (2.14)$$

L'équation (2.14) montre que la formule d'itération de mean shift est invariante à la transformation d'échelle des poids. Par conséquent, BWH n'améliore pas réellement le suivi par mean shift en transformant la représentation du modèle cible et du modèle candidat cible.

5. Suivi par Mean shift via CBWH :

Bien que l'idée de BWH soit bonne, mais il n'améliore pas la localisation de la cible. Pour atteindre ce que le BWH souhaite réaliser, Ning et al. [33] ont proposé une nouvelle méthode de transformation, à savoir l'algorithme corrigé du BWH (CBWH (*Corrected Background-Weighted Histogram*)) qui est utilisé pour transformer uniquement le modèle cible mais pas le modèle candidat cible (voir la figure 2.6).

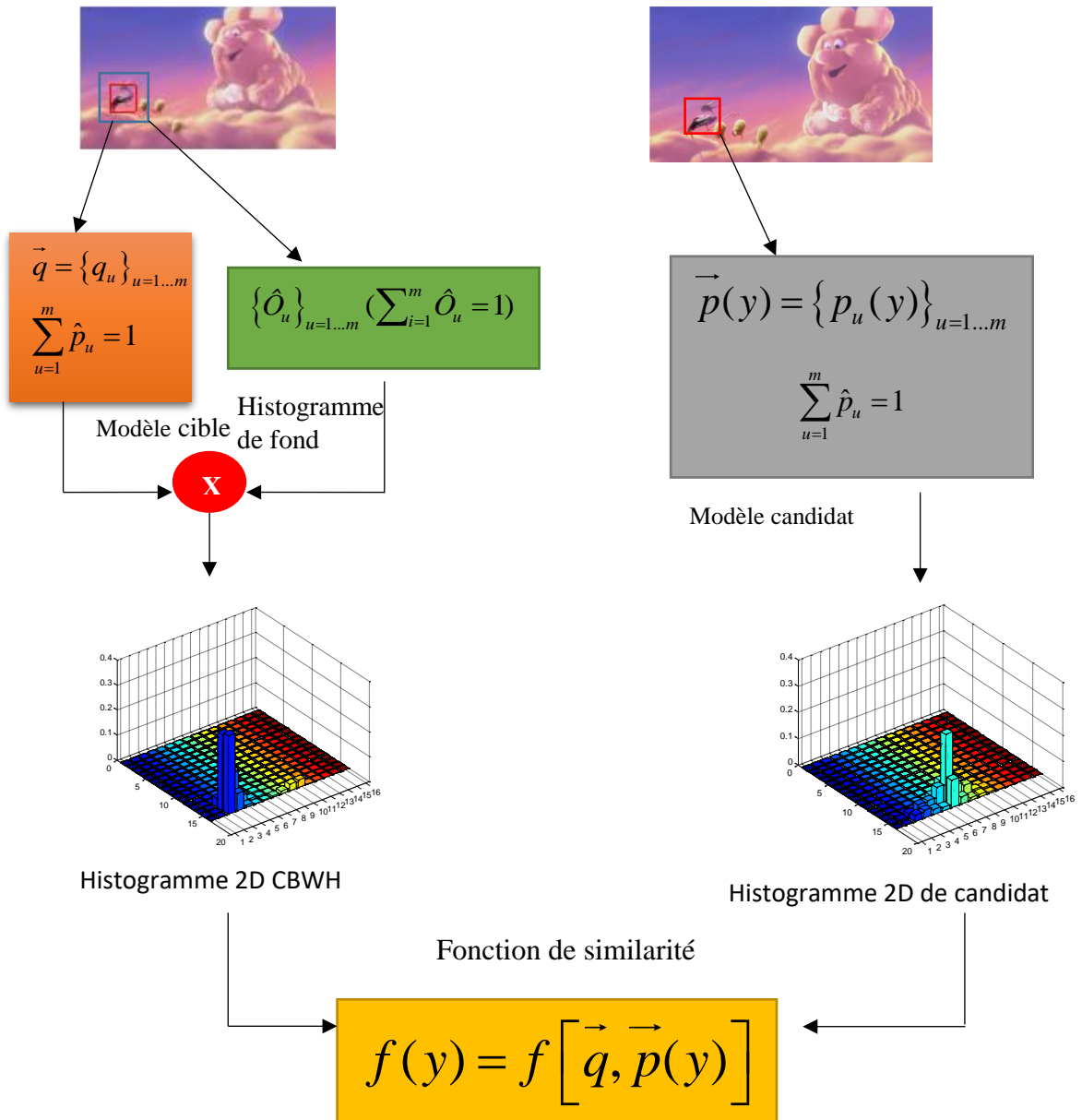


Figure 2.6 : la similarité de la méthode (CBWH).

Dans CBWH, l'équation (2.14) est utilisé pour transformer uniquement le modèle cible mais pas le modèle candidat cible. En d'autres termes, nous réduisons les caractéristiques de fond importantes uniquement dans le modèle cible mais pas dans le modèle candidat cible.

La nouvelle formule de poids est donnée par :

$$w_i'' = \sqrt{\hat{q}_u' / \hat{p}_u'(y)} \tag{2.15}$$

On note que le dénominateur dans l'équation ci-dessus est différent de celui de l'éq (10)., on peut facilement obtenir :

$$w_i'' = \sqrt{C'/C} \cdot \sqrt{v_u'} \cdot w_i \quad (2.16)$$

Puisque $\sqrt{C'/C}$ est un constant, il n'a pas d'influence sur le processus de suivi mean shift. On peut le supprimer et simplifier l'équation (14) comme :

$$w_i'' = \sqrt{v_u'} \cdot w_i \quad (2.17)$$

L'équation (2.17) indique clairement la relation entre le poids calculé en utilisant la représentation cible habituelle (c'est-à-dire w_i) et le poids calculé en exploitant les informations de fond (c'est-à-dire w_i''). Si la couleur du point i dans la région de fond est importante, la valeur correspondante de v_u' est faible. Par conséquent, dans l'équation (2.17), le poids de ce point est réduit et sa pertinence pour la localisation de la cible est réduite, ce qui accélérera la convergence du mean shift vers les caractéristiques saillantes de la cible.

Notez que si nous n'utilisons pas les informations de fond, v_u' sera égal à 1 et

$\{w_i''\}_{i=1...n_h}$ se dégradera en w_i avec la représentation cible habituelle.

5.1 Mise à jour du model de fond dans CBWH :

Dans CBWH, un modèle de couleur de fond $\{\hat{O}_u\}_{u=1...m}$ est utilisé et initialisé au début du suivi. Toutefois, dans le processus de suivi, le fond change souvent en raison des variations d'éclairage, du point de vue, d'occultation et du contenu de la scène, etc. Si le modèle de couleur d'arrière-plan d'origine est toujours utilisé sans mise à jour, la précision de suivi peut être réduite car le fond actuel peut être très différent du modèle de fond précédent. Par conséquent, il est nécessaire de mettre à jour dynamiquement le modèle de fond pour obtenir des performances de suivi CBWH robustes.

Ning et al, ont proposé ici une méthode simple de mise à jour du modèle de fond. Tout d'abord, les caractéristiques de fond $\{\hat{O}_u'\}_{u=1...m}$ et $\{v_u'\}_{u=1...m}$ de la trame actuelle sont calculées. Ensuite, la similarité de Bhattacharyya entre $\{\hat{O}_u'\}_{u=1...m}$ et l'ancien modèle de fond

$\{\hat{O}_u\}_{u=1...m}$ est calculée par :

$$\rho = \sum_{u=1}^m \sqrt{\hat{O}_u \hat{O}'_u} \quad (2.18).$$

Si ρ est inférieur à un seuil, cela signifie qu'il y a des changements considérables dans le fond, puis on met à jour $\{\hat{O}_u\}_{u=1\dots m}$ de $\{\hat{O}'_u\}_{u=1\dots m}$ et mettre à jour $\{v_u\}_{u=1\dots m}$ par $\{v'_u\}_{u=1\dots m}$. Le modèle cible transformé $\{\hat{q}'_u\}_{u=1\dots m}$ est ensuite calculé par (2.9) en utilisant $\{v'_u\}_{u=1\dots m}$. Sinon, nous ne mettons pas à jour le modèle de fond.

L'algorithme de suivi mean shift basé sur CBWH, peut être résumé comme suit :

Algorithme de suivi mean shift par CBWH

1. Calculer le modèle cible \hat{q} avec (2.1) et l'histogramme de fond $\{\hat{O}_u\}_{u=1\dots m}$ après calculer $\{v_u\}_{u=1\dots m}$ avec (2.6) et le modèle cible transformé \hat{q}' avec (2.7).

Initialisez la position y_0 de la région candidate cible dans la trame précédente.

2. Mettre $k \leftarrow 0$.
3. Calculer le modèle candidat $\hat{p}(y_0)$ avec (2.2) dans la courante trame.
4. Calculer les poids $\{w_i^*\}_{i=1\dots n_h}$ avec (2.13).
5. Calculer la position y_1 de la région candidate cible avec (2.5).
6. Mettre $d \leftarrow \|y_1 - y_0\|$, $y_0 \leftarrow y_1$, $k \leftarrow k + 1$. Poser le seuil d'erreur $\varepsilon_1 = 0.1$ (valeur par défaut), N le nombre maximal d'itérations et le seuil de mise à jour du modèle de fond $\varepsilon_2 = 0.5$ (valeur par défaut)

Si $d < \varepsilon_1$ ou $k \geq N$.

Calculer $\{\hat{O}_u\}_{u=1\dots m} \leftarrow \{\hat{O}'_u\}_{u=1\dots m}$ et $\{v_u\}_{u=1\dots m} \leftarrow \{v'_u\}_{u=1\dots m}$ basée sur les résultats du suivi de trame courante.

Si ρ de (16) est inférieur à ε_2 donc $\{\hat{O}_u\}_{u=1\dots m} \leftarrow \{\hat{O}'_u\}_{u=1\dots m}$ et

$\{v_u\}_{u=1\dots m} \leftarrow \{v'_u\}_{u=1\dots m}$ et $\{\hat{q}'_u\}_{u=1\dots m}$ est mis à jour par (2.7).

Arrêter l'itération et passer à l'étape 2 pour la prochaine trame.

Sinon revenir à l'étape 3.

6 Conclusion :

Dans ce chapitre, nous avons présenté la méthode de suivi d'objet mean shift qui est basée sur l'histogramme de couleur aussi vu son développement et ses améliorations. Malgré sa robustesse mais nous avons trouvé que si la corrélation entre la cible et le fond est élevée, la précision de localisation de l'objet sera diminuée dû l'interférence des caractéristiques saillantes de fond dans la localisation de la cible. Pour surmonter ce problème nous avons utilisé une nouvelle méthode de transformation, à savoir l'algorithme corrigé du BWH (CBWH) qui est utilisé pour transformer uniquement le modèle cible mais pas le modèle candidat cible. En d'autres termes, nous réduisons les caractéristiques de fond importantes uniquement dans le modèle cible mais pas dans le modèle candidat cible.

Chapitre 3 :

Simulation, résultats expérimentaux et discussion

1 Introduction :

Après avoir vu dans le chapitre précédant la théorie du principe de base de l'algorithme de suivi mean shift dans une séquence d'images, et aussi les deux améliorations BWH et CBWH on a pour but dans ce chapitre de tester ces deux algorithmes qui sont basés sur l'information du fond. Tous d'abord, on présente l'architecture globale de l'algorithme, à l'aide d'un schéma illustré leur différentes étapes de la méthode de l'algorithme mean shift utilisant l'histogramme CBWH. Ensuite on applique cet algorithme sur plusieurs séquences d'images, pour le valider. Enfin, on présente les résultats obtenus avec leurs discussions.

2 Structure de l'algorithme :

Le suivi par mean shift via CBWH s'effectue en plusieurs étapes. Le suivi commence par l'acquisition de l'image initiale et détecter ou sélectionner l'objet d'intérêt manuellement ou automatiquement. Ensuite, on modélise cet objet par un rectangle sur laquelle on calcule sa distribution des caractéristiques de couleur HSV puis on calcule la distribution des caractéristiques du fond (Background). Enfin, on calcule l'histogramme CBWH de modèle et comparais avec l'histogramme de candidats par le coefficient de Battacharyya. La figure résume la structure de l'algorithme du tracker amélioré mean shift à l'aide de CBWH.

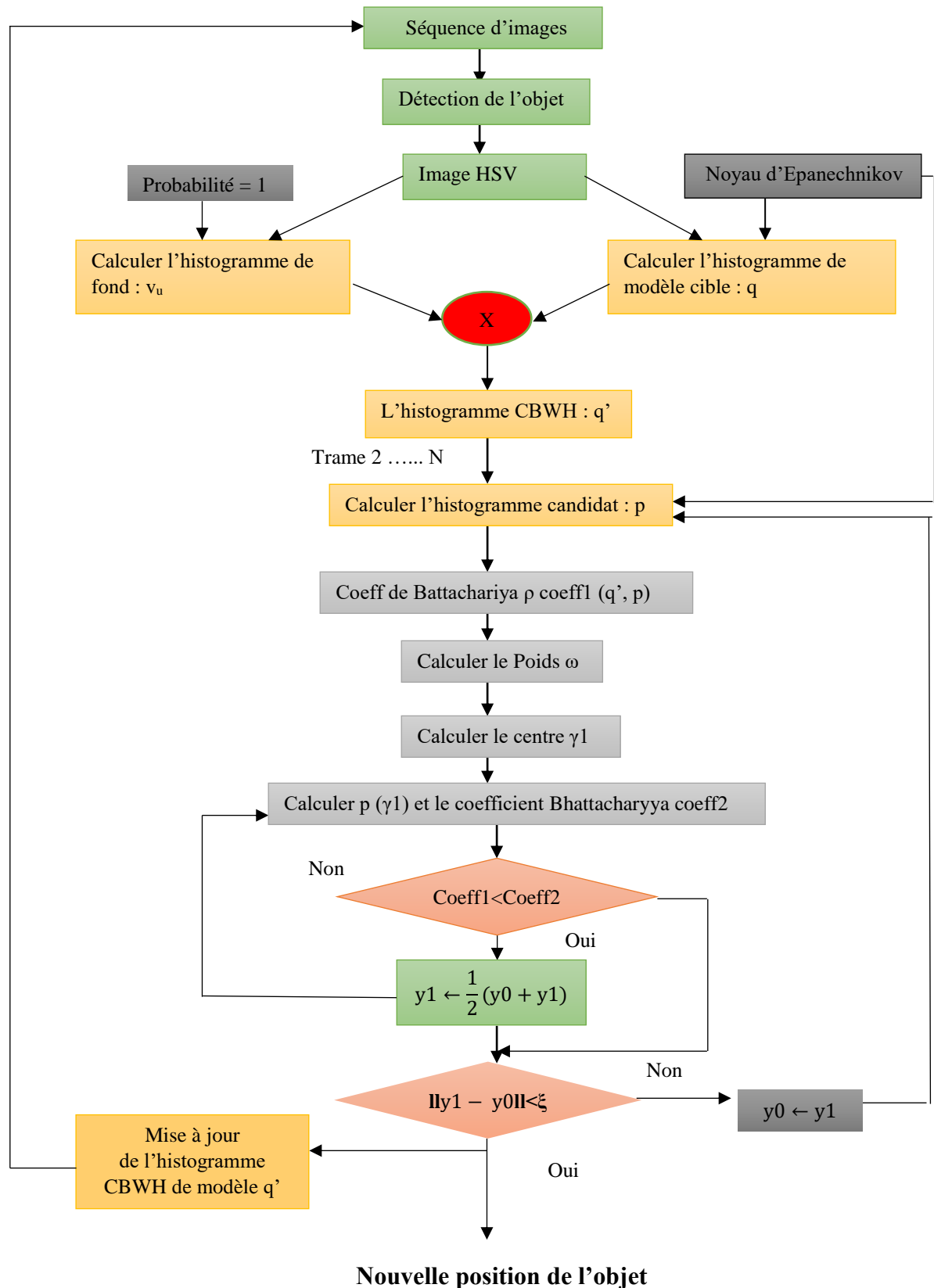


Figure3.1 : La structure de l'algorithme de suivi par Mean shift via CBWH

2.1 Détection de l'objet :

Pour chaque méthode de suivi d'objets exige un mécanisme de détection d'objet dans chaque trame ou quand l'objet apparaît d'abord dans la vidéo. Il existe plusieurs façons de détection d'objet parmi eux soustraction de fond, Viola et Jones, ...etc. Dans ce travail, nous avons sélectionnée la position initiale de l'objet à partir de la base de données utilisées (OTB) et les informations initiale du fond (background).

2.2 Transformation RGB vers HSV :

Dans ce travail nous utilisons l'espace de couleur HSV au lieu de l'espace RGB. Le principe de l'espace HSV est de caractériser les couleurs de façon plus intuitive, conformément à la perception naturelle des couleurs en termes de :

- **Teinte (Hue) :** c'est le nom qu'on utilisera pour désigner la couleur "verte", "mauve", "orange".
- **Saturation (Saturation) :** c'est le taux de pureté de la couleur, qui doit varier entre la pureté maximale (couleur éclatante) et l'achromatisme (niveau de gris).
- **Valeur (Value) :** c'est la mesure de l'intensité lumineuse de la couleur, qui doit varier entre le noir absolu et le blanc.

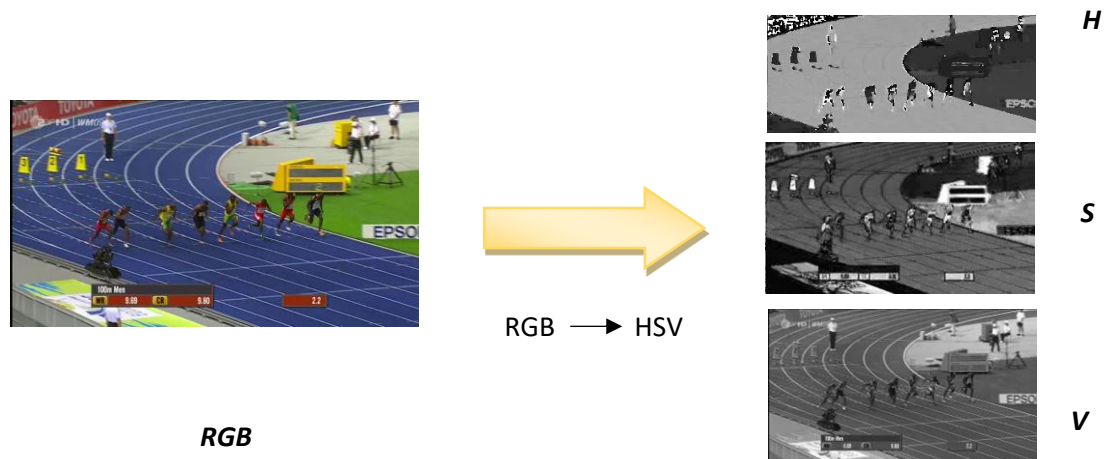


Figure 3.2 : Exemple de la transformation d'espace couleur RGB vers HSV

2.3 Histogramme CBWH :

Dans ce travail, nous utilisons les caractéristiques de couleur de l'espace HSV pour représenter le modèle d'apparence de l'objet cible par l'histogramme pondéré HSV et nous représentons aussi les caractéristiques du fond. Dans le tracker Mean shift on calcule la distribution de couleur pondéré par le noyau d'Epanechnikov et la distribution des caractéristiques du fond par une probabilité = 1. Le produit de ces deux histogrammes est l'histogramme de modèle CBWH.

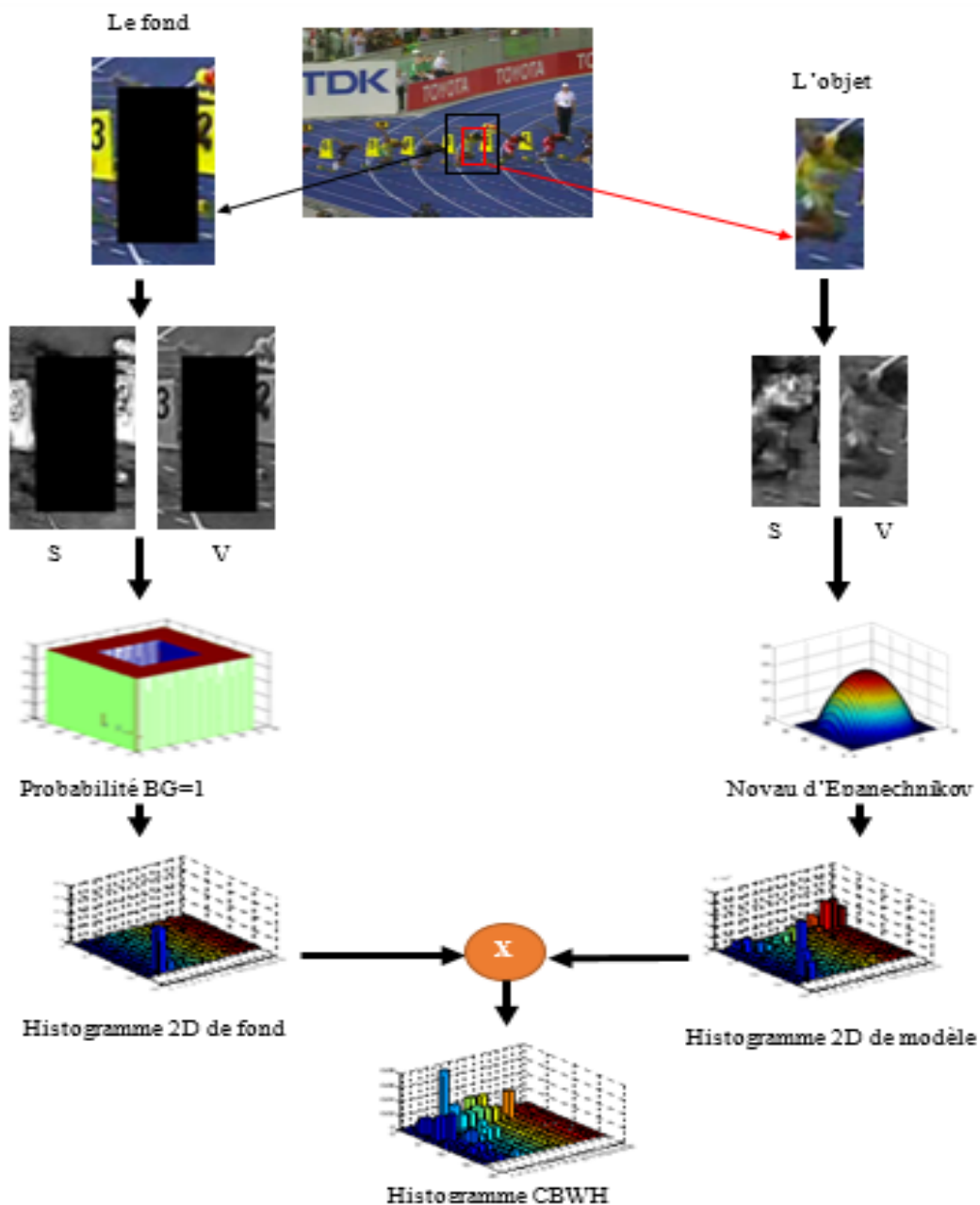


Figure 3.3 : Exemple sur l'histogramme 2D CBWH de composante S, V de l'espace HSV.

2.3.3 Poids et le nouveau centre :

Pour chaque trame, le poids de chaque pixel est calculé en fonction de l'histogramme de modèle CBWH de l'objet et de l'histogramme pondéré de candidat. Ce poids représente le nombre de pixels ayant la couleur de ce pixel. Ainsi, l'image de poids (voir la figure 3.4) représente des valeurs élevées au niveau des régions ayant une couleur proche de la couleur de l'objet suivi. Un poids est associé à la mesure de similarité et on peut alors extraire et calculer le vecteur Mean Shift, qui a pour but de fournir la nouvelle position estimée de la cible dans l'image courante.



Figure 3.4: Exemple sur l'image de poids.

2.4 Environnement de travail :

Dans cette section, nous présenterons les environnements matériel et logiciel de notre travail.

2.4.1 Environnement matériel :

L'ensemble de matériel pour réaliser ce travail :

PC DELL Inspiron 15 series 3000 dont les caractéristiques :

- Procésseur : Intel core i5
- Ram : 8 go
- Carte graphique : NVIDIA GEFORCE 920 2go
- Disque dur : 500 go HDD
- Système d'exploitation : Windows 10 64bits







2.4.2 Langage de programmation :

Dans ce travail, nous utilisons logiciel MatlabR2007b qui il est un langage de haut niveau qui permet l'exécution de tâches nécessitant une grande puissance.

2.4.3 Bases de données :

Les séquences d'images utilisées pour évaluer les performances de notre algorithme existent dans la base de données OTB qui est disponible sur le site : <http://www.visual-tracking.net>. Chacune des séquences vidéo est associées une vérité terrain qui capture l'interprétation vraie de la scène en termes d'objets à suivre et est liée à un fichier TXT décrivant la position de la cible dans chaque image (gauche, haut, largeur, hauteur). Le tableau (3.1) illustre les séquences d'images utilisées :

Tableau 3.1 : Séquences d'images utilisées

<i>Séquence d'image</i>	<i>1^{er} image</i>	<i>Nombre de trames</i>	<i>Taille de l'image</i>	<i>Défis</i>
<i>Bolt</i>		350	640x360	OCC, DEF, IPR, OPR
<i>Bird2</i>		99	720x400	OCC, DEF, FM, IPR, OPR
<i>Human8</i>		128	320x240	IV, SV, DEF
<i>Tiger1</i>		354	640x480	IV, OCC, DEF, MB, FM, IPR, OPR
<i>BlurOwl</i>		631	640x480	SV, MB, FM, IPR
<i>David3</i>		252	640x480	OCC, DEF, OPR

- OCC : Occlusion
- DEF : Déformation
- IPR : Rotation dans le plan
- OPR : Rotation hors le plan
- FM : Mouvement rapide
- IV : Variation d'éclairage
- SV : Variation d'échelle
- MB : Flou de mouvement

2.5 Mesure des performances d'un système de suivi :

L'évaluation de performance des systèmes de suivi exige une comparaison des résultats des algorithmes avec des résultats "optimaux" la vérité terrain. Les principales métriques de performance dans le suivi d'objet sont [18] [71] : l'erreur de localisation du centre et le taux de recouvrement.

2.5.1 Erreur de localisation du centre (ELC):

L'erreur de localisation du centre (figure 3.5(a)) est une mesure courante consistant à mesurer la distance moyenne entre les centres des boîtes prédites $\{p_i\}_{i=1}^M$ et de la vérité terrain $\{g_i\}_{i=1}^M$

$$ELC = \frac{1}{M} \sum_{i=1}^M \|p_i - g_i\|$$

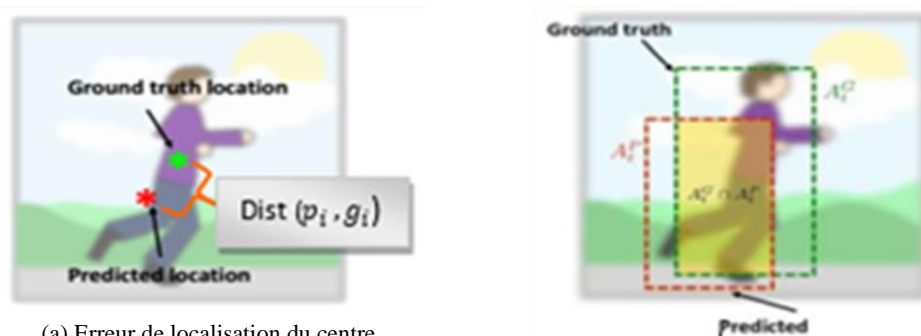
Cette mesure ne rend pas compte de la précision en taille des boîtes prédites. Le seuil habituellement utilisé pour comparer la précision entre différents trackers est de 20 pixels.

2.5.2 Taux de recouvrement (TR) :

Le taux de recouvrement (figure 3.5(b)) entre la boîte prédite B et la vérité terrain B' est définie comme étant le rapport des aires d'intersection et d'union des boîtes :

$$TR(B, B') = \frac{|B \cap B'|}{|B \cup B'|}$$

TR, est une mesure d'erreur plus précise que l'erreur de localisation du centre puisqu'elle tient compte de la taille des boîtes. Le taux de recouvrement moyen consiste alors à moyenner TR sur l'ensemble des images de la base.



(a) Erreur de localisation du centre

(b) Taux de recouvrement des boîtes prédite (rouge) et de vérité terrain (vert)

Figure 3.5 : Métriques de performance : (a) Erreur de localisation du centre, (b) Taux de recouvrement.

2.6 Résultats Expérimentaux :

Dans notre travail, nous avons choisi six séquences d'images pour illustrer les résultats obtenus avec des algorithmes de suivi d'objet : Mean shift, BWH, CBWH. Ces séquences contiennent plusieurs défis de suivi. L'objet cible dans les séquences d'images est représenté par un rectangle et on sélectionne la cible dans la première trame à partir des résultats obtenus sur le terrain. Les résultats expérimentaux sont exprimés sur deux plans, qualitatif (visuellement) et quantitatif (Erreur de localisation du centre, Taux de recouvrement). Dans ce travail, nous avons utilisé le nombre d'itération = 20 et $\epsilon = 0.5$ et le seuil de mise à jour de modèle de fond égale à 0.5.

2.6.1 Comparaison entre Mean shift et BWH et CBWH:

Dans cette partie, nous allons comparer les différents résultats obtenus dans chacune des séquences par : Mean shift qui utilise l'histogramme de couleur HSV, BWH qui utilise l'histogramme de couleur HSV et les informations de fond pour le modèle cible et le candidat cible et enfin CBWH qui utilise l'histogramme de couleur HSV et l'information de fond pour le modèle cible uniquement.

2.6.1.1 La séquence Bolt:

L'objet à suivre dans la séquence Bolt est un athlète et les défis présent dans cette séquence sont : l'occultation, la déformation de l'objet, la rotation dans le plan et la rotation hors plan de l'objet. La figure (3.6) montre les résultats de suivi par MS, BWH et CBWH pour les trames 68, 117, 180, et 252. On peut voir que, les trois trackers peuvent suivre l'objet avec précision dans les premiers trames (trame 68), mais la précision diminue dans la trame 117 à cause d'existence d'un autre objet de même couleur sur le bord. Dans les trames 180 et 252, MS et BWH ont carrément perdu l'objet et suivent un autre objet qui a les mêmes caractéristiques de couleurs. Cependant, CBWH peut à suivre la cible tout au long de la séquence de manière très précise et robuste à cause d'utilisation des informations de fond pour représenter le modèle cible et la mise à jour de ce modèle.

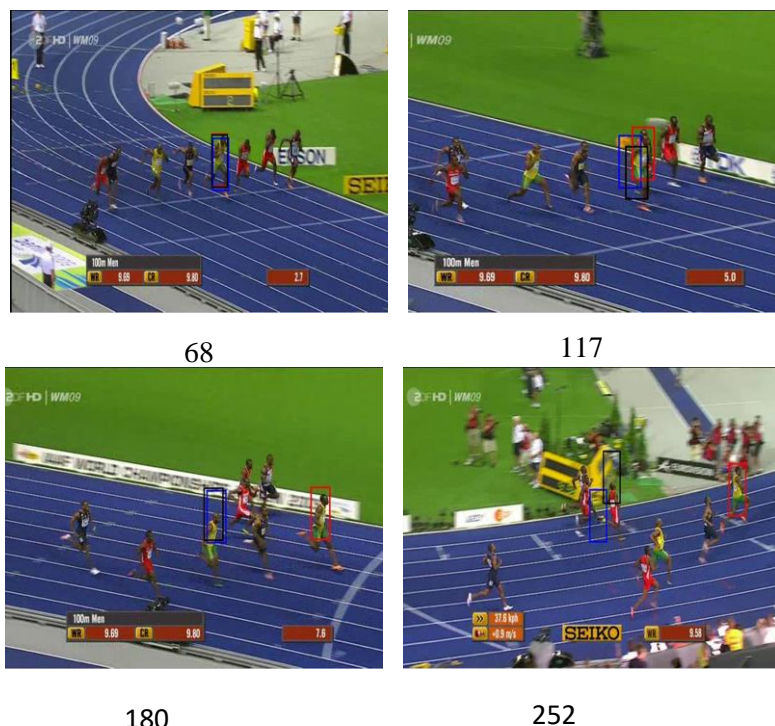


Figure 3.6 : Résultats de suivi sur la séquence Bolt avec Mean shift (rectangle bleu), BWH (rectangle noir) et CBWH (rectangle rouge).

Pour la comparaison quantitative, la figure (3.7) représente les courbes de l'erreur de localisation du centre et le taux de recouvrement en fonction de nombre trame pour les trackers MS, BWH et CBWH. On trouve que, les courbes de l'erreur de location du centre et taux de recouvrement pour MS et BWH sont presque identiques. Ils peuvent suivre l'objet jusqu'à la trame 110 (ELC inférieur à 20 et TR supérieur à 0,5 dans la plupart), puis perdent l'objet à partir 120 (ELC supérieur à 20 et TR égale à 0), alors que CBWH peut suivre l'objet avec précision jusqu'à la fin de la séquence comme montre dans les deux courbes.

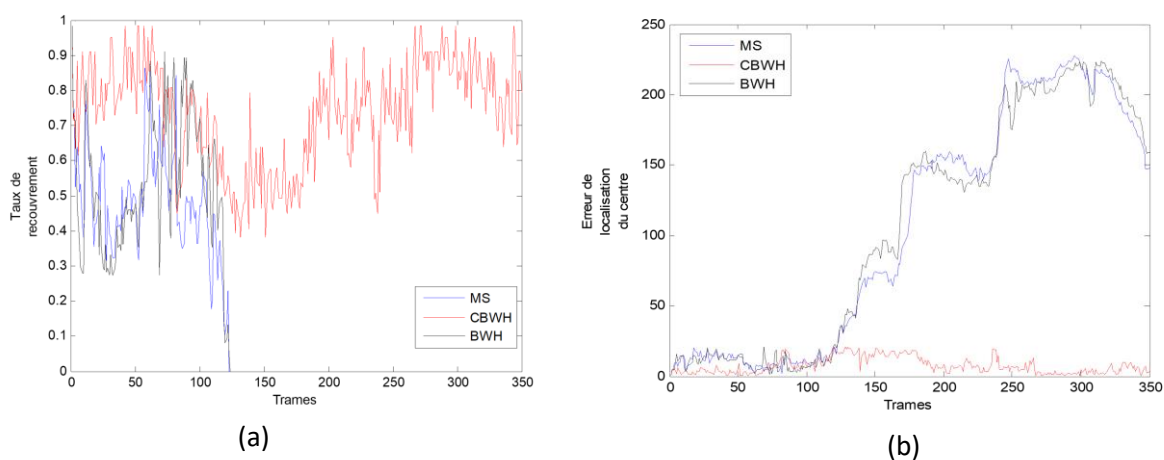


Figure 3.7: Comparaison quantitative entre MS, BWH et CBWH la séquence Bolt : (a) Taux de recouvrement. (b) Erreur de localisation du centre

2.6.1.2 Séquence Bird2 :

La séquence Bird2 représente un oiseau qui est l'objet à suivre et les défis présent dans cette séquence sont: OCC, DEF, FM, IPR et OPR. La figure (3.8) montre que les résultats de suivi par MS, BWH et CBWH pour les trames 8, 19, 48 et 86. On remarque que, le tracker CBWH suit l'objet avec précision après dans la trame 19 MS perd l'objet dû à la ressemblance entre l'objet et l'arrière-plan après dans la trame 62 BWH aussi perd l'objet à cause du même problème tandis que CBWH continu à suivre l'objet avec précision et robustess. En arrivant à la trame 86 on voie que MS et BWH retrouvent l'objet mais par hasard et continuent à le suivre jusqu'à la fin de la séquence tout comme CBWH.



Figure 3.8: Résultats de suivi sur la séquence Bird2 avec Mean shift (rectangle bleu), BWH (rectangle noir) et CBWH (rectangle rouge).

Pour la comparaison quantitative, la figure (3.9) représente les courbes de l'erreur de localisation du centre et le taux de recouvrement en fonction de nombre de trame pour les trackers MS, BWH et CBWH. On trouve que, les deux courbes pour MS et BWH sont presque identiques sauf dans quelques trames. Les valeurs de courbe ELC est supérieur à 20 pixels et dans le TR inférieur à 0,5 dans la plus part des trames en raison que la couleur de fond et de cible sont similaire. Cela signifie que l'objet à suivre par MS et BWH est perdu après les premiers trames jusqu'à les dernières trames, où il le retrouve par hasard à la fin de la séquence. Cependant, en regardant les courbes pour CBWH on trouve qu'il suit l'objet avec grande précision tout au long de la séquence.

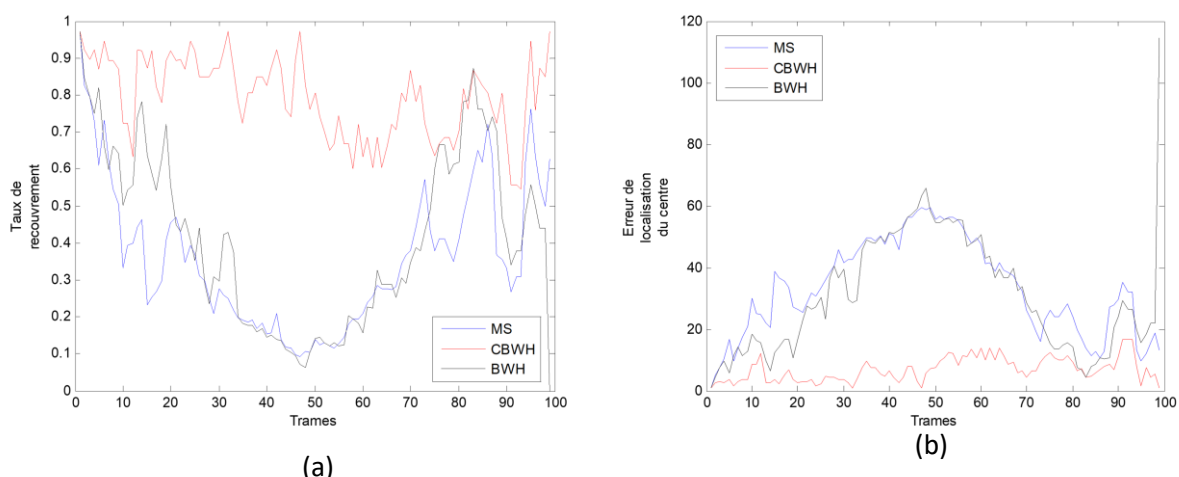


Figure 3.9 : Comparaison quantitative entre MS, BWH et CBWH la séquence Bird2 : (a) Taux de recouvrement. (b) Erreur de localisation du centre.

2.6.1.3 Séquence Human8 :

L'objet à suivre dans cette séquence est un homme et les défis présent dans cette séquence sont: IV et la SV puis DEF. La figure (3.10) montre que les résultats de suivi obtenus par MS, BWH et CBWH pour les trames 4, 47, 68 et 118. On peut voir que, de la trame 1 jusqu'à la trame 40 les 3 trackers suivent l'objet avec précision et une certaine robustesse pour CBWH. aussi, entre les trames 40 et 128, le tracker BWH perd complètement l'objet en raison de changement de fond et de leur similarité avec l'objet. Cependant, les deux trackers MS et CBWH continuent à suivre l'objet malgré tous les problèmes.

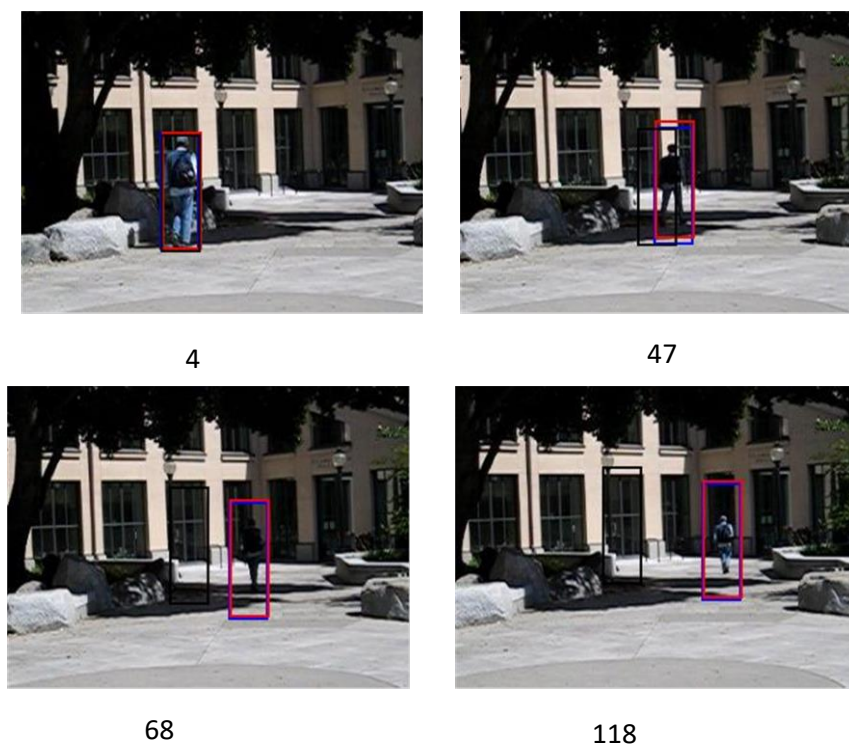


Figure 3.10: Résultats de suivi sur la séquence Human8 avec Mean shift (rectangle bleu), BWH (rectangle noir) et CBWH (rectangle rouge).

La figure (3.11) représente les courbes de l'erreur de location du centre et taux de recouvrement en fonction nombre trame pour les trackers de suivi MS, BWH et CBWH. On remarque que, les courbes de l'erreur de location du centre et taux de recouvrement pour MS, BWH et CBWH se ressemblent entre les trames 1 et 40. Après la trame 40, les courbes de tracker BWH change de façon considérable, le TR commence à diminuer et le ELC commence à augmenter jusqu'à la trame 50. Puis il atteint la valeur 0 pour TR et la valeur maximal pour ELC, cela veut dire que BWH a complètement perdu l'objet. Les courbes pour MS et CBWH sont presque identiques jusqu'à la fin de la séquence, où les deux trackers peuvent à suivre l'objet avec précision (ELC inférieur à 10 pixel et TR supérieur à 0,5 dans la plupart de séquence).

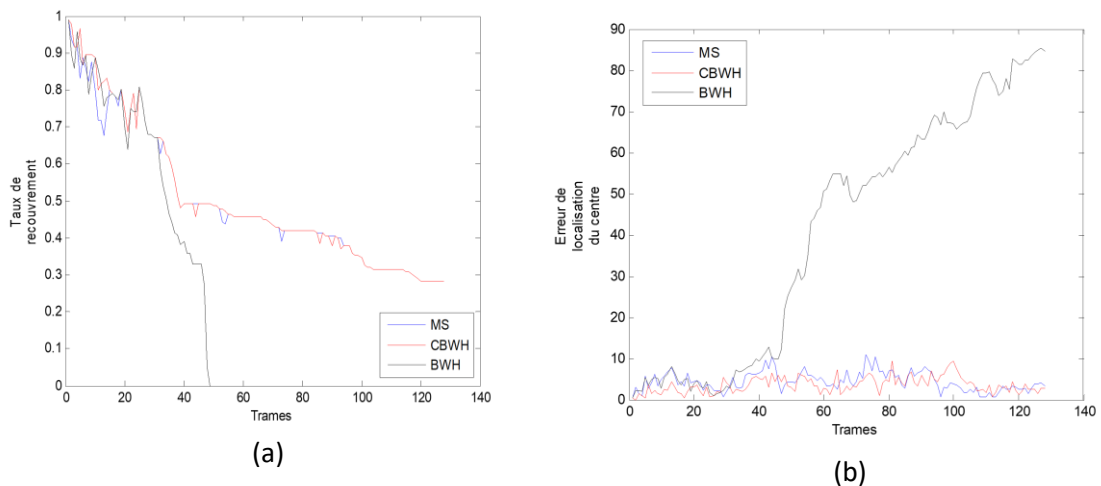


Figure 3.11: Comparaison quantitative entre MS, BWH et CBWH la séquence Human8 : (a) Taux de recouvrement. (b) Erreur de localisation du centre.

2.6.1.4 Séquence Tiger1 :

Le tigre en peluche est l'objet que nous allons suivre dans la séquence tiger1 et les défis présent dans cette séquence sont: IV, OCC, DEF, MF, MB, IPR et OPR. La figure (3.12) montre les résultats de suivi par MS, BWH et CBWH pour les trames 10, 152, 221 et 250. Dans les trames présentées 10 et 350 nous voyons que les 3 trackers peuvent suivre l'objet. Cependant, dans les trames 152 et 221 MS et BWH perdent l'objet à cause du changement d'illumination et le mouvement rapide, alors que CBWH continue à le suivre avec plus de précision. En arrivant à la trame 350 un problème d'occultation partielle se pose mais n'affecte pas les trackers.



Figure 3.12 : Résultats de suivi sur la séquence Tiger1 avec Mean shift (rectangle bleu), BWH (rectangle noir) et CBWH (rectangle rouge).

Pour la comparaison quantitative, la figure (3.13) représente les courbes de l'erreur de location du centre et taux de recouvrement en fonction nombre trame pour les trackers de suivi MS, BWH et CBWH. On remarque que, les courbes de l'erreur de location du centre et taux de recouvrement pour MS, BWH et CBWH sont presque identiques pour les trois trackers sauf sur quelques trames ou CBWH est plus précis et robuste par rapport à MS et BWH.

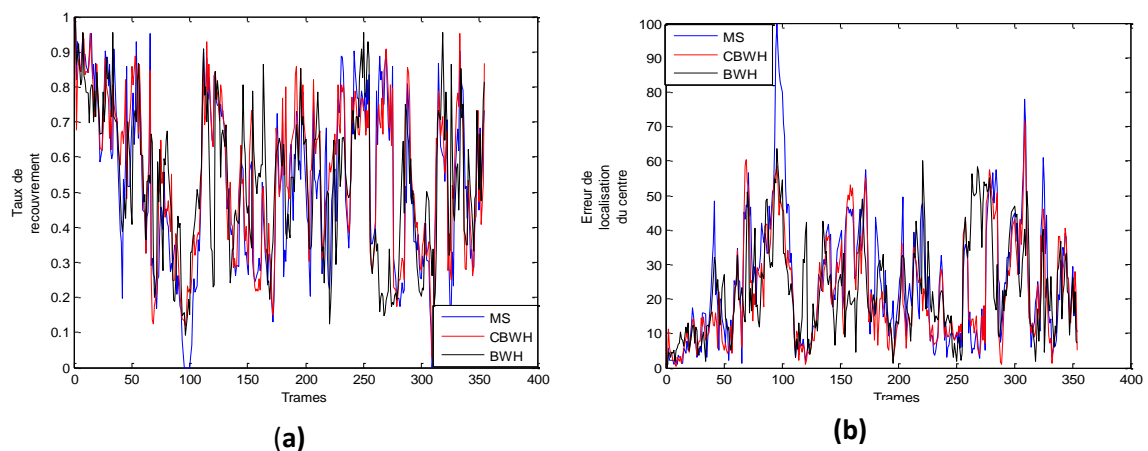


Figure 3.13 : Comparaison quantitative entre MS, BWH et CBWH pour la séquence Tiger1 : (a) Taux de recouvrement. (b) Erreur de localisation du centre.

2.6.1.5 Séquence David3:

Dans la séquence David3 l'objet à suivre est un homme qui marche et les défis présent dans cette séquence sont: mouvement de camera, déformation et occultation partielle. La figure (3.14) montre les résultats de suivi par MS, BWH et CBWH pour les trames 15, 96, 153 et 240. On peut voir que les 3 trackers peuvent suivre l'objet tout au long de la séquence mais CBWH est le plus précis et le plus robuste surtout dans les parties où il y a une grande ressemblance entre l'objet et le fond.



Figure 3.14: Résultats de suivi sur la séquence David3 avec Mean shift (rectangle bleu), BWH (rectangle noir) et CBWH (rectangle rouge).

La figure (3.15) représente les courbes de l'erreur de location du centre et taux de recouvrement en fonction de nombre trame pour les trackers de suivi MS, BWH et CBWH. On remarque que les courbes de MS et BWH sont similaires et ils peuvent suivre l'objet sauf entre les trames 85 et 110 (ELC supérieur à 20 pixels et TR inférieur 0,5). Par contre, les deux courbes de CBWH sont plus précises (ELC inférieur à 20 pixel et TR supérieur à 0,5) dans le suivi de notre objet sur toute la séquence.

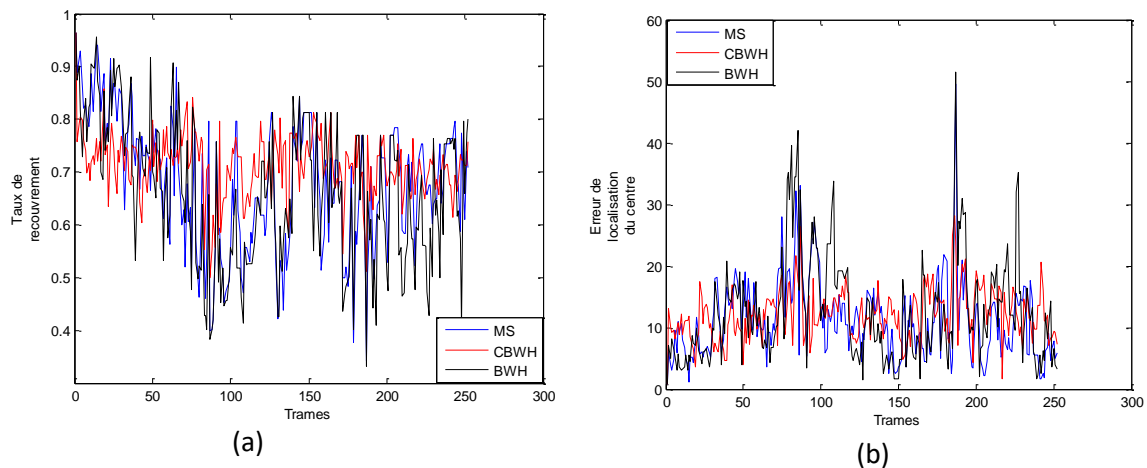


Figure 3.15: Comparaison quantitative entre MS, BWH et CBWH pour la séquence David3 :
 (a) Taux de recouvrement. (b) Erreur de localisation du centre.

2.6.1.6 Séquence BlurOwl:

Dans cette séquence l'objet qu'on va suivre est un dessin sur un poster et les défis présent sont: mouvement rapide, flou du mouvement, variation d'échelle et Rotation dans le plan. La figure (3.16) montre les résultats de suivi par MS, BWH et CBWH pour les trames 12, 49, 392 et 547. On remarque qu'au début de la séquence les trois trackers suivent l'objet (trame 12). Ensuite dans la trame 49 on voit la présence du flou de mouvement du coup, MS perd partiellement l'objet et BWH le perd carrément en ce temps CBWH peut suivre avec précision. Après en arrivant à la trame 392 on remarque que MS et BWH ont perdu l'objet dû au mouvement rapide mais CBWH n'est pas affecté. Enfin, dans la trame 547 qui montre la rotation dans le plan et le flou de mouvement, MS et BWH n'arrivent pas à suivre l'objet tandis que CBWH suit avec précision et robustesse.



Figure 3.16 : Résultats de suivi sur la séquence BlurOwl avec Mean shift (rectangle bleu), BWH (rectangle noir) et CBWH (rectangle rouge).

Pour la comparaison quantitative, la figure (3.17) représente les courbes de l'erreur de location du centre et taux de recouvrement en fonction de nombre trame pour les trackers de suivi MS, BWH et CBWH. On remarque que, pour la courbe de l'erreur de localisation du centre MS dépasse 20 pixels qui sont notre limite dans la plupart des trames. cependant, BWH dépasse 20 pixels dans les trames 60, 100, 120, 150, 230, 320, 380, 400, 500, 550 et 580 donc il suit moyennement l'objet. Pour le tracker CBWH les valeurs de ELC inférieur à 20 pixels pour toutes les trames, cela signifie que ce tracker est robuste. Dans la courbe de taux de recouvrement on voit que MS BWH sont inférieurs à 0,5 dans la plupart des trames alors que CBWH est supérieur à 0,5 pour toutes les trames, donc CBWH réussit à suivre tout au long de la séquence.

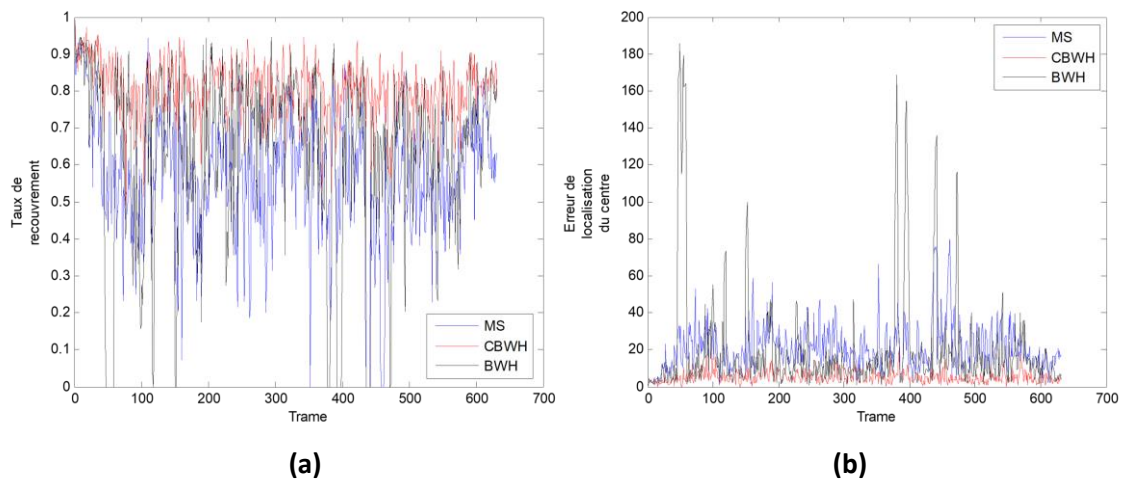


Figure 3.17: Comparaison quantitative entre MS, BWH et CBWH pour la séquence BlurOwl : (a) Taux de recouvrement. (b) Erreur de localisation du centre.

2.7 La comparaison entre les moyennes d'erreur de localisation du centre et le taux de recouvrement pour MS, BWH et CBWH :

Le tableau ci-dessous résume la comparaison entre les trois trackers MS, BWH et CBWH, en utilisant les métriques ELC et TR pour chaque séquence testée. Nous pouvons clairement voir que l'erreur de localisation du centre minimum et le taux de recouvrement maximal dans toutes les séquences testées sont obtenus par CBWH (9,454 / 0,684), en particulier dans la séquence Bolt (7,391 / 0.734). Cela signifie que, CBWH a atteint de meilleures performances par rapport aux autres trackers. A travers ces résultats, on peut dire que, le tracker CBWH peut suivre l'objet cible de manière plus robuste et plus précision que le Mean shift et BWH.

Tableau 3.2 : Les moyennes des résultats obtenus de l'erreur de localisation du centre et du taux de recouvrement.

Séquence	Erreur de localisation du centre			Taux de recouvrement		
	MS	CBWH	BWH	MS	CBWH	BWH
Bolt	105.7526	7.3914	106.5718	0.1694	0.7349	0.1851
Bird2	33.3489	6.8306	30.5743	0.3594	0.7943	0.4148
Human8	4.3818	3.6384	40.7770	0.4995	0.5067	0.2427
Tiger1	24.8466	21.6953	22.9445	0.5220	0.5550	0.5296
David3	11.2271	12.1122	12.5255	0.6789	0.7125	0.6606
BlurOwl	20.6718	5.1610	20.1832	0.5676	0.8017	0.6372
Moyens	33.37146	9.45481	38.92938	0.46613	0.68418	0.445

3 Conclusion:

Dans ce chapitre, au début nous avons vu l'algorithme de suivi mean shift via CBWH qui utilise un histogramme de couleur HSV et les informations du fond. Ensuite on a commencé à tester notre algorithme sur des séquences d'images on l'on trouve des différents défis qui s'oppose au suivi d'objet (les problèmes du fond, déformation de l'objet, mouvement rapide et flou du mouvement...). Après on a présenté les résultats obtenus pour chaque séquence pour les algorithmes MS, BWH et CBWH et on a comparé entre les trois par les deux métriques : erreur de localisation du centre et taux de recouvrement. Les résultats obtenus ont montré que CBWH est plus précis et plus robuste que MS et BWH pour le suivi d'objet face aux diverses problématiques et précisément dans les problèmes du fond. Donc on a trouvé que prendre en compte les informations du fond pour représenter le modèle de l'objet cible dans le tracker mean shift ajoute une grande précision et une grande robustesse.

Conclusion générale

Conclusion générale :

Le suivi d'objets est une tâche importante dans le domaine de la vision par ordinateur. La prolifération des ordinateurs de grande puissance, la disponibilité de haute qualité et peu coûteux des caméras vidéo, et la nécessité croissante pour l'analyse vidéo automatisée a suscité un grand intérêt dans les algorithmes de suivi de l'objet [72]

A ce jour, aucun tracker n'est véritablement capable de maîtriser toutes les situations difficiles pouvant apparaître lors du suivi d'un objet : changements d'apparence, d'illumination, occultations, mouvement de la caméra, etc. [18]. Le but de notre travail était de voir à quel point la combinaison entre l'histogramme de couleur de l'objet cible et les informations du fond pourrait être avantageuse pour améliorer la précision et robustesse de l'algorithme de suivi Meanshift.

Le travail de ce mémoire s'articule autour de l'étude de l'algorithme de suivi mean shift qui maximise la similarité d'apparence itérativement en comparant l'histogramme de couleur de l'objet cible et une fenêtre autour de la position hypothèse d'un objet candidat via CWBH qui exploite les caractéristiques de couleurs du fond.

Dans ce mémoire, on a implémenté les algorithmes Meanshift, BWH et CBWH sur des séquences d'images de base de données OTB qui contient plusieurs défis (les problèmes du fond, déformation de l'objet, mouvement rapide et flou du mouvement...) et a fait une étude comparative entre les trois méthodes du point de vue qualitatif (visuel) et de vue quantitatif (erreur de localisation de centre et taux de recouvrement). Les résultats obtenus montrent que l'utilisation de l'histogramme de couleurs HSV avec les informations du fond pour le modèle cible et le modèle candidat c.-à-d. BWH n'améliorent pas la précision et la robustesse de Mean shift et d'un autre coté diminue la rapidité du suivi d'une manière considérable. Par contre CBWH a amélioré la précision du suivi et a réduit le nombre d'itérations par rapport à Mean shift. Aussi, l'un de ses avantages importants réside dans le fait qu'il réduit la sensibilité du suivi du mean shift à l'initialisation de la cible de sorte que CBWH puisse suivre de manière robuste la cible même s'il n'est pas bien initialisé. Par conséquence, on a trouvé que prendre en compte les informations du fond pour représenter le modèle de l'objet cible dans le tracker mean shift ajoute une grande précision et une grande robustesse.

En perspectives, nous proposons :

- Utiliser d'autres caractéristiques (HOG (Histogram of Oriented Gradients) et CNN (caractéristiques des réseaux convolutifs)) avec les caractéristiques de couleur pour calculer l'histogramme conjoint.
- Développer l'algorithme de suivi Mean shift et le rendre le plus intelligent ; adaptable aux mouvements rapides et à la variation d'apparence de l'objet, détecter l'objet cible après l'occultation complexe et plus distinctif entre l'objet et le fond, surtout lorsqu'ils sont similaires.

Bibliographie

Bibliographie :

- [1] A. Yilmaz, O. Javed. Object Tracking : “A Survey “, ACM Computing Surveys, Vol. 38, No. 4, Article 13, Publication date : December 2006. pp. 1-45
- [2] D. Comaniciu, V. Ramesh et P. Meer. Kernel-Based Object Tracking, “IEEE Transactions” On Pattern Analysis and Machine Intelligence, vol. 25, no. 5, pp. 564-575, May 2003.
- [3] Shweta, k. Talmale, Nitin et j. janwe. “Object tracking using modified Mean-shift algorithm.” International Journal of Pure and Applied Research in Engineering and Technology « IJPRET, 2016; Volume 4 (9): 560-570.
- [4] P. Hingane and S. Shirsat. "Object Tracking Using Joint Color-Texture Histogram." International Journal of Science and Research (2013): 2603-2606.
- [5] S. Medouakh, M. Boumehraz, and N. Terki. "Improved object tracking via joint color-LPQ texture histogram based mean shift algorithm." Signal, Image and Video Processing 12.3 (2018) : 583-590.
- [6] M. Kristan, J. Matas, A. Leonardis, T. Vojir, R. Pflugfelder, G. Fernandez, G. Nebehay, F. Porikli, et L. Cehovin, (2015b). « A novel performance evaluation methodology for single target trackers ».
- [7] H. Lazhari. Suivi d'objets dans une séquence d'images par l'algorithme "Camshift". Thèse de Master, université de Mohamed Khider Biskra .Mai 2016. juin 2017
- [8] M. Mikram Suivi d'objets dans une séquence d'images par modèle d'apparence : « conception et évaluation, Thèse de doctorat. Université de Toulouse, 2008 ».
- [9] Bouchir Suivi d'objet par caractéristique locales encadrant la structure. « Thèse de Doctorat, Université Montréal, 2014 ».
- [10] R. Trichet. Suivi d'objet pour la télévision interactive, « thèse de doctorat, L'école télécom Paris Tech, 08 décembre 2008 ».
- [11] T. PENNE, Développement d'un système de tracking vidéo sur caméra robotisée. « Thèse de doctorat. Université Blaise Pascal-Clermont-Ferrand II. 2011 ».
- [12] M. Brulin, « Analyse sémantique d'un trafic routier dans un contexte de vidéoSurveillance. Diss. Bordeaux 1, 2012 ».
- [13] T. Brox, et J. Malik, (2010). « Object segmentation by long term analysis of point trajectories. In European conference on computer vision, pages 282–295. Springer »
- [14] B. Babenko, M. Yang, H. Belongie, S. Robust : « object tracking with online multiple instance learning. IEEE Trans. Pattern. Anal. Mach. Intell. 33(8), 1619–1632 (2011) ».

- [15] A.YILMAZ, X. LI, AND M. SHAH, 2004. « Contour based object tracking with occlusion handling in video acquired using mobile cameras.IEEE Trans. Patt. Analy. Mach. Intell. 26, 11, 1531–1536 ».
- [16] G. MOSTAFAOUI, C. ACHARD, et M. MILGRAM, Suivi temps réel de personnes dans des séquences d'images couleur. In : « 20^e Colloque sur le traitement du signal et des images, FRA, 2005. GRETSI, Groupe d'Etudes du Traitement du Signal et des Images, 2005 ».
- [17] X. Li, W. Hu, C. Shen, Z. Zhang, A. Dick, and A. V. D. Hengel. « A survey of appearance models in visual object tracking »,ACM Transactions on Intelligent Systems and Technology, 4(4) :58, 2013b. 2 and 10 ».
- [18] I. Leang. Fusion en ligne d'algorithmes de suivi visuel d'objet. « Thèse de doctorat, Université Pierre et Marie Curie .Décembre 2016 ».
- [19] M. Pressigout and E. Marchand, "Real time planar structure tracking for visual servoing : a contour and texture approach," 2005 IEEE/RSJ International Conference on Intelligent Robots and Systems, Edmonton, Alta., 2005, pp. 251-256.
- [20] C. Huang, B. Wu, et R. Nevatia, "Robust object tracking by hierarchical association of detection responses", dans Computer Vision–ECCV 2008. Springer, 2008, pp. 788–801.
- [21] S. J. McKenna, S. Jabri, Z. Duric, A. Rosenfeld, et H. Wechsler, "Tracking groups of people", Computer Vision and Image Understanding, vol. 80, no. 1, pp. 42–56, 2000.
- [22] C. Harris, et M. Stephens, (1988). A combined corner and edge detector. In Alvey vision conference, volume 15, page 50. Citeseer.
- [23]D.Lowe, G. (2004). Distinctive image features from scale-invariant keypoints.International journal of computer vision, 60(2) :91–110.
- [24] H. Bay, A. Ess, T. Tuytelaars, et L. Van Gool, (2008). « Speeded-up robust features (surf). Computer vision and image understanding, 110(3) :346–359 ».
- [25]T.Ljouad, "Suivi d'objets mobiles dans les séquences vidéo moyennant les métaheuristiques." (2015).
- [26] R. Sekkal, « Techniques visuelles pour la détection et le suivi d'objets 2D. Diss. INSA de Rennes » 2014.
- [27] M. Isard et A. Blake. « Condensation - Conditional Density Propagation for Visual Tracking. International Journal of Computer Vision » vol. 29, no. 1, pages 5–28, 1998.
- [28] K. Mikolajczyk et C. Schmid. « A Performance Evaluation of Local Descriptors. IEEE Transaction on Pattern Analysis and Machine Intelligence » vol. 27, no. 10, pages 1615–1630, Octobre 2005
- [29] D. Comaniciu, V.Ramesh, and P. Meer, (2000). « Real-time tracking of non-rigid objects using mean shift. In Computer Vision and Pattern Recognition », 2000. Proceedings. IEEE Conference on, volume 2, pages 142–149 vol.2. (Cité en pages 14 et 15.)

- [30] B. Lucas, D. Kanade, T. al. (1981). « An iterative image registration technique with an application to stereo vision » In IJCAI, volume 81, pages 674–679
- [31] Z. Kalal, K.Mikolajczyk, and J.Matas, (2012). « Tracking-learning-detection. Pattern Analysis and Machine Intelligence », IEEE Transactions on, 34(7) :1409–1422.
- [32] J. Matas, et T. Vojir, (2011). « Robustifying the flock of trackers. In 16th Computer Vision Winter Workshop ». Citeseer, page 91. Citese
- [33] R.E. Kalman. « A new approach to linear filtering and prediction problems. Journal of basic Engineering », vol. 82, no. Series D, pages 35-45, 1960.
- [34] D. Riahi, Suivi multi-objets par la détection : « Application à la vidéo surveillance. Diss ». Ecole Polytechnique, Montréal (Canada), 2016
- [35] N.G. Gordon, D.J. Salmond, et A.F. Smith, (1993). « Novel approach to nonlinear/nongaussian bayesian state estimation ». In IEE Proceedings F-Radar and Signal Processing, volume 140, pages 107–113. IET.
- [36] G. Kitagawa, (1996). « Monte carlo filter and smoother for non-gaussian nonlinear state space models ». Journal of computational and graphical statistics, 5(1) :1–25.
- [37] WU, Yi, LIM, Jongwoo, et YANG, Ming-Hsuan. Online object tracking : A benchmark. In : Proceedings of the IEEE conference on computer vision and pattern recognition. p. 2411-2418. 2013
- [38] X. Mei, H. Ling : « Robust visual tracking using l1 minimization ».In IEEE 12th International Conference on Computer Vision, pp. 1436–1443 (2009).
- [39] D.A Ross, J. Lim, R.S. Lin, M.H Yang, « Incremental learning for robust visual tracking ». Int. J. Comput. Vis. 77(1–3), 125–141(2008).
- [40]. S. Avidan, "Support vector tracking," IEEE Transactions On Pattern Analysis And Machine Intelligence, PAMI, 26(8) :1064-1072, 2004.
- [41] S. Hare, A. Saffari, et P.H Torr, (2011). Struck : « Structured output tracking with kernels ». In ICCV, 2011 IEEE International Conference on, pages 263–270.IEEE.
- [42] M. Chao, Y. Xiaokang , Z. Chongyang, and Y. Ming-Hsuan, "Learning A Temporally Invariant Representation for Visual Tracking," in IEEE International Conference on Image Processing (ICIP), 2015
- [43] B.D Lucas, T. Kanade, et al. (1981). « An iterative image registration technique with an application to stereo vision ». In IJCAI, volume 81, pages 674–679
- [44] T. Vojir, J. Noskova, et J. Matas, (2014). « Robust scale-adaptive meanshift for tracking ». Pattern Recognition Letters, 49:250 – 258.
- [45] Ning, Jifeng, et al. "Robust object tracking using joint color-texture histogram." International Journal of Pattern Recognition and Artificial Intelligence 23.07 (2009) : 1245-1263.

- [46] L. Si, Z. Tianzhu, C. Xiaochun, X. Changsheng : « Structural Correlation Filter for Robust Visual Tracking ». CVPR 2016 : 4312-4320
- [47] C. Ma, X. Yang, C. Zhang, and M.-H. Yang. « Long-term correlation tracking ». In CVPR, 2015.
- [48] D. S Bolme, J. R Beveridge, B. A Draper, et Y. M. Lui, (2010). « Visual object tracking using adaptive correlation filters ». In Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on, pages 2544–2550. IEEE
- [49] M. Danelljan, G. Häger, F. Khan, et M. Felsberg, (2014). « Accurate scale estimation for robust visual tracking ». In British Machine Vision Conference, Nottingham, September 1-5, 2014. BMVA Press
- [50] Ma, X. Yang, C. Zhang, and M.-H. Yang. Long-term correlation tracking In CVPR, 2015b.
- [51] A. Krizhevsky, I. Sutskever, and G. E. Hinton. ImageNet classification with deep convolutional neural networks. In NIPS, 2012.
- [52] R. Girshick, J. Donahue, T. Darrell, and J. Malik. « Rich feature hierarchies for accurate object detection and semantic segmentation ». In CVPR, 2014.
- [53] Taigman, M. Yang, M. Ranzato, and L. Wolf. DeepFace : « Closing the gap to human-level performance in face verification ». In CVPR, 2014.
- [54] N. Wang, S. Li, A. Gupta, and D.-Y. Yeung. « Transferring rich feature hierarchies for robust visual tracking ». arXiv preprint arXiv : 1501.04587, 2015b.
- [55] H. Nam and B. Han. « Learning multi-domain convolutional neural networks for visual tracking ». In CVPR, 2016.
- [56] L. Wang, W. Ouyang, X. Wang, and H. Lu. « Visual tracking with fully convolutional networks ». In ICCV, 2015a.
- [57] Fukunaga, Keinosuke, and L. Hostetler. "The estimation of the gradient of a density function, with applications in pattern recognition." IEEE Transactions on information theory 21.1 (1975) : 32-40
- [58] B.Soufian. Détection et suivi d'objets par l'approche "Mean Shift ".Thèse master, Université de Mohamed Khider Biskra .2016
- [59] Cheng, Yizong. "Mean shift, mode seeking, and clustering." IEEE transactions on pattern analysis and machine intelligence 17.8 (1995) : 790-799.
- [60] E. Choi and P. Hall, "Data Sharpening as a Prelude to Density Estimation," Biometrika, vol. 86, pp. 941-947, 1999.
- [61] C.K. Chu, I.K. Glad, F. Godtliebsen, and J. S. Maron, "Edge Preserving Smoothers for Image Processing," J. Am. Statistical Assoc., vol. 9S, pp. 526-541, 1998

- [62] D. Comaniciu, "Nonparametric robust Methods for Computer vision," PHD theses Dept of Electrical and Computer Eng., Rutgers Univ., 1999.
- [63] D. Comaniciu and P. Meer, "Robust Analysis of Leature Spaces: Color Image Segmentation," Proc. 1997 IPPP Conf. Computer Vision and Pattern Recognition, pp. 750-755, June 1997.
- [64] A.K. Jain, R.P.W. Duin, and J. Mao, "Statistical Pattern Recogntion: A Review," IPPP Trans. Pattern Analysis and Machine Intelligence, vol. 22, no. 1, pp. 4-S7, Jan. 2000.
- [65] L. Kauffman and P. Rousseeuw, Finding Groups in Data: An Introduction to Cluster Analysis. J. Wiley & Sons, 1990.
- [66] D.T. Kuan, A.A. Sagchuk, T.C. Strand, and P. Chavel, "Adaptive Noise Smoothing Lilter for Images with Signal Dependent Noise," IPPP Trans. Pattern Analysis and Machine Intelligence, vol. 7, no. 2, pp. 165-177, Mar. 1985.
- [67] Ning, Jifeng, et al. "Robust object tracking using joint color-texture histogram." International Journal of Pattern Recognition and Artificial Intelligence 23.07 (2009): 1245-1263.
- [68] J.Ning, L. Zhang, D. Zhang and W.Chengke. "Biometrics Research Center", Dept. of Computing, the Hong Kong Polytechnic Univ.2010.
- [69] Lim, H.Youn, and D-S. Kang. "Object tracking system using a VSW algorithm based on color and point features." EURASIP Journal on Advances in Signal Processing 2011.1 (2011): 60
- [70] V. KARAVASILIS, C. NIKOU, ET A. LIKAS, "Visual tracking by adaptive Kalman filtering and mean shift". In: Artificial Intelligence: Theories, Models and Applications. Springer Berlin Heidelberg p. 153-162, 2010.
- [71] Black, J. Michael and A- D. Jepson. "Eigentracking: Robust matching and tracking of articulated objects using a view-based representation." European conference on computer vision. Springer, Berlin, Heidelberg, 1996.
- [72] G. Vincent, « suivi d'objets d'intérêt dans une séquence d'image des point saillants aux mesures », thèse de doctorat, l'Université de Nice 30avril 2009