

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 : Télécommunication

Option : Réseaux et Télécommunication

Réf:.....

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

MASTER

Thème

*Suivi d'objets robuste en utilisant un
histogramme conjoint de couleur et de texture*

Présenté par :
Djenaidia Azeddine
Soutenu le : 24 Juin 2018

Devant le jury composé de :
Mr BAARIR Zine-Eddine
Mme MEDOUAKH Saadia
Mme FEDIAS Meriem

Prof
MAA
MCB

Président
Encadreur
Examineur

Année universitaire : 2017 / 2018

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 : Telecommunication
Option : Réseaux et Télécommunication

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

MASTER

Thème

*Suivi d'objets robuste en utilisant un histogramme
conjoint de couleur et de texture*

Présenté par :

DJENAIIDIA Azeddine

Avis favorable de l'encadreur :

MEDOUAKH Saadia

Avis favorable du Président du Jury

Mr BAARIR Zine-Eddine

Cachet et signature

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 : Télécommunication
Option : Réseaux et Télécommunication

Thème :

Suivi d'objets robuste en utilisant un histogramme conjoint de couleur et de texture

Proposé par : MEDOUAKH Saadia

Dirigé par : MEDOUAKH Saadia

RESUMES (Français et Arabe)

Dans le suivi d'objets, l'algorithme Mean Shift est largement utilisé en raison de sa rapidité, l'efficacité et la simplicité. Dans ce travail, nous présentons l'utilisation d'un histogramme conjoint de couleur et texture pour représenter le modèle de l'objet cible puis l'appliquer au cadre de Mean Shift. L'algorithme Mean Shift a été prouvé être robuste à une occlusion partielle, à l'échelle, la rotation et la déformation non-rigide de la cible. Cependant, l'utilisation d'un histogramme de couleur uniquement dans cet algorithme présente certains problèmes. Tout d'abord, l'information spatiale de l'objet cible est perdue. Deuxièmement, il a provoqué un suivi des objets de manière imprécise ou échoué facilement lorsque l'objet cible a apparence similaire à l'arrière-plan ou des autres objets. Pour surmonter ces limitations nous utilisons les caractéristiques de texture LBP qui reflètent la structure spatiale de l'objet combiné avec les caractéristiques de couleur HSV pour créer un histogramme pondéré conjoint HSV couleur - LBP texture de l'objet cible. Cet histogramme permet d'améliorer la précision et la robustesse de cet algorithme. Les algorithmes Mean shift et Mean shift avec LBP 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 qualitatif (visuel) et quantitatif (distance euclidienne et superposition) a été effectuée.

Mots clés :

Vision par ordinateur, Suivi d'objet, Mean shift, Histogramme conjoint couleur-texture, LBP.

المخلص

في تتبع الكائنات، يتم استخدام خوارزمية Mean Shift على نطاق واسع نظرًا لسرعتها وكفاءتها وبساطتها. في هذا العمل، نقدم استخدام المدرج التكراري المشترك للون والملمس لتمثيل نموذج الكائن المستهدف ثم تطبيقه في إطار Mean Shift. لقد أثبتت خوارزمية Mean Shift أنها قوية للانسداد الجزئي، التدرج، الدوران، والتشوه غير الثابت للهدف. ومع ذلك، استخدام الرسم البياني للألوان فقط في هذه الخوارزمية لديه بعض المشاكل. أولاً، يتم فقدان المعلومات المكانية للكائن الهدف. وثانياً، تسبب في تعقب الكائنات بطريقة غير دقيقة أو فشلها بسهولة عندما يكون الكائن المستهدف مشابهاً للخلفية أو لكائنات أخرى. للتغلب على هذه القيود، نستخدم مميزات نسيج LBP التي تعكس البنية المكانية للجسم مع خصائص الألوان HSV لإنشاء مدرج تكراري مشترك للون HSV والملمس LBP للكائن المستهدف. هذا المدرج التكراري يحسن دقة ومثانة هذه الخوارزمية. طبقت الخوارزميات Mean Shift و Mean Shift مع LBP على قاعدة البيانات OTB التي تحتوي على القيم الحقيقية² ثم تتم المقارنة المتبادلة فيما يتعلق بأداء كل منها في جهات النظر النوعية (البصرية) والكمية (المسافة الإقليدية والتداخل بين الأطر).

الكلمات المفتاحية:

رؤية الكمبيوتر، تتبع الهدف، Mean Shift، المدرج التكراري المشترك، LBP.

Dédicace

✚ *Je dédie ce mémoire :*

✚ *A mon père et ma mère avec toute ma reconnaissance,*

✚ *A mon frère et ma sœur*

✚ *A tous mes professeurs et enseignants*

✚ *A tous mes amis pour leurs soutiens et leurs encouragements.*

✚ *A toute personne ayant contribué à ce travail de près ou de loin.*

Azeddine

Remerciements

Avant tous, nous remercions Allah tout puissant qui nous donné la force et la volonté pour pouvoir finir ce mémoire de master.

*Nous tenons à remercier vivement notre encadreur **MEDOUAKH Saadia** d'avoir accepté de diriger ce projet et pour la confiance qu'il nous accordées, ses encouragements et ses précieux conseils.*

Nous vaudrions remercier également membre de jury de soutenance de ce mémoire :

***Prof. BAARIR Zine-Eddine**, d'avoir accepté la présider de jury de soutenance.*

***Dr. FEDIAS Meriem** pour avoir accepté d'examiner notre travail et faire partie du jury.*

Nous tenons à remercier, tous ceux qui nous enseignés durant toutes notre études et en particuliers nos enseignants à l'université de Mohamed Khider Biskra.

Nous tenons aussi à remercier tous les personnes qui nous ont encouragés pendant la réalisation de ce travail, famille, collègue, amis, sans exception.

Liste des tableaux

Chapitre 3 : Simulation, Résultats et discussions

Tableau 3.1 : Séquences d'images utilisées.....	47
Tableau 3.2 : Résultats obtenus de l'erreur de localisation du centre et du taux de recouvrement pour les différents histogrammes pondérés conjoints.....	51
Tableau 3.3 : Les moyennes des résultats obtenus de l'erreur de localisation du centre et du taux de recouvrement.....	59

Chapitre 1 : Généralités sur le suivi d'objets

Figure 1.1 : Exemple de suivi d'objets.....5

Figure 1.2: Processus de suivi (a) étape de représentation, (b) étape de localisation.....6

Figure 1.3: Exemple de domaine d'application.....6

Figure 1.4 : Exemple d'occultation : (a) : occultation partielle, (b) : occultation totale.....7

Figure 1.5 : Exemple de changement des conditions d'éclairage lors d'un suivi de visage sur la séquence vidéo David.....8

Figure 1.6 : Exemple de changement d'échelle.....8

Figure 1.7 : Exemples de Changement du modèle d'apparence dû aux rotations.....9

Figure 1.8 : Exemple de mouvement de camera.....9

Figure 1.9 : Exemple de mouvement rapide.....10

Figure 1.10 : Exemple de représentatif 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.....11

Figure 1.11 :Illustration des modèles utilisés pour la reconstruction L1.....19

Figure 1.12 : Le principe de tracker MIL.....20

Chapitre 2 : Suivi d'objets robuste via l'algorithme Mean shift en utilisant un histogramme conjoint de couleur et de texture

Figure 2.1 : l'itération du Processus Mean Shift24

Figure 2.2: Exemple de construction d'un histogramme pondéré par un Noyau gaussienne d'une image en niveau de gris. L'intensité est quantifiée en 4 niveaux.....26

Figure 2.3 .la similarité de méthode Mean shift.....28

Figure 2.4 : Construction d'un motif local binaire et calcul du code LBP.....32

Figure 2.5 : Représente un exemple d'image, LBP et l'histogramme de LBP.....33

Figure 2.6 : Exemples de l'opérateur LBP étendus : le circulaire voisinage (a) :(8, 1), (b) :(16, 2), (c) : (8, 2).....34

Figure 2.7 : Construction et uniformité d'un motif LBP. (a) le motif construit ici est non uniforme, (c) exemple de motifs respectivement uniforme et non uniforme.....35

Figure 2.8 : Exemple d'histogramme conjoint couleur HSV et texture LPB.....37

Chapitre 3 : Simulation, Résultats et discussions

Figure 3.1 : La structure de l’algorithme de suivi par Mean shift avec LBP.....41

Figure 3.2 : Exemple de la transformation d'espace couleur RGB vers HSV.....43

Figure 3.3. Exemple d’opération LBP8,1.....43

Figure 3.4 : Exemple sur l’histogramme conjoint 2D de composante V de l’espace HSV.....44

Figure 3.5 : Les résultats de l’opération sur l’image RGB (a) image en niveau de gris, (b) à chaque composante de l’image HSV.....45

Figure 3.6 : Noyau d’Epanechnikov.....46

Figure 3.7 : Exemple représenté l’image de poids.....46

Figure 3.8.Métriques de performance :(a) Taux de recouvrement, (b) Erreur de localisation du centre.....49

Figure 3.9 : Résultats de suivi sur les séquences, Divid3,Crossing,Couple, DragonBaby



 MS_LBP (niveau de gris), MS_LBP(niveau de gris de (V)),
 MS_LBP (niveau de gris de (S)), MS_LBP(niveau de gris de (H))50

Figure 3.10 : Résultats de suivi sur la séquence Divid3 avec Mean shift (rectangle jeune) et Mean Shift avec LBP (rectangle rouge).....52

Figure 3.11 : Comparaison quantitative avec Mean shift et Mean shift pour séquence Divd3 : (a) erreur de location du centre (b) Taux de recouvrement.....52

Figure 3.12 : Résultats de suivi sur la séquence Crossing avec Mean shift (rectangle jeune) et Mean Shift avec LBP (rectangle rouge).....53

Figure 3.13 : Comparaison quantitative avec Mean shift et Mean shift pour séquence Crossing: (a) erreur de location du centre (b) Taux de recouvrement.....54

Figure 3.14 : Résultats de suivi sur la séquence DragonBaby avec Mean shift (rectangle jeune)et Mean shift avec LBP (rectangle rouge).....54

Figure 3.15 : Comparaison quantitative avec Mean shift et Mean shift pour séquence DragonBaby : (a) erreur de location du centre (b) taux de recouvrement.....55

Figure 3.16 : Résultats de suivi sur la séquence Couple avec Mean shift (rectangle jeune) et Mean Shift avec LBP (rectangle rouge).....56

Figure 3.17 : Comparaison quantitative avec Mean shift et Mean shift pour séquence Couple : (a) erreur de location du centre (b) taux de recouvrement.....57

Liste des figures

Figure 3.18 : Résultats de suivi sur la séquence Blurcar4 avec Mean shift (rectangle jaune) et Mean Shift avec LBP (rectangle rouge).....	58
Figure 3.19 : Comparaison quantitative avec Mean shift et Mean shift avec LBP pour la séquence BlurCar 4: (a) erreur de location du centre (b) taux de recouvrement.....	58

Liste des abréviations

q : Distribution du modèle.

$P(y)$: Distribution du candidat.

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.

LPB :Local Binary Pattern.

MS : Mean shift.

MSLBP : Mean shift avec LBP.

HSV :Hue, Saturation, Value.

RGB : Red ,Green ,Blue.

HOG: Histogram of Oriented Gradients.

CamShift : Continuously Adaptive Mean Shift.

Résumé

Dans le suivi d'objets, l'algorithme Mean Shift est largement utilisé en raison de sa rapidité, l'efficacité et la simplicité. Dans ce travail, nous présentons l'utilisation d'un histogramme conjoint de couleur et texture pour représenter le modèle de l'objet cible puis l'appliquer au cadre de Mean Shift. L'algorithme Mean Shift a été prouvé être robuste à une occlusion partielle, à l'échelle, la rotation et la déformation non-rigide de la cible. Cependant, l'utilisation d'un histogramme de couleur uniquement dans cet algorithme présente certains problèmes. Tout d'abord, l'information spatiale de l'objet cible est perdue. Deuxièmement, il a provoqué un suivi des objets de manière imprécise ou échoué facilement lorsque l'objet cible a apparence similaire à l'arrière-plan ou des autres objets. Pour surmonter ces limitations nous utilisons les caractéristiques de texture LBP qui reflètent la structure spatiale de l'objet combiné avec les caractéristiques de couleur HSV pour créer un histogramme pondéré conjoint HSV couleur - LBP texture de l'objet cible. Cet histogramme permet d'améliorer la précision et la robustesse de cet algorithme. Les algorithmes Mean shift et Mean shift avec LBP 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 qualitatif (visuel) et quantitatif (distance euclidienne et superposition) a été effectuée.

Mots clés :

Vision par ordinateur, Suivi d'objet, Mean shift, Histogramme conjoint couleur-texture, LBP.

المخلص

في تتبع الكائنات، يتم استخدام خوارزمية Mean Shift على نطاق واسع نظرًا لسرعتها وكفاءتها وبساطتها. في هذا العمل، نقدم استخدام المدرج التكراري المشترك للون والملمس لتمثيل نموذج الكائن المستهدف ثم تطبيقه في إطار Mean Shift. لقد أثبتت خوارزمية Mean Shift أنها قوية للانسداد الجزئي، التدرج، الدوران، والتشوه غير الثابت للهدف. ومع ذلك، استخدام الرسم البياني للألوان فقط في هذه الخوارزمية لديه بعض المشاكل. أولاً، يتم فقدان المعلومات المكانية للكائن الهدف. وثانياً، تسبب في تعقب الكائنات بطريقة غير دقيقة أو فشلها بسهولة عندما يكون الكائن المستهدف مشابها للخلفية أو لكائنات أخرى. للتغلب على هذه القيود، نستخدم مميزات نسيج LBP التي تعكس البنية المكانية للجسم مع خصائص الألوان HSV لإنشاء مدرج تكراري مشترك للون HSV والملمس LBP للكائن المستهدف. هذا المدرج التكراري يحسن دقة ومثانة هذه الخوارزمية. طبقت الخوارزميات Mean Shift و Mean Shift مع LBP على قاعدة البيانات OTB التي تحتوي على القيم الحقيقية. ثم تتم المقارنة المتبادلة فيما يتعلق بأداء كل منها في وجهات النظر النوعية (البصرية) والكمية (المسافة الإقليدية والتداخل بين الأطر).

الكلمات المفتاحية:

رؤية الكمبيوتر، تتبع الهدف، Mean Shift، المدرج التكراري المشترك، LBP.

Sommaire

Introduction générale	1
Chapitre 1 : Généralités sur le suivi d'objets	
1.1 Introduction.....	4
1.2 Le suivi d'objet	4
1.2.1 Principe des algorithmes de suivi.....	5
1.2.2 Domaine d'application.....	6
1.2.3 Les défis du suivi.....	7
1.2.3.1 Occultations.....	7
1.2.3.2 Changement d'illumination.....	8
1.2.3.3 Changement d'échelle.....	8
1.2.3.4 Changement du modèle d'apparence.....	9
1.2.3.5 Mouvements de caméra.....	9
1.2.3.6 Objet en mouvement rapide.....	10
1.3 Représentation d'objet.....	10
1.3.1 Représentation de la forme d'un objet.....	10
1.3.2 Représentation de l'apparence d'un objet.....	12
1.4 Les caractéristiques.....	13
1.4.1 Couleur.....	13
1.4.2 Texture.....	14
1.4.3 Gradient	14
1.4.4 Deep features.....	15
1.5 Localisation de l'objet.....	15
1.5.1 Méthodes d'optimisation.....	15
1.5.2 Filtrage de Kalman.....	16
1.5.3 Le filtrage particulière.....	16
1.6 Méthodes de suivi d'objets.....	17
1.6.1 Méthodes basées sur les modèles génératifs.....	18
1.6.2 Méthodes basées sur les modèles discriminatifs.....	19
1.7 Conclusion.....	21

Chapitre 2 : Suivi d'objets robuste via l'algorithme Mean shift en utilisant un histogramme conjoint de couleur et de texture

2.1 Introduction.....	23
2.2 Suivi par Mean shift.....	23
2.3 Procédure du tracker Mean shift.....	24
2.3.1 La représentation de la cible modèle et candidats.....	25
2.3.1.1 Histogramme pondéré.....	26
2.3.2 Mesure de la similarité entre les distributions.....	27
2.3.3 Localisation de cible.....	29
2.4 Suivi d'objet par Mean shift en utilisant un histogramme conjoint couleur-texture.....	31
2.4.1 Modélisation de l'apparence de l'objet.....	31
2.4.2 Motif Binaire Local (LBP Local Binary Pattern).....	32
2.4.2.1 Principe de L'opérateur LBP basique.....	32
2.4.2.2 Dérivation de l'opérateur (LBP étendus).....	33
2.4.2.3 LBP invariant par rotation.....	34
2.4.3 La représentation de l'objet cible avec l'histogramme conjoint HSV couleur-LBP texture.....	36
2.5 Conclusion.....	38

Chapitre 3 : Simulation, Résultats et discussions

3.1 Introduction.....	40
3.2 Structure de l'algorithme.....	40
3.2.1 Détection d'objet.....	42
3.2.2 Transformation RGB vers HSV.....	42
3.2.3 Calcule l'image LBP.....	43
3.2.4 La distribution de la cible modèle et candidats.....	44
3.2.4.1 Histogramme conjoint HSV couleur-LBP texture.....	44
3.2.4.2 Noyau d'Epanechnikov.....	45
3.2.4.3 Poids et le nouveau centre.....	46
3.3 Environnement de travail.....	46
3.3.1 Environnement matériel.....	46
3.3.2 Langage de programmation.....	47

Sommaire

3.4 Bases de données.....	47
3.5 Mesure des performances d'un système de suivi.....	48
3.5.1 Taux de recouvrement.....	48
3.5.2 Erreur de localisation du centre.....	48
3.6 Résultats Expérimentaux.....	49
3.6.1 Les résultats des évaluations de tracker Mean shift avec des différentes LBP images.....	49
3.6.2 Comparaison entre Mean shift et Mean Shift avec LBP.....	51
3.6.2.1 La séquence Divid3.....	51
3.6.2.2 La séquence Crossing.....	53
3.6.2.3 La séquence DraganBaby.....	54
3.6.2.4 La séquence Couple.....	55
3.6.2.5 La séquence BlurCar 4.....	57
3.6.2.6 Comparaison entre les moyennes d'erreur de localisation du centre et le taux de recouvrement pour Mean Shift et Mean shift avec LBP.....	59
3.7 Conclusion.....	60
Conclusion générale	62
Bibliographie	64

Introduction générale

Introduction générale

Le suivi d'objet dans une séquence d'images, en vision par ordinateur est, depuis ces dernières années, un thème de recherche très actif. Le suivi d'objets s'est révélé être un outil indispensable pour des applications aussi diverses que la vidéosurveillance, l'interaction homme machine, l'analyse de séquences sportives et la compression vidéo, ...etc. Il joue également un rôle fondamental dans de nombreuses tâches de vision informatique de haut niveau, par exemple, la reconnaissance d'action et la compréhension vidéo.

Le suivi d'objets en mouvement est parmi les problèmes les plus étudiés ces dernières années, il est à la base de tout système de vision [1]. Différentes perturbations visuelles sont à l'origine des dérives telles que les changements d'apparence de l'objet, les changements d'illumination de la scène, le contraste dans l'image, les mouvements de caméra, les occultations, etc. [2].

Comme l'un des sujets de recherche les plus importants et les plus difficiles dans la vision par ordinateur, le suivi des objets visuels vise « le problème d'estimer la trajectoire d'un objet dans le plan de l'image alors qu'il se déplace autour d'une scène » [3]. De nombreux algorithmes de suivi ont été proposés dans la littérature [4][5][6][7]. 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. Bien que de nombreux algorithmes aient été proposés au cours des dernières décennies [8], il est encore difficile de construire un système de suivi robuste et efficace pour faire face aux tous défis.

Parmi les différents algorithmes de suivi proposés, l'algorithme Mean shift [4] est récemment devenu populaire en raison de sa simplicité, sa rapidité et son efficacité. Mean shift est une méthode de suivi basée sur le noyau qui utilise des modèles d'apparence basés sur la densité de couleur de l'objet pour représenter les cibles. La méthode suit les cibles en trouvant le modèle de distribution le plus similaire dans une séquence de trames avec son modèle d'échantillon par recherche itérative [9]. Le tracker Mean shift repose sur l'histogramme pondéré de couleur afin de modéliser l'objet cible. Cependant, l'utilisation d'histogramme de couleurs seulement dans ce tracker pose certains problèmes. Premièrement, l'information spatiale de la cible est perdue. Deuxièmement, lorsque la cible a une apparence similaire à l'arrière-plan, l'histogramme des couleurs deviendra invalide pour les distinguer [10]. Pour cette raison, plusieurs chercheurs [10] [11] [12] affirment que la combinaison de plusieurs caractéristiques visuelles peut améliorer la convergence de cet algorithme dans des

conditions complexes, mais le choix des caractéristiques reste un problème au cœur de la recherche.

Dans ce mémoire nous nous intéressons à l'étude de l'algorithme de suivi d'objet Mean shift avec LBP qui se base sur l'histogramme conjoint de couleur et de texture. Cet algorithme utilise les caractéristiques de texture LBP qui reflètent la structure spatiale de l'objet combiné avec les caractéristiques de couleur HSV pour créer un histogramme pondéré conjoint de couleur HSV-LBP de manière plus distinctive et efficace. Aussi, une étude comparative entre les algorithmes de suivi d'objets Mean shift et Mean shift avec LBP, puis entre les différents histogrammes conjoints qui utilisent les images LBP obtenus par les composantes H, S et V et l'image en niveau de gris.

Nous décomposons ce travail en trois chapitres :

Dans le premier chapitre, nous décrivons dans un premier temps les difficultés que les algorithmes de suivi sont susceptibles de rencontrer, puis nous présentons l'état de l'art des méthodes de suivi d'objets dans une séquence d'images. Ces méthodes se basent sur le modèle d'apparence d'objet cible, elles sont classées en deux catégories : les méthodes génératives et les méthodes discriminatives.

Le deuxième chapitre est consacré à l'étude un nouvel algorithme de suivi d'objet en utilisant un histogramme conjoint couleur-texture pour représenter l'objet cible, puis en l'appliquant au cadre de Mean shift pour améliorer sa précision et sa robustesse. Puis nous présentons l'opération LBP basique et la méthode de combinaison entre les caractéristiques de couleur HSV et de texture LBP pour créer l'histogramme conjoint.

Enfin, le chapitre trois présentera les résultats expérimentaux obtenus par les algorithmes Mean shift avec LBP et Mean shift, suivi d'une discussion avec interprétation des résultats.

Nous terminerons ce mémoire par une conclusion générale.

Chapitre 1 :
Généralités sur le suivi
d'objets

1.1 Introduction

Le suivi d'objet dans des séquences d'images est, depuis ces dernières décennies, un thème de recherche très actif en vision par ordinateur pour son importance dans différents domaines de la vie, avec le développement des ordinateurs et des caméras haute qualité et le besoin croissant d'analyser les vidéos engendré un grand intérêt pour les algorithmes de suivi. Le suivi consiste à localiser un objet sur chaque trame d'une séquence vidéo en utilisant le modèle d'apparence d'objet cible qui modélise par des caractéristiques extraites d'objets, tel que couleur, texture..., etc. Dans la littérature, les méthodes de suivi d'objet sont en général classées selon la représentation utilisée pour décrire les objets cibles ou selon le modèle d'apparence d'objets.

Dans ce chapitre, nous présentons tout d'abord les difficultés que les algorithmes de suivi sont susceptibles de rencontrer. Ensuite, nous décrivons la représentation de la forme et de l'apparence de l'objet en utilisant les différentes caractéristiques présentées. Aussi, les techniques de la localisation d'objets sont décrites. Enfin, nous présentons l'état de l'art des méthodes de suivi d'objets dans une séquence d'images. Ces méthodes se basent sur le modèle d'apparence d'objet cible, elles sont classées en deux catégories : les méthodes génératives et les méthodes discriminatives.

1.2 Le suivi d'objet

Le suivi d'objet dans sa forme la plus simple, est l'opération consistant à estimer l'état d'objets à travers le temps [13][14]. 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. Le procédé de la 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. [15]. L'objet d'intérêt 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 [16].

Le but du suivi est de générer la trajectoire d'un objet dans le temps en déterminant sa position dans chaque image d'une séquence vidéo. Le suivi peut aussi fournir la région de

l'image occupée par l'objet à chaque instant [17]. Le problème du suivi d'objet peut s'exprimer en termes de détection de l'objet au sein de chaque image [18].

1.2.1 Principe des algorithmes de suivi

Le but d'un algorithme de suivi d'objet est d'estimer, dans chaque image de la séquence, les paramètres de l'objet qui peuvent être divers, mais comprennent principalement une composante géométrique, indiquant la position dans l'image du centre de l'objet, comme illustré dans la figure (1.1) [4].

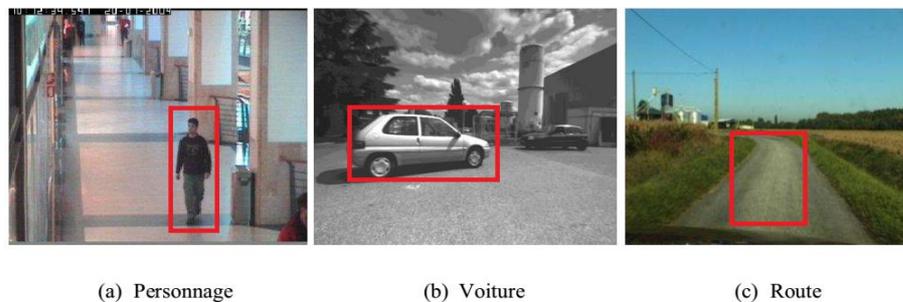


Figure 1.1 : Exemple de suivi d'objets.

D'un point de vue fonctionnel, le suivi d'objet se décompose en deux étapes principales : La représentation et la localisation [15].

- **L'étape de représentation** consiste à associer à l'objet suivi des caractéristiques de forme et/ou d'apparence permettant de le reconnaître dans des images successives. Les caractéristiques sélectionnées sont destinées à former le modèle de référence ; il s'agit d'une description numérique ou symbolique qui permet de caractériser l'apparence visuelle de l'objet dans l'image ou la séquence. Ceci peut être accompli de façon classique par l'initialisation du modèle sur une image de référence (figure 1.2.a) bien que certaines méthodes prennent en compte plusieurs images de référence afin de transcrire une plus grande variabilité dans l'apparence de l'objet [19][15].
- **L'étape de la localisation** est le procédé numérique réalisant le suivi qui a pour but de fournir la position de l'objet dans l'image courante de la séquence, sur la base de son modèle de représentation et de son état dans l'image précédente (figure 1.2.b)[15].

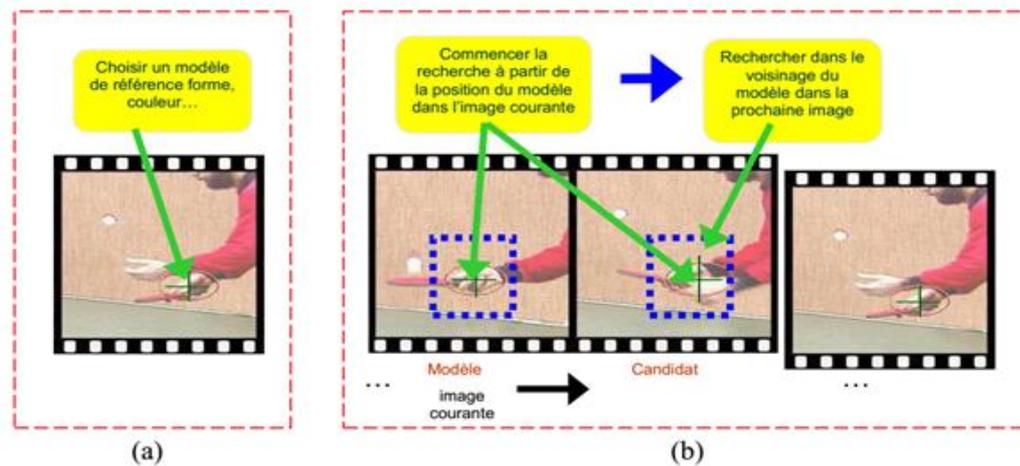


Figure 1.2: Processus de suivi (a) étape de représentation, (b) étape de localisation.

1.2.2 Domaine d'application

Le suivi visuel d'objets dans des séquences d'images est devenu une des tâches les plus utilisées, parmi celles développées pour la vision par ordinateur. Il existe d'innombrables techniques mises en œuvre pour réaliser le suivi d'objet [20], appliquées à des domaines aussi [21] comme la vidéosurveillance (détection, suivi, reconnaissance du comportement de personnes, d'intrus), la vidéoconférence (suivi des interlocuteurs), la robotique, [22] l'imagerie médicale, interaction homme-machine, [21] militaire (suivi des cibles ou guidage de missiles), surveillance de la circulation des véhicules (voir la figure 1.3).



Figure 1.3: Exemple de domaine d'application.

1.2.3 Les défis du suivi

Malgré les avancées importantes réalisées en vision par ordinateur (notamment dans la modélisation des régions d'intérêt, le calcul des caractéristiques visuelles, la modélisation des mouvements et la détection d'objets), la conception d'un modèle d'apparence robuste et d'une stratégie de recherche efficace sans connaissance à priori du modèle d'apparence demeure un défi de taille. Les difficultés émanent à la fois des environnements non contraints et des apparences et mouvements complexes des objets suivis [23]. Les difficultés du suivi d'objets dépendent de plusieurs facteurs relatifs aux données ou à l'application (figure 3)[22][24][25]:

1.2.3.1 Occultations

Occultations, l'observation d'un objet peut être cachée par un autre objet ou un autre composant de la scène. Deux types d'occultations dans une scènes ont possibles :

(a) Occultation partielle, quand un objet déplace devant un mur, une table et d'autre objet statiques. Ce type d'occultation affecte uniquement une petite partie de l'observation de l'objet d'intérêt (figure 1.4 (a)).

(b) Occultation totale affecte toute la région de l'objet d'intérêt où l'observation de l'objet disparaître complètement certain temps (figure 1.4 :(b) [26].



(a)

(b)

Figure 1.4 : Exemple d'occultation : (a) : occultation partielle, (b) : occultation totale.

1.2.3.2 Changement d'illumination

Les images enregistrées par une caméra dépendent du contenu physique de la scène aussi bien que des conditions d'éclairage. Où la signature visuelle d'un objet change avec le changement de l'environnement de capture, par exemple : la couleur d'un objet dans une salle est totalement différente de la couleur du même objet à l'extérieur de cette salle. [22].



Figure 1.5 : Exemple de changement des conditions d'éclairage lors d'un suivi de visage sur la séquence vidéo David.

1.2.3.3 Changement d'échelle

Cette transformation change la taille de l'objet recherché et donc éventuellement du patron associé. La majeure partie des applications ne traitent pas ce cas de figure et considèrent une fenêtre de taille fixe. Le suivi est alors moins précis (figure 1.6) [16].

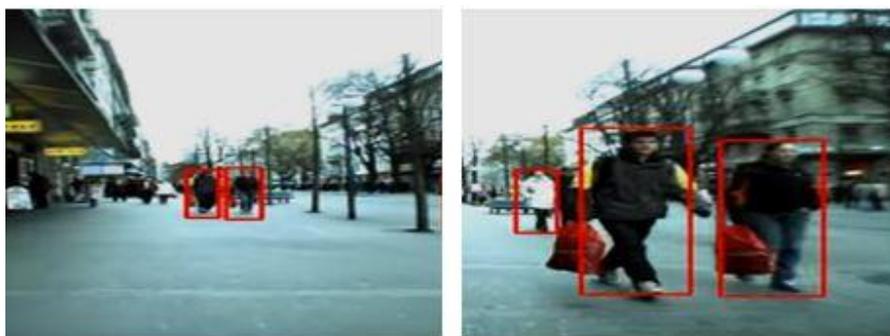


Figure 1.6 : Exemple de changement d'échelle.

1.2.3.4 Changement du modèle d'apparence

Les objets non rigides peuvent subir des changements au niveau de leurs apparences. En fait, les objets cibles sont généralement représentés par des formes géométriques avec une orientation initiale. Dans le monde réel, ces objets peuvent subir des variations de rotations importantes soit sur le plan de l'image ou des rotations sur le plan réel tridimensionnel (par exemple, une personne vue de face initialement après une rotation, on aura sa vue de dos) [27].



Figure 1.7 : Exemples de Changement du modèle d'apparence dû aux rotations.

1.2.3.5 Mouvements de caméra

Certains systèmes de suivi supposent le statisme de la caméra pour détecter les objets en mouvement [18,21]. En effet, il est plus facile de modéliser l'arrière-plan lorsque la caméra est fixe. Les régions de l'image où des changements se produisent sont ainsi extraites et considérées comme des cibles potentielles (car ne suivant pas le modèle de l'arrière-plan). Cependant, la contrainte de la caméra stationnaire n'est souvent pas valide. Par exemple, un suivi hors-ligne peut être effectué sur une séquence vidéo enregistrée par une caméra déplacée et orientée par un opérateur humain [23].



Figure 1.8 : Exemple de mouvement de camera.

1.2.3.6 Objet en mouvement rapide

Il résulte d'un mouvement rapide le phénomène de «flou de bougé » gommant les contours et diluant les couleurs et textures. Les applications se basant sur ces traits seront donc pénalisées. De plus, lors d'un mouvement rapide, un suivi imprécis perdra plus facilement l'objet (figure 1.9) [16].



Figure1.9 : Exemple de mouvement rapide.

1.3 Représentation d'objet

Le suivi d'objets consiste à associer les objets détectés dans l'image courante avec ceux détectés aux images précédentes. Il s'agit de maintenir l'identité des objets et l'évolution temporelle de leurs positions (ou d'une autre caractéristique). Ce problème peut être vu comme un problème de localisation spatiale et temporelle des objets présents dans la scène. De nombreuses approches de suivi d'objets se basent sur l'apparence d'un objet [29].

Dans le suivi d'objets, un objet est décrit par la forme et les caractéristiques qui le représentent. La plupart des méthodes de suivi d'objet diffèrent dans la manière de la représentation en termes de forme et l'apparence (voir la figure 1 .10) [15].

1.3.1 Représentation de la forme d'un objet

- **Points** Un objet peut être représenté par un point (figure 1.10 (a ,b)) [28][32]. Il peut s'agir par exemple de son centre de masse, du centre de sa boîte englobant, ou tout autre point caractéristique de la forme. Il s'agit d'une représentation simple de la localisation 2D (ou 3D) de l'objet. Cette représentation se généralise à un ensemble de points auxquels peuvent être associés des descripteurs locaux de couleur, de texture ou de mouvement.[3][17].
- **Formes primitive géométrique** L'objet peut être représenté par une forme rectangulaire ou elliptique (figure 1.10 (c,d)) [30] ,etc. .Le mouvement d'objets pour de telles

représentations est généralement modélisé par une transformation de translation, affine ou projective (homographie). Bien que les formes primitives géométriques soient plus appropriées pour représenter de simples objets rigides, ils sont également utilisés pour le suivi d'objets non rigides [31].

- **Contours** La représentation d'un objet par son contour permet une description plus complète de la forme d'un objet (figure 1.10 (g, h)). 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 silhouette de l'objet et peut être utilisée conjointement à l'information de contour pour le suivi d'objets [3], [29].
- **Squelette** Le squelette d'un objet peut être extrait pour caractériser la forme d'un objet ou d'une forme géométrique (figure 1.10 (i)) [33]. Ce modèle est utilisé en tant que descripteur de forme pour la reconnaissance d'objets. Cette représentation peut être aussi bien utilisée sur des objets déformables que des objets rigides. Dans 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 [3], [29].

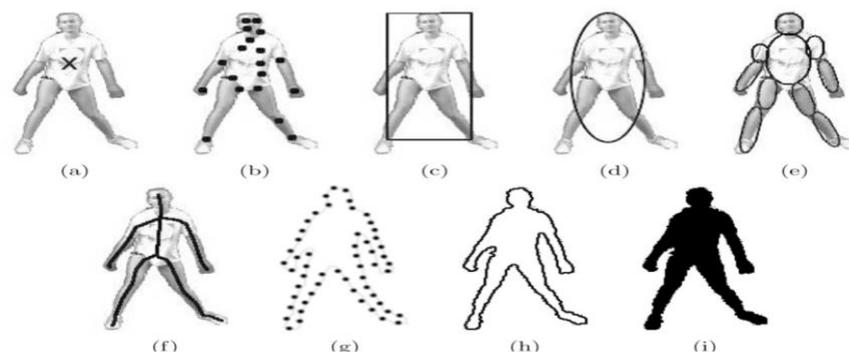


Figure 1.10 : Exemple de représentations de la forme d'objet. (a) : un point, (b) : points d'intérêt, (c) : rectangle, (d) : ellipse, (e) : multiple blocs, (f) : squelette, (g)-(h) contour, (i) : Silhouette[3].

1.3.2 Représentation de l'apparence d'un objet

La représentation des cibles basée sur l'apparence est définie comme un modèle de projection de la signature de l'apparence d'objet d'intérêt sur le plan de l'image[26]. Il ne s'agit pas de coder toute l'information véhiculée par l'objet mais de se concentrer sur l'information qui permet de définir un critère objectif efficace pour déterminer la position réelle de l'objet suivi dans l'image [15]. Les caractéristiques d'apparence sont généralement utilisées conjointement aux caractéristiques de formes dans l'objectif de compléter la représentation de l'objet à suivre[3][29]. Parmi les méthodes existantes, on retrouve les fonctions de densité de probabilité (estimateurs à noyaux, histogrammes ou modèle de mélange de gaussiennes) les patrons (template) ou encore les modèles dynamiques d'apparence (Active Appearance Models) [3][29].

- **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), 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.
- **Patrons (Template)**: cette représentation considère directement le signal lumineux dans les images. Dans le cadre du suivi d'objet, les méthodes basées sur un patron (Template based matching) 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.
- **Modèle dynamique d'apparence** : les modèles dynamiques d'apparence (Active appearance models) modélisent généralement les objets à travers des caractéristiques de formes et d'apparence des objets. Ces méthodes ont pour objectif de prendre en compte les variations d'apparence d'un objet. Ces variations peuvent être de deux types différents, intrinsèque ou extrinsèque. La déformation de la forme ou le changement de pose d'un objet est considéré comme une variation intrinsèque, tandis que les variations causées par le changement de luminosité, le mouvement de la caméra ou l'occlusion sont considérés comme une variation extrinsèque. Ces modèles nécessitent généralement une phase

d'apprentissage permettant d'apprendre à partir d'un jeu d'exemples la forme et l'apparence d'un objet.

1.4 Les caractéristiques

Habituellement extraites d'un Template objet, d'une partie d'objet ou du voisinage d'un point (descripteurs locaux), les caractéristiques visuelles ont été largement exploitées par les trackers. Elles fournissent une description plus robuste de l'apparence de l'objet par rapport à l'information d'intensité des pixels. Cette dernière étant sensible à la plupart des perturbations de l'image (bruit, illumination, réflectance), Cependant, certaines caractéristiques le sont plus ou moins : les caractéristiques de forme sont moins sensibles aux changements de luminosité et aux dégradations de l'image, tandis que les caractéristiques de couleur sont moins sensibles aux transformations géométriques. [2].

1.4.1 Couleur

La couleur, de par son pouvoir discriminant, est un des attributs les plus utilisés en représentation d'objets[33],Il existe de très nombreux espaces de couleurs possédant chacun leurs caractéristiques propres mais partageant le même principe de base[17],mais les auteurs se concentrent autour de quelques grands thèmes tels que l'identification de l'espace couleur le plus discriminant[33].Chaque espace créé une décomposition, selon ses propres paramètres, de toutes les couleurs en un ensemble de composantes dites primaires[17].Chaque point d'une image possédant une couleur spécifique, Le premier point concerne le choix de l'espace colorimétrique. La littérature dense sur cet aspect montre qu'il n'y a pas d'espace couleur idéal pour la représentation [15].Dans l'espace colorimétrique RVB, chaque couleur apparait comme une combinaison des couleurs primaires : rouge (R), vert (V) et bleu (B).Si certains auteurs [4] utilisent l'espace RVB, d'autres font le choix de projeter les images dans des espaces colorimétriques alternatifs caractérisés par des propriétés différentes. Ainsi, l'espace HSV séparant l'information relative à la teinte, la saturation et l'intensité est communément utilisé [29],la composante Hue de l'espace HSV offre une meilleure invariance. Les espaces définis comme perceptiblement uniformes tels que CIE Lab sont aussi considérés [33] [3] [15].Les espaces couleur sont sensibles au bruit [22].

1.4.2 Texture

La texture est une mesure de la variation d'intensité d'une surface qui quantifie propriétés telles que la douceur et la régularité. Comparé à la couleur, la texture nécessite une étape de traitement pour générer les descripteurs [3]. Les caractéristiques de texture sont moins sensibles aux changements d'éclairage par rapport à la couleur. Ce qui en fait une caractéristique essentielle à considérer dans les algorithmes de détection ou de reconnaissance, ce type de caractéristiques pour représenter un objet est courant et discriminant dans de nombreux cas. La texture présente de nombreuses qualités parmi lesquelles on peut citer sa faible sensibilité aux d'illumination. Elle peut être décrite en termes de direction, de grain, de contraste et changements bien d'autres encore. Cette diversité fait de la texture une facette particulièrement intéressante de l'image qui résulte en une multitude de représentations possibles [17]. Les modèles binaires locaux (LBP en anglais pour Local Binary Pattern) est base des caractéristique du texture.

1.4.3 Gradient

L'information de gradient spatial des objets a été largement utilisée pour caractériser la forme et le contour d'un objet. 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 (détecteur de Moravec, Harris). L'algorithme CONDENSATION (Conditionnal density propagation) consiste à initialiser une courbe spline sur les contours, et un filtre à particules est utilisé pour mettre à jour les paramètres de la courbe paramétrée. Des techniques de minimisation d'énergie le long des contours des objets ont également été proposées pour suivre les objets sous certaines contraintes de régularisation (snakes et contours actifs). Les histogrammes d'orientations de gradient (HOG) ont été utilisés en tant que primitives pour la construction de certains descripteurs. Par exemple, les descripteurs SIFT (Scale Invariant Feature Transform) qui combinent un détecteur et un descripteur invariants à l'échelle basés sur la distribution du gradient. [29].

1.4.4 Deep features

Récemment, l'utilisation des caractéristiques calculées par des réseaux profonds, appelées *deep features*, est devenue très populaire. Connus pour leur grande capacité de représentation des objets, les réseaux profonds issus de « l'apprentissage profond » (*Deep learning*) sont utilisés comme fournisseurs de caractéristiques pour le suivi d'objet [33, 35, 36.]. Le réseau est tout d'abord pré-entraîné sur une grande base d'images, par exemple ImageNet [37] (plusieurs millions d'images), puis utilisé en tant que boîte noire pour fournir les cartes de caractéristiques des images passées en entrée du réseau [2].

1.5 Localisation de l'objet

Estimer la meilleure localisation de l'objet dans chaque image (position, taille et parfois, angle et pose), soulève le problème du coût de calcul nécessaire pour le réaliser. Pour cela, différentes approches ont été utilisées pour estimer la position ou l'état de l'objet [2].

1.5.1 Méthodes d'optimisation

Parmi les algorithmes d'optimisation les plus utilisés dans la littérature, un processus d'optimisation récursif est adapté afin d'avoir les meilleures associations possibles entre les trajectoires qui existent déjà et l'ensemble des objets candidats à chaque instant. Cette optimisation est basée sur le principe de maximiser ou de minimiser d'une façon globale la fonction de coût total pour toutes les assignations possibles (ça dépend de la valeur de similarité ou de dissemblance utilisé) [27].

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 du flot optique [38] ou de l'algorithme Mean shift [39]. Le flot optique 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 [40] [94].

Le Mean shift [39] 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. CamShift (*Continuously Adaptive Mean Shift*) [41] est une extension du Mean shift avec une adaptation de taille de la fenêtre [2].

1.5.2 Filtrage de Kalman

Une autre stratégie de recherche est d'utiliser un modèle de mouvement pour restreindre l'espace de recherche de l'objet[2].Le filtrage de Kalman [42] a été utilisé de façon intensive dans les algorithmes de suivi d'objets [43], [44], [45].Il s'agit d'un filtre linéaire prédictif qui fournit une solution simple et efficace (dans des conditions particulières) pour estimer et prédire la position d'un objet[29]. Celui-ci réalise la prédiction d'état de l'objet (localisation) en supposant que l'état de l'objet est normalement distribué (gaussien) et le modèle de mouvement linéaire (par exemple, un modèle à vitesse constante [46] [2] Le filtre de Kalman se fait sur deux étapes : une étape de prédiction puis de correction de l'état. 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 correction consiste à corriger l'estimation de l'état à l'instant courant afin d'obtenir plus de précision. [27].

1.5.3 Le filtrage particulaire

Le filtrage particulaire (Ou méthodes de Monte-Carlo séquentielles) [47, 48]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 [2].. La méthode du Filtre de Particules consiste à estimer la localisation de l'objet cible en fonction de particules pondérées [27].L'objectif est d'approximer la distribution de filtrage de l'état à partir de N particules échantillonnées aléatoirement dans l'espace d'état. Chaque particule simule une trajectoire possible de l'objet, l'importance de chacune d'elle dépend de la vraisemblance observée (relative au modèle d'apparence).[2]. Le temps d'exécution peut être contrôlé par le nombre de particules générées à chaque itération. Cette propriété permet à la méthode de fonctionner en temps réel. Les résultats de cette méthode vont permettre une meilleure interprétation d'une séquence vidéo (exemple : identification des comportements des personnes, reconnaissance des activités, etc.). L'approche du Filtre de Particules est utilisée souvent pour les méthodes de suivi multi objets appelées *tracking-by-detection*. [27].

1.6 Méthodes de suivi d'objets

Les méthodes de suivi d'objets permettent d'estimer au fil du temps les paramètres d'un objet cible présent dans le champ de vision de la caméra et initialement détecte sur la première image de la séquence vidéo [22]. De nombreuses approches de suivi d'objets ont été proposées et la différence entre ces méthodes réside en partie dans le choix de la représentation et de la forme des objets, des caractéristiques de l'image utilisées, de la nature du mouvement estimé, etc. Ce choix dépend de l'application ainsi que de la vidéo traitée [50].

Plusieurs revues de l'état de l'art ont été publiées et rendent compte des approches de suivi proposées. Yilmaz et al, [3] proposent une classification des méthodes de suivi en fonction de la représentation d'objet utilisée : suivi par points, suivi par noyaux et suivi par silhouette.

- **Suivi par point.** L'objet cible est représenté par son centroïde ou par un ensemble de points caractéristiques.
- **Suivi par noyaux.** L'objet cible est représenté par des formes géométriques telles qu'un rectangle, une ellipse ou un cercle (peut être une forme qui entoure ou qui est à l'intérieur de l'objet cible).
- **Suivi par silhouette.** L'objet cible est représenté par son contour ou sa silhouette.

Récemment, les méthodes de suivi peuvent être classées selon le modèle d'apparence utilisé [51] : méthodes basées sur un modèle génératif, méthodes basées sur un modèle discriminatif [2].

- **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 autre terme, le suivi avec la modélisation générative se concentre généralement sur l'apprentissage d'un modèle pour représenter l'objet cible, puis l'utilise pour trouver la région la plus similaire dans les futurs cadres. De plus, la correspondance des méthodes peuvent être appliquées pour estimer le mouvement de l'objet cible entre les images consécutives, puis suivre l'objet au fil du temps [52] [27].En général, cette catégorie de méthode ne nécessite pas un grand ensemble de données pour l'apprentissage.
- **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. Les méthodes discriminantes de classification ont été exploitées pour gérer les changements d'apparence

lors du suivi visuel, où un classificateur est formé et mis à jour en ligne pour distinguer l'objet de l'arrière-plan. Cette méthode est également appelée par détection, dans lequel un objet cible identifié par l'utilisateur dans la première image est décrite par un ensemble de caractéristiques. Un ensemble distinct de fonctionnalités décrit l'arrière-plan, et un classificateur binaire sépare la cible de fond dans les trames réussis. Pour gérer l'apparence changements, le classificateur est mis à jour de façon incrémentielle au fil du temps [53].

Dans ce qui suit, nous allons détailler la classification des algorithmes de suivi qui se base sur le modèle d'apparence.

1.6.1 Méthodes basées sur les modèles génératifs

Mean shift (MS).L'algorithme de suivi Mean shift [4] a été proposé par Comaniciu et al, en 2003. Le Mean shift est une technique d'optimisation utilisée dans plusieurs applications. Il est sans doute la méthode la plus populaire qui utilise une représentation de l'apparence d'un objet sous forme d'histogrammes. Il s'agit d'une méthode non paramétrique qui maximise de façon itérative la similarité entre l'apparence d'un objet et celle d'un candidat autour d'une position estimée. Le Mean shift estime le taux de déplacement du noyau dans les séquences d'images. L'algorithme de suivi par Mean shift [39] se base essentiellement sur deux étapes : (1) la description des cibles par un modèle de couleur non-paramétrique (Histogramme). (2) la convergence du Mean shift, où une recherche itérative de la nouvelle position de l'objet dans l'image courante est effectuée à travers la maximisation d'une métrique de similarité entre deux histogrammes de couleurs (histogramme de modèle cible et histogramme de candidat) [29].

ASMS (Adaptive Scalemean-shift) .Vojir et al. [94] ont proposé l'algorithme ASMS pour traiter le problème de l'adaptation de l'échelle dans mean shift. ASMS propose deux améliorations : (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

KLT (Lucas-Kanade Tracker) [38] est la méthode la plus efficace qui basée sur la mise en correspondance de Templates. Le tracker trouve la correspondance affine transformé entre la zone de délimitation de la cible et les fenêtres candidates autour de la localisation précédente. La transformation affine est calculée par alignement d'image incrémentiel basé sur spatiotemporel dérivés et déformations capables de traiter avec échelle, rotation et traduction. La localisation de la cible est déterminée en mappant la position cible dans l'image précédente à la localisation dans

la trame actuelle en utilisant transformation affine calculée. Nous utilisons le calcul version efficace dans [54] [49].

Le tracker L1. Mei et Ling. [55] ont proposé une méthode de suivi robuste, appelée le tracker L1, en traitant 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 (Figure 1.11). Le tracker L1 nécessite des ressources de calcul élevé en raison de nombreux calculs de minimisation L1.

IVT (Incremental Visual Tracking) .Ross et al. [56] ont proposé un algorithme de suivi utilise un modèle sous-espace incrémental pour décrire l'objet cible afin d'adapter les changements d'apparence (IVT). Le tracker IVT 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. Un aspect important de la mise à jour concerne le choix des exemples utilisés dans la mise à jour et leur labellisation. La localisation de l'objet dans ce tracker est le seul exemple positif utilisé pour la mise à jour de leur modèle. Cette méthode est peu robuste en particulier lorsque la localisation de l'objet est imprécise [2].

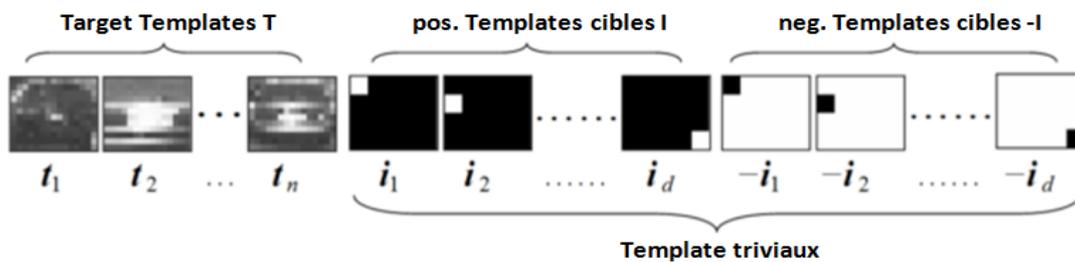


Figure 1.11 : Illustration des modèles utilisés pour la reconstruction L1[55].

1.6.2 Méthodes basées sur les modèles discriminatifs

Aviden (SVM) Avidan [57] a proposé une approche discriminative, où le classifieur SVM (machine à vecteurs supports) et un tracker basé sur le flot optique sont combinés pour suivre des véhicules sur de longues séquences vidéo. La classe des exemples positifs (objet) est peuplée d'images de l'objet à suivre alors que la classe des négatifs est peuplée d'images de fond (régions

pouvant être confondues avec l'objet). La méthode de suivi d'Avidan consiste alors à maximiser le score de classification du SVM sur des régions de l'image afin d'estimer la position de l'objet. Ce type d'approche intègre explicitement la connaissance du fond de la scène au sein du suivi.[3] [17][95] .

TLD (Tracking Learning and Detection). Kalal et al. [58] ont proposé un algorithme de suivi visuel robuste, appelé TLD. Ils décrivent le TLD comme la coopération d'un tracker basé flot optique et d'un détecteur appris en ligne, leur combinaison et mise à jour sont gérées par un autre modèle construit à partir de templates basse résolution de l'objet qui évalue la qualité de chaque estimée. Lorsque le tracker échoue, il est corrigé par le détecteur. Le détecteur est mis à jour lorsque le tracker et lui-même sont validés par le 3ième modèle. TLD utilise une fonction de redétection long-terme combinée à une fonction de suivi à très court-terme. TLD est robuste aux grands mouvements de caméra ou de l'objet [49] [2].

Le tracker MIL. Babenko et al. [59] ont proposé d'appliquer l'apprentissage en ligne à instance multiple (MIL) au lieu des méthodes traditionnelles d'apprentissage supervisé dans le suivi d'objets visuels, afin de permettre au classificateur de sélectionner un nombre d'échantillons positifs potentiels en fonction de son état actuel. Comparé à d'autres méthodes traditionnelles (voir la figure 1.12), en sélectionnant un seul échantillon positif imparfait ou en utilisant plusieurs échantillons bruyants positifs. 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. Le tracker MIL conserve suffisamment d'échantillons d'entraînement et tolère le bruit d'étiquetage lors de la mise à jour de son modèle. [52].

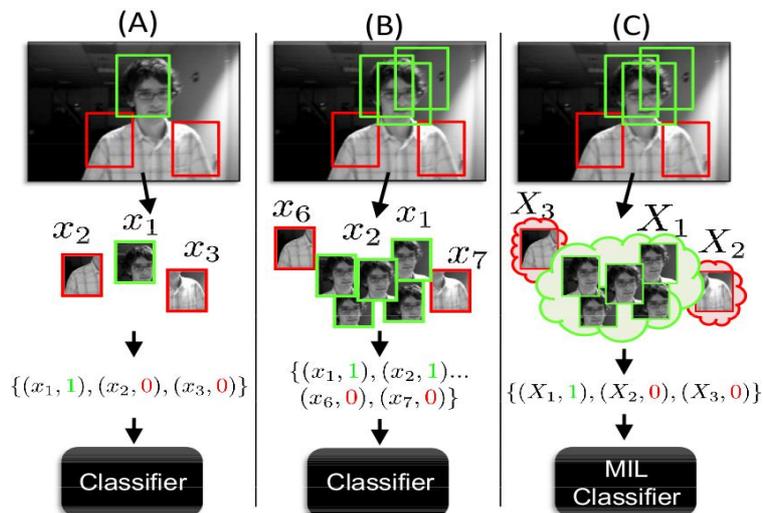


Figure 1.12 : Le principe de tracker MIL[52].

Le tracker Struck. Hare et al. [60] ont proposé le tracker Struck, qui est résout le problème du suivi d'objet par l'apprentissage en ligne d'un modèle SVM (en anglais Support Vector Machine) [61] à sorties structurées. Struck intègre les étapes de la localisation cible (ie, étiquetage des échantillons dans une région de recherche) 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.

Le filtre de corrélation (KCF). Récemment, les méthodes de suivi discriminatives basées sur le filtre de corrélation [62] [63] 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. SAMF et DSST sont des extensions des méthodes basées sur le filtre de corrélation [52].

1.7 Conclusion

Dans ce chapitre nous avons vu dans un premier temps le principe de suivi d'objet, la représentation de l'objet à la fois en termes de leurs formes et de leurs apparences, les caractéristiques visuelles et les différentes méthodes de la localisation d'objets. Tandis que dans un deuxième temps, nous avons présenté l'état de l'art des méthodes de suivi d'objets. La différence entre ces méthodes réside en partie dans le choix de la représentation et de la forme des objets, des caractéristiques de l'image utilisées, de la nature du mouvement estimé, etc. Ce choix dépend de l'application ainsi que de la vidéo traitée. Ces méthodes divisent en deux catégories basées sur le modèle d'apparence d'un objet : méthodes génératives utilisant un modèle d'apparence qui décrit seulement l'objet d'intérêt, visant à en trouver la meilleure correspondance sur l'image courante et méthodes discriminatives qui décrit l'objet par rapport à l'arrière-plan, en transformant le problème de suivi en un problème de classification binaire pour distinguer la cible de l'arrière-plan. 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'objets robuste via
l'algorithme Mean shift en
utilisant un histogramme
conjoint de couleur et de
texture*

2.1 Introduction

De nombreuses méthodes de suivi d'objet sont proposées dans la littérature. Le suivi d'objet demeure un problème non résolu à cause du nombre élevé de facteurs environnementaux. Il existe dans la littérature beaucoup d'algorithmes permettant le suivi des objets en mouvement dans une séquence vidéo. Parmi les différents algorithmes de suivi, l'algorithme Mean shift est l'un des algorithmes de suivi les plus robustes et les plus efficaces pour les applications en temps réel et les objets non rigides, en raison de sa simplicité et de sa robustesse. Mean shift n'est pas robuste aux grands mouvements, aux occultations importantes et l'algorithme peut échouer en présence d'un autre objet de couleurs similaires, en particulier lorsque l'objet destiné à être suivi et son arrière-plan respectif ont la même couleur [53], parce qu'il repose sur l'histogramme pondéré de couleur afin de modéliser l'objet cible.

Dans ce chapitre, nous avons présenté un nouvel algorithme de suivi d'objet en utilisant un histogramme conjoint couleur-texture pour représenter l'objet cible, puis en l'appliquant au cadre de Mean shift pour améliorer la précision de suivi et résoudre le problème de la similarité de l'objet et l'arrière-plan. Les caractéristiques de texture de l'objet sont également extraites en utilisant la technique du modèle binaire local (LBP) pour représenter l'objet.

2.2 Suivi par Mean shift

Le Mean shift est une technique non-paramétrique d'analyse d'un espace de caractéristiques. A l'origine, la procédure Mean shift est une procédure itérative de recherche de maximum local, basée sur une montée de gradient. Parmi les applications de la méthode Mean shift, nous pouvons citer la segmentation, la classification et le suivi [20]. Mean shift a d'abord été proposé par Fukunaga et Hostetler[65], adapté plus tard par Cheng [68] dans le but d'analyse d'images, et plus récemment prolongé par Comaniciu, Meer et Ramesh à des problèmes de vision bas niveau, y compris, la segmentation [69], adapté lissage [69] et suivi [4][67][70]

Dans l'approche de suivi Mean shift, Comaniciu et al, [4] utilisent un histogramme pondéré calculé sur une région rectangulaire pour représenter l'objet. Au lieu de réaliser une recherche exhaustive pour localiser l'objet. La procédure de suivi Mean shift maximise la similarité d'apparence itérativement en comparant les histogrammes de cible modèle et une fenêtre autour de la position hypothèse cible candidat.

La similarité entre deux histogrammes est définie en termes de coefficient de Bhattacharya. À chaque itération, le vecteur Mean shift est calculé tel que la similarité entre les histogrammes est augmentée [15]. Ce processus est répété jusqu'à ce que la convergence soit réalisée, qui s'effectue habituellement cinq à six itérations [64] [66] [4].

L'algorithme de suivi par Mean shift se base essentiellement sur deux étapes (figure 2.1) : (1) la description des cibles par un modèle de couleur non-paramétrique (Histogramme de couleur). (2) la convergence du Mean shift (localisation d'objet). Le suivi est effectué à partir de sa position initiale dans la première image. L'initialisation s'effectue pour détecter de l'objet d'intérêt est modélisé par forme géométrique, sur laquelle on calcule sa distribution de couleur. La distribution de couleur initiale est référencée en tant que modèle, et ensuite comparée à celle des sites candidats pour déterminer la position la plus probable dans l'image suivante. [15].

2.3 Procédure du tracker Mean shift

Mean shift s'appuie sur la dérivation du vecteur de densité de probabilité de l'histogramme de couleur dans son processus de suivi, après une initialisation de la position d'objet d'intérêt et la fenêtre correspondante, il calcule la distribution de couleur pondéré par un noyau dans ce rectangle ou ellipse généralement. Il évalue la similarité avec le modèle de référence. Un poids est associé a cette mesure de similarité et on peut alors calculer le vecteur Mean shift, qui a pour but de fournir la nouvelle position estimé de la cible dans l'image courante. On réitère ces étapes jusqu'à trouver le maximum local, comme illustré dans la figure 2.1. [4] [15] [71] [72] [78].

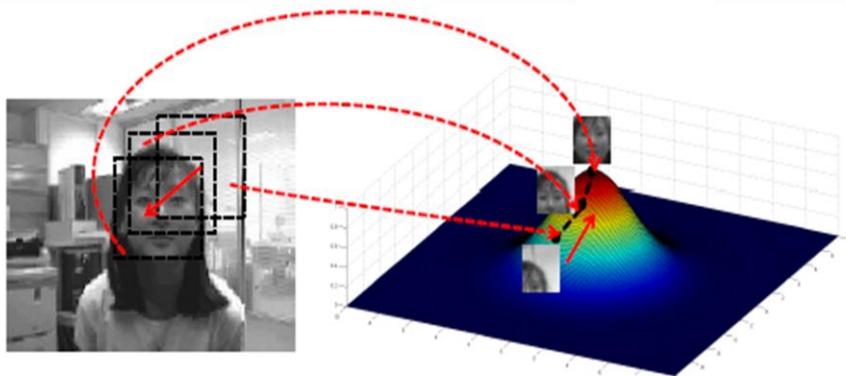


Figure 2.1 : l'itération du Processus Mean Shift [71]

2.3.1 La représentation de la cible modèle et candidats

Le tracker Mean shift [4] est un algorithme en essayant de localiser l'objet en trouvant le maximum local d'une fonction. Il prend généralement l'histogramme de couleur comme fonction de recherche et construit l'estimation de la densité du noyau du modèle cible en une fonction du noyau. Le modèle de la cible est représenté par une région rectangulaire de taille (h_x, h_y) dans l'image. 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 $k(x)$, $k : [0; \infty[\rightarrow R$, 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 $b : R^2 \rightarrow \{1 \dots m\}$ 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 bin su 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 bin su 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 $\sum_{u=1}^m \hat{p}_u = 1$.

2.3.1.1 Histogramme pondéré

Lorsque la région sélectionnée initiale contient quelques pixels de l'extérieur de l'objet (pixels d'arrière-plan), la distribution de probabilité de modèle cible sera influencée. 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. L'histogramme de modèle cible pour un bin u est donné par l'équation 2.1.

Le profil du noyau le plus simple utilisé pour générer l'histogramme pondéré est le noyau Epanechnikov. Ce noyau est une fonction de pondération utilisée dans les techniques d'estimation non-paramétrique. Le poids du noyau Epanechnikov pour la fenêtre d'objet en forme rectangulaire [4][73][6]. La fonction de densité (Epanechnikov) est donnée par :

$$K = \begin{cases} \frac{1}{2} c_d^{-1} (d + 2) (1 - \|x\|^2) & \text{if } \|x\| < 1 \\ 0 & \text{otherwise} \end{cases} \quad (2.3)$$

Où d est le nombre de dimensions, et c_d est le volume de la dimension.

L'exemple suivant (figure 2.2) 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.

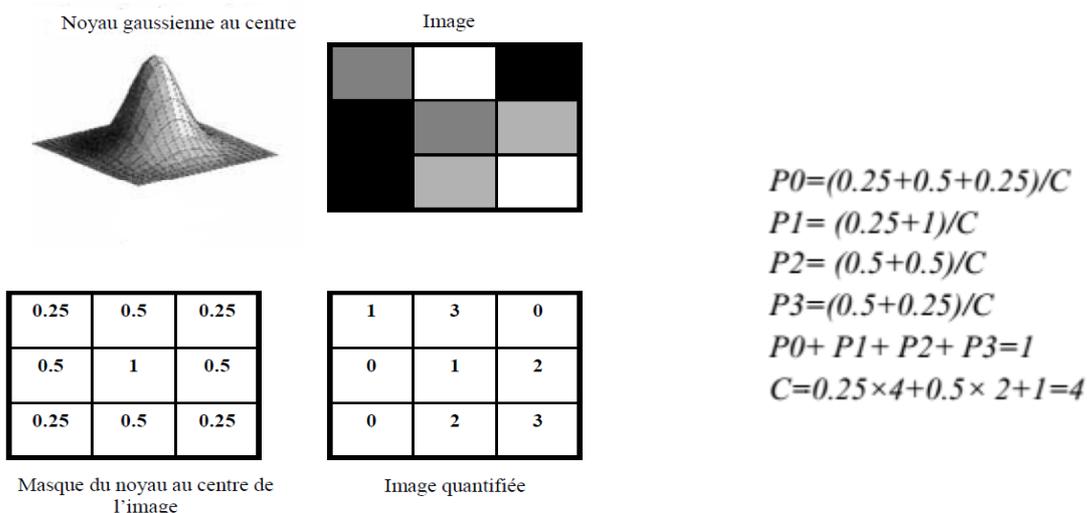


Figure 2.2: Exemple de construction d'un histogramme pondéré par un Noyau gaussienne d'une image en niveau de gris. L'intensité est quantifiée en 4 niveaux. [15]

2.3.2 Mesure de la similarité entre les distributions

Le système de suivi Mean shift maximise la similarité d'apparence itérativement en comparant les histogrammes de l'objet q et une fenêtre autour de la position hypothèse d'objet, $p(y)$. 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 la convergence soit réalisée. Le coefficient de Bhattacharyya [15][74][75][76][78] est défini par :

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

La distance de Bhattacharyya est définie par :

$$d(y) = \sum_{u=1}^m \sqrt{1 - \rho[\hat{p}_u(y), \hat{q}_u]} \quad (2.5)$$

Chapitre 2 : Suivi d'objets robuste via l'algorithme Mean shift en utilisant un histogramme conjoint de couleur et de texture

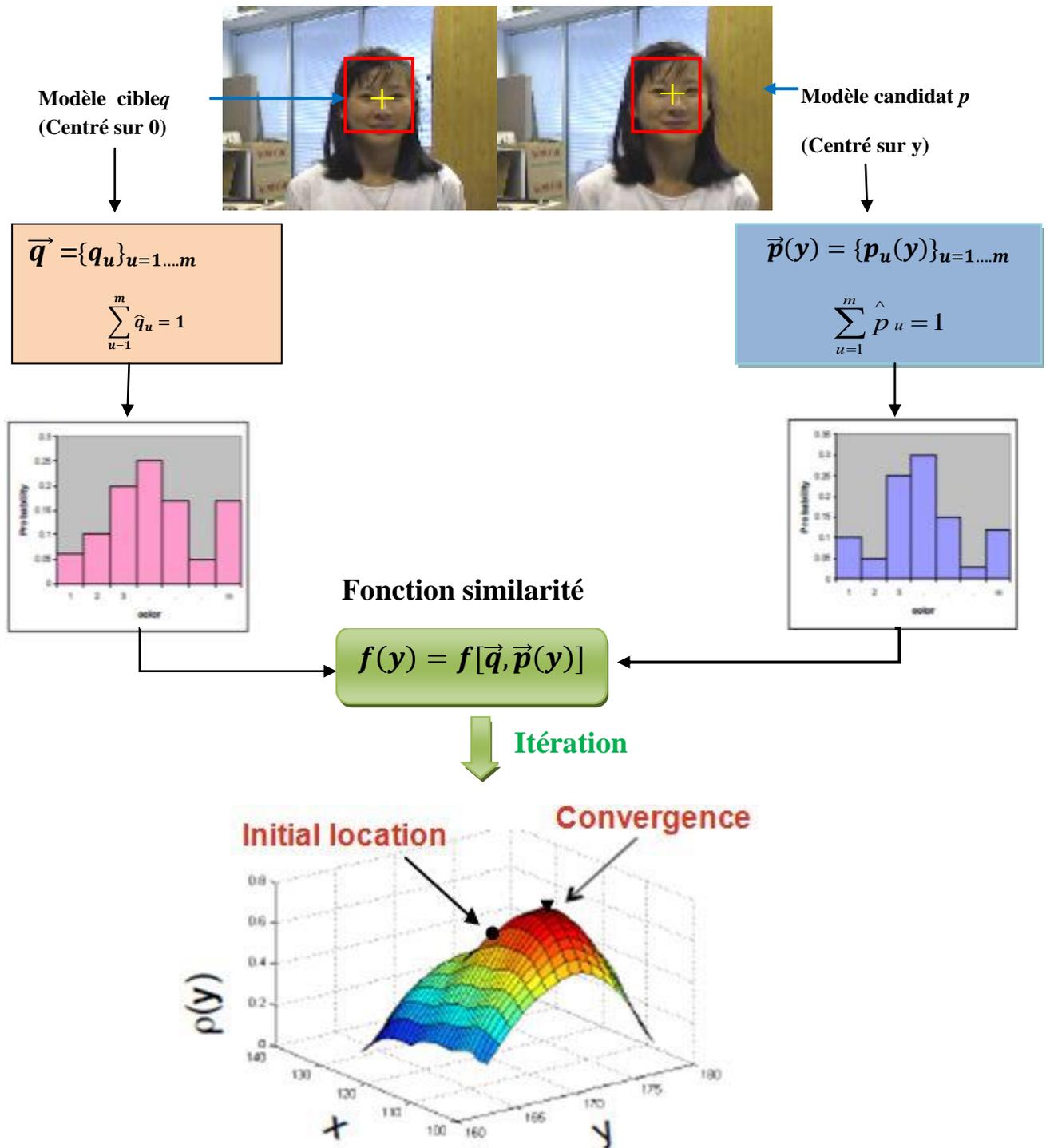


Figure 2.3 : la similarité de méthode Mean shift

2.3.3 Localisation de cible

La localisation y la plus probable de l'objet dans l'image courante est obtenue en minimisant la distance (2.5), ce qui est équivalent à maximiser le coefficient de Bhattacharyya $p(y)$ (2.4). La recherche de la nouvelle position de l'objet cible dans la trame courante commence à partir de la position de l'objet cible dans la trame précédente, c.-à-d. 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 (y_0) dans l'image courante: $\{p_u(y_0)\}_{u=1...m}$ et en utilisant l'approximation linéaire de Taylor du $\rho(y)$ autour de ces valeurs [4][77][78].

$$\rho(\hat{p}_u(y), \hat{q}_u) \approx \frac{1}{2} \sum_{u=1}^m \sqrt{\hat{p}_u(y) \hat{q}_u} + \frac{C}{2} \sum_{u=1}^{n_h} w_i k\left(\left\|\frac{y - x_i}{h}\right\|^2\right) \quad (2.6)$$

$$\text{Où : } w_i = \sum_{u=1}^m \sqrt{\frac{\hat{q}_u}{\hat{p}_u(y)}} \delta[b(x_i) - u] \quad (2.7)$$

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:

$$y_1 = \frac{\sum_{i=1}^n x_i w_i g\left(\|y_0 - x_i\|^2\right)}{\sum_{i=1}^n w_i g\left(\|y_0 - x_i\|^2\right)} \quad (2.8)$$

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

Chapitre 2 : Suivi d'objets robuste via l'algorithme Mean shift en utilisant un histogramme conjoint de couleur et de texture

La maximisation peut être effectuée efficacement en utilisant les itérations Mean shift, à l'aide de l'algorithme suivant [4][15][79][78]:

Algorithme de Mean shift

On dispose de l'estimation de la distribution du modèle $\langle q_u \rangle_{u=1\dots m}$ aussi que l'estimation de position y_0 de l'objet dans l'image précédente. Il faut ensuite

- 1) Calculer la distribution $p(y_0) = \langle p_u(y_0) \rangle_{u=1\dots m}$
- 2) Calculer $\rho(\hat{p}(y), \hat{q}) = \sum_{u=1}^m \sqrt{\hat{p}_u(y) \hat{q}_u}$
- 3) Calculer les poids $\{w_i\}_{i=1\dots m}$ avec $w_i = \sum_{u=1}^m \sqrt{\frac{\hat{q}_u}{\hat{p}_u(y)}} \delta[b(x_i) - u]$
- 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 $p(y_1) = \{p_u(y_1)\}_{u=1\dots m}$ puis évaluer

$$p(p(y_1, q)) = \sum_{u=1}^m \sqrt{p_u(y_1) q_u}$$

- 6) Tant que $p(p(y_1, q)) < p(p(y_0, q))$ faire $y_1 \leftarrow \frac{1}{2}(y_0 + y_1)$
 - 7) Si $\|y_1 - y_0\| < \varepsilon$ stop Sinon $y_0 \leftarrow y_1$ et retourner à l'étape 1.
-

2.4 Suivi d'objet par Mean shift en utilisant un histogramme conjoint couleur-texture

2.4.1 Modélisation de l'apparence de l'objet

L'étape de modélisation de l'objet cible est une étape cruciale pour le système de suivi multi objets. La modélisation est le processus d'extraction des caractéristiques discriminantes qui permet de décrire et de distinguer un objet d'intérêt dans la séquence vidéo [27]. 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 [2]. Diversifier les modèles d'apparence en combinant des caractéristiques de couleur, de forme ou de texture, augmente la représentativité des apparences des objets et du contexte, permettant de mieux gérer les variations d'apparence rencontrées au cours du suivi.

Plusieurs algorithmes de suivi prennent en compte les informations de couleur pour la représentation de l'objet cible .Le tracker Mean shift repose sur l'histogramme pondéré de couleur afin de modéliser l'objet cible. Malgré, il est robuste à une occultation partielle, à l'échelle, la rotation et la déformation non-rigide de la cible, mais il ne permet pas la conservation de l'information spatiale de chaque pixel dans l'image [10].Pour cette raison, plusieurs chercheurs [10][11][5][6][7] affirment que la combinaison de plusieurs caractéristiques visuelles peuvent améliorer la convergence de cet algorithme dans des conditions complexes, mais le choix des caractéristiques reste un problème au cœur de la recherche. Afin d'améliorer le tracker Mean shift plusieurs caractéristiques ont été utilisées pour qu'ils combinent avec un histogramme couleur; telles que les caractéristiques spatiotemporelles, les caractéristiques en gradient, les caractéristiques de texture [83].Les motifs de texture, qui reflètent la structure spatiale de l'objet sont des caractéristiques efficaces pour représenter et reconnaître les cibles. En raison de son le pouvoir discriminant et la simplicité de calcul du descripteur de texture LBP dans diverses applications de vision par ordinateur, on utilise un nouveau modèle d'apparence en combinant l'histogramme pondéré de couleur HSV et la texture LBP [10][5][6], a pour but de rendre le tracker Mean shift plus robuste et plus précision. Dans ce travail, nous utilisons le descripteur LBP pour représenter les caractéristiques de textures de l'objet cible, puis nous combinons ces caractéristiques avec l'histogramme pondéré de couleur pour créer un histogramme conjoint de couleur HSV-LBP de manière plus distinctive et efficace.

2.4.2 Motif Binaire Local (LBP Local Binary Pattern)

Le motif binaire local est un opérateur de texture simple mais très efficace utilisé pour extraire les caractéristiques les plus utiles à partir d'images (prétraité). L'opérateur LBP a été initialement proposé par T. Ojala et al [81] en 2002 afin d'exprimer la texture des patches de l'image [82]. Il a été largement appliqué avec divers algorithmes de systèmes de reconnaissance de visage comme une méthode d'extraction de caractéristiques locales. Ce descripteur se base sur la comparaison de niveau de luminance d'un pixel à analyser avec celles de ses proches voisins et la valeur qui le caractérise est calculé par leur somme pondérée par un certains poids (code binaire). Selon l'échelle du voisinage utilise, certaines zones d'intérêts tels des coins ou des bords peuvent être détectées par ce descripteur [1].

2.4.2.1 Principe de L'opérateur LBP basique

L'opérateur LBP de base attribue à chaque pixel un motif binaire. Cet opérateur fonctionne avec un bloc de 3 x 3 pixels d'une image, en utilisant la valeur de pixel central tant que seuil. Les pixels de ce bloc sont seuillés par sa valeur de pixel central. Si un pixel voisin a une valeur de niveau de gris plus élevé que le pixel central (ou la même valeur) un est affecté à ce pixel, sinon il obtient un zéro. Un code de LBP est produit, en multipliant ainsi les valeurs seuillées avec des poids donnés (puissances de deux) par les pixels correspondant, et en additionnées. Le résultat sert d'obtenir un label pour le pixel central, comme le voisinage se compose de huit pixels, on obtient donc, comme pour une image à niveaux de gris, une image des valeurs LBP contenant des pixels dont l'intensité se situe entre 0 et 255 [84] [85][83]. La figure 2.4 donne une illustration de l'opérateur LBP de base.

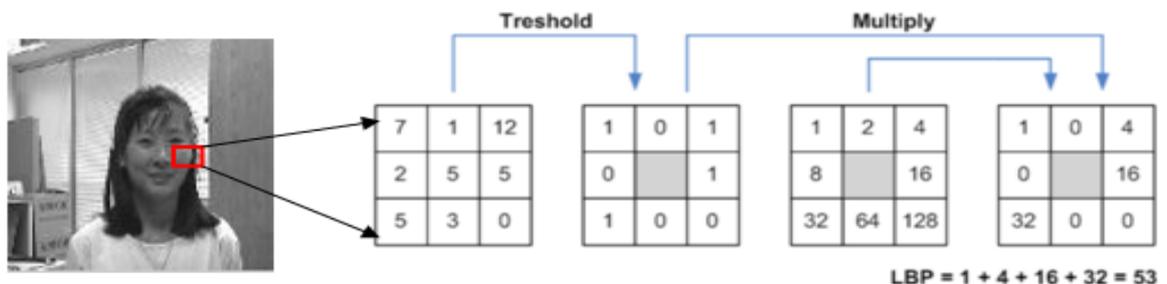


Figure 2.4 : Construction d'un motif local binaire et calcul du code LBP [1].

Pour calculer un code LBP dans un voisinage de P pixels, dans un rayon R, on compte simplement les occurrences de niveaux de gris g_p plus grands ou égaux la valeur centrale.

$$LBP_{P,R}(x_c, y_c) = \sum_{P=0}^{P-1} S(g_c - g_i) \times 2^P \text{ avec } S(x) = \begin{cases} 1 & x \geq 0 \\ 0 & x < 0 \end{cases} \quad (2.9)$$

où g_i et g_c sont respectivement les niveaux de gris d'un pixel voisin et du pixel central.[2][83].

La Figure 2.5 présente un exemple de calcul de l'opérateur LBP basique sur une image au niveau de gris.

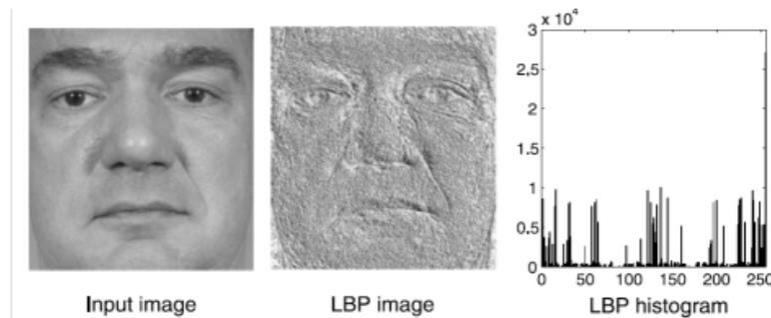


Figure 2.5 : Représente un exemple d'image, LBP et l'historgramme de LBP.

Suite à cette version plusieurs variantes ont été développées et utilisées pour la classification des images texturées, la segmentation et le suivi d'objet en mouvement. Ci-après, nous présentons quelques méthodes les plus utilisées et nous exposons deux nouvelles approches de LBP pour le suivi d'objet en mouvement dans une séquence d'image. La première méthode développée par Ojala [86] traite l'image texturée à différentes échelles en considérant différentes tailles de la fenêtre de voisinage, et la seconde définie ce que l'on appelle les LBP uniformes [86][1]. Le travail de Ning et al.[10] concerne l'utilisation de LBP invariant par rotation afin d'améliorer la robustesse de tracker Mean shift en combinant les caractéristiques de couleur et de texture LBP pour modéliser l'objet cible.

2.4.2.2 Dérivation de l'opérateur (LBP étendus)

Plusieurs années après sa publication de LBP originale .La première méthode de Ojala, dite LBP multi-échelle est fondé sur le choix du voisinage afin de calculer un code LBP pour pouvoir traiter les textures a différentes échelles [86] [1]. Contrairement à la LBP de base en utilisant huit pixels dans un bloc pixels de 3 x 3 ne peuvent pas capturer les traits dominants avec grandes structures. La dérivation de la LBP générique (étendus) ne met pas de limites à la taille du voisinage.[83].

Le LBP étendu sélectionne les pixels voisins comme un ensemble de points d'échantillonnage P réparti uniformément le long d'un cercle avec comme centre le point g_c (pixel central) et un rayon R comme représenté dans la figure 2.6. Dans la littérature la notion

LBP est généralement utilisée pour désigner l'opérateur LBP basique, tandis que la notion LBPP.R est utilisée pour représenter l'étendue LBP où, l'indice P représente le nombre des points d'échantillonnage et l'indice R représente le rayon du cercle. [82].

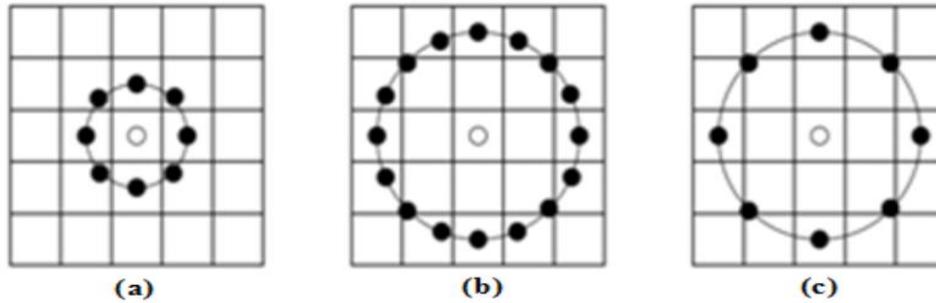


Figure 2.6 : Exemples de l'opérateur LBP étendus : le circulaire voisinage (a) : (8, 1), (b) : (16, 2), (c) : (8, 2)

2.4.2.3 LBP invariant par rotation

Quant à la seconde méthode dite, LBP invariant par rotation ou motif binaire uniforme, est basée sur la notion d'uniformité. Cette notion permet de représenter les informations de primitives structurelles comme les arêtes (coins) et les contours. L'opérateur LBP (P, R) produit 2^P valeurs de sortie différentes, correspondant à 2^P motifs binaires différents formés par les pixels voisinages P . Il a été démontré que certains motifs contiennent plus d'informations que les autres, il est possible d'utiliser uniquement un sous-ensemble de motifs binaires 2^P pour décrire la texture des images. Ojala et al [87], nommé ces motifs des motifs uniformes, notée $LBP_{(P,R)}^{U2}$, où l'exposant U2 signifie l'utilisation que des motifs uniformes.

En effet, l'uniformité est assurée par 0 ou 2 transitions (01 ou 10) dans un parcours circulaire autour du pixel central. Par exemple, 00000000 (0 transitions) et 01110000 (2 transitions) sont uniformes alors 11001001 (4 transitions) et 01010011 (6 transitions) ne sont pas uniformes [88] [83]. La Figure 2.7 illustre un exemple de motif binaire uniforme (10000001 ou 00011000) et non uniforme (00101010) [1].

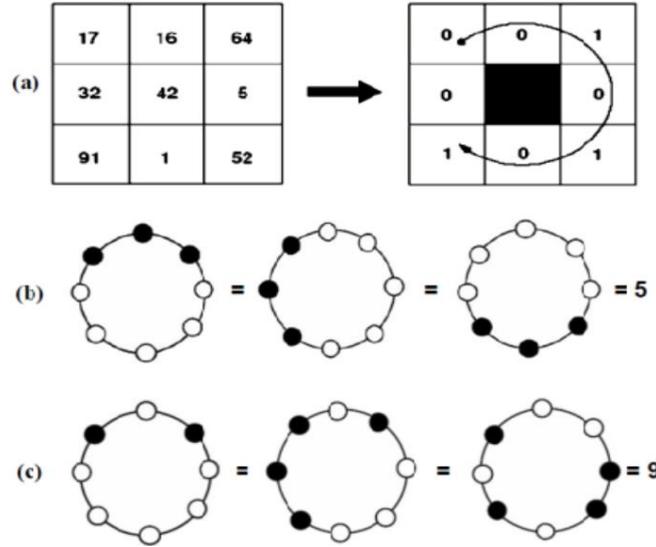


Figure 2.7 : Construction et uniformité d'un motif LBP. (a) le motif construit ici est non uniforme, (c) exemple de motifs respectivement uniforme et non uniforme.

Pour tout pixel (x, y) , le calcul du motif est donné par Ojala [86]:

$$LBP_{P,R}^{ri}(x, y) = \min\{ROR(LBP_{P,R}(x, y), i / i \in [0, p - 1])\} \quad (2.10)$$

Où ROR (a, i) correspond au résultat de i décalages circulaires successifs vers la droite des bits du nombre a (codé sur P bits). $LBP_{P,R}^{ri}$ quantifie ainsi les statistiques d'occurrence de modèles individuels invariants en rotation correspondant à certaines micro caractéristiques de l'image. Pour une valeur $R = 1$, $LBP_{P,R}^{ri\mu 2}$ est l'équivalent de LBP_{ROT} proposé par [89] qui est moins performant. Pour remédier ce problème, une mesure est effectuée sur les motifs pour déterminer leur uniformité. L'opérateur est appelé $LBP_{P,R}^{ri\mu 2}$ et défini de la manière suivante :

$$LBP_{P,R}^{ri\mu i} = \begin{cases} \sum_P^{P-1} \mu(g_0 - g_c) & \text{si } U(LBP_{P,R}) \leq 2 \\ P + 1 & \text{autrement} \end{cases} \quad (2.11)$$

Où U est une fonction qui retourne le nombre de transitions spatiales. Avec :

$$U(LBP_{P,R}) = |\mu(g_{P-1} - g_c) - \mu(g_0 - g_c)| + \sum_P^{P-1} |\mu(g_P - g_c) - \mu(g_{P-1} - g_c)| \quad (2.12)$$

2.4.3 La représentation de l'objet cible avec l'histogramme conjoint HSV couleur-LBP texture

Les histogrammes de couleur constituent des informations très utiles sur les images en couleur. Comme les caractéristiques de la texture introduisent de nouvelles informations que l'histogramme des couleurs ne transmet pas, l'utilisation de l'histogramme conjoint couleur-texture pour la représentation de la cible est plus fiable et robuste que l'utilisation de l'histogramme de couleur uniquement pour le suivi des scènes complexes [10]. Dans notre travail, La représentation de l'objet cible utilisée dans le cadre de l'algorithme de suivi Mean shift est l'histogramme de HSV couleur - LBP texture .L'algorithme de suivi Mean shift traditionnel utilise l'histogramme de couleur RGB pour la représentation de l'objet cible. Pour améliorer les performances de cet algorithme nous avons appliqué l'histogramme de couleur HSV, qui permet de suivi robuste des conditions d'éclairage, car HSV possède un certain degré d'invariance contre les changements d'illumination.

Pour représenter conjointement le modèle cible nous utilisons l'espace HSV et la caractéristique de texture LBP de chaque pixel dans la région cible pour obtenir la région cible LBP, dont la valeur est comprise entre 0 et 255. Pour obtenir la distribution des couleurs et des textures de la région cible, nous utilisons (eq 2.1) pour calculer la distribution des couleurs et des textures du modèle cible q , où $u = 16 \times 16 \times 16 \times 16$. Les trois premières dimensions (c'est-à-dire $16 \times 16 \times 16$) représentent les bins quantifiées des canaux de couleur et la quatrième dimension (à savoir 16) est la bins des motifs de texture LBP. De même, le modèle candidat cible $p(y)$ est calculé avec (eq2.3).La figure 2.8 montre le modèle d'apparence par l'histogramme conjoint HSV couleur-LBP texture dans le tracker Mean shift. Après la modélisation de l'objet cible par l'histogramme conjoint HSV-LBP, on utilise la distance de Bhattacharyya afin de mesurer la similarité entre l'histogramme du modèle et les histogrammes des régions candidats. Cette opération est répétée jusqu'à trouver le maximum local.

Chapitre 2 : Suivi d'objets robuste via l'algorithme Mean shift en utilisant un histogramme conjoint de couleur et de texture

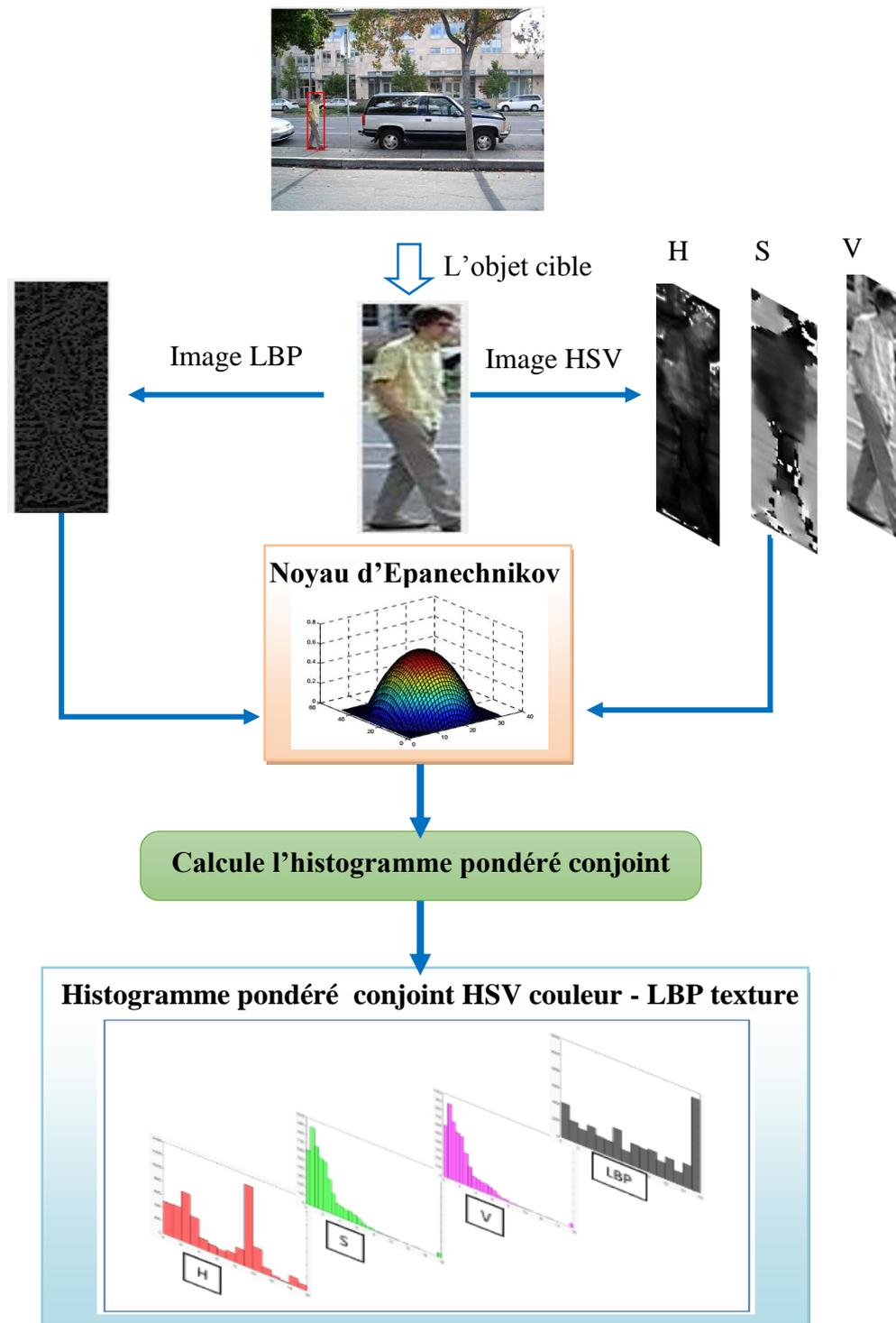


Figure 2.8 : Exemple sur un histogramme conjoint couleur HSV et texture LPB

2.5 Conclusion

Dans ce chapitre, nous avons présenté l'algorithme de suivi d'objet par Mean shift qui se base sur l'histogramme de couleur. Nous avons discuté des fonctionnalités et les différentes étapes de cet algorithme. Malgré, il est robuste à une occultation partielle, à l'échelle, la rotation et la déformation non-rigide de la cible, mais il ne permet pas la conservation de l'information spatiale de chaque pixel dans l'image. Pour surmonter ce problème nous avons utilisé les caractéristiques de texture LBP qui reflètent la structure spatiale de l'objet combiné avec les caractéristiques de couleur HSV pour créer un histogramme conjoint de couleur HSV-LBP de manière plus distinctive et efficace. Cet histogramme permet d'obtenir de meilleures performances de suivi et un modèle d'apparence plus robuste à différents défis en particulier où la similarité entre l'objet et le fond.

Chapitre 3 :
Simulation, Résultats et
discussions

3.1 Introduction

Ce chapitre est consacré d'évaluer les performances de l'algorithme de suivi d'objet Mean shift en utilisant l'histogramme conjoint HSV couleur- LBP texture, dans une séquence d'images. Cet histogramme permet d'améliorer la robustesse et la précision de tracker Mean shift. Tous d'abord, on présente l'architecture globale de l'algorithme à l'aide d'un schéma illustré leur déférente étapes du tracker Mean shift, en particulier la méthode de la combinaison entre les caractéristiques de couleurs et de textures pour représenter l'objet cible. Ensuite, Afin de valider ce travail, nous avons testé notre algorithme sur différentes séquences d'images. Enfin, les résultats obtenus ainsi que leurs discussions seront présentées. Aussi, une étude comparative avec le tracker Mean shift traditionnel qui utilise l'histogramme de couleur et entre l'amélioration de ce tracker qui utilise l'histogramme conjoint couleur-texture.

3.2 Structure de l'algorithme

Le suivi par l'algorithme Mean shift en utilisant l'histogramme conjoint couleur-texture s'effectue en deux étapes détection de cible et le suivre de cibles. Le suivi commence par l'acquisition de l'image initiale et détecter ou sélectionner l'objet d'intérêt manuellement ou automatiquement ensuite modélisé par une rectangle sur laquelle on calcule sa distribution des caractéristiques de couleur HSV et de texture LBP. La figure 3.1 résume la structure de l'algorithme de tracker amélioré Mean shift à l'aide d'un histogramme HSV couleur- LBP texture.

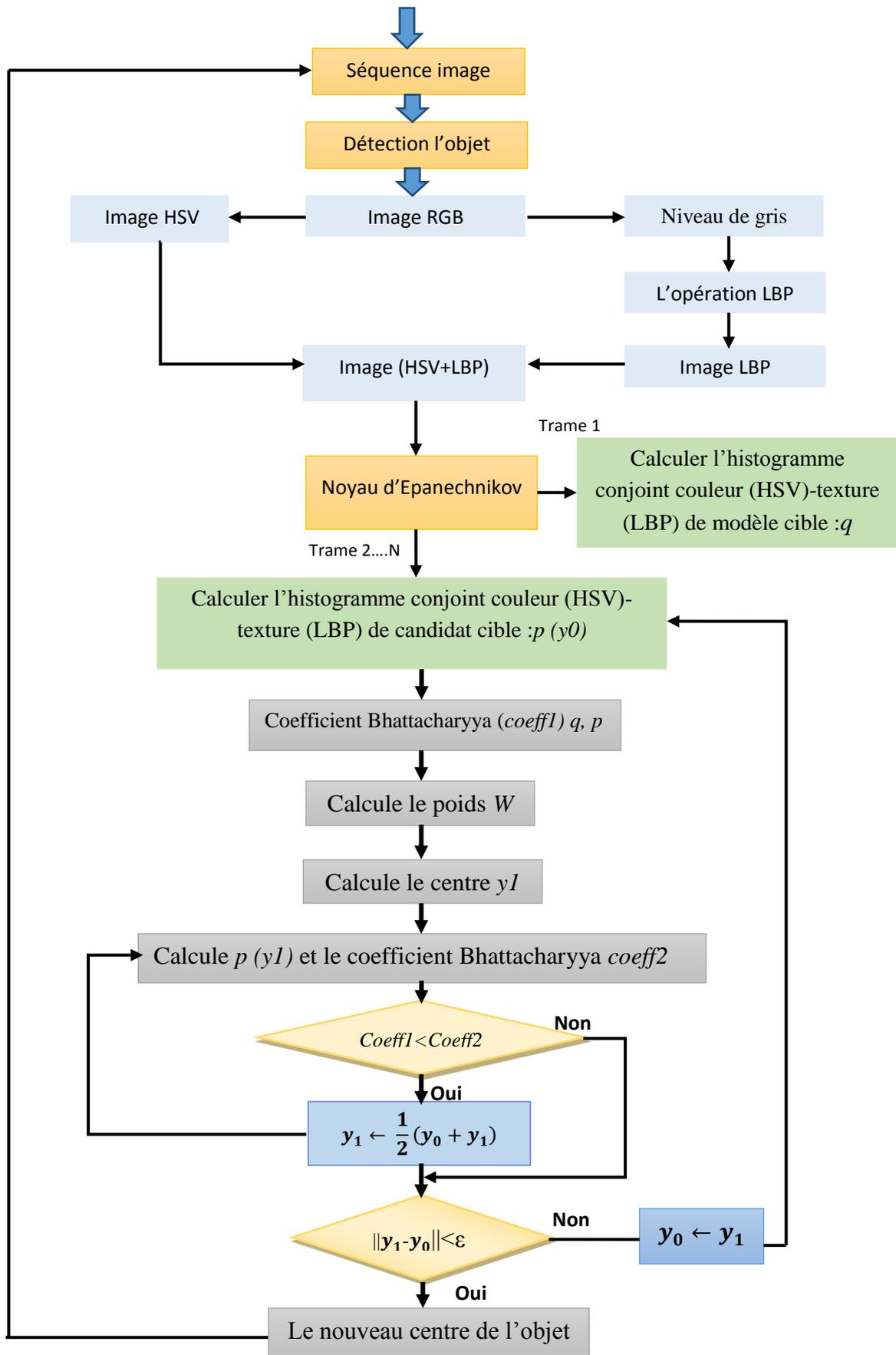


Figure 3.1 : La structure de l'algorithme de suivi par Mean shift avec LBP.

3.2.1 Détection d'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.

3.2.2 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):** intuitivement, c'est le nom qu'on utilisera pour désigner la couleur "verte", "mauve", "orange", etc. Idéalement associé à une longueur d'onde, donc à une position sur le cercle de Newton.
- **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.

Dans plusieurs applications on s'intéresse surtout aux canaux H et S qui représentent l'information chromatique (e.q. 3.1).[90].

Avec :

$$H \in [0,360^\circ] \text{ et } S \in [0,1] \text{ et } V \in [0,1]$$

H : teinte, S : saturation, V : valeur

Le passage de RGB (R : rouge, G : vert, B : bleu) à HSV se fait par une transformation non linéaire. Plusieurs opérateurs ont été proposés pour la conversion (e.q.3.1)

$$\begin{aligned}
 V &\leftarrow \max(R, G, B) \\
 S &\leftarrow \begin{cases} V - \min(R, G, B)/V & \text{if } V \neq 0 \\ 0 & \text{Otherwise} \end{cases} \\
 H &\leftarrow \begin{cases} 60 \frac{G - B}{S} & \text{if } V = R \\ 120 + \frac{60(B - R)}{S} & \text{if } V = G \\ 240 + \frac{60(R - G)}{S} & \text{if } V = B \end{cases}
 \end{aligned} \tag{3.1}$$

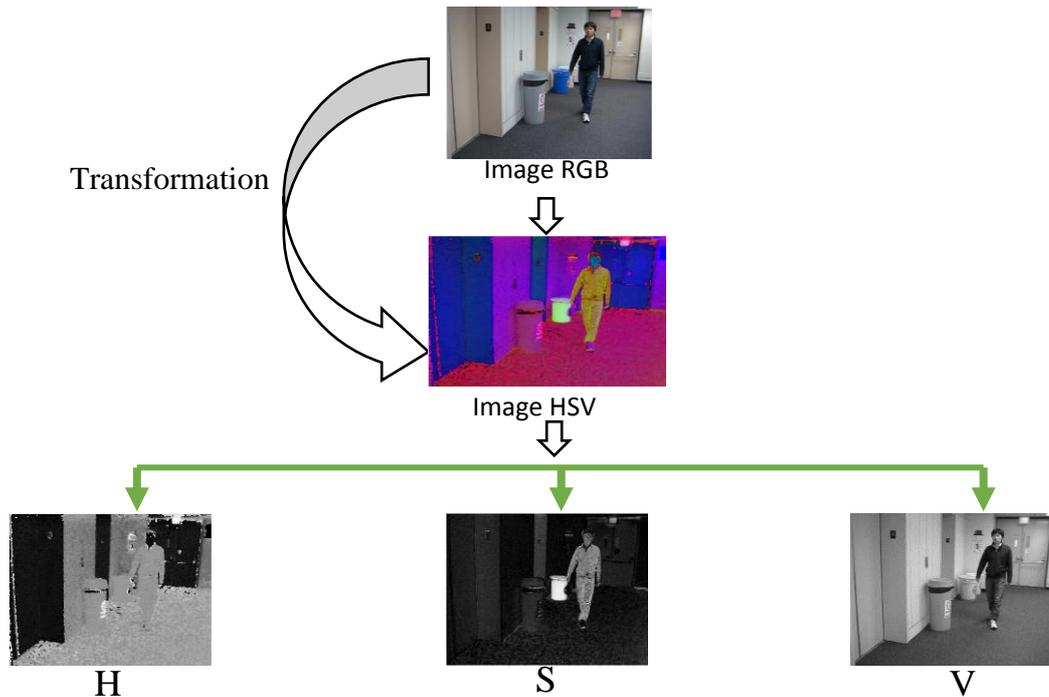


Figure 3.2 : Exemple de la transformation d'espace couleur RGB vers HSV.

3.2.3 Calcule l'image LBP

Dans ce travail, nous utilisons l'opérateur LBP basique, c'est une technique locale d'extraction des caractéristiques au niveau des pixels. Le descripteur LBP encode le voisinage local d'un pixel en stockant la différence entre le pixel examiné et ses voisins.

Pour l'opération LBP, nous transformons l'image RGB à niveaux de gris, puis nous utilisons l'opération $LBP_{8,1}$, $P=8$ et $R=1$. L'image LBP extraite à d'image de visage illustre dans la figure suivante.

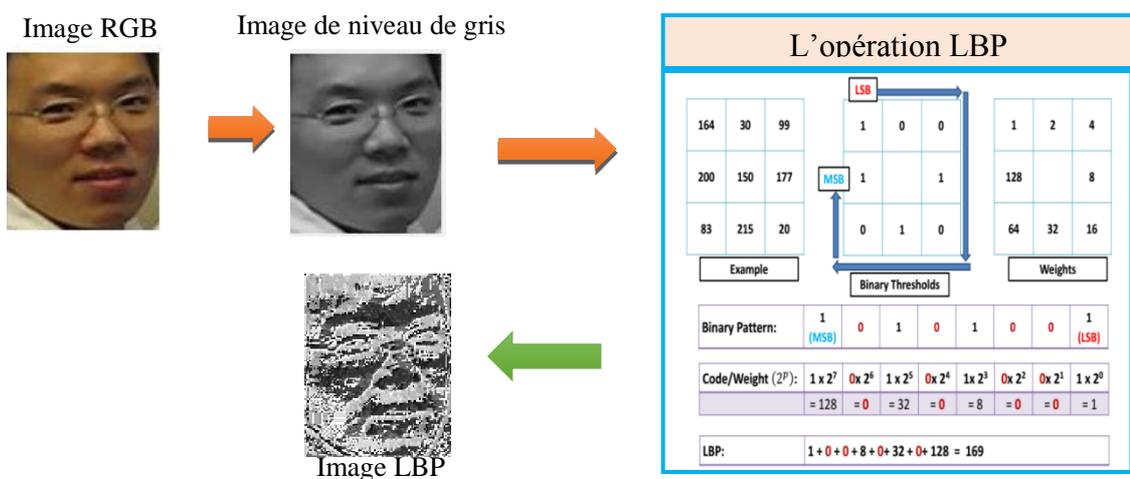


Figure 3.3. Exemple d'opération $LBP_{8,1}$.

3.2.4 La distribution de la cible modèle et candidats

3.2.4.1 Histogramme conjoint HSV couleur-LBP texture

Dans ce travail, nous combinons les caractéristiques de couleur de l'espace HSV et de texture de l'opérateur basique LBP pour représenter le modèle d'apparence de l'objet cible par l'histogramme pondéré conjoint HSV couleur-LBP texture. Dans le tracker Mean shift on calcule la distribution de couleur et de texture pondérée par le noyau d'Epanechnikov. Nous utilisons pour calculer la distribution de couleur et de texture du modèle cible et les candidats trois composantes d'espace de couleur HSV et l'image de texture LBP. Ces images quantifiées en 16 bins et la taille de l'histogramme conjoint est $16 \times 16 \times 16 \times 16$, où $u = 16 \times 16 \times 16 \times 16$. Les trois premières dimensions représentent les bins quantifiées des canaux de couleur HSV et la quatrième dimension est la bins des motifs de texture LBP. La figure 3.4 montre l'histogramme conjoint V couleur- texture LBP en utilisant le composant V de l'espace HSV.

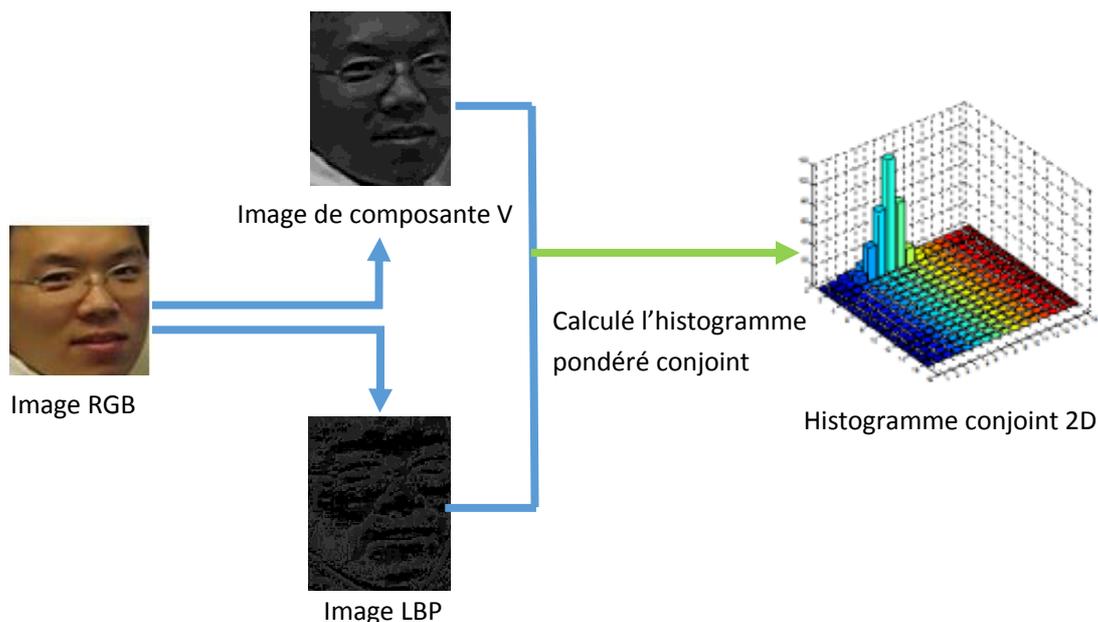


Figure 3.4 : Exemple sur l'histogramme conjoint 2D de composante V de l'espace HSV.

Dans ce travail, nous étudierons l'histogramme conjoint en utilisant une composante de couleur H, S, V de l'espace HSV au lieu de l'image en niveau de gris pour obtenir l'information de texture par le descripteur LBP (c.-à-d. l'image LBP). Ces images combinées avec les trois images H, S, V pour créer l'histogramme conjoint HSV couleur-LBP texture.

La figure ci-dessous montre la différence entre les images LBP à partir des composantes H, S, V et l'image en niveau de gris.

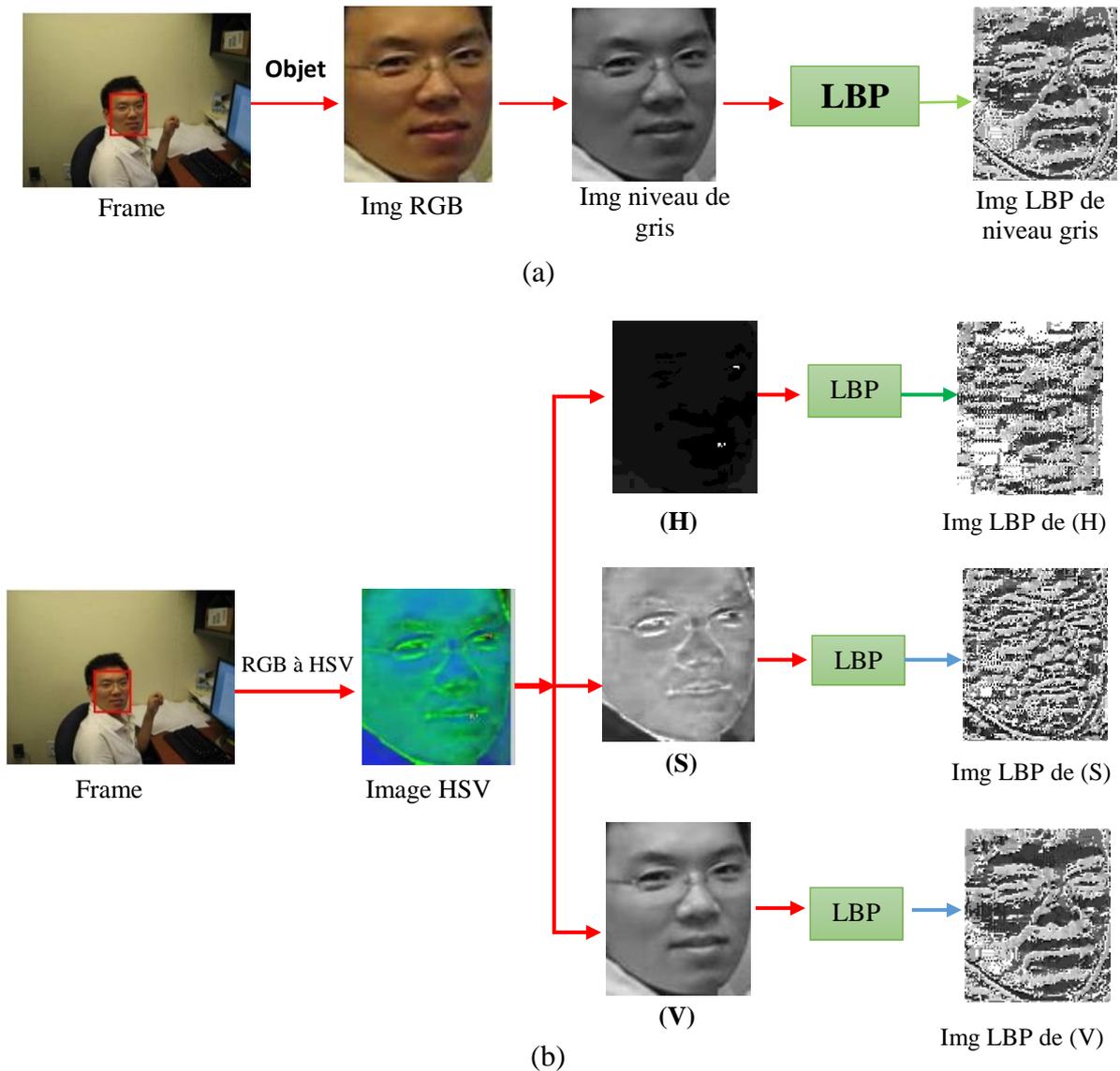


Figure 3.5 : Les résultats de l'opération sur l'image RGB (a) image en niveau de gris, (b) à chaque composante de l'image HSV.

3.2.4.2 Noyau d'Epanechnikov

Noyau d'Epanechnikov est une fonction de pondération utilisée dans les techniques d'estimation non-paramétrique. Les noyaux interviennent dans l'estimateur par noyau pour estimer la densité de probabilité. Le poids du noyau Epanechnikov pour la fenêtre d'objet en forme rectangulaire (voir figure 3.3).

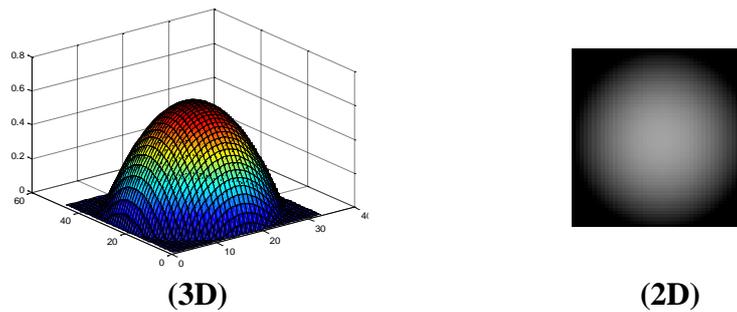


Figure 3.6 : Noyau d'Epanechnikov.

3.2.4.3 Poids et le nouveau centre

Pour chaque trame, pour chaque pixel de l'image, un poids est calculé en fonction de l'histogramme conjoint la couleur et texture de l'objet. Ce poids représente le nombre de pixels ayant la couleur de ce pixel. Ainsi, l'image de poids 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 extraction et calculer le vecteur Mean Shift, qui a pour but de fournir la nouvelle position estimée de la cible dans l'image courante, la figure 3.7 représente cette étape. 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 $\rho(y)$. Cette maximisation peut être effectuée efficacement en utilisant les itérations Mean shift.



Figure 3.7 : Exemple représenté l'image de poids.

3.3 Environnement de travail

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

3.3.1 Environnement matériel

Afin de mener ce travail, il a été mis à notre disposition un ensemble de matériels suivants :

- Ordinateur HP 630 avec les caractéristiques suivantes :
- Processeur : Intel Pentium Inside
- RAM : 2,0 GB.

- Disque Dur : 300 GB HDD.
- carte graphique : Intel graphique.
- Type du système : système d’exploitation 64 bits.
- Edition de Windows : Windows 7 Professional.

3.3.2 Langage de programmation

Pour réaliser le travail, nous utilisons logiciel MatlabR2014a qui il est un langage de haut niveau qui permet l’exécution de tâches nécessitant une grande puissance.

3.4 Bases de données

Pour évaluer les performances de cet algorithme, nous utilisons les séquences d’images qui existent dans la base de données OBT sur le site : http://cvlab.hanyang.ac.kr/tracker_benchmark/datasets.html, Chacune des séquences vidéo est liée à un fichier TXT décrivant les entités de chaque image avec leur position.

Le tableau (3.1) illustre les séquences d’images utilisées.

Tableau 3.1 : Séquences d’images utilisées

Séquence d’image	1 ^{ère} image	Nombre de trames	Taille d’image	Défis
Divid3		252	640X480	OCC, DEF, MB
DragonBaby		113	640X360	OCC, MB, SV
Crossing		120	360X240	IV, SV, BC
Couple		140	320X240	SV, BC, FM
BleuCar4		379	640X480	FM, IV, BC

(OCC : Occultation, DEF : Déformation, MB : Mouvement de caméra, SV : Changement d’échelle, IV : Changement d’illumination, BC : L’arrière-plan clutter, FM : Mouvement rapide).

3.5 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. Il existe de nombreuses métriques de performance en suivi d'objet. Les principales utilisées sont : erreur de localisation du centre et taux de recouvrement [2]:

3.5.1 Taux de recouvrement

Le taux de recouvrement 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 :

$$IoU(B, B') = \frac{|B \cap B'|}{|B \cup B'|} \quad (3.2)$$

IoU , illustré dans la figure 3.8 (a), est une mesure d'erreur plus précise que l'erreur de localisation du centre puisqu'il tient compte de la taille des boîtes. Le taux de recouvrement moyen consiste alors à moyennner IoU sur l'ensemble des images de la base [91].

3.5.2 Erreur de localisation du centre

L'erreur de localisation du centre [92], [93] 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 $\{g_i\}_{i=1}^M$:

$$ELC = \frac{1}{M} \sum_{i=1}^M \|p_i - g_i\| \quad (3.3)$$

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 (voir la figure 3.8.(b)).

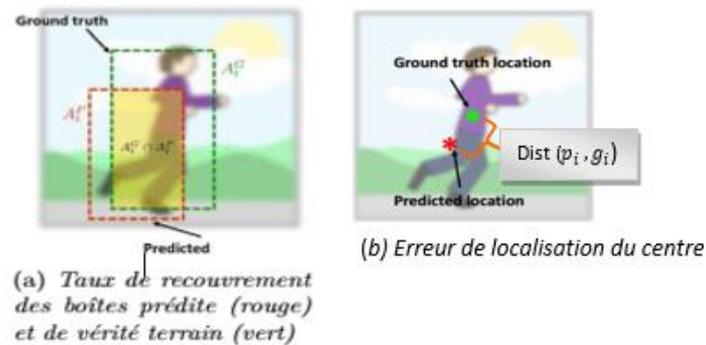


Figure 3.8. Métriques de performance : (a) Taux de recouvrement, (b) Erreur de localisation du centre.

3.6 Résultats Expérimentaux

Dans ce travail, nous avons choisi cinq séquences d'images pour illustrer les résultats obtenus avec algorithmes de suivi d'objet Mean shift en utilisant l'histogramme de couleur (MS) et l'histogramme conjoint HSV couleur- LBP texture. Ces séquences contiennent la plupart des 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 de la vérité 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 $\epsilon = 0.5$.

3.6.1 Les résultats des évaluations de tracker Mean shift avec des différentes LBP images

Dans cette partie, nous présentons une étude comparative entre les différents histogrammes pondérés conjoints qui représentent l'objet cible du tracker Mean shift. Ces histogrammes combinaient entre les composantes de couleur HSV et les images de texture LBP qui obtient en utilisant l'image en niveau de gris d'une part et d'autre part chaque composante de l'espace HSV (H, S et V). La figure ci-dessous représente les résultats de suivi en utilisant les différents histogrammes conjoints dans le tracker Mean shift.

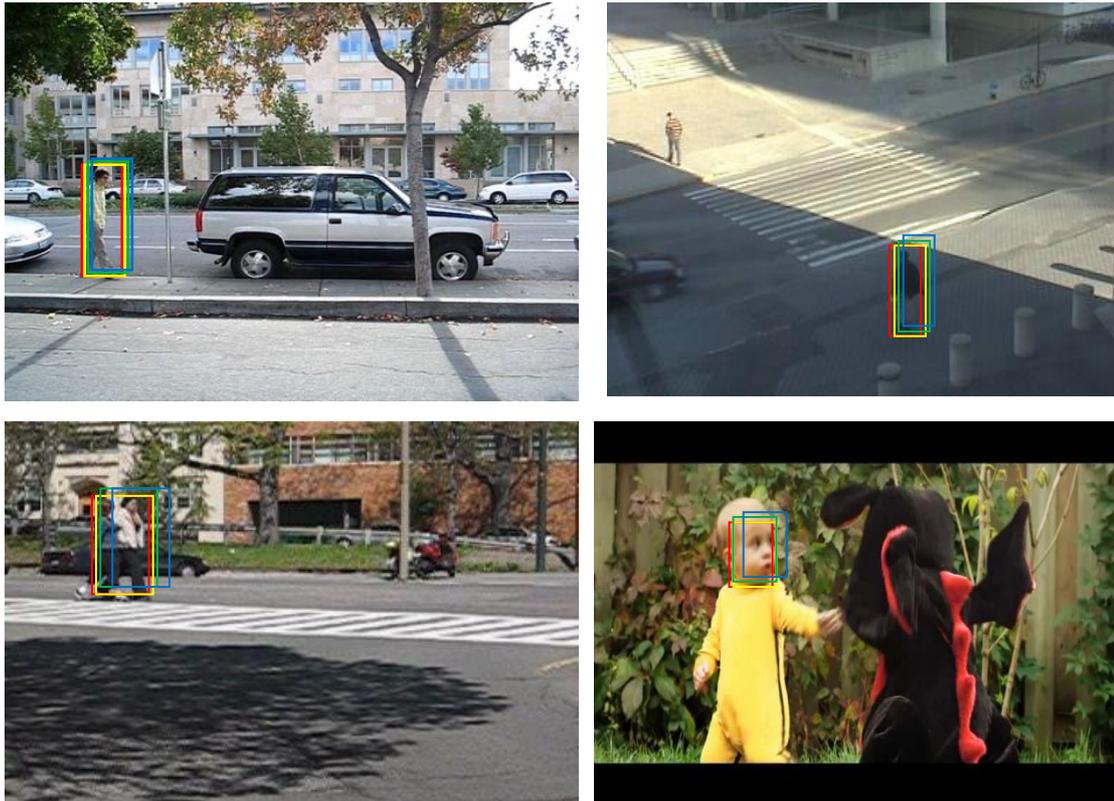


Figure 3.9 : Résultats de suivi sur les séquences, Divid3, Crossing, Couple, DragonBaby .

— MS_LBP (niveau de gris), — MS_LBP(niveau de gris de (V)),
 — MS_LBP (niveau de gris de (S)), — MS_LBP(niveau de gris de (H)) .

Le tableau (3.2) illustre les moyennes de l'erreur de localisation du centre (AveErrCenter) et taux de recouvrement (AveOver) de tracker Mean shift en utilisant les différents histogrammes pondérés conjoints pour les séquences d'images utilisées. On remarque que, la moyenne de l'erreur de localisation du centre et la moyenne du taux de recouvrement du tracker Mean shift qui a utilisé l'image LBP obtenu par l'image en niveau de gris pour créer l'histogramme conjoint égales à 11,285 et 0,6626, respectivement, sont supérieurs aux autres moyennes qui ont utilisé les histogrammes conjoints les images LBP obtenus par les composantes H, S et V. Après les résultats obtenus, on peut dire que le tracker Mean Shift par l'histogramme conjoint HSV couleur- LBP texture est robuste en utilisant l'image LBP obtenu par l'image en niveau de gris.

Tableau 3.2. Résultats obtenus de l'erreur de localisation du centre et du taux de recouvrement pour les différents histogrammes pondérés conjoints.

Séquence d'image	MS_LBP		MS_LBP(H)		MS_LBP(S)		MS_LBP(V)	
	AveErrCenter	AveOver	AveErrCenter	AveOver	AveErrCenter	AveOver	AveErrCenter	AveOver
Divid3	10,9174	0,7002	11,19	0,6941	10,8968	0,6994	11,0767	0,7025
Crossing	5,5344	0,6527	5,8086	0,6402	6,0002	0,6383	5,5227	0,6501
Couple	11,7745	0,5399	28,7357	0,3902	17,048	0,4505	14,9154	0,4931
DragonBaby	21,4967	0,5691	35,8656	0,5593	18,1387	0,6172	18,1427	0,6134
BlurCar 4	21,2613	0,7116	23,6326	0,6909	23,1147	0,6934	21,9644	0,7062
Moyenne	11,2858	0,6626	15,31954	0,62422	12,8447	0,6364	12,0742	0,65284

3.6.2 Comparaison entre Mean shift et Mean Shift avec LBP

Dans la partie précédente, selon les résultats obtenus nous avons conclu que le tracker Mean Shift qui utilise LBP sur l'image en niveau de gris exploite les caractéristiques spatiales de texture d'objet de façon mieux. Pour cela, nous comparons le tracker Mean Shift traditionnel (MS) qui utilise l'histogramme de couleur HSV et Mean shift avec LBP (MS-LBP) qui utilise l'histogramme conjoint HSV couleur-LBP texture dans cette partie.

3.6.2.1 La séquence Divid3

Dans Cette séquence, l'objet à suivre est un «homme» avec des problématiques (mouvement de camera, déformation et occultation partielle). La figure 3.10 montre les résultats obtenus de suivi par les deux algorithmes Mean shift et Mean Shift avec LBP pour les trames 7, 55, 78, 165,210 et 246. On remarque que, les algorithmes Mean shift et Mean Shift avec LBP peuvent suivre l'objet cible avec précision dans la plupart des trames, mais MS-LBP est plus robuste que MS parce qu'il ajoute les caractéristiques spatial de l'objet cible.

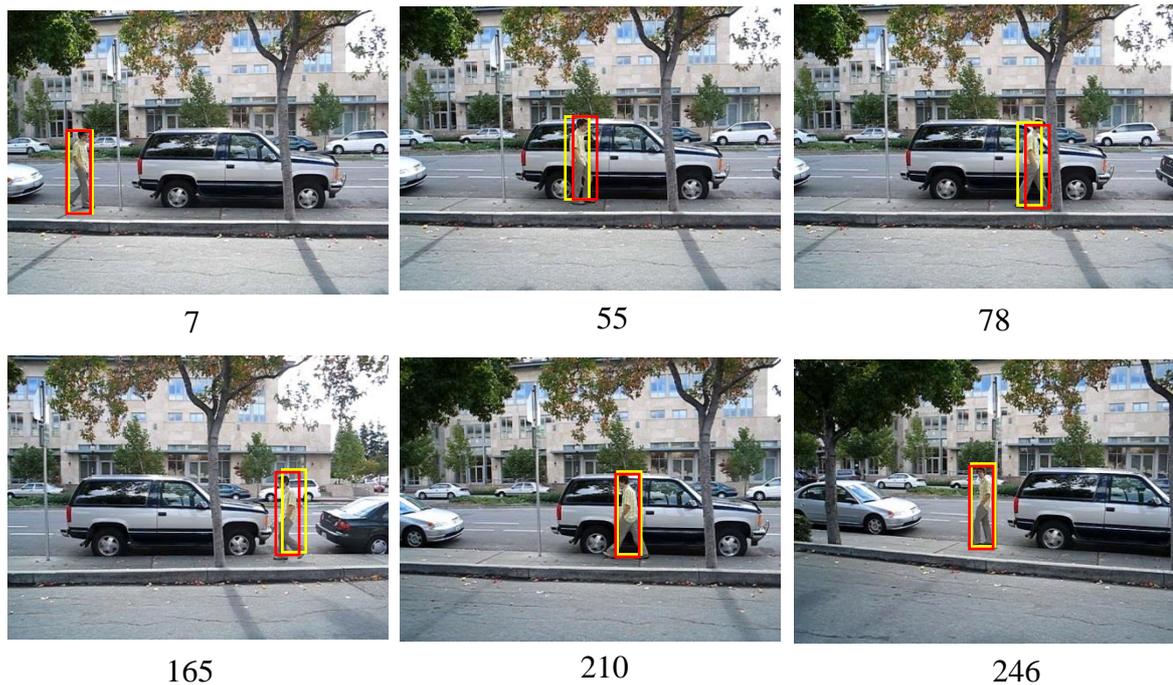


Figure 3.10 : Résultats de suivi sur la séquence Divid3 avec Mean shift (rectangle jaune) et Mean Shift avec LBP (rectangle rouge).

Pour la comparaison quantitative, la figure la figure 3.11 représente les courbes de l’erreur de location du centre et taux de recouvrement en fonction nombre trame avec les trackers de suivi MS et MS-LBP. On remarque que, les courbes de l’erreur de location du centre et taux de recouvrement pour MS et MS-LBP sont presque identiques dans la plupart des trames, sauf entre les trames de 74 à 88 les résultats par MS-LBP est meilleur que MS.

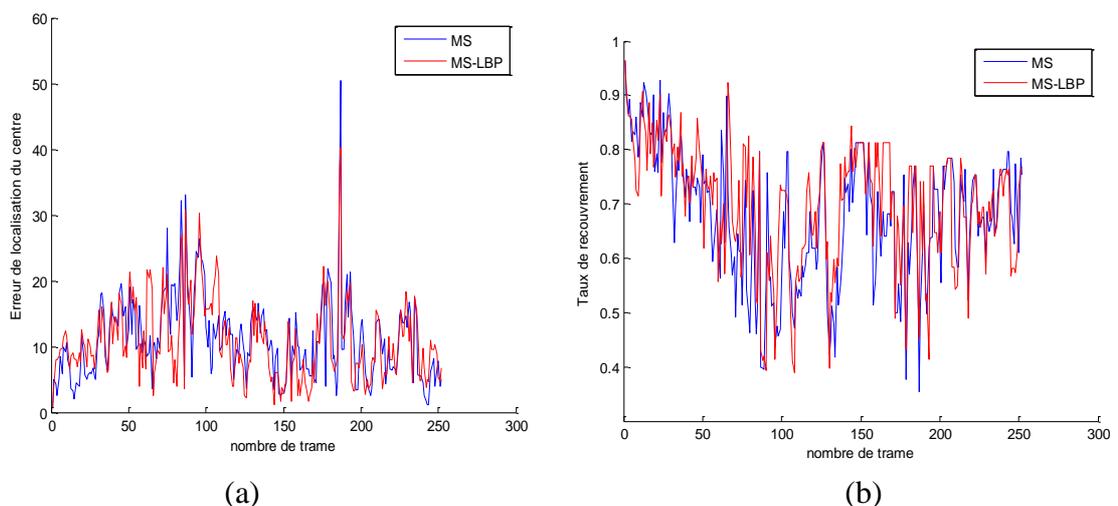


Figure 3.11 : Comparaison quantitative avec Mean shift et Mean shift pour séquence Divd3 : (a) erreur de location du centre (b) Taux de recouvrement.

3.6.2.2 La séquence Crossing

La séquence Crossing, l'objet à suivre est un homme (objet non rigide) en petit taille avec changement d'échelle, changement d'illumination ainsi qu'en présence de similarité des couleurs entre l'objet et l'arrière-plan. La figure 3.12 montre les résultats obtenus de suivi par les deux algorithmes Mean shift et Mean shift avec LBP pour les trames 8, 31, 45, 87, 94 et 113. On peut voir que, les deux algorithmes peuvent suivre la cible avec précision en particulier en début trames, comme illustré dans les trames présentées 8, 31. Cependant, MS et MS-LBP sont moins précision dans les dernières trames, parce qu'il existe une variance d'illumination et d'échelle et en présence d'un autre objet (voiture) a de couleur similaire à l'objet cible et à l'arrière-plan. Malgré ça, MS-LBP est meilleur que MS.

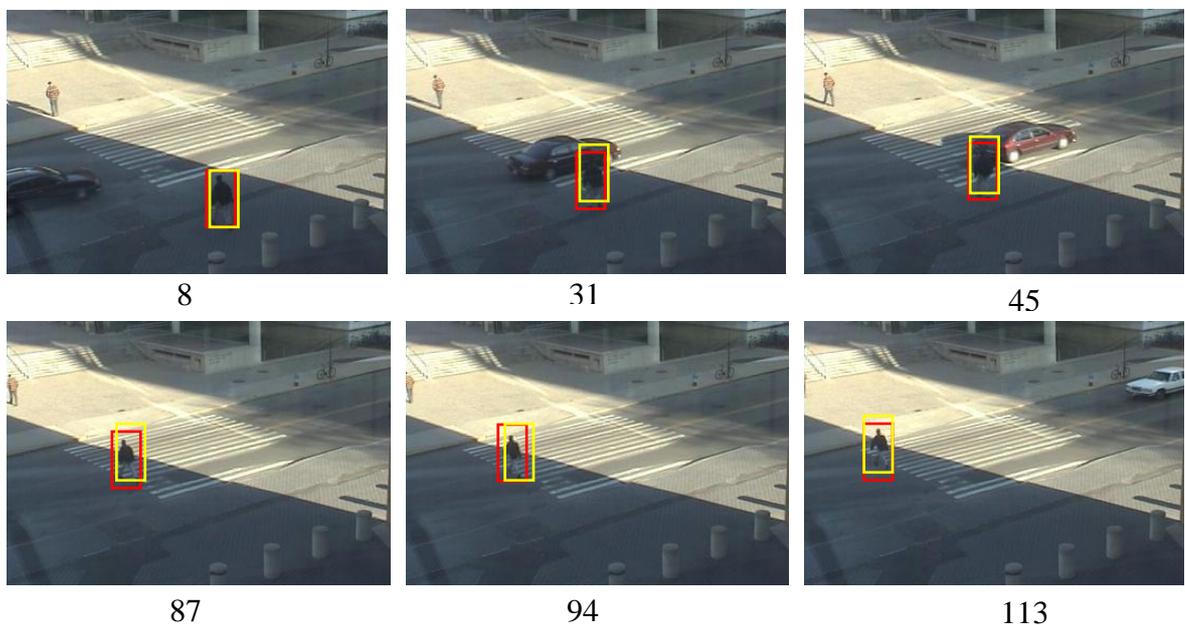


Figure 3.12 : Résultats de suivi sur la séquence Crossing avec Mean shift (rectangle jaune) et Mean Shift avec LBP (rectangle rouge).

Dans la figure ci-dessous, on remarque que, les valeurs de courbe de l'erreur de localisation du centre (a) de MS-LBP sont inférieures aux valeurs de MS dans toutes les trames. Cependant, les valeurs de courbe de taux de recouvrement (b) de MS-LBP dans toutes les trames supérieurs au MS. Cela signifie que MS-LBP peut suivre la cible avec précision et robustesse que MS, en présence des défis difficiles.

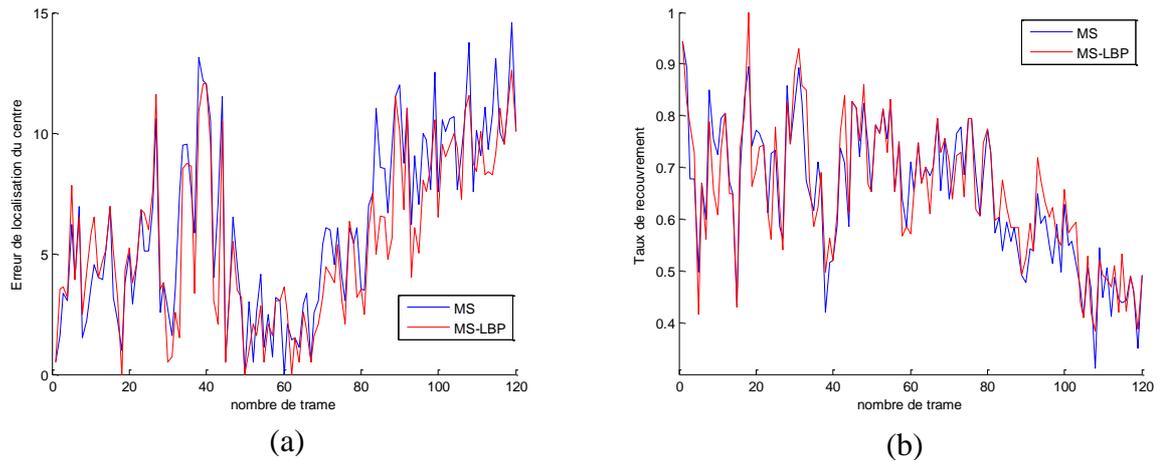


Figure 3.13 : Comparaison quantitative avec Mean shift et Mean shift pour séquence Crossing: (a) erreur de location du centre (b) Taux de recouvrement

3.6.2.3 La séquence DraganBaby

L’objet à suivre dans la séquence Dragan Baby est un « visage de bébé » avec des problématiques : occultation, mouvement de caméra, changement d’échelle et l’objet en mouvement rapide (flou de bougé)). La figure 3.14 montre les résultats obtenus de suivi par MS et MS-LBP pour les trames 12, 33, 53, 76, 83 et 95. On peut voir que, les algorithmes Mean shift et Mean Shift avec LBP peuvent suivre l’objet cible dans 12, 33, 76, 83 et 95 avec précision. Cependant, dans la trame 45 il y a une occultation totale, l’objet occulté complètement. Le tracker MS-LBP peut retrouver rapidement la cible quand il réapparaît après occultation comme montré dans la trame 53. Par contre, le tracker MS ne peut pas retrouver la cible rapidement. On peut dire que MS-LBP est plus robuste que MS aux occultations à cause d’utilisation les informations spatial de l’objet cible.



Figure 3.14 : Résultats de suivi sur la séquence DraganBaby avec Mean shift (rectangle jaune) et Mean shift avec LBP (rectangle rouge).

La figure ci-dessous représente les résultats obtenus pour la comparaison quantitative de l'erreur de localisation du centre et du taux de recouvrement de la séquence DragonBaby par les deux algorithmes MS et MS-LBP. On remarque que, les courbes de l'erreur de location du centre et les courbes de taux de recouvrement pour MS et MS-LBP sont presque identiques dans les trames de 0 à 42 et de 58 à 113. Après les trames 43 à 54 et 43 à 58 pour MS-LBP et MS respectivement. Les valeurs de taux de recouvrement descendent jusqu'à 0. Par contre, les valeurs de l'erreur de location du centre changent vers au maximum. Cela signifie que l'objet à suivre par les deux algorithmes est perdu parce qu'il y a une occultation totale. Après la trame 54, on peut voir que le taux de recouvrement et l'erreur de location du centre du tracker MS-LBP changent vers plus près à 0.8 et 10 pixel, respectivement. Par rapport à MS les valeurs des deux courbes changent après la trame 58.

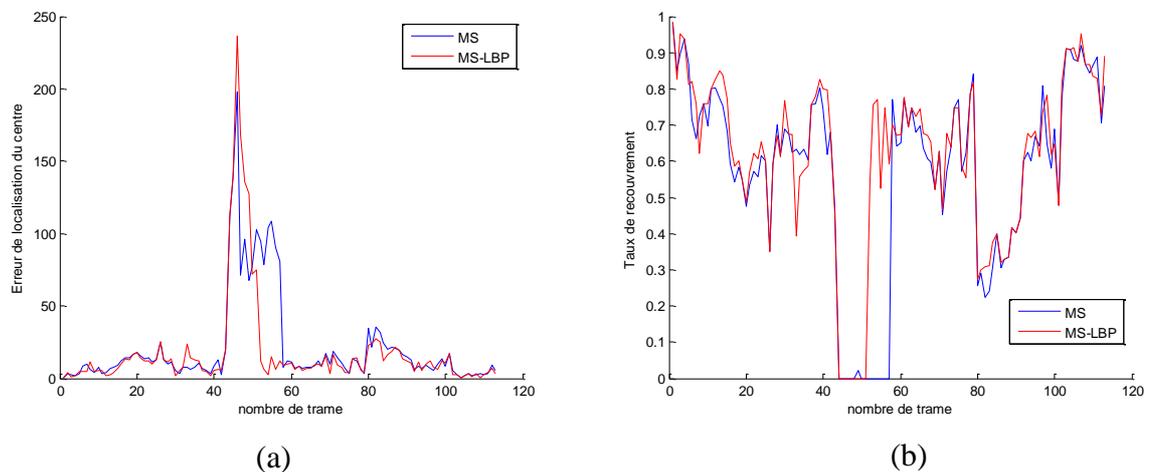


Figure 3.15 : Comparaison quantitative avec Mean shift et Mean shift pour séquence DragonBaby : (a) erreur de location du centre (b) taux de recouvrement.

3.6.2.4 La séquence Couple

L'objet à suivre dans la séquence Couple sont « deux femmes » avec des problématiques (changement d'échelle, l'arrière plan et l'objet en mouvement rapide (flou de bougé)). La figure 3.16 montre les résultats obtenus de suivi par les algorithmes Mean shift et Mean Shift avec LBP pour les trames 5, 30, 88, 92, 108 et 122. On peut voir que, les deux trackers peuvent suivre l'objet cible dans 5, 30 et 88 avec précision. Mais, ils échouent et suivent un autre objet (voiture) (voir la trame 92) parce qu'il existe un autre objet (voiture) similaire que la première femme qui occulte la deuxième femme. Le tracker MS-LBP peut retrouver l'objet cible après quelques trames comme montré dans les trames 108 et 122, car il utilise les informations spatiales. Par contre, le tracker MS ne peut pas retrouver la cible.

On peut dire que MS-LBP est plus robuste que MS en présence la similarité entre l'objet cible et des autres objets.

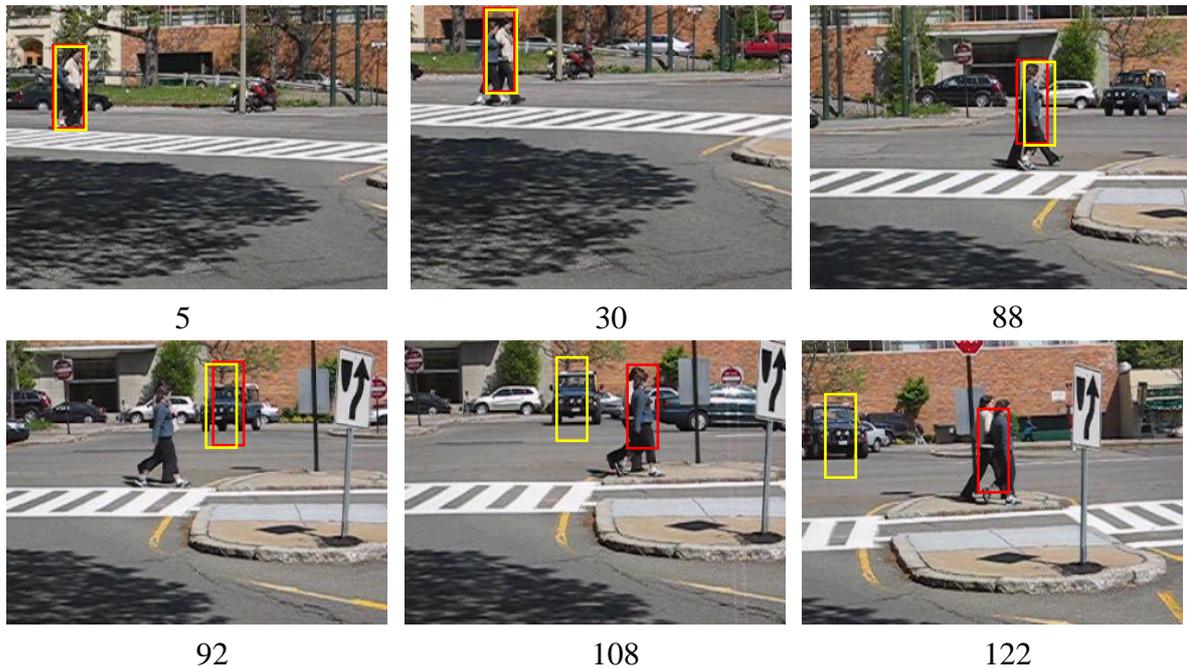


Figure 3.16 : Résultats de suivi sur la séquence Couple avec Mean shift (rectangle jeune) et Mean Shift avec LBP (rectangle rouge).

D'après la figure ci-dessous nous remarquons que dans les trames de 0 à 90 le suivi par les deux algorithmes MS et MS-LBP sont presque identiques et robuste sauf dans les trames de 45 à 50. Après la trame 98, les deux trackers ne peuvent pas suivre bien l'objet cible car il existe un autre objet voiture a de couleur similaire et le deuxième objet est occulté par l'autre cible. Cependant, après la trame 108 les valeurs de l'erreur et le taux de recouvrement pour MS sont supérieurs à 20 pixels et égale à 0, respectivement. Cela signifie que l'objet à suivre par MS est perdu.

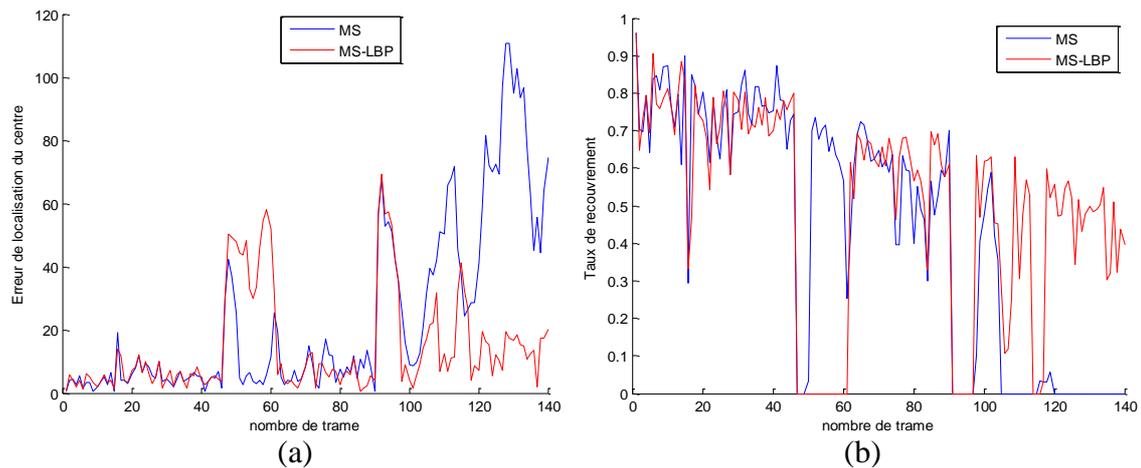


Figure 3.17 : Comparaison quantitative avec Mean shift et Mean shift pour séquence Couple : (a) erreur de location du centre (b) taux de recouvrement.

3.6.2.5 La séquence BlurCar 4

Dans cette séquence, l'objet à suivre est une « Voiture » avec des problématiques (mouvement de caméra rapide (flou de bougé) et changement d'échelle ainsi qu'en présence de similarité des couleurs entre l'objet et l'arrière-plan). La figure (3.16) représente les résultats visuels de suivi de voiture dans les trames 18, 117, 146, 220, 314 et 363. Malgré il y a un mouvement de caméra (image flou) comme montre dans la plupart des trames les deux trackers MS-LBP et MS peuvent suivre la voiture avec précision. Cependant, MS-LBP est plus robuste que MS pour tous les trames, en particulier dans la trame 363 où MS est dérive (comme illustré dans la trame 363) puisqu' il y a un problème de changement d'échelle et la similarité des couleurs entre l'objet et l'arrière-plan. On peut dire que la combinaison les informations de couleur et de texture pour modéliser l'objet cible rend le tracker Mean shift plus robuste.

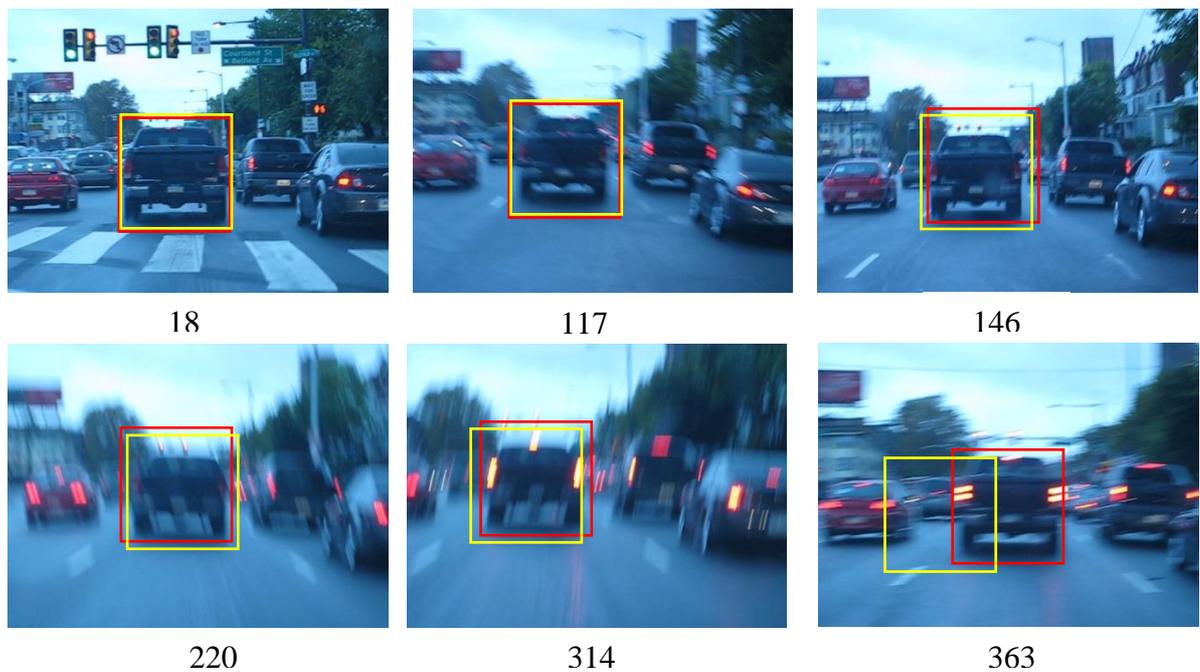


Figure 3.18 : Résultats de suivi sur la séquence Blurcar4 avec Mean shift (rectangle jeune) et Mean Shift avec LBP (rectangle rouge).

La figure ci-dessous représente la comparaison quantitative de l'erreur de localisation et le taux de recouvrement de la séquence blurCar. On remarque que dans les trames de 1 à 255, les courbes d'erreur de localisation pour MS et MS-LBP et les courbes de taux de recouvrement pour MS et MS-LBP sont presque identiques. Mais, après la trame 256, les valeurs de l'erreur de location du centre de MS-LBP est inférieur à MS, et les valeurs du taux de recouvrement de MS-LBP est supérieure à MS. cela signifie que la robustesse de tracker MS-LBP par rapport MS.

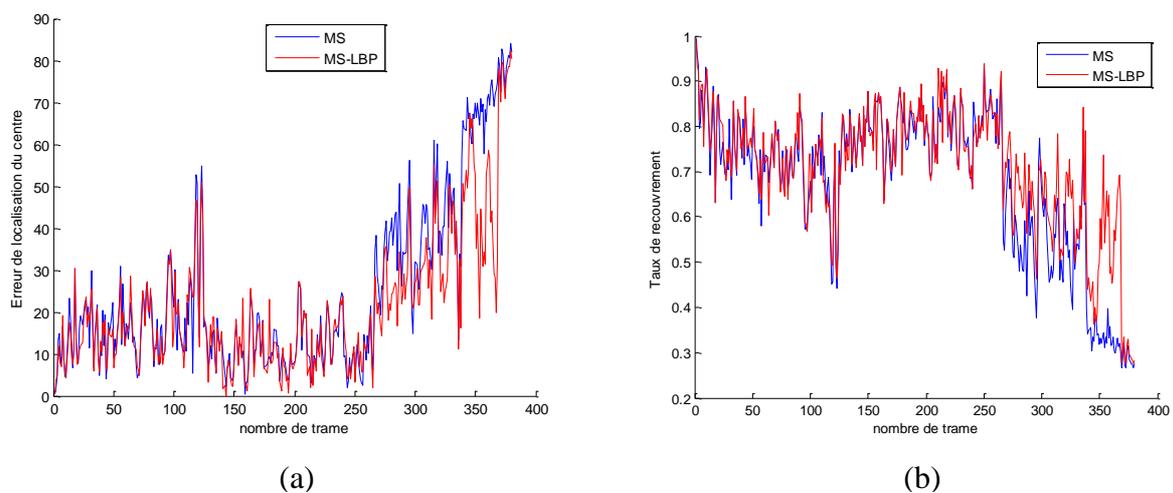


Figure 3.19 : Comparaison quantitative avec Mean shift et Mean shift avec LBP pour la séquence BlurCar 4: (a) erreur de location du centre (b) taux de recouvrement.

3.6.2.6 Comparaison entre les moyennes d'erreur de localisation du centre et le taux de recouvrement pour Mean Shift et Mean shift avec LBP

Le tableau 3.1 est présenté le résultat obtenus avec les moyennes de l'erreur de localisation du centre et du taux de recouvrement par les deux algorithmes Mean shift et Mean Shift avec LBP, pour les séquences d'images utilisées. On remarque que la moyenne de l'erreur de localisation du centre de tracker Mean Shift avec LBP égale à 13,52526 inférieur à MS (17,54934), et la moyenne du taux de recouvrement égale à 0,64432 supérieur à MS (0,609825), malgré toutes les problématiques (le changement d'échelle et la mouvement de camera et en présence d'un autre objet de couleurs similaires...). A travers ces résultats, on peut dire que, le tracker Mean shift avec LBP peut suivre l'objet cible de manière plus robuste et plus précision que le Mean shift.

Tableau 3.3 : Les moyennes des résultats obtenus de l'erreur de localisation du centre et du taux de recouvrement.

Séquence d'image	Erreur de localisation du centre		Taux de recouvrement	
	Mean Shift	Mean Shift avec LBP	Mean Shift	Mean Shift avec LBP
Divid3	11,1818	10,9174	0,6763	0,7002
Crossing	6,1421	5,5344	0,6468	0,6527
Couple	24,0402	11,7745	0,4372	0,5399
DragonBaby	21,4967	18,1387	0,5691	0,6172
BleurCar 4	24,8859	21,2613	0,6790	0,7116
Moyenne	17,5493	13,5252	0,6098	0,6443

3.7 Conclusion

Dans ce chapitre, nous avons présenté l'algorithme de suivi Mean shift qui utilise l'histogramme conjoint HSV couleur- LBP texture pour modéliser l'objet cible. On a implémenté cet algorithme sur un groupe des séquences d'images, chacune séquence contient les problématiques de suivi (occultation, mouvement de caméra rapide (flou de bougé) et changement d'échelle ainsi qu'en présence de similarité des couleurs entre l'objet et l'arrière-plan, ...). On a présenté aussi les résultats obtenus de Mean shift et Mean shift avec LBP de chacune séquence. Les résultats obtenus montrent que l'utilisation de l'histogramme conjoint HSV couleur- LBP texture dans le tracker Mean shift pour représenter le modèle de la cible donne robuste suivi comparé à l'histogramme de couleur. On peut dire que, la combinaison entre les caractéristiques de couleur HSV et de texture LBP peut améliorer la précision et la robustesse de tracker Mean shift et le rend la plus efficace aux occultations totale, mouvement rapide (flou bougé) et en présence la similarité entre objet et fond ou des objets similaires. On peut conclure que, l'histogramme conjoint HSV couleur- LBP texture est plus robuste que l'histogramme de couleur à cause d'utiliser l'information spatiale avec l'information couleur.

Conclusion générale

Conclusion générale

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.[2]. L'objectif de ce travail était d'étudier dans quelle mesure il était possible de tirer avantage de la combinaison des caractéristiques visuelles de couleur et de texture pour créer un histogramme conjoint plus robuste.

Dans le cadre de ce mémoire, nous avons intéressé à l'étude de l'algorithme de suivi d'objet Mean shift avec LBP dans une séquence d'images. L'algorithme Mean shift utilise l'histogramme pondéré conjoint en combinant entre les caractéristiques de couleur HSV et de texture LBP afin modéliser l'objet cible. Cet histogramme permet d'améliorer la précision et la robustesse de cet algorithme.

Dans ce mémoire on a implémenté les algorithmes Meanshift avec LBP et Meanshift, puis sont appliqués sur de nombreuses séquences d'images (base de données OTB) qui contiennent les problèmes du changement d'illumination et d'échelle, d'occlusion partielle, etc. L'étude effectuée a permis d'obtenir des résultats satisfaisants des points de vue qualitatif et quantitatif. Les résultats obtenus montrent que l'utilisation de l'histogramme conjoint HSV couleur- LBP texture dans le trackerMeanshift pour représenter le modèle de la cible donne des performances meilleures comparé à l'histogramme de couleur. On peut dire que, la combinaison entre les caractéristiques de couleur HSV et de texture LBP peut améliorer la précision et la robustesse de trackerMeanshift et le rend la plus efficace aux occultations totale, mouvement rapide (flou bougé) et en présence la similarité entre objet et fond ou des objets similaires.

En perspectives, nous proposons l'utilisation des autres caractéristiques (HOG, filtre de gabor,...) pour calculer l'histogramme conjoint combiné avec les caractéristiques de couleurs pour améliorer la robustesse de tracker Mean shift, ainsi que l'utilisation le filtre de kalman ou le filtrage particulière avec ce tracker pour améliorée le suivi dans le problème d'occultation totale.

Bibliographie

- [1] H.RAMI. Contribution au développement de descripteurs LBP pour le suivi d'objet en mouvement. Thèse doctorat, université Mohamed V. Avril 2016.
- [2] Leang, Isabelle. Fusion en ligne d'algorithmes de suivi visuel d'objet. Thèse de doctorat, université Pierre et Marie Curie .Décembre 2016.
- [3] Yilmaz A, Omar Javed, "Object Tracking: A Survey", ACM Computing Surveys, Vol. 38, No. 4, Article 13, Publication date: December 2006. pp. 1-45.
- [4] 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.
- [5] Shweta k. Talmale, Nitin 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.
- [6] Hingane, P., and S. Shirsat. "Object Tracking Using Joint Color-Texture Histogram." International Journal of Science and Research (2013): 2603-2606.
- [7] Medouakh, Saadia, Mohamed Boumehraz, and Nadjiba Terki. "Improved object tracking via joint color-LPQ texture histogram based mean shift algorithm." Signal, Image and Video Processing 12.3 (2018): 583-590.
- [8] Zhang, Kaihua, Lei Zhang, and Ming-Hsuan Yang. "Real-time object tracking via online discriminative feature selection." *IEEE Transactions on Image Processing* 22.12 (2013): 4664-4677.
- [9] 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.
- [10] 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.
- [11] Xiaorong, Pu, and Zhou Zhihu. "A more robust mean shift tracker on joint color-CLTP histogram." International Journal of Image, Graphics and Signal Processing 4.12 (2012): 34.
- [12]Sliti, Oumaima, Habib Hamam, and Hamid Amiri. "CLBP for scale and orientation adaptive mean shift tracking." Journal of King Saud University-Computer and Information Sciences(2017).
- [13] S.S. Blackman and R. Popoli. Design and analysis of modern tracking systems. Artech House,Norwood, MA, USA, 1999.
- [14] Hachour, Samir. *Suivi et classification d'objets multiples: contributions avec la théorie des fonctions de croyance*. Diss. Artois, 2015.

Bibliographie

- [15] M, Mikram. Suivi d'objets dans une séquence d'images par modèle d'apparence : conception et évaluation, Thèse de doctorat. Université de Toulouse, 2008
- [16] Trichet. R, suivi d'objet pour la télévision interactive, thèse de doctorat, L'école télécom Paris Tech, 08 décembre 2008.
- [17] Penne, Thomas. Développement d'un système de tracking vidéo sur caméra robotisée. Diss. Université Blaise Pascal-Clermont-Ferrand II, 2011.
- [18] Wang, Qing, et al. "An experimental comparison of online object-tracking algorithms." SPIE Optical Engineering+ Applications. International Society for Optics and Photonics, 2011.
- [19] Black, Michael J., and Allan D. Jepson. "Eigentracking: Robust matching and tracking of articulated objects using a view-based representation." European conference on computer vision. Springer, Berlin, Heidelberg, 1996.
- [20] B.Soufian. Détection et suivi d'objets par l'approche "Mean Shift ".Thèse master, université de Mohamed Khider Biskra .2016.
- [21] 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.
- [22] Ait fares W. F, Détection et suivi d'objets par vision fondes sur segmentation par contour actif base regio. Thèse doctorat, université de Toulouse. 2013.
- [23] W.Bouachir. Suivi d'objets par caractéristiques locales encodant la structure. Suivi d'objets par caractéristiques locales encodant la structure. Thèse de doctorat, université de Montréal. Décembre 2014.
- [24] Housni. K, Analyse d'événements dans une vidéo par les coupes de graphes, thèse de doctorat université ibn zohr, 22 décembre 2012.
- [25] Guillot. C, détection d'objets stationnaires par une paire de caméras PTZ. Autre, thèse doctorat Université Blaise Pascal - Clermont-Ferrand II, 2012.
- [26] F. Bousetouane. Traitement décentralisé des informations visuelles dans les réseaux de capteurs sans fil multimédia. Thèse doctorat , Université Badji Mokhtar –Annaba.2015.
- [27] Riahi, Dorra. Suivi multi-objets par la détection: Application à la vidéo surveillance. Diss. Ecole Polytechnique, Montreal (Canada), 2016.
- [28] Veenman et al C.J.Veenman.M.Reinders et E.Backer 2001.Resolving.Motion Correspondence for Densely Moving Points.IEEE Transaction on Patten Analysis and Machine Intelligence, vol pages 54-72.

Bibliographie

- [29] Brulin, Mathieu. Analyse sémantique d'un trafic routier dans un contexte de vidéo-surveillance. Diss. Bordeaux 1, 2012.
- [30] R. Cucchiara, C. Grana, M. Piccardi, and A. Prati. Detecting moving objects, ghosts, and shadows in video streams. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(10) :1337–1342, 2003
- [31] Huynh, Huu Hung. Vidéosurveillance pour appartements intelligents: application à la détection de prise de médicaments. Diss. Aix Marseille 2, 2010.
- [32] ALI, Anjum et AGGARWAL, J. K. Segmentation and recognition of continuous human activity. In : *Detection and recognition of events in video*, 2001. Proceedings. IEEE Workshop on. IEEE, 2001. p. 28-35.
- [33] EL MAADI, A. Vision infrarouge appliquée à la vidéosurveillance extérieure automatisée : reconnaissance, suivi, classification et détection d'événements. Thèse de doctorat. Université Laval. 2006.
- [34] Wang, N. et Yeung, D.-Y. (2013). Learning a deep compact image representation for visual tracking. In *Advances in neural information processing systems*, pages 809–817
- [35] Zhong, B., Yao, H., Chen, S., Ji, R., Chin, T.-J. et Wang, H. (2014). Visual tracking via weakly supervised learning from multiple imperfect oracles. *Pattern Recognition*, 47(3):1395 – 1410.
- [36] Danelljan, M., Häger, G., Khan, F. et Felsberg, M. (2014). Accurate scale estimation for robust visual tracking. In *British Machine Vision Conference*, Nottingham, September 1-5, 2014. BMVA Press.
- [37] Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., Huang, Z., Karpathy, A., Khosla, A., Bernstein, M., Berg, A. C. et Fei-Fei, L. (2015). ImageNet Large Scale Visual Recognition Challenge. *International Journal of Computer Vision (IJCV)*, 115(3):211–252.
- [38] Lucas, B. D., Kanade, T. et al. (1981). An iterative image registration technique with an application to stereo vision. In *IJCAI*, volume 81, pages 674–679.
- [39] Comaniciu, D., Ramesh, V. et Meer, P. (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. IEEE.
- [40] Kalal, Z., Mikolajczyk, K. et Matas, J. (2010b). Forward-backward error : Automatic detection of tracking failures. In *Pattern Recognition (ICPR), 2010 20th International Conference on*, pages 2756–2759. IEEE.

Bibliographie

- [41] Allen, J. G., Xu, R. Y. et Jin, J. S. (2004). Object tracking using camshift algorithm and multiple quantized feature spaces. In *Proceedings of the Pan-Sydney area workshop on Visual information processing*, pages 3–7. Australian Computer Society, Inc.
- [42] 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.
- [43] X. Zou, D. Li et J. Liu. Real-time vehicles tracking based on Kalman filter in an ITS. In *International Symposium on Photoelectronic Detection and Imaging : Technology and Applications 2007*, pages 662306-662306. International Society for Optics and Photonics, 2007.
- [44] Y. Du et F. Yuan. Real-time vehicle tracking by Kalman filtering and Gabor decomposition. In *Information Science and Engineering (ICISE), 2009 1st International Conference on*, pages 1386-1390. IEEE, 2009
- [45] A. Alin, M.V. Butz et J. Fritsch. Tracking moving vehicles using an advanced grid-based Bayesian filter approach. In *Intelligent Vehicles Symposium (IV), 2011 IEEE*, pages 466-472. IEEE, 2011.
- [46] Čehovin, L., Kristan, M. et Leonardis, A. (2011). An adaptive coupled layer visual model for robust visual tracking. In *2011 International Conference on Computer Vision*, pages 1363–1370. IEEE.
- [47] Gordon, N. J., Salmond, D. J. et Smith, A. F. (1993). Novel approach to nonlinear/non-gaussian bayesian state estimation. In *IEE Proceedings F-Radar and Signal Processing*, volume 140, pages 107–113. IET.
- [48] Kitagawa, G. (1996). Monte carlo filter and smoother for non-gaussian nonlinear state space models. *Journal of computational and graphical statistics*, 5(1):1–25.
- [40] Smeulders, Arnold WM, et al. "Visual tracking: An experimental survey." *IEEE transactions on pattern analysis and machine intelligence* 36.7 (2014): 1442-1468.
- [50] Vincent, G, suivi d'objets d'intérêt dans une séquence d'image des points saillants aux mesures, thèse de doctorat, l'Université de Nice .30 avril 2009.
- [51] Li, Xi, et al. "A survey of appearance models in visual object tracking." *ACM transactions on Intelligent Systems and Technology (TIST)* 4.4 (2013): 58.
- [52] Hua, Yang. Vers un suivi robuste d'objets visuels: sélection de propositions et traitement des occlusions. Diss. Grenoble Alpes, 2016.
- [53] Yang, Hanxuan, et al. "Recent advances and trends in visual tracking: A review." *Neurocomputing* 74.18 (2011): 3823-3831.

Bibliographie

- [54] H. Ghazouani. Suivi d'objets basé forme et couleur pour la navigation robotique en temps réel. 2014. <https://hal-lirmm.ccsd.cnrs.fr/lirmm-01077566>
- [55] Mei, X., Ling, H.: Robust visual tracking using l1 minimization. In: IEEE 12th International Conference on Computer Vision, pp. 1436–1443 (2009).
- [56] Ross, D.A., Lim, J., Lin, R.S., Yang, M.H.: Incremental learning for robust visual tracking. *Int. J. Comput. Vis.* 77(1–3), 125–141(2008).
- [57]. S. Avidan, “Support vector tracking,” *IEEE Transactions On Pattern Analysis And Machine Intelligence, PAMI*, 26(8):1064-1072, 2004.
- [58] Kalal, Z., Mikolajczyk, K., Matas, J.: Tracking-learning-detection. *IEEE Trans. Pattern Anal. Mach. Intell.* 34(7), 1409–1422 (2012).
- [59] Babenko, B., Yang, M.H., Belongie, S.: Robust object tracking with online multiple instance learning. *IEEE Trans. Pattern. Anal. Mach. Intell.* 33(8), 1619–1632 (2011).
- [60] Hare, S., Saffari, A. et Torr, P. H. (2011). Struck : Structured output tracking with kernels. *In ICCV, 2011 IEEE International Conference on*, pages 263–270. IEEE.
- [61] Vapnik, V. N. (1995). *The Nature of Statistical Learning Theory*. Springer-Verlag New York, Inc., New York, NY, USA.
- [62] Si Liu, Tianzhu Zhang, Xiaochun Cao, Changsheng Xu: Structural Correlation Filter for Robust Visual Tracking. *CVPR 2016*: 4312-4320
- [63] C. Ma, X. Yang, C. Zhang, and M.-H. Yang. Long-term correlation tracking. In *CVPR*, 2015.
- [64] C. Teulière, approches déterministes et bayésiennes pour un suivi robuste, application à l'asservissement visuel d'un drone, these de doctorat, Université Rennes 1, 2010.
- [65] Fukunaga, Keinosuke, and Larry 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
- [66] Lim, Hye-Youn, and Dae-Seong 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.
- [67] Comaniciu, D. & Meer, P, Mean shift: A robust approach toward feature space analysis, 2003.
- [68] Cheng, Yizong. "Mean shift, mode seeking, and clustering." *IEEE transactions on pattern analysis and machine intelligence* 17.8 (1995): 790-799.

Bibliographie

- [69] Comaniciu, Dorin, and Peter Meer. "Mean shift: A robust approach toward feature space analysis." *IEEE Transactions on pattern analysis and machine intelligence* 24.5 (2002): 603-619.
- [70] M.Radja. Développement d'un Système de Dénombrement de Cellules Sanguines Thèse de Mater, université Abou Bakr Belkaid– Tlemcen, 2013.
- [71] Li, X., Hu, W., Shen, C., Zhang, Z., Dick, A., & Hengel, A. V. D. (2013). A survey of appearance models in visual object tracking. *ACM transactions on Intelligent Systems and Technology (TIST)*, 4(4), 58.
- [72] LIM, Hye-Youn et KANG, Dae-Seong. Object tracking system using a VSW algorithm based on color and point features. *EURASIP Journal on Advances in Signal Processing*, 2011, vol. no 1, p. 1-12, 2011.
- [73] Artner, N. M. (2008, April). A comparison of mean shift tracking methods. In *12th Central European Seminar on computer graphics* (pp. 197-204).
- [74] Debeir. O, Poursuite temps réel d'objets dans une séquence d'images par algorithmes du type "Mean shift", mémoire de fin d'études (MFE), université libre de Bruxelles, 2005.
- [75] Jharna. M, kavya. G. Registration and Tracking by using Mean Shift Algorithm, *ages*, International Journal of Science and Research (IJSR). Volume 3 Issue 8, pp 358-363, August 2014.
- [76] CHIA, Y. S., KOW, Wei Yeang, KHONG, Wei Leong, et al. Kernel-based object tracking via particle filter and mean shift algorithm. In : *Hybrid Intelligent Systems (HIS)*, 2011 11th International Conference on. p. 522-527. IEEE, 2011.
- [77] Bugeau.A, Détection et suivi d'objet en mouvement dans des scènes complexes application à la surveillance des conducteurs, thèse de doctorat, université de rennes1,2007.
- [78] O.Mokhtar. Poursuite d'objets mobiles appliquée à la vidéo surveillance. Thèse de Magister, universite des sciences et de la technologie d'Oran, 2010.
- [79] KARAVASILIS, V, NIKOU, C, et LIKAS, A. Visual tracking by adaptive Kalman filtering and mean shift. In : *Artificial Intelligence: Theories, Models and Applications*. Springer Berlin Heidelberg p. 153-162, 2010.
- [80] Yang, H., Shao, L., Zheng, F., Wang, L., Song, Z.: Recent advances and trends in visual tracking: a review. *Neurocomputing* 74, 3823–3831 (2011).
- [81]T. Ojala, M. Pietikäinen, and T. Mäenpää, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 24, pp. 971-987, 2002.

Bibliographie

- [82] C. Ammar. Analyse d'images d'expressions faciales et orientation de la tête basée sur la profondeur. Thèse de doctorat , université Mohamed Khider Biskra , 2016.
- [83] B. Mohcene. Extraction de Caractéristiques d'images faciales par les méthodes LBP, BSIF, LPQ, et HOG, Application à la Biométrie. Thèse de Master, , université Mohamed Khider Biskra ,2015.
- [84] PIETIKÄINEN, Matti, HADID, Abdenour, ZHAO, Guoying, et al. Computer vision using local binary patterns. Springer Science & Business Media, 2011.
- [85] RAHIM, Abdur, AZAM, Md Shafiul, HOSSAIN, Nazmul, et al. Face Recognition using Local Binary Patterns (LBP). Global Journal of Computer Science and Technology5, 2013, vol. 13, no 4.
- [86]OJALA T, Pietikainen M and Harwood D (1996) A comparative study of texture measures with classification based on featured distribution. Pattern Recognition, 29(1) :51-59.
- [87] HUANG, Di, SHAN, Caifeng, ARDEBILIAN, Mohsen, et al. Facial Image Analysis Based on Local Binary Patterns : A Survey. IEEE Transactions on Image Processing, 2011.
- [88] HUANG, Di, SHAN, Caifeng, ARDABILIAN, Mohsen, et al. Local binary patterns and its application to facial image analysis : a survey. Systems, Man, and Cybernetics, Part C : Applications and Reviews, IEEE Transactions on, 2011, vol. 41, no 6, p. 765-781.
- [89] Pietikainen, M., Hadid, A., Zhao, G., Ahonen, T. (2011). Description of Interest Regions. In Computer Vision Using Local Binary Patterns (pp. 81-88). Springer London.
- [90] Y. GAFOUR. Détection de vidéos similaires. Thèse de Magister, université des Sciences et des Technologies Mohamed Boudiaf d'Oran, 2014.
- [91] Zhang, K., Zhang, L. et Yang, M.-H. (2012). Real-time compressive tracking. In Computer Vision–ECCV 2012, pages 864–877. Springer.
- [92] Babenko, B., Yang, M.-H. et Sivic, J. (2009). Visual tracking with online multiple instance learning. In Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on, pages 983–990. IEEE.
- [93] Kwon, J. et Lee, K. M. (2010). Visual tracking decomposition. In Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on, pages 1269–1276. IEEE.
- [94] Vojir, T., Matas, J. et Noh, J. (2015). Online adaptive hidden markov model for multi-tracker fusion. arXiv preprint arXiv : 1504.06103.
-

Bibliographie
