

**REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE**  
**Ministère de l'Enseignement Supérieur et de la Recherche Scientifique**  
**Université des Sciences et de la Technologie d'Oran U. S. T. O.**

**« Mohamed Boudiaf »**

**Faculté de Génie Electrique**

**Département d'Electronique**

**Laboratoire Signaux, Systèmes et Données**

## **Mémoire en vue de l'obtention du diplôme de Magister**

**Spécialité : Electronique**

**Option : Signaux et Systèmes**

Présenté par :

Melle MEDJAHED Fatiha

# **Détection et Suivi d'Objets en Mouvement Dans Une Séquence d'Images**

**SOUTENU LE**

**DEVANT LE JURY COMPOSE DE :**

	<b>Nom&amp;Prénom</b>	<b>Grade</b>	<b>Etablissement</b>
<b>PRESIDENT</b>	<b>Mr. N. BOUGHANMI</b>	<b>Professeur</b>	<b>USTO</b>
<b>RAPPORTEUR</b>	<b>Mr. M.F. BELBACHIR</b>	<b>Professeur</b>	<b>USTO</b>
<b>CO-RAPPOTEUR</b>	<b>Mme. N. BENMOUSSAT</b>	<b>M.C</b>	<b>USTO</b>
<b>EXAMINATEUR</b>	<b>Mr. A.DJEBBARI</b>	<b>Professeur</b>	<b>USBA</b>
<b>EXAMINATEUR</b>	<b>Mr. T. ZOUAGUI</b>	<b>M.C</b>	<b>USTO</b>

Année universitaire 2011/2012



## Sommaire

Remerciements.....	iv
Liste des figures.....	v
Introduction Générale.....	1
<b>CHAPITRE I : Généralités sur l'analyse d'image</b>	
I.1 Introduction.....	4
I.2 Image numérique.....	5
I.3 Bruit.....	6
a. Bruit lié aux dispositifs d'acquisition.....	6
b. Bruit lié à la scène .....	6
c. Bruit lié à la transmodulation .....	6
d. Bruit lié à l'échantillonnage .....	6
I.4 Filtrage .....	7
I.4.1 Notion du Filtre.....	7
I.4.2 Principe du filtrage.....	7
a. Filtrage global.....	7
b. Filtrage local.....	7
I.4.3 Types de filtrage.....	8
a. Les filtres passe-bas.....	8
b. Les filtres passe-haut.....	8
I.4.4 Techniques du filtrage.....	8
a. Filtres linéaires.....	8
b. Filtres non linéaires.....	10
I.4.5 Avantages du filtrage.....	11
I.5 Segmentation d'image.....	11
I.6 Détection de contour.....	12
I.6.1 Technique du gradient.....	13
I.7 Segmentation par morphologie mathématique.....	15
a. Dilatation.....	15
b. Erosion.....	16
c. Ouverture.....	16
d. Fermeture.....	17
I.8 Seuillage.....	17
I.9 Définition d'une vidéo.....	18
I.10 Résolution temporelle d'une vidéo.....	18
I.11 Séquence d'image.....	19
I.12 Analyse du mouvement dans une séquence d'images.....	19
I.12.1 Détection de mouvement.....	20
I.12.2 Définition d'un objet vidéo.....	22
I.12.3 Suivi d'objets dans la vidéo.....	22
I.13 Conclusion.....	23
<b>CHAPITRE II : Détection d'objets en mouvement</b>	
II.1 Introduction .....	24
II.2 Détection d'objets en mouvement par la méthode de Segmentation .....	24

a. Simplification d'image .....	25
b. Approximation du gradient .....	26
c. Détection des lignes de partage des eaux (watersheds).....	26
d. Fusion des régions.....	27
II.3 Détection par différence d'images .....	29
II.3.1 Approche région.....	29
a. Avec image de référence.....	29
b. Sans image de référence.....	31
II.3.2 Approche contour.....	33
a. Proposition de HAYNES.....	33
b. Proposition de STELMASZYK.....	35
c. Proposition de VIEREN.....	35
d. Proposition d'ORKISZ .....	37
II.4 Détection des contours en mouvement par (mcm).....	37
II.5 Conclusion.....	39
<b>CHAPITRE III : Suivi d'objets en mouvement</b>	
III.1 Introduction.....	40
III.2 Algorithmes de suivi.....	40
III.3 Classification des algorithmes de suivi.....	40
III.3.1 Les algorithmes orientés mouvement.....	40
III.3.2 Les algorithmes orientés modèle.....	41
III.3.3 Les méthodes basées sur les frontières.....	41
III.3.4 Les méthodes basées sur les régions .....	42
III.4 Structure général des algorithmes de suivi.....	42
III.4.1 Phase de prédiction.....	42
a. Prédiction de mouvement nul.....	42
b. Prédiction à court-terme simple.....	43
c. Prédiction à court-terme lissée.....	43
d. Prédiction à long-terme.....	43
III.4.2 Phase d'ajustement.....	43
a. Ajustement basé contour.....	43
b. Ajustement basé région .....	43
III.4.3 Phase d'estimation.....	44
a. Les méthodes par mise en correspondance.....	44
b. Les méthodes par transformées .....	44
c. Les méthodes différentielles.....	44
III.5 Suivi de trajectoire d'objet en mouvement.....	44
III.6 Le filtre de Kalman.....	46
III.7 Suivi de mouvement par filtre de Kalman.....	47
III.7.1 Principe du filtre de Kalman.....	47
III.7.2 Calcul de l'estimateur de Kalman.....	48
III.7.3 Expression des paramètres du filtre de Kalman.....	50
III.7.4 Dérivation de l'expression de la matrice de covariance.....	52
III.8 Optimalité .....	55
III.9 Application du filtre.....	56
III.9.1 Le modèle de mouvement.....	57
III.9.2 Le modèle de mouvement à vitesse constante .....	57
III.10 Conclusion .....	60
<b>CHAPITRE IV : Implémentation pratique et résultats expérimentaux</b>	
IV.1 Introduction.....	61

IV.2	Détection de mouvement dans une séquence d'image.....	61
IV.2.1	Détection par la méthode de segmentation.....	61
IV.2.2	Détection par la méthode de l'image de référence.....	65
IV.2.3	Détection sans l'image de référence .....	66
IV.2.4	Détection des contours en mouvement .....	67
IV.2.5	Détection des contours en mouvement par l'opérateur mcm .....	69
IV.3	Suivi d'objets en mouvement par filtre de Kalman.....	76
IV.4	Conclusion.....	86
	Conclusion Générale ET Perspectives.....	87

## ***Liste des figures***

### **CHAPITRE I : Généralités sur l'analyse d'image**

Figure I.1 application du filtre moyenneur, (a) : image d'origine, (b) : filtre 3×3, (c) : filtre 9×9

Figure I.2 application du filtre median, (a) : image d'origine, (b) : image bruitée (poivre et sel), (c) : filtre 3×3

Figure I.3 Exemple d'une segmentation d'image, (a) : image d'origine, (b) : image segmentée

Figure I.4 Exemple d'application du filtre Robert, (a) : image d'origine, (b) : contour de Robert

Figure I.5 exemple d'application du filtre Sobel, (a) : image d'origine, (b) : contour de Sobel

Figure I.6 exemple d'application de la dilation, (a) : image binaire, (b) : image dilatée

Figure I.7 exemple d'application de l'érosion, (a) : image binaire, (b) : image érodée

Figure I.8 Opération de seuillage. (a) : une image originale en 256 niveaux de gris, (b) et (c)

Figure I.9 Exemple d'une séquence d'images

Figure I.10 Exemple de vidéosurveillance, (a) : les cameras de surveillance, (b) : dans salle de contrôle : les hommes sont aidés par la machine dans la tâche de détection

### **CHAPITRE II : Détection d'objets en mouvement**

Figure II.1 : Schéma synoptique de la segmentation spatiale

Figure II.2 : Les minima locaux  $m_i$  et les bassins de captures  $BC_i$  associés

Figure II.3 Partitions et graphes d'adjacences associés

Figure II.4 Exemples de fusion de deux régions à partir d'un GAR de la partition

Figure II.5 : Technique de différence avec image de référence, (a) l'image à l'instant  $t$ , (b) l'image de référence, (c) l'image de différence

Figure II.6 : Technique de différence sans image de référence, (a) image précédente, (b) image courante, (c) image de différence

Figure II.7: extraction des contours en mouvement par la proposition de Haynes, (a) image précédente, (b) image courante, (c) image de différence, (d) contour d'image courante, (e) contour en mouvement

Figure II.8: Extraction des contours en mouvement par la proposition de Vieren.

Figure II.9 schéma synoptique de l'opérateur mcm

### **CHAPITRE III : Suivi d'objets en mouvement**

Figure III.1 Exemple de trajectoire mesurée et estimée

Figure III.2 le cycle du filtre de Kalman

Figure III.3 le cycle du filtre de Kalman avec les équations de prédiction et de correction.

Figure III.4 (a) Trame à l'instant  $t$ , (b) l'objet en mouvement détecté avec son centre de gravité, (c) trajectoire mesurée sur 5 trames, (d) représentation conjointe des trajectoires mesurée et estimée, (e) représentation conjointe des trajectoires mesurée et estimée sur l'image

### **CHAPITRE IV : Implémentation pratique et résultats expérimentaux**

Figure IV.1. Simplification d'image (Ouverture, Fermeture). (a) 5<sup>ème</sup> trame de la séquence Mother & Daughter, (b) Ouverture de la 5<sup>ème</sup> trame de la séquence, (c) Fermeture de la 5<sup>ème</sup> trame de la séquence

Figure IV.2. Simplification d'image (Ouverture, Fermeture). (a) 2<sup>ème</sup> trame de la séquence Akiyo, (b) Ouverture de la 2<sup>ème</sup> trame de la séquence, (c) Fermeture de la 2<sup>ème</sup> trame de la séquence

Figure IV.3. (a) La 5<sup>ème</sup> trame de la séquence Mother & Daughter dilatée, (b) La 5<sup>ème</sup> trame de la séquence érodée, (c) Image gradient de la 5<sup>ème</sup> trame de la séquence

Figure IV.4. (a) La 2<sup>ème</sup> trame de la séquence Akiyo dilatée, (b) La 2<sup>ème</sup> trame de la séquence érodée, (c) Image gradient de la 2<sup>ème</sup> trame de la séquence

Figure IV.5. Image segmentée par LPE, (a) 5<sup>ème</sup> trame de la séquence Mother & Daughter (b) 2<sup>ème</sup> trame de la séquence Akiyo

Figure IV.6. Graphe d'adjacence des régions, (a) 5<sup>ème</sup> trame de la séquence Mother & Daughter, (b) 2<sup>ème</sup> trame de la séquence Akiyo

Figure IV.7. Fusion des régions (a) 5<sup>ème</sup> trame de la séquence Mother & Daughter (b) 2<sup>ème</sup> trame de la séquence Akiyo

Figure IV.8. (a) Trame de séquence scène routière, (b) image de référence, (c) image de différence, (d) image filtré par filtre médian, (e) les objets détectés après seuillage

Figure IV.9. (a) Trame à l'instant  $t$ , (b) trame à l'instant  $t+1$ , (c) l'image de différence seuillée

Figure IV.10. (a) 1<sup>re</sup> Trame de la séquence « chercheur », (b) 18<sup>em</sup> trame de la séquence « chercheur », (c) image de différence, (d) image de différence seuillée

Figure IV.11. (a) Trame à l'instant  $t$ , (b) trame à l'instant  $t+1$ , (c) la différence entre les deux images, (d) gradient de l'image  $t+1$ , (e) contour détecté

Figure IV.12. Trame à l'instant  $t$ , (b) trame à l'instant  $t+1$ , (c) la différence entre les deux images, (d) gradient de l'image  $t+1$ , (e) gradient de l'image de différence, (f) contour détecté.

Figure IV.13. (a) 5<sup>ème</sup> Trame de la séquence « chercheur », (b) 15<sup>ème</sup> Trame de la séquence « chercheur », (c) 25<sup>ème</sup> Trame de la séquence « chercheur », (d) gradient de (a), (e) gradient de (b), (f) gradient de (c), (g) différence entre (d) et (e), (h) différence entre (e) et (f), (i) contour en mouvement

Figure IV.14. (a) 14<sup>ème</sup> Trame de la séquence « chercheur », (b) 15<sup>ème</sup> Trame de la séquence « chercheur », (c) 16<sup>ème</sup> Trame de la séquence « chercheur », (d) gradient de (a), (e) gradient de (b), (f) gradient de (c), (g) différence entre (d) et (e), (h) différence entre (e) et (f), (i) contour en mouvement

Figure IV.15. (a) Trame à l'instant  $t-1$ , (b) trame à l'instant  $t$ , (c) trame à l'instant  $t+1$ , (d) gradient de (a), (e) gradient de (b), (f) gradient de (c), (g) différence entre (d) et (e), (h) différence entre (e) et (f), (i) contour en mouvement

Figure IV.16. (a) Trame à l'instant  $t-1$ , (b) trame à l'instant  $t$ , (c) trame à l'instant  $t+1$ , (d) gradient de (a), (e) gradient de (b), (f) gradient de (c), (g) différence entre (d) et (e), (h) différence entre (e) et (f), (i) contour en mouvement

Figure IV.17. Détection d'objet en mouvement de séquence Mother & Daughter, (a) trame à l'instant  $t$ , (b) trame à l'instant  $t+1$ , (c) différence entre la trame (a) et (b) seuillée, (d) amélioration par les filtres de la morphologie mathématique, (e) objet mobile

Figure IV.18. Détection d'objet en mouvement de séquence scène routière, (a) image originale, (b) détection des contours en mouvement, (c) amélioration par les filtres de la morphologie mathématique, (d) masque des objets en mouvement, (e) objet mobile

Figure IV.19. Colonne 1 : 10 échantillons d'image (scène routière), colonne 2 :  
Objet détecté avec son centre de gravité

Figure IV.20. Trajectoire mesurée de l'objet en mouvement

Figure IV.21. Position mesurée/estimée de l'objet mobile

Figure IV.22. Trajectoire mesurée/estimée de l'objet en mouvement

Figure IV.23. Colonne 1 : 10 échantillons d'image de la séquence « chercheur », colonne 2 :  
Objet détecté avec son centre de gravité

Figure IV.24. Trajectoire mesurée de l'objet en mouvement

Figure IV.25. Position mesurée/estimée de l'objet mobile

Figure IV.26. Trajectoire mesurée/estimée de l'objet en mouvement

Figure IV.27. Colonne 1 : 2 échantillons d'image (Akiyo), colonne2 : Objet détecté avec son centre de gravité

Figure IV.28. Position mesurée/estimée de l'objet

## *Remerciements*

*Je tiens tout d'abord à remercier ALLAH le tout puissant de m'avoir donné le courage et la patience pour mener à bien ce modeste travail, qu'il soit béni et glorifié.*

*Le travail présenté dans ce mémoire a été élaboré au niveau du laboratoire de recherche signaux et systèmes et des données LSSD dirigé par Monsieur le Professeur BELBACHIR M.F, à qui je tiens à exprimer ma profonde gratitude pour la confiance qu'il m'a accordé en m'intégrant au sein de son laboratoire ainsi qu'avoir accepté d'être mon encadreur et dont le soutien permanent m'a permis de mener à bien ce mémoire.*

*J'adresse également mes remerciements à Mme BENMOUSSAT N, pour avoir accepté d'être mon co-encadreur. je le remercie aussi pour ses conseils, ses corrections et ses orientations.*

*Je tiens aussi à exprimer mes remerciements les plus respectueux à messieurs les membres de jury :*

*Merci au Pr. BOUGHANMI N de m'avoir fait l'honneur de présider mon jury.*

*Je suis très honoré que Messieurs Pr. DJEBBARI A et Dr. ZOUAGUI T, aient accepté de rapporter sur l'ensemble de mon travail.*

*J'associe à ces remerciements mes collègues du laboratoire LSSD.*

# ***Introduction générale***

# Introduction générale

Avec la généralisation de l'utilisation des images numériques, l'analyse du mouvement dans les vidéos s'est révélée être un outil indispensable pour des applications aussi diverses que la vidéo surveillance, la compression vidéo, l'imagerie médicale, la robotique, l'interaction homme-machine, l'analyse de séquences sportives...etc. En effet, les zones de mouvement d'une séquence d'images correspondent souvent à des événements sur lesquels un système de vision doit se focaliser.

L'analyse du mouvement est un vaste sujet qui englobe un certain nombre de problématiques. On peut notamment citer :

- la détection du mouvement, qui consiste à étiqueter chaque pixel d'une image suivant si il correspond ou non à une région en mouvement dans la scène,
- la détection des objets en mouvement, c'est-à-dire la détection d'un ensemble de régions d'intérêt en mouvement dans la scène tridimensionnelle observée,
- la segmentation basée mouvement de la scène, pour laquelle chaque région de l'image ayant un mouvement distinct des autres est détectée et segmentée,
- l'estimation du mouvement, qui consiste à estimer, à partir d'une séquence d'images, le mouvement apparent des objets composants une scène tridimensionnelle,
- le suivi de primitives ou de régions, dont le but est de déterminer la position de chaque primitive ou région dans l'image à chaque instant,
- la reconnaissance et la modélisation d'activités ou de gestes.

Les trois premières problématiques (détection du mouvement, détection des objets en mouvement et segmentation basée mouvement), sont en général une première étape pour des outils automatiques de vision par ordinateur. Ces outils peuvent avoir pour vocation, soit uniquement de détecter, soit de détecter et reconnaître, soit de détecter et suivre des objets pour, par exemple, analyser le comportement ou la trajectoire de ces objets.

Les autres problématiques sont toutes aussi importantes et nécessitent la mise en place de méthodes simples et robustes. Tous ces sujets font l'objet d'un grand nombre de travaux, mais il n'existe pas, à l'heure actuelle, d'algorithmes aboutis s'adaptant à n'importe quelle situation.

Le but de ce travail est la détection et le suivi d'objets en mouvement dans une séquence d'image.

Nous allons donc, dans un premier temps, mettre en place un algorithme de détection des objets en mouvement. On définit un objet comme une région fermée de mouvement quasi constant.

Le détecteur d'objets en mouvement que nous proposons n'exploite pas la cohérence temporelle pouvant exister entre les détections à des instants différents. Il n'y a donc aucune corrélation entre les détections à deux instants consécutifs. L'exploitation des détections pour un grand nombre d'applications (surveillance, analyse de séquences sportives, compression ...) nécessite l'ajout de cette cohérence temporelle. Nous avons donc, dans un deuxième temps, proposé une méthode de suivi.

Le problème de suivi temporel bidimensionnel en vision numérique est de générer une inférence au sujet du mouvement perçu d'un objet dans une séquence d'images successives.

Les données connues pour résoudre le problème consistent en un modèle de déplacement de l'objet à suivre et certaines mesures à partir d'une séquence d'images. Ces mesures sont le plus souvent des points caractéristiques ou des régions, mais à peu près n'importe quoi de mesurable dans l'image peut être utilisé. Ces mesures ne sont pas nécessairement bonnes, en effet, du bruit dans l'image peut créer de faux points caractéristiques, par exemple, et l'algorithme de détection peut s'y méprendre. Le problème du suivi est résolu en pratique en utilisant l'inférence statistique. L'objet en mouvement est conceptualisé comme ayant un état interne propre à lui à chaque quantum temporel (à chaque trame dans la séquence). Le problème revient donc à estimer l'état de l'objet en combinant à son état les mesures effectuées. Si la dynamique du mouvement et de la mesure sont linéaires, le problème a une solution efficace et éprouvée, le filtre de Kalman.

Nous décomposons ce travail en quatre chapitres.

Chapitre 1 : Dans le premier chapitre, on présente des généralités sur l'image numérique et l'analyse de mouvement dans une séquence d'images.

Chapitre 2 : Le deuxième chapitre présente un état l'art des méthodes de détection et de segmentation du mouvement. Nous différencions les méthodes de détection basées sur la segmentation par ligne de partage des eaux (LPE) et les méthodes basées sur la différence inter-images ; celles reposant sur une image de référence et celles extrayant les contours d'objets en mouvement. Ceux-ci sont extraits à l'aide d'un opérateur de mesure des contours en mouvement

(mcm) basé sur la différenciation des gradients spatiaux de trois images successives dans la séquence. Toutes ces différentes techniques ont leurs propres avantages et leurs limites.

Chapitre 3 : le troisième chapitre traite de l'estimation du mouvement servant comme complément à la détection ; un bref état de l'art des méthodes de suivi de mouvement est fourni, et nous détaillons la méthode proposée qui est le suivi de la trajectoire d'objet par filtre de kalman.

Le chapitre 4 conclut ce mémoire en présentant l'ensemble des résultats expérimentaux obtenus pour la détection et le suivi d'objets en mouvement.

Une conclusion synthétisera les contributions et les résultats obtenus et proposera des perspectives à nos travaux.

# *Chapitre I*

*Généralités sur  
l'analyse d'image*

## **I.1 Introduction**

Le XXI<sup>e</sup> siècle est celui de l'image numérique, vecteur d'information pour la presse, les artistes, le monde médical mais aussi celui de l'industrie. Le traitement d'images est une science récente qui a pour but d'offrir aux spécialistes de différents domaines, comme au grand public, des outils de manipulation de ces données digitales issues du monde réel. L'histoire commence dans les années cinquante avec les rayons X. Puis vint l'étude de méthodes d'extraction automatique d'informations avec de nombreuses applications en reconnaissance de caractères et en contrôle qualité. [1]

Dans les années quatre-vingt, les efforts se poursuivirent avec l'introduction des systèmes experts qui devaient remplacer l'homme dans ses différents diagnostics ! Enfin dans le courant des années quatre-vingt et quatre-vingt-dix, le traitement d'images passe du 2D au  $nD$  grâce aux nouvelles possibilités technologiques. Les supports de stockage, les processeurs permettent désormais de traiter des informations multidimensionnelles de couleur et de mouvement.

Puis dans les années deux mille, l'intérêt se porte de plus en plus sur les modèles perceptifs afin d'optimiser les modèles usuels. Par ailleurs, plutôt que de développer des algorithmes utilisables dans toutes les situations, des méthodes dédiées sont développées, en intégrant de la connaissance à priori sur les contenus et surtout sur la finalité de l'exploitation des données.

Le traitement d'images est donc une science en pleine évolution qu'il n'est pas possible de décrire pleinement dans ces quelques pages. Nous allons tenter de donner les bases de cette science qui a pour objectifs principaux :

- de saisir une scène réelle pour constituer un fichier informatique, c'est l'étape d'acquisition.
- de mettre en évidence les informations intéressantes : c'est l'étape de traitement dans laquelle on distinguera tout particulièrement les étapes de débruitage puis la segmentation.

- d'interpréter ces informations pour décider d'un diagnostic ou d'une action à engager : c'est l'étape d'analyse.
- sans oublier les étapes de stockage des informations, à différentes étapes de la chaîne, et de visualisation et/ou d'impression des résultats.

## **I.2 Image numérique :**

Une image numérique est une matrice de pixels repérés par leur coordonnées (x, y). S'il s'agit d'une image couleur, un pixel est codé par 3 composantes (r, g, b) (chacune comprise au sens large entre 0 et 255), représentant respectivement les doses de rouge, vert et bleu qui caractérisent la couleur du pixel.

S'il s'agit d'une image en niveau de gris, il est codé par 1 composante comprise au sens large entre 0 et 255, représentant la luminosité du pixel.

Exemples de couleurs :

- (0, 0,0)=noir.
- (255, 0,0)=rouge.
- (0, 255,0)=vert.
- (0, 0,255)=bleu.
- (127, 127,127)=gris moyen.
- (255, 255,255)=blanc.

Pour transformer une image couleur en niveau de gris on remplace les composantes de chaque pixel par sa valeur de luminosité : (r, g, b) devient (l, l, l) où  $l=(r+g+b)/3$ .

### **I.3 Bruit [2]**

En traitement d'image, toute brusque fluctuation d'un pixel par rapport à ses voisins étant considéré comme un bruit provoquant ainsi la dégradation de la qualité d'image que se soit de luminance ou de contraste. Ses sources les plus fréquentes sont :

#### **a. Bruit lié aux dispositifs d'acquisition**

La caméra, les câbles et tous autres blocs de connexion engendrent un bruit qui a un effet multiplicatif, pour cela, il serait plus judicieux de mettre les sources qui produisent plus de bruit dans les étages inférieurs pour diminuer leur contribution générale au bruit général.

#### **b. Bruit lié à la scène**

Il se trouve parfois que la scène soit polluée de poussière, ou couverte de nuages, brouillard ou buée qui dégrade la qualité des images. Cependant, là encore, la connaissance a priori du phénomène perturbateur permet d'envisager une modélisation et donc une correction systématique. Ce type de bruit a un effet de bruit additif.

#### **c. Bruit lié à la transmodulation**

Les images qui sont capturées à partir d'une séquence vidéo qui regroupe le son et l'image dans un même signal sont atteintes par un bruit du à la transmodulation qui se produit du fait qu'on utilise une double modulation : modulation audio et vidéo. Il arrive, ainsi, des cas où des composantes du signal audio brulent notre signal vidéo.

#### **d. Bruit lié à l'échantillonnage**

Une image est un signal discret. Il est donc nécessaire de passer du domaine continu au domaine discret au niveau de la scène et de l'intensité lumineuse émanant de celle ci. De manière générale, le spectre des intensités lumineuses noir et blanc est quantifié sur 256 niveaux de gris différents si on admet que chaque pixel est codé sur un octet.

Pour éliminer le bruit et rendre l'image plus nette, on utilise les techniques de filtrage, qui seront abordées dans la section suivante.

## I.4 Filtrage

### I.4.1 Notion du Filtre

Un filtre est souvent une transformation mathématique (appelée produit de convolution) permettant pour chaque pixel de la zone à laquelle il s'applique de modifier sa valeur en fonction des valeurs des pixels avoisinants, affectées de coefficients.

Le filtre est représenté par un tableau (matrice), caractérisé par ses dimensions et ses coefficients, dont le centre correspond au pixel concerné. Les coefficients du tableau déterminent les propriétés du filtre. Voici un exemple de filtre 3 x 3 :

$$\begin{matrix} 1 & 1 & 1 \\ 1 & 4 & 1 \\ 1 & 1 & 1 \end{matrix}$$

### I.4.2 Principe du filtrage:

Le filtrage consiste à modifier la valeur des pixels de tout ou partie d'une image numérique, généralement dans le but d'améliorer son aspect lorsque l'image ne peut pas être exploitée directement. En pratique, il s'agit de créer une nouvelle image en se servant des valeurs des pixels de l'image d'origine entièrement ou partiellement :

#### a. Filtrage global

Où chaque pixel de la nouvelle image est calculé en prenant en compte la totalité des pixels de l'image de départ.

#### b. Filtrage local

Où chaque pixel de la nouvelle image est calculé en prenant en compte seulement un voisinage du pixel correspondant dans l'image d'origine. Il est d'usage de choisir un voisinage carré et symétrique autour du pixel considéré. Ces voisinages sont donc assimilables à des tableaux à deux dimensions (matrices) de taille impaire.

N'entrent pas dans la catégorie du filtrage toutes les transformations de l'image d'origine : zoom, découpage, projections.

### I.4.3 Types de filtrage [3]

On distingue généralement les types de filtres suivants :

#### a. Les filtres passe-bas

Consistant à atténuer les composantes de l'image ayant une fréquence haute (pixels foncés). Ce type de filtrage est généralement utilisé pour atténuer le bruit de l'image, c'est la raison pour laquelle on parle habituellement de "lissage", "débruitage" ou encore "filtre antibruit". Les coefficients du masque sont positifs et leur somme est égale à 1.

#### b. Les filtres passe-haut

À l'inverse des passe-bas, atténuent les composantes de basse fréquence de l'image et permettent notamment d'accentuer les détails et le contraste, et donc les différences entre les pixels voisins. C'est la raison pour laquelle le terme de "filtre d'accentuation" est parfois utilisé. Les coefficients du masque peuvent être de signe quelconque mais leur somme est égale à 0.

### I.4.4 Techniques du filtrage

Il en existe plusieurs, Chaque filtre cherche à atténuer un type de défaut bien précis. Il n'y a pas de filtre universel capable de corriger tous les défauts. Il faut choisir les bons filtres suivant les défauts que nous désirons corriger.

#### a. Filtres linéaires

Le filtre local est dit linéaire si la valeur du nouveau pixel est une combinaison linéaire des valeurs des pixels du voisinage y compris le pixel en question.

##### ➤ Filtre moyenneur

Cette technique est la plus directe pour adoucir une image trop heurtée dans ses contours. Le niveau de gris de chaque pixel est remplacé par la moyenne des niveaux de gris de ses huit voisins. Ce filtrage pourra se faire à l'aide du filtre :

$$\mathbf{H} = \frac{1}{9} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

Le facteur  $1/9$  sert à normaliser le filtre de manière que celui-ci n'influe pas sur l'intensité globale de l'image.

Cette technique a pour défaut de créer un flou sur les bords des objets.

Pour une image d'entrée  $I_e$  on a :

Is image de sortie tel que :

$$\begin{aligned}
 I_s &= 1/9 \{I_e(x-1,y-1) + I_e(x-1,y) + I_e(x-1,y+1) + I_e(x,y-1) + I_e(x,y) + I_e(x,y+1) \\
 &+ I_e(x+1,y-1) + I_e(x+1,y) + I_e(x+1,y+1)\}. \\
 &= 1/9 \sum \sum H(i+1,j+1) \cdot I_e(x+i,y+j).
 \end{aligned}
 \tag{I.1}$$

et  $i \in [-1,+1] ; j \in [-1,+1]$

De manière plus générale, pour un filtre  $H$  de taille  $n+1 \times n+1$  et la somme des coefficients vaut  $K$  (facteur de normalisation) :

$$I_s = 1/k \sum \sum H(i+n/2, j+n/2) \cdot I_e(x+i, y+j).
 \tag{I.2}$$

Avec  $n$  : pair,  $i \in [-n/2,+n/2] , j \in [-n/2,+n/2]$ .

La formule précédente n'est autre que le produit de convolution discrète.



Figure I.1 application du filtre moyennneur, (a) : image d'origine, (b) : filtre 3×3, (c) : filtre 9×9

### b. Filtres non linéaires:

Si le filtre ne peut pas être exprimé par une combinaison linéaire, il est appelé " non linéaire

". Les filtres non-linéaires sont plus complexes à mettre en œuvre que les filtres linéaires (en termes de complexité algorithmique et de temps de traitement). Cependant les résultats obtenus avec les filtres non-linéaires sont très souvent de meilleure qualité que ceux obtenus par les filtres linéaires.

#### ➤ Filtre median

Consiste à remplacer l'intensité de chaque pixel par la valeur médiane de ses voisins choisis à l'avance. C'est un filtre passe bas qui élimine le bruit de type neige, mais il entraîne cependant une petite perte de résolution. son principe est :

Pour chaque pixel on classe les pixels voisins du pixel courant, par valeurs croissantes ,on prend la valeur médiane des pixels classés et on l'affecte au pixel courant .

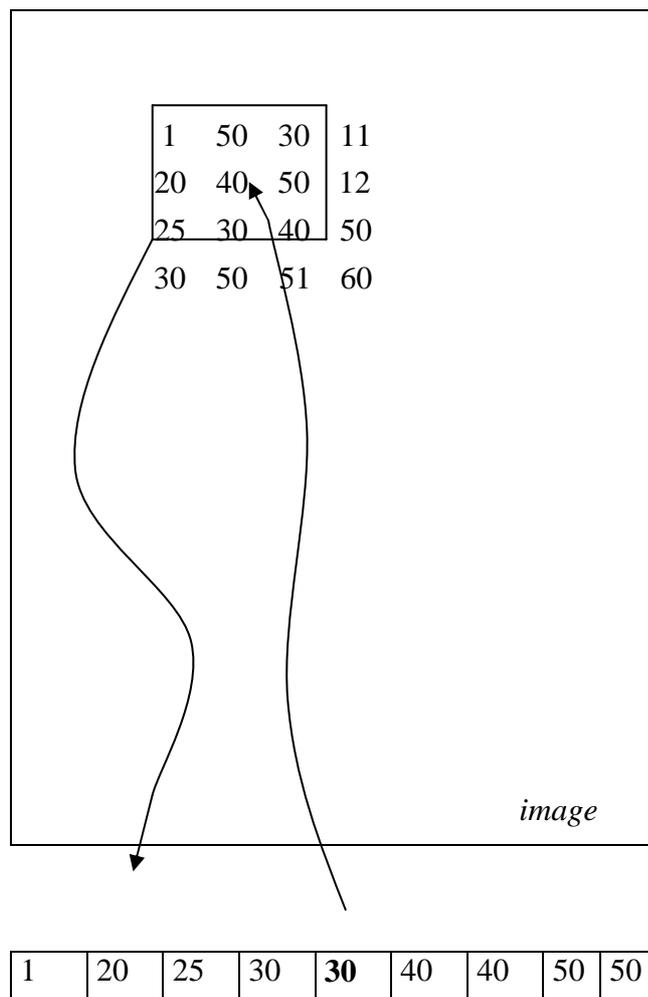




Figure I.2 application du filtre median, (a) : image d'origine, (b) : image bruitée (poivre et sel), (c) : filtre 3×3.

#### I.4.5 Avantages du filtrage

On citera comme avantage :

- Accentuer les variations d'intensité de l'image : renforcer les différences entre les pixels appartenant à des régions adjacentes (rehaussement de contraste).
- Détection de contours et préservation de la forme des régions.
- Renforcement de l'homogénéité à l'intérieur des régions (réduction de bruit). [3][4]

Cependant, Le résultat obtenu suite à un filtrage dépend énormément de la qualité du signal de l'image d'origine. Si le signal est très détérioré, le filtrage seul sera très peu efficace. Dans ce genre de situation le filtrage est juste une première étape dans un traitement plus complexe qui fera intervenir des informations sémantiques sur l'image.

#### I.5 Segmentation d'image

Les images sont composées de régions possédant des propriétés locales qui peuvent être la répartition des niveaux de gris. En regroupant des points de l'image qui possèdent une même propriété donnée, on obtient des régions uniformes. Cette opération s'appelle «segmentation».

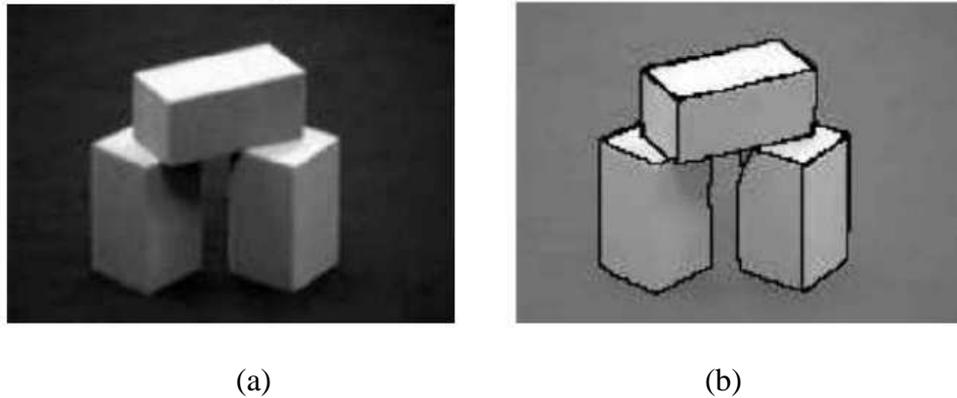


Figure I.3 Exemple d'une segmentation d'image, (a) : image d'origine, (b) : image segmentée.

Une autre méthode de segmentation basée sur les discontinuités des niveaux de gris, celles-ci indiquent les frontières entre les régions de l'image. La variation brusque des niveaux de gris s'appelle «contour» (un contour correspond à une forte variation de la fonction intensité).

### I.6 Détection de contour

La détection de contours est très utilisée comme étape de prétraitement pour la détection d'objets, pour trouver les limites de régions. En effet, un objet peut être localisé à partir de l'ensemble des pixels de son contour. De plus, trouver cet ensemble permet d'obtenir une information sur la forme de l'objet. Du point de vue théorique, un contour est défini par un changement marqué de l'intensité d'un pixel à l'autre. Comme en chaque point, le contour est considéré comme perpendiculaire à la direction du gradient, ce dernier est utilisé pour la détection : selon les modèles classiques de traitement d'image, en tout point de l'image les dérivées partielles sont estimées. Ceux qui fournissent les réponses les plus fortes correspondent à des pixels de contour.

En pratique, le calcul de gradient est effectué par la convolution de filtres linéaires. Ces filtres peuvent estimer les dérivées premières, les dérivées secondes, etc., et se focaliser sur des contours de directions différentes. Il existe ainsi de nombreux filtres, qui ont des fondements mathématiques différents. Cependant, quelle que soit leur justification théorique, leur objectif reste le même : mettre en évidence les pixels qui ont une valeur très différente de leurs voisins. Nous ne présentons ici, à titre d'exemple, que le filtre gradient et les opérateurs de Robert et Prewitt parmi les plus couramment employés.

**I.6.1 Technique du Gradient**

Soit la fonction image  $F(x, y)$ , le gradient de  $F$  en un point  $(x, y)$  est un vecteur bidimensionnel qui s'écrit :

$$\vec{G} = \begin{bmatrix} \frac{\partial F}{\partial X} \\ \frac{\partial F}{\partial Y} \end{bmatrix} \tag{I.3}$$

Le gradient  $\vec{G}$  est calculé en chaque point  $(x, y)$  de l'image, son orientation va dans le sens de la variation maximale de  $F$  et sa norme est généralement donnée par :

$$|\vec{G}| = \left( \left( \frac{\partial F}{\partial X} \right)^2 + \left( \frac{\partial F}{\partial Y} \right)^2 \right)^{1/2} \tag{I.4}$$

Localement, la norme du gradient est d'autant plus élevée que les courbes de niveau sont rapprochées et de valeurs différentes.

La direction de  $\vec{G}$  par rapport à l'axe des X est donnée par l'angle :

$$\vartheta = \text{arctg} \left( \frac{\partial F / \partial Y}{\partial F / \partial X} \right) \tag{I.5}$$

Le calcul du gradient est utile pour mettre en évidence les contours. Le point  $(x, y)$  est un contour si La norme est supérieure à un seuil.

➤ **Opérateurs de Robert**

$$G_x = \begin{bmatrix} -1 & 0 & 0 & -1 \\ 0 & 1 & 1 & 0 \end{bmatrix} \quad G_y = \begin{bmatrix} 0 & 1 & 1 & 0 \\ -1 & 0 & 0 & -1 \end{bmatrix}$$

On fait la convolution de ces masques avec l'image originale et on trouve les résultats suivants.



Figure I.4 Exemple d'application du filtre Robert, (a) : image d'origine, (b) : contour de Robert

➤ Opérateurs de Sobel et de Prewitt

$$\begin{array}{r}
 \begin{array}{ccc} 1 & 0 & -1 \\ c & 0 & -c \\ 1 & 0 & -1 \end{array} & \begin{array}{ccc} 1 & c & 1 \\ 0 & 0 & 0 \\ -1 & -c & -1 \end{array} \\
 G_x = & G_y =
 \end{array}$$

$c = 2$  pour *Sobel*,  $c=1$  pour *Prewitt*.



Figure I.5 exemple d'application du filtre Sobel, (a) : image d'origine, (b) : contour de Sobel.

Les opérateurs de *Sobel* et *Prewitt* sont fréquemment utilisés, pour leurs simplicité et leur faible coût d'évaluation, ils sont très performants dans les scènes de faible bruit et texture.

### I.7 Segmentation par morphologie mathématique

Le principe de la morphologie mathématique repose sur le principe suivant :

On considère un élément structurant à centre  $G$  cet élément structurant et formé d'un ensemble de points forme une forme géométrique. L'élément structurant constitue le concept de base de la morphologie mathématique. Il est défini par un ensemble de points constituant une forme géométrique plane. Le centre  $G$  est l'élément principal qu'on pose sur le pixel pour faire appliquer la morphologie binaire. Donc la segmentation par morphologie mathématique consiste à reconstruire l'objet à segmenter (isolé) par l'élément structurant.

#### a. Dilatation

On fait passer le centre de l'élément structurant sur tout les pixels de l'image et seuls les pixels pour les quels l'intersection entre l'élément structurant et la zone à dilater seront garder pour construire la zone dilatée.

Elle consiste à éliminer les points noirs isolés au milieu des parties blanches.

Soit  $E^N$  un espace euclidien à  $N$  dimension et soit  $I$  et  $S$  deux sous ensembles de  $E^N$ , la dilatation de  $I$  par  $S$  est définie à partir de l'addition de Minkowski par :

$$I \oplus S = \{ c \in E^N / c = i + s, i \in I, s \in S \} \quad (I.6)$$



(a)



(b)

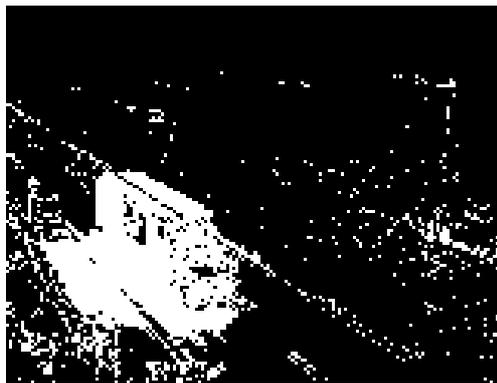
Figure I.6 exemple d'application de la dilatation, (a) : image binaire, (b) : image dilatée.

### b. Erosion

Dans l'opération de construction on pose le centre G de l'élément structurant sur les pixels (appartenant à la région à segmenter) et on vérifie si tout l'élément structurant est inclus totalement dans l'objet à segmenter. L'opération d'érosion consiste à prendre que les pixels pour lesquels l'élément structurant est inclus dans l'objet.

Dans l'opération d'érosion les points blancs isolés disparaissent. Soit  $E^N$  un espace euclidien à N dimension et soit I et S deux sous ensembles de  $E^N$ , l'érosion de I par S est définie à partir de la soustraction de Minkowski par :

$$I \ominus S = \{ c \in E^N / i = c + s, i \in I, \forall s \in S \} \quad (I.7)$$



(a)



(b)

Figure I.7 exemple d'application de l'érosion, (a) : image binaire, (b) : image érodée.

### c. Ouverture

L'ouverture de I par S notée  $I \circ S$  est le résultat d'une érosion de I suivie d'une dilatation de l'ensemble érodé par le même élément structurant.

$$I \circ S = (I \ominus S) \oplus S \quad (I.8)$$

L'ouverture adoucit les contours et élimine les pics aigus.

#### d. Fermeture

C'est l'opération duale de l'ouverture, notée  $I \bullet S$ , c'est le résultat d'une dilatation suivie d'une érosion en utilisant le même élément structurant.

(I.9)

$$I \bullet S = (I \oplus S) \ominus S$$

La fermeture fusionne les coupures étroites, élimine les petits trous, et comble les vides sur les contours.

#### I.8 Seuillage

L'opération non linéaire dite de seuillage simple consiste à mettre à zéro tous les pixels ayant un niveau de gris inférieur à une certaine valeur (appelée seuil, en anglais *threshold*) et à la valeur maximale (souvent la valeur 1) les pixels ayant une valeur supérieure.

Ainsi le résultat du seuillage est une image binaire contenant des pixels noirs et blancs, c'est la raison pour laquelle le terme de binarisation est parfois employé. Le seuillage permet de mettre en évidence des formes ou des objets dans une image. Toutefois la difficulté réside dans le choix du seuil à adopter.



(a)



(b)



(c)

Figure I.8 Opération de seuillage. (a) : une image originale en 256 niveaux de gris, (b) et (c)

Cette convention de seuillage s'appelle **seuillage supérieur** (*threshold above*), ainsi, on pourra définir le seuillage inverse dit **inférieur** (*threshold below*). Il est encore possible de définir deux valeurs de seuil, Respectivement borne inférieure et borne supérieure, afin de mettre à la valeur maximale les pixels ayant une valeur comprise entre les bornes et à zéro l'ensemble des autres valeurs. Il s'agit d'un seuillage intérieur (*threshold inside*), et par opposition, on définit le **seuillage extérieur** (*threshold outside*) qui est son inverse.

### I.9 Définition d'une vidéo

La vidéo est une succession d'images animées défilant à une certaine cadence afin de créer une illusion de mouvement pour l'œil humain. Elle peut être analogique (signal continu d'intensité de luminance) ou numérique (suite de trames ou images).

On peut distinguer deux types de vidéo:

- **la vidéo entrelacée:** où chaque image est formée de deux champs entrelacés. Le premier champ contenant uniquement les lignes impaires de l'image, le second champ, quant à lui, contenant les lignes paires (les signaux de télévision nord-américain et européen sont entrelacés).
- **la vidéo progressive:** dans une vidéo progressive, contrairement au mode entrelacé, toutes les lignes d'une image sont lues en une passe (vidéo numérique).

### I.10 Résolution temporelle d'une vidéo

La résolution temporelle d'une vidéo est définie par le nombre d'images défilant par seconde.

Afin d'éviter les désagréments dus aux papillotements, et prenant en considération les spécificités de l'œil humain qui garde une image environ 10 micro secondes au niveau de sa rétine ; les images doivent défiler à une certaine cadence:

- environ 24 images/s pour un film de cinéma.
- 25 images/s pour la télévision européenne.
- environ 30 images/s pour la télévision nord-américaine et japonaise.

### I.11 Séquence d'image

Une séquence d'images est une succession d'images bidimensionnelles qui montre l'évolution temporelle d'une scène. La cadence est de 25 images par seconde, ce qui correspond au seuil à partir duquel l'œil humain perçoit la séquence comme un stimulus continu, grâce à la persistance rétinienne. Par la suite, nous appellerons « trame » ou « plan » chaque image bidimensionnelle correspondant à un instant donné de la séquence.

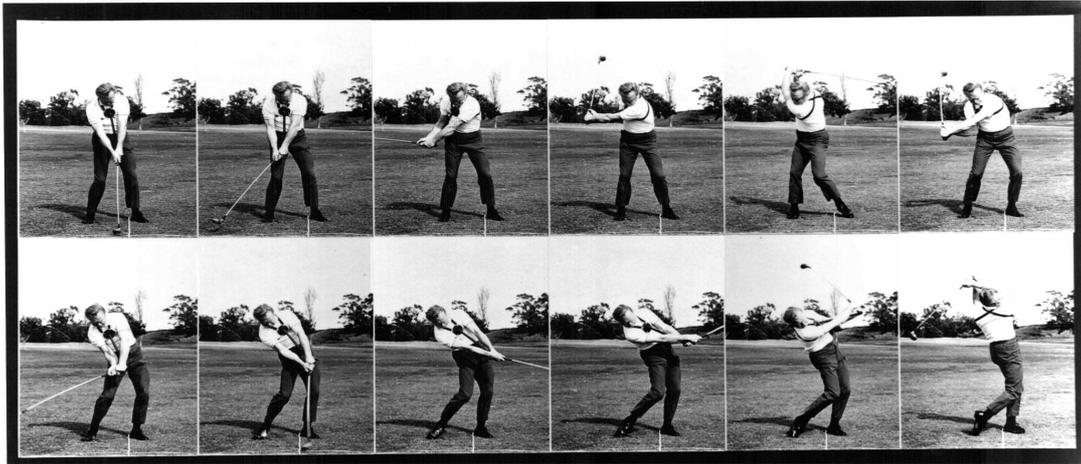


Figure I.9 Exemple d'une séquence d'images.

La considération de séquences d'images numériques ordonnées dans le temps introduit un champ spécifique d'investigation en analyse d'image, à savoir l'extraction d'informations de mouvement et plus généralement d'informations spatio-temporelles. Ce type d'études a émergé dans les années 70 à travers trois applications principales : le codage d'images de télévision et notamment la conception de codeurs avec compensation de mouvement [5-6], l'imagerie météorologique avec l'estimation de champs de vent via des traceurs nuageux [7], le suivi de cibles mobiles dans un contexte militaire [8].

### I.12 Analyse du mouvement dans une séquence d'images

L'analyse des séquences vidéo est l'un des sujets de recherche les plus tentants. Peut être que ceci découle de la richesse en informations que présente une scène ou encore de la variété des applications (médicales, sécuritaires, militaires, ...etc.). Les spécialistes du domaine ont tenté, et tentent encore d'imiter via le système caméra/calculateur le comportement humain en termes de vision. Pour y parvenir différents axes ont été développés. Ainsi apparaît le terme de "groupement spatio-temporel" [5].

C'est le terme qui englobe à la fois l'identification des objets et la détection du mouvement.



(a)



(b)

Figure I.11 Exemple de vidéosurveillance, (a) : les cameras de surveillance, (b) : dans salle de contrôle : les hommes sont aidés par la machine dans la tâche de détection.

L'analyse de mouvement dans une séquence d'images doit être posée comme un problème d'estimation et de segmentation, puisqu'il s'agit d'appréhender des informations partielles observables et discontinues. C'est un problème particulièrement difficile, mais dont la résolution est cruciale pour la plupart des tâches en analyse de scène dynamique. Ce problème peut se présenter en fait sous plusieurs variantes, suivant que l'objectif prioritaire se trouve être l'obtention d'une mesure dense ou paramétrique du mouvement, d'une partition de l'image en régions, ou l'extraction d'entités pertinentes

L'analyse du mouvement dans une séquence d'images repose soit sur la détection du mouvement soit sur la segmentation basée mouvement soit sur la poursuite de mouvement.

### I.12.1 Détection de mouvement

La détection du mouvement dans une séquence d'images consiste à distinguer les zones fixes et mobiles d'une scène. Comme le mouvement d'objets induit des différences temporelles entre images, sa détection s'appuie sur l'étude des variations temporelles de la fonction de luminance : si le capteur est fixe et que l'éclairage de la scène varie peu, alors nous supposons que toute variation temporelle de l'intensité est liée au mouvement d'un objet ou à du bruit.

➤ **Détection de mouvement (Motion détection)**

Ce terme est le plus Générique, il indique uniquement que l'on parle d'une méthode qui a pour objet de trouver en quels points de l'image un mouvement a eu lieu. Un algorithme ayant cet objectif fournit en sortie une variable quantitative (quantité de mouvement) ou qualitative (booléenne) pour tout pixel de chaque image d'entrée. Toutes les méthodes présentées ci-après rentrent dans cette catégorie.

➤ **Estimation du mouvement (Motion estimation)**

Cette notion inclut la précédente en y ajoutant la contrainte que le résultat fourni doit être quantitatif. Ce terme est surtout utilisé par les auteurs travaillant sur l'estimation du flot optique ; dans ce cas, le résultat de l'estimation est un champ de vecteurs à deux dimensions représentant la projection sur le plan de l'image du mouvement réel tridimensionnel ayant lieu dans la scène. Pour un état de l'art détaillé et une étude comparative des méthodes classiques de calcul du flot optique, voir [9].

➤ **Modélisation de l'arrière-plan (Background Modelling)**

Cette catégorie regroupe toutes les méthodes de détection de mouvement qui consistent à créer un modèle de l'arrière-plan de la scène filmée (sans aucun objet mobile). Ce modèle peut être une image créée à partir des pixels observés à différents instants de la séquence vidéo [10].

➤ **Soustraction de l'arrière-plan (Background subtraction)**

La soustraction de l'arrière-plan est l'opération qui suit logiquement la modélisation de l'arrière-plan afin d'obtenir une détection de mouvement. Si le modèle de l'arrière-plan est une image, une différence en valeur absolue entre ce modèle et l'image courante est effectuée afin d'obtenir une détection de mouvement.

➤ **Segmentation de (par le) mouvement (Motion[-based] segmentation)**

Cette tâche va au-delà de la détection de mouvement puisqu'il s'agit de segmenter chaque image en régions qui présentent une homogénéité du mouvement apparent. Cette opération est généralement réalisée à partir d'une

estimation du flot optique [11] ou des dérivées spatiotemporelles de l'intensité lumineuse [12].

La détection du mouvement contient les méthodes de différences d'images (approche directe), de corrélation ou de recherche dans l'espace des paramètres.

Les méthodes basses niveau exploitent la comparaison pixel à pixel, ou petite région à petite région entre deux images consécutives d'une séquence. Elles permettent de déterminer les régions de variations de l'image dans le temps. Elles nécessitent soit une caméra fixe, soit un recalage préalable dans le cas d'un observateur mobile, afin de détecter uniquement les zones de mouvement dans la scène. Elles sont généralement limitées aux mouvements d'objets rigides et au cas de petits déplacements.

### **I.12.2 Définition d'un objet vidéo**

Définir un objet vidéo est une chose difficile. Intuitivement, un objet vidéo est défini par une forme, une texture et un mouvement (rigide ou non rigide). Cependant, la notion d'objet vidéo est beaucoup plus descriptive qu'une simple région, un objet vidéo peut par exemple être un modèle 3D.

La notion d'objet ne fait pas forcément référence à un objet du monde réel. En effet, dans le domaine de la vidéo, un objet n'est pas nécessairement un objet d'une scène 3D mais plutôt le résultat de l'analyse de la projection d'un monde 3D sur un plan. Ainsi, un objet vidéo est défini comme une région de la vidéo conforme à un modèle. On peut par exemple avoir pour modèle :

- un modèle de mouvement.
- un modèle d'objet physique.

La notion d'objets est alors subjectivement définie par rapport aux modèles utilisés. On constate que pour pouvoir trouver des objets vidéo, il est nécessaire de proposer des modèles auxquels ces objets répondront [13].

### **I.12.3 Suivi d'objets dans la vidéo**

Le suivi d'objets quelconques dans une séquence vidéo réelle est une tâche très délicate. Surtout quand ces objets sont non-rigides, le fond de la scène n'est pas fixe et dans le cas de

plusieurs objets mobiles dans la même scène. Lorsqu'il s'agit de suivre des objets particuliers comme le visage, la main, le bras. La tâche est plus simple car le modèle et les contraintes de variation de ces objets sont connus a priori. D'autres critères de type géométrique et/ ou statique peuvent aussi être introduits dans l'identification des ces objets.

### **I.13 Conclusion**

Dans ce chapitre nous avons abordé certaines définitions primaires concernant le traitement et la transformation d'image d'un état à l'autre comme le filtrage, la segmentation...etc. et nous avons vu les notions de base de l'analyse du mouvement dans une séquence d'image.

L'objectif de notre travail est la détection et suivi d'objet en mouvement dans une séquence d'image, les prochains chapitres sont consacrés à la présentation de cette dernière.

# *Chapitre II*

*Détection d'objets en  
mouvement*

**II 1.Introduction**

Les applications de la détection d'objets sont nombreuses. On peut notamment citer le suivi d'objets, l'analyse du comportement, la compression vidéo, la reconstruction 3D...etc.

Les techniques de détection du mouvement peuvent être considérées comme un prétraitement permettant de réduire la quantité d'informations à analyser. Un objet en mouvement est détecté si sa position change relativement à celle d'un ensemble d'objets statiques ou s'il est localisé dans une image acquise à l'instant  $t$  de la séquence à une position différente de celle qu'il occupait dans l'image précédente. Le mouvement d'un objet dans la séquence d'images peut donc être mis en évidence en comparant le contenu d'au moins deux images issues de la séquence.

Dans ce chapitre, nous nous intéressons à la détection d'objets en mouvement dans une séquence d'image, par conséquent aux méthodes de séparation des objets mobiles du fond de l'image (l'arrière plan). Dans ce contexte, la méthode la plus utilisée est la méthode de détection des contours en mouvement. On présente également dans ce chapitre des autres méthodes : la méthode de segmentation par Ligne de Partage des eaux (LPE) et la méthode de la différence inter images (avec et sans l'image de référence).

**II 2. Détection d'objets en mouvement par méthode de Segmentation**

Que cela soit dans le domaine spatial, temporel ou spatio-temporel, la segmentation a pour objectif de partitionner une image ou un groupe d'images de sorte que chacune des régions ait une caractéristique qui la distingue des autres. De nombreuses techniques de segmentation ont été proposées. Ces techniques ont toutes en commun la minimisation d'une fonctionnelle d'énergie. Ainsi quelque soit l'algorithme utilisé, on peut s'attendre à obtenir sensiblement les mêmes résultats. Le choix d'un algorithme est alors fonction de ses avantages et de ses inconvénients (sensibilité par rapport à l'initialisation, vitesse d'exécution, complexité, convergence ...).

La segmentation spatiale divise l'image en régions homogènes en termes d'intensité. Les différentes homogénéités des régions sont distinguées par leurs frontières entourant. SALEMBIER et al [14] ont proposé une méthode de segmentation spatiale basée sur une étape de simplification d'image utilisant les filtres morphologiques (ouverture et fermeture) et

l'algorithme de ligne de partage des eaux (minima locaux). Cette méthode a montré des problèmes de sur-segmentations d'une même région. Afin de remédier à ce problème une étape de fusionnement des régions est alors proposée. Le schéma synoptique de la figure (II.1) résume les différentes étapes de la segmentation spatiale qui seront détaillées par la suite.

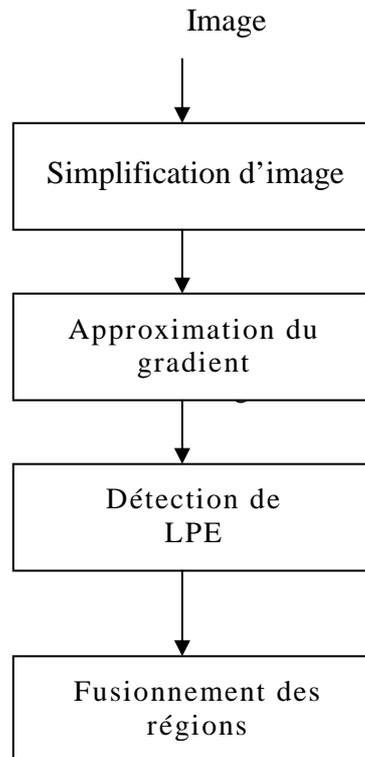


Figure. II.1 : Schéma synoptique de la segmentation spatiale

### a. Simplification d'image

Pour la segmentation spatiale, les images sont d'abord simplifiées. Des filtres morphologiques [15] sont conçus pour lisser les images bruitées par des opérateurs d'ouverture et de fermeture utilisant un élément structurant donné. On peut noter qu'afin de simplifier l'image originale l'opération de l'ouverture enlève les composants clairs et l'opération de fermeture élimine les composants foncés.

## b. Approximation du gradient

Le gradient spatial de l'image simplifiée est appliqué par l'utilisation d'un opérateur morphologique gradient qui consiste à détecter les brusques variations d'intensité lumineuse dans l'image.

## c. Détection des lignes de partage des eaux (watersheds)

La détection des lignes de partage des eaux (LPE) est une technique classique de segmentation d'image [16]. Cette approche considère l'image à niveaux de gris comme un relief topographique ou l'altitude étant la valeur du niveau de gris. Ce relief comporte des minima locaux. Ce sont des zones connexes de l'image dont tous les pixels sont à même niveau de gris et dans lesquelles il est impossible d'atteindre un pixel d'altitude inférieure sans avoir une montée.

Considérons maintenant le problème hydraulique suivant : Supposant qu'une goutte d'eau tombe sur le relief, elle va s'écouler, le long de celui-ci jusqu'à atteindre un minimum local. On appelle bassin de capture, l'ensemble des pixels tel que si une goutte d'eau y tombait, elle se dirigerait vers un même minimum local. Ceci est illustré sur la figure (II.2) ; les parties en gras sont les minima locaux  $m_i$  les lignes en pointillées (lignes de partage des eaux) délimitent les bassins de captures  $BC_i$

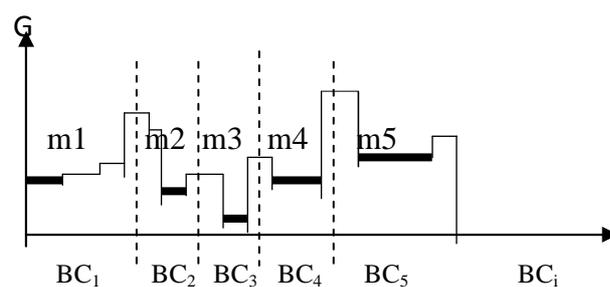


Figure. II.2 : Les minima locaux  $m_i$  et les bassins de captures  $BC_i$  associés

Les bassins de captures sont connexes et forment une partition de l'image. La frontière de ces bassins est appelée ligne de partage des eaux (watersheds). Cette frontière est souvent définie à un pixel près.

Pour segmenter une image, on utilise en règle générale les watersheds sur l'image gradient ou sur toute autre fonction d'énergie. Par exemple les variations d'intensités ou de texture localisée se traduisent par les maxima locaux du gradient, des crêtes en termes de relief. La valeur des maxima du gradient est faible dans des parties homogènes de l'image. Ces parties correspondent à des minima locaux.

L'algorithme de ligne de partage des eaux LPE est extrêmement sensible au gradient bruité qui rapporte beaucoup de bassins de capture, ayant pour résultat une sur-segmentation. Pour cela un algorithme de fusionnement des régions (Region merging) basé sur l'homogénéité d'intensité devrait être introduit pour résoudre le problème de sur-segmentation.

#### **d. Fusion des régions**

Une fusion revient à supprimer une frontière et à faire absorber une région par l'autre. Que ce soit dans les approches de fusion-division ou pour d'autres techniques, il est nécessaire de définir une mesure de similarité ou de dis similarité entre deux régions. Cela revient à modéliser le problème par une fonctionnelle d'énergie qu'il faut minimiser. Le choix de fusionner deux régions peut alors être fonction de la variation d'énergie avant - après fusion ou fonction de la similarité entre deux régions. Il y a donc une notion de seuil qui intervient pour stopper les fusions. Différents seuils peuvent être utilisés comme le seuil sur la variation d'énergie, celui sur la similarité ou celui sur le nombre total de fusions.

##### **❖ Intérêt de la fusion**

- Le but de la fusion est de diminuer le nombre de régions en vu de certains traitements.
- l'objectif est d'obtenir des régions homogènes.

##### **❖ Fusion sur un graphe d'adjacence des régions (GAR)**

Les techniques par graphes d'adjacence [17] sont beaucoup utilisées sur les régions sur-segmentées. Ces sur segmentations sont par exemple le résultat d'un prédicat trop sévère dans la phase de segmentation ou d'un algorithme très sensible aux variations locales comme la technique des lignes de partage des eaux.

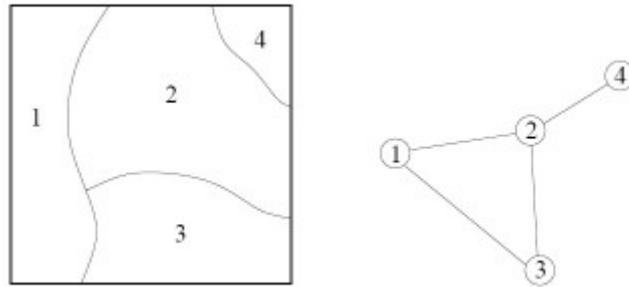


Figure II.3 Partitions et graphes d'adjacences associés

L'idée de base consiste à plonger les régions obtenues dans une structure de graphe où une région est un nœud et un arc est une relation d'adjacence. Puis on définit une fonction de similarité entre 2 nœuds. On trie tous les couples de nœuds adjacents dans une liste ordonnée. On regroupe les 2 candidats sur un critère pré défini. On remet à jour la liste et on itère.

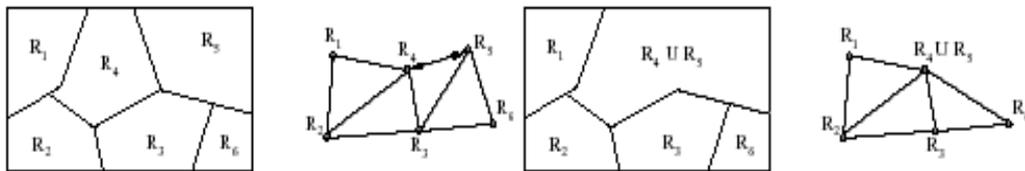


Figure II.4 Exemples de fusion de deux régions à partir d'un GAR de la partition

Dans le domaine spatio-temporel, la fusion sur un GAR est empirique et requiert de calcul lourd. Afin de remédier à ce problème on procède à la détection de mouvement par la différence inter images puis on applique un nouveau schéma de détection des contours en mouvement.

## II.3 Détection par différence d'images

### II.3.1 Approche région

#### a. Avec image de référence

Dans certaines applications (acquisition avec caméra fixe en particulier), il peut être possible de disposer d'une image dans laquelle seuls les éléments stationnaires sont présents. Cette image est alors utilisée comme une image de référence  $R(x, y)$  notée  $R$ .

Cette image doit posséder les caractéristiques suivantes :

- Être exempte de tout objet mobile.
- Avoir été acquise sous les mêmes conditions d'éclairements que les images avec lesquelles elle sera comparée.

L'image de différence notée  $D$  devient alors :

$$D = |I - R| \quad (\text{II.1})$$

Où  $I$  représente l'image à l'instant  $t$

et  $R$  est l'image de référence

Les régions en mouvement sont obtenues après seuillage de l'image différence (nécessaire à cause du bruit). Cette image de différence fait apparaître 2 types de régions :

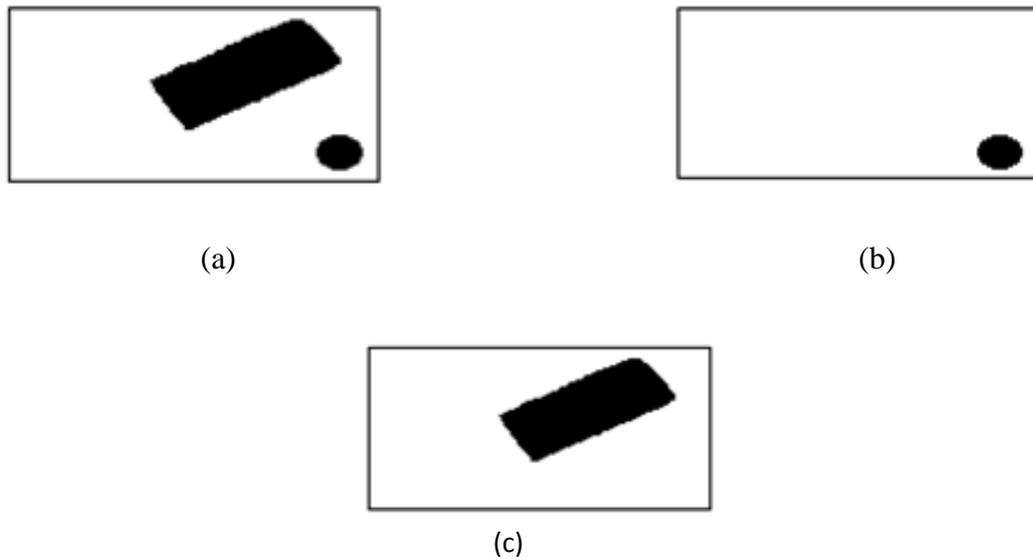


Figure. II.5 : Technique de différence avec image de référence, (a) l'image à l'instant  $t$ , (b) l'image de référence, (c) l'image de différence.

- La première zone est constituée des points appartenant aux objets mobiles dans leur position courante.
- La seconde zone est constituée des points stationnaires présents dans les deux images  $I(x, y, t)$  et  $R(x, y)$ .

Cette méthode est très utilisée dans le domaine de détection d'objet en mouvement mais la plus grande difficulté consiste à obtenir l'image de la référence (du fond). Dans certains cas cette image peut être obtenue en sélectionnant manuellement une image où aucun objet n'apparaît, mais ce n'est pas toujours le cas. La construction d'une image de référence est donc une question délicate abordée dans la littérature par de nombreux auteurs [18, 19, 20]. L'image de référence doit en permanence être identique à l'environnement statique perçu dans chaque image de la séquence. Dans [21], Jain propose un algorithme qui repose sur l'analyse de l'image différence. Dans [22], Makarov construit l'image de référence en se fondant sur l'analyse de la moyenne du niveau de gris d'un pixel donné sur une séquence de  $N$  images. À chaque acquisition d'une nouvelle image  $I$  de la séquence, une nouvelle image de référence  $R$  est construite. Le niveau de gris de chaque pixel de l'image de référence est déterminé par la relation suivante :

$$R(x, y, t) = \frac{1}{N} \sum_{i=t-n}^{t-1} I(x, y, t) \quad (\text{II.2})$$

Ou  $N$  : le nombre des images dans la séquence

Cette méthode nécessite une mémorisation de  $N$  images successives de la séquence, mais cette opération est assez coûteuse. Une autre difficulté vient des changements de luminosité ; même si l'hypothèse du faible changement entre deux images successives est respectée, le changement à long terme peut être important sur une séquence très longue.

### b. Sans l'image de référence

La différence temporelle s'exprime par :

$$F_D(IC, IP) = |IC - IP| \quad (\text{II.3})$$

Ou :  $F_D$  : image de différence

Et  $IC$  image courante, et  $IP$  image précédente.

L'image  $F_D(IC, IP)$  ainsi obtenue est nulle en tout points où le signal d'entrée  $I$  est constant. Les points où un changement temporel s'est produit sont détectés par simple seuillage de  $F_D$ . Le seuillage est nécessaire à cause du bruit présent dans les 2 images. Le résultat du seuillage est une image binaire indiquant les zones en mouvement.

Selon l'exemple de la figure (II.6) la différence fait apparaître quatre zones distinctes.

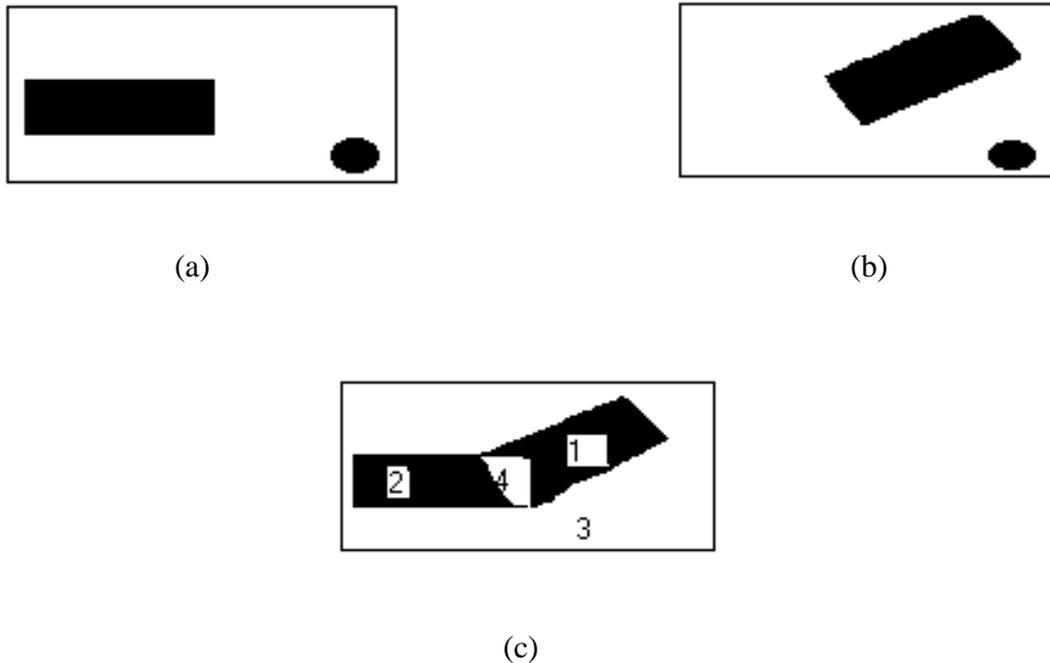


Figure. II.6 : Technique de différence sans image de référence, (a) image précédente, (b) image courante, (c) image de différence.

- La zone '1' située à l'avant de l'objet dans le sens du déplacement est constituée des points qui appartenaient au fond de l'image précédente et qui sont recouverts par l'objet dans l'image courante. L'amplitude des différences de niveaux de gris y est fonction du contraste entre l'objet et le fond.
- La zone '2' située à l'arrière de l'objet est constituée des points qui appartenaient à l'objet dans l'image précédente et qui ont été découverts suite au mouvement. Ils appartiennent au fond de l'image courante. L'amplitude des différences de niveaux de gris y est fonction du contraste entre l'objet et le fond.
- La zone '3' constituée des points appartenant au fond et aux objets stationnaires dans les deux images. L'amplitude des différences de niveaux de gris y est faible.

- La zone '4' qui n'existe que lorsqu'il y a chevauchement des positions de l'objet et qui est constituée des points appartenant à l'objet en mouvement dans les deux images. L'amplitude des différences de niveaux de gris y est faible.

Ainsi, les régions '1' et '4' retracent l'objet dans l'image à l'instant  $t$ , et les régions '2' et '4' à l'instant  $t - dt$ , les régions '1' et '2' sont détectées comme des régions en mouvement, par contre, la région '4' l'est rarement et le problème est d'extraire cette région.

### II.3.2 Approche contour

Jusqu'à maintenant, la détection des changements temporels a été abordée sous l'aspect extraction de région en mouvement. On propose dans ce paragraphe une approche légèrement différente, qui consiste à extraire les contours en mouvement.

Nous noterons  $G(I)$  l'image des contours (gradients) de l'image  $I$ , et  $CM(I)$  l'image des contours en mouvement de l'image  $I$ .

On distingue plusieurs variantes de la méthode

#### a. Proposition de HAYNES [23]

Il propose d'effectuer le produit simple (AND) entre l'image de différence  $F_D(IC, IP)$  et l'image des contours  $G(IC)$  soit :

$$F_D(IC, IP) = |IC - IP| \quad (\text{II.4})$$

$$CM(IC) = F_D(IC, IP) \cdot G(IC). \quad (\text{II.5})$$

La différence inter-image est considérée comme un détecteur de mouvement, l'opérateur  $G(IC)$  est considéré comme un détecteur de contour, et le produit est alors un détecteur de contours en mouvement.

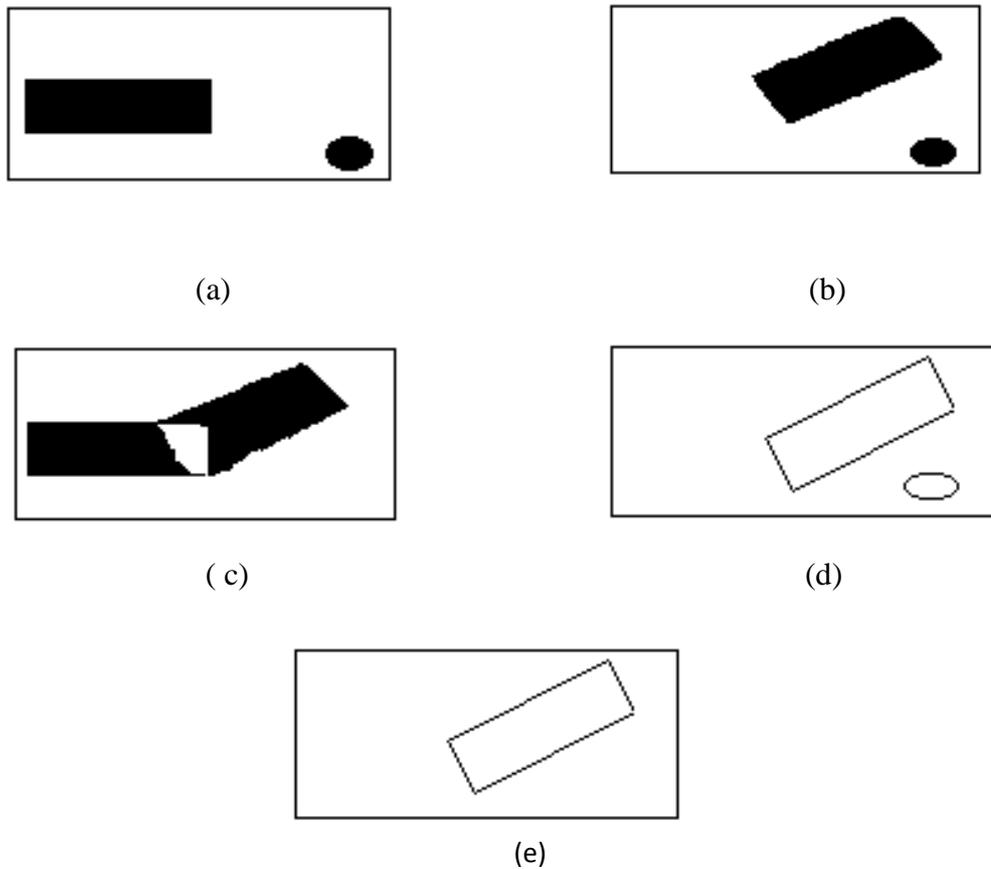


Figure II.7: extraction des contours en mouvement par la proposition de Haynes, (a) image précédente, (b) image courante, (c) image de différence, (d) contour d'image courante, (e) contour en mouvement.

Bien que sensible au bruit, un seuillage approprié de  $CM(IC)$  est néanmoins nécessaire afin de ne conserver que les points des contours en mouvement. Cette technique rencontre aussi des difficultés lorsqu'un objet se déplace parallèlement à l'un de ses contours. Dans ce cas, malgré la présence d'un contour ( $G(IC)$  fort), la différence inter-images est voisine de zéro et par conséquent,  $CM(IC)$  est trop faible pour permettre de détecter le contour en mouvement.

Une excellente critique de ce détecteur démontre que pour un déplacement donné, le détecteur n'est fonction que de la différence inter-images. Lorsque celle-ci est trop faible, la réponse du détecteur est faible. Ce détecteur peut aussi répondre à des contours variant temporellement qui ne sont pas nécessairement des contours en mouvement. Un contour

variant temporellement peut être dû à un changement d'éclairage, à un contour statique caché par un objet en mouvement ou bien à un objet se déplaçant.

### **b. proposition de STELMASZYK [24]**

Elle est basée sur le fait que les contours en mouvement sont les seuls à être présents à la fois dans l'image des contours courants  $G(IC)$  et dans l'image des contours de la différence  $G(F_D(IC, IP))$ .

On obtient alors :

$$CM(IC) = G(IC) \cdot G(F_D(IC, IP)). \quad (\text{II.6})$$

Cette méthode est élégante et homogène car l'opérateur d'extraction de contours est le même pour l'image courante et pour l'image des différences. D'autre part, le produit s'effectuant entre deux images de contours, les valeurs combinées correspondent forcément au même point de l'image courante. Par conséquent, tous les points de contours en mouvement répondront fortement alors que pour les détecteurs de Haynes et Jain beaucoup de points de contours en mouvement sont perdus si l'amplitude de la différence est faible. Par contre, ce détecteur pose un problème délicat lorsque les objets se déplacent sur un fond non uniforme. Pour ce cas, l'image  $G(IC)$  contient les contours du fond qui étaient cachés par l'objet dans sa position précédente. L'image  $G(F_D(IC, IP))$  contient aussi ces contours puisque  $F_D(IC, IP)$  contient les points du fond en ces endroits là, par conséquent, ces points du contour du fond, communs aux deux images, apparaissent dans  $CM(IC)$ .

Pour éviter cela, Vieren propose de mettre en jeu une troisième image.

### **c. Proposition de VIEREN [25]**

Le détecteur proposé est le suivant :

$$CM(IC) = G(F_D(IC, IP)) \cdot G(F_D(IC, IS)) \quad (\text{II.7})$$

Où  $IS$  représente l'image suivante

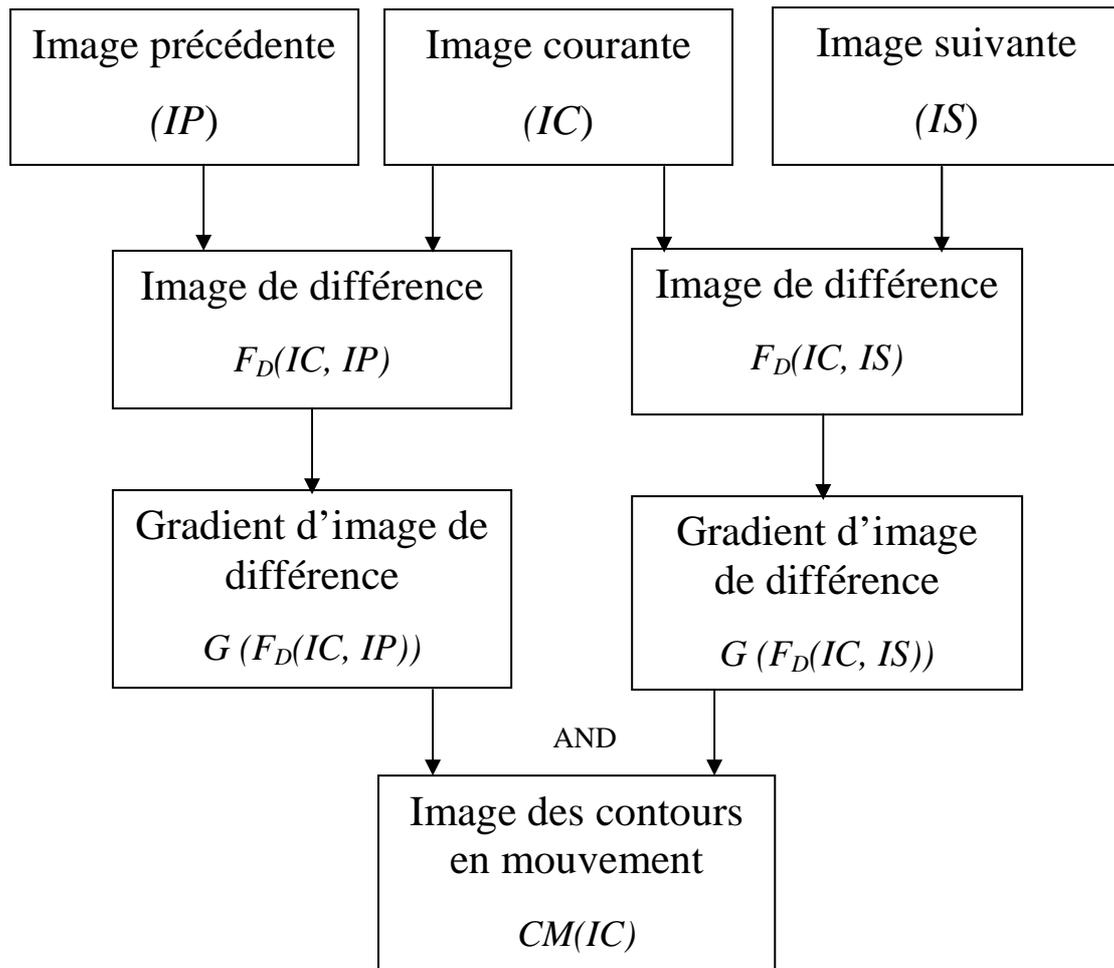


Figure. II.8: Extraction des contours en mouvement par la proposition de Vieren.

La suppression des fausses détections est assurée, car il ne peut pas y avoir de zone de l'objet commune aux images  $IP$  et  $IS$  qui soit extérieure aux contours de l'objet dans sa position courante.

Néanmoins, l'inconvénient majeur de toutes les méthodes ci-dessus provient du fait qu'elles mettent toutes en jeu une multiplication d'un gradient qui ne peut fonctionner correctement que dans les parties bien contrastées. La multiplication par une petite valeur de gradient affaiblit le résultat.

A partir de ces considérations, Orkisz propose de comparer les images des gradients.

**d. Proposition d'ORKISZ [26]**

La soustraction du gradient de l'image précédente à celui de l'image courante élimine les contours stationnaires et affecte peu les contours des objets en mouvement dans l'image courante. Par conséquent, pour les contours stationnaires, le résultat de cette soustraction est différent des valeurs du gradient de l'image courante, par contre, il est très proche pour les contours en mouvement.

Pour détecter les objets en mouvement, il faut donc chercher les endroits où le gradient de l'image est suffisamment grand (présence de contours) et à peu près égal à la différence des gradients des images courante et précédente.

On compare donc :

$$G(IC) \approx |G(IC) - G(IP)| \quad (\text{II.8})$$

Partant du même principe de cette proposition (éviter le produit gradient), une autre méthode basée sur la mesure des contours en mouvement (mcm) qui est proposé.

**II.4 Détection des contours en mouvement par (mcm)**

L'objectif de cette étape est de détecter les contours des régions en mouvement. Donc la méthode proposée [27, 28, 29], prend en compte trois images successives  $IP$ ,  $IC$ ,  $IS$  et on calcule le module du gradient de chaque image. Ensuite, on applique symétriquement la différence absolue sur les deux couples successifs de module du gradient. La mesure des contours en mouvement (mcm) à un temps  $t$  donné est alors définie par le minimum (AND) de ces deux différences.

$$\text{mcm} = \min ( |G(IP) - G(IC)|, |G(IC) - G(IS)| ) \quad (\text{II.9})$$

la figure II.9 résume les différentes étapes nécessaires du calcul du mcm :

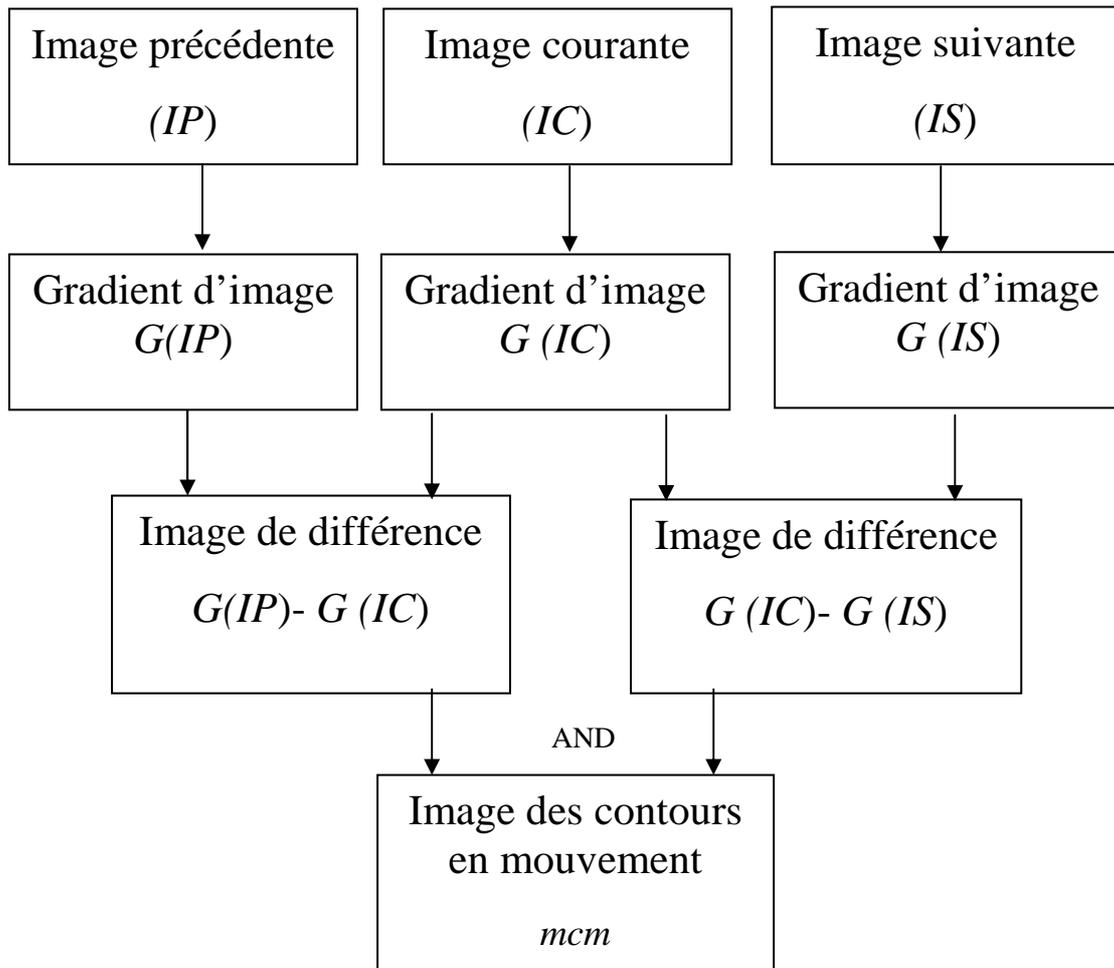


Figure II.9 schéma synoptique de l'opérateur mcm

Les contours des objets en mouvement sont ainsi extraits à l'aide d'un opérateur basé sur la différenciation des gradients spatiaux de trois images successives. Les contours forts de chaque image sont obtenus par le gradient.

L'opérateur mcm est robuste à l'amplitude du déplacement des objets (mouvement rapide) et au bruit aléatoire présent dans l'image. Mais l'obtention de contours dans le cas d'un mouvement de faible amplitude (mouvement lent) par le mcm sont incomplet probablement non satisfait, donc dans ce cas la on applique la méthode de différence entre deux images suivi par un seuillage.

Une étape supplémentaire est alors nécessaire pour agglomérer les différentes zones issues de la détection des changements et pour filtrer les détections parasites. Des traitements de type dilatation, érosion sont appliqués à la sortie des détecteurs des changements temporels. Ils

permettent d'une part de nettoyer les images résultats et d'autre part, ils peuvent être utiles à la reconstruction des masques des objets en mouvement. Ces concepts font partis des techniques de morphologie mathématique.

## **II.5 Conclusion**

Dans ce chapitre, nous avons présenté les techniques de détection des objets en mouvement. Nous différencions la méthode de détection basées sur la segmentation par LPE et les méthodes de la différence inter-images, celles reposant sur une image de référence et celles sans image de référence. Dans ce contexte, on a proposé une méthode de différence entre deux images successives suivi par un seuillage et certaines techniques de morphologie mathématique lorsque le mouvement est faible (lent). Dans le deuxième cas si le mouvement est rapide, on opère une méthode pour extraire les contours des objets en mouvement basée sur la différenciation des gradients spatiaux de trois images successives.

# *Chapitre III*

*Suivi d'objets en  
mouvement*

### **III.1 Introduction**

Après avoir détecté les objets dans une séquence d'image, nous nous intéressons dans ce chapitre à la phase de suivi.

La problématique que nous nous proposons consiste à suivre un objet de forme et de texture quelconques soumis à un mouvement quelconque pouvant être de forte amplitude, dans des séquences d'images en 256 niveaux de gris. Un bref aperçu des algorithmes de suivi sera donc présenté, ainsi que les modèles de mouvement les plus utilisés.

Nous présenterons ensuite l'algorithme de suivi que nous avons choisi.

### **III.2 Algorithmes de suivi**

Le suivi d'objets dans une séquence d'images est un problème classique et largement étudié dans le domaine de l'analyse du mouvement, du fait de ses nombreuses applications parmi lesquelles on peut noter la surveillance vidéo [30] la robotique, la compression de vidéos, l'étude de la déformation d'objets, . . . Ainsi, de nombreux algorithmes de suivi ont été proposés ces dernières années. Nous allons, dans les sections suivantes, présenter une classification possible de ces algorithmes, ainsi que leur structure générale.

### **III.3 Classification des algorithmes de suivi**

Les algorithmes de suivi peuvent globalement être classés en deux catégories :  
Les algorithmes orientés mouvement, et ceux orientés modèle

#### **III.3.1 Les algorithmes orientés mouvement**

S'appuient sur des méthodes de regroupement de régions homogènes au sens du mouvement. Ils ont l'avantage d'être rapides, cependant, les régions suivies n'ont en général pas de signification physique globale puisqu'elles ont été obtenues à partir d'un critère d'homogénéité et non d'une représentation d'objet. Le principe de ces algorithmes est de partir d'une segmentation spatio-temporelle initiale [31,32]. Cette dernière peut être obtenue de différentes façons : à partir de pixels individuels (techniques Markoviennes fournissant un champ dense [33]) de petits blocs de pixels (approche MPEG), d'une grille arbitraire [34,35], de l'image entière, ou encore d'une segmentation spatiale [36,37]. Les paramètres de mouvement sont ensuite estimés à partir de cette segmentation, puis des fusions ou des divisions de régions sont réalisées selon un critère d'uniformité du mouvement. Cette segmentation est ensuite suivie au cours de la séquence. Ce type d'algorithme est utilisé pour

la compression, le codage, l'interpolation et l'indexation de vidéos et la détection et le suivi de fouillis radar..

### **III.3.2 Les algorithmes orientés modèle**

Contrairement aux précédents, ces algorithmes se basent sur le suivi des régions correspondant aux projections 2D des objets auxquels on s'intéresse, en les représentant soit par des modèles déformables 3D [38], soit par des modèles déformables 2D [39]. Les régions suivies sont donc plus significatives, mais la complexité calculatoire est plus importante du fait de la nécessité de déterminer d'éventuelles rotations, translations, homothéties ou autres déformations.

Dans les deux cas, le suivi peut être réalisé par des méthodes basées soit sur les frontières, soit sur les régions. Il est à noter que les premières recherches dans le domaine s'appuyaient sur le suivi de primitives simples telles que des segments de contours ou des points de référence dans les objets [40]. Cependant, ces techniques étaient peu fiables du fait de données trop éparées, et sensibles aux occlusions. C'est pourquoi les recherches se sont orientées vers des primitives plus complexes et plus globales telles que les contours et les régions.

### **III.3.3 Les méthodes basées sur les frontières**

Elles reposent sur l'information apportée par le contour de l'objet à suivre. Leur avantage est d'être indépendant du type de mouvement ainsi que de l'objet. De plus, elles sont efficaces pour extraire les contours de l'objet avec précision. Elles sont donc particulièrement bien adaptées au suivi d'objets déformables. Ce sont en général des modèles de contours déformables tels que les contours actifs (snakes) [41], les contours actifs géodésiques [42] ou encore des modèles statistiques [43] qui sont utilisés. Ainsi, les méthodes proposées dans [44] sont basées sur ces modèles. Les principaux inconvénients de ces modèles sont leur instabilité due à la prise en compte d'une information locale, ainsi que leur sensibilité à l'initialisation. En effet, le principe de ces algorithmes est de partir d'un contour initial et de converger de façon itérative vers le contour recherché, par minimisation d'une fonction d'énergie. Ils sont donc performants dans le cas de mouvements lents, mais mal adaptés aux déplacements importants.

### III.3.4 Les méthodes basées sur les régions

Utilisent, quant à elles, l'information apportée par la région entière, comme par exemple des propriétés de texture ou de mouvement. Bien qu'ayant des difficultés à suivre avec précision les contours, ces méthodes sont plus robustes puisqu'elles utilisent une information plus globale et qu'elles sont moins dépendantes de l'initialisation. Elles sont ainsi capables de gérer des déplacements importants. De tels modèles ont été développés dans [45] en utilisant des modèles de régions déformables géodésiques, et dans [46] en utilisant des contours actifs basés région. De plus, en associant au suivi de région une segmentation de régions basée mouvement [47], il est même possible de s'affranchir de toute initialisation puisque cette segmentation détecte les objets en mouvement dans l'image.

Ces deux types de méthodes possèdent donc des qualités complémentaires. En les combinant, certains algorithmes permettent un suivi optimisé [48].

### III.4 Structure général des algorithmes de suivi

Le suivi des primitives présentées précédemment d'un instant  $t$  vers un instant  $t + I$ , est généralement composé d'un minimum de trois étapes, après avoir calculé une segmentation spatiale ou spatio-temporelle initiale, selon le type d'algorithme de suivi considéré. Ces trois étapes sont les suivantes :

- Prédiction du mouvement des primitives de l'instant  $t$  à l'instant  $t + I$
- Ajustement de la segmentation prédite
- Estimation du mouvement des régions suivies de  $t$  vers  $t + I$

Nous allons brièvement considérer chacune de ces trois étapes

#### III.4.1 Phase de prédiction

Elle a pour objectif de relier la segmentation de l'image à l'instant  $t$  à celle de l'image à l'instant  $t + I$ . Quatre modèles de prédiction peuvent être envisagés [49]:

##### a. Prédiction de mouvement nul

On considère que la primitive ne se déplace pas entre  $t$  et  $t + 1$ . Sa position prédite est donc identique à celle qu'elle avait à l'instant précédent.

**b. Prédiction à court-terme simple**

On prend uniquement en compte l'information du mouvement entre les deux images précédentes. Le terme court-terme vient du fait que dans ce cas, on n'utilise aucun modèle temporel d'évolution.

**c. Prédiction à court-terme lissée**

On utilise alors un modèle d'évolution, tel que le filtre de Kalman. Le mouvement considéré est celui fourni par les équations de filtrage (ou lissage) de ce filtre.

**d. Prédiction à long-terme**

Dans ce cas, on prend en compte l'information du mouvement entre toutes les images précédentes. Les valeurs prédites sont alors issues de l'équation de prédiction du filtre de Kalman. Nous préciserons ultérieurement de manière plus détaillée, ces deux dernières méthodes.

**III.4.2 Phase d'ajustement**

Une fois la phase de prédiction réalisée, nous disposons d'une segmentation prédite de l'image à l'instant  $t + 1$ . Cependant, cette dernière n'est en général pas parfaite. Il est donc nécessaire d'ajuster les frontières des régions de cette segmentation sur les frontières réelles. C'est le rôle de la phase d'ajustement. Selon le type de primitives suivies, deux types de méthodes d'ajustement peuvent être envisagés :

**a. Ajustement basé contour**

Ces méthodes utilisent des modèles de contours déformables afin d'ajuster les frontières des régions prédites, en se basant uniquement sur des informations spatiales, telles que le gradient spatial, dans l'image courante.

**b. Ajustement basé région**

Cette approche peut s'appuyer, par exemple, sur une modélisation markovienne des zones d'occlusion [50], ou encore sur des modèles de régions déformables [51]. Elle peut prendre en considération à la fois des informations spatiales dans l'image courante (gradient spatial, texture,...), et des informations de mouvement entre les images courante et précédente.

### III.4.3 Phase d'estimation

Elle est réalisée après la phase d'ajustement et permet de déterminer quel a été le mouvement effectif des régions suivies entre les images aux instants  $t$  et  $t + I$ . Il existe trois grands types de méthodes d'estimation du mouvement [52]:

#### a. Les méthodes par mise en correspondance

Elles cherchent à déterminer des similarités entre des points, des contours ou des régions, présents dans deux images successives. Pour cela, une fonction de corrélation doit être maximisée. Ce sont en général des régions constituées de petits blocs carrés qui sont considérées. Seules les translations sont tolérées, car ce modèle suppose que tous les pixels appartenant à une même région subissent un déplacement identique. Les normes MPEG, par exemple, ont recours à ces méthodes.

#### b. Les méthodes par transformées

Elles utilisent les effets du mouvement dans l'image sur les composantes ou les caractéristiques fréquentielles de transformées telles que la transformée de Fourier ou encore la décomposition en séries de Gabor. En effet, une translation spatiale, par exemple, modifie la phase de la transformée de Fourier de l'image.

#### c. Les méthodes différentielles

Elles reposent sur l'hypothèse d'invariance de la luminance des points au cours du temps, lors de leur déplacement. L'estimation est alors réalisée par la minimisation d'une fonction d'énergie mesurant la différence de luminance d'un point d'une image et son projeté dans l'image suivante.

### III.5 Suivi de trajectoire d'objet en mouvement

Dans le vidéo tracking [53], un objet peut être de nature quelconque. Tous les scénarios sont possibles à imaginer : un bateau en pleine mer, un joueur dans un stade avec son ballon, un véhicule en autoroute, un avion dans le ciel, une pièce à usiner dans une chaîne de fabrication, etc. Selon la nature de l'objet et l'objectif de l'application à réaliser, de différentes représentations de la forme (points, formes géométriques de bases [54], modèle articulé, silhouette, etc.) et d'apparence (le modèle d'apparence active (AAM), etc.) de l'objet sont utilisées permettant, ainsi, plusieurs algorithmes de suivi d'objet en mouvement.

Dans ce travail nous aborderons la représentation de la forme points dans le suivi des objets en mouvement, où chaque objet est représenté par un point qui est son centre de gravité. La trajectoire d'objet en mouvement dans ce cas est l'interpolation entre les centres de gravité de l'objet détecté dans chaque trame.

Lorsqu'on segmente un objet dans l'image, ou lorsqu'on mesure sa position il y a toujours lieu aux erreurs. L'objectif est alors de calculer la trajectoire de l'objet en diminuant ces erreurs. Il s'agit alors de l'estimation et de la correction de la trajectoire de l'objet en mouvement (figure III.1).

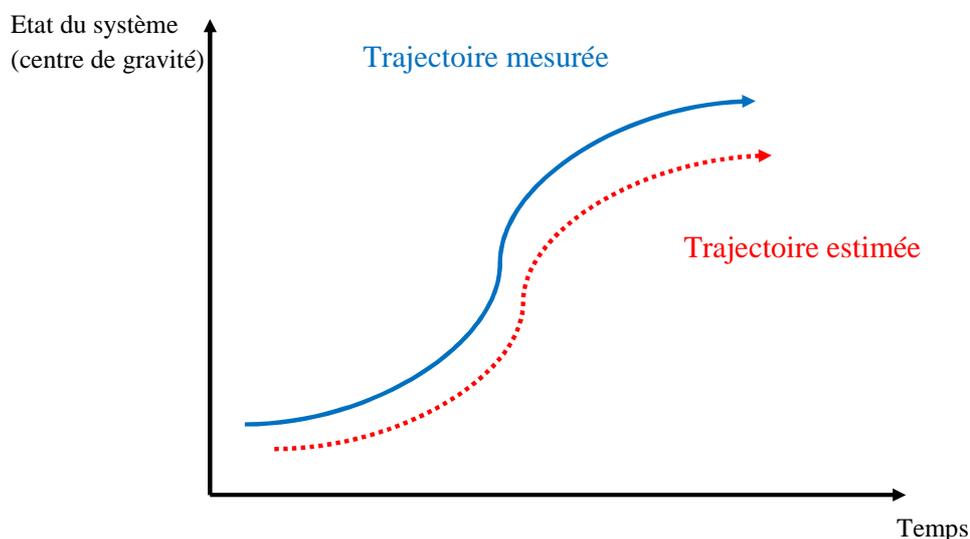


Figure III.1 Exemple de trajectoire mesurée et estimée.

L'estimation de l'état en cours d'un système dynamique peut aussi être appelée filtrage, le terme filtrage est utilisé car le processus d'obtention de la meilleure estimée à partir de données bruitées consiste à éliminer ou à filtrer ce bruit. L'état du filtre peut être composé de la position du point, ainsi que d'éventuelles informations complémentaires telles que sa vitesse ou son modèle d'intensité. L'estimation du nouvel état est divisée en deux étapes majeures. Premièrement, une prédiction du nouvel état est obtenue à partir du modèle dynamique. Ce modèle peut être défini *a priori*, ou déduit d'une étape préalable d'apprentissage. Ensuite, la seconde étape consiste à prendre en compte une mesure de l'état à partir de la séquence d'images. Cette mesure, qui se révèle souvent bruitée ou incomplète, permet de corriger la prédiction. Les filtres stochastiques proposent des solutions pour estimer

l'état d'un système conditionnellement à l'ensemble des mesures passées. Parmi ces filtres, on propos d'utiliser le filtre de Kalman.

Les deux idées qui motivent l'utilisation d'un filtre de Kalman sur les paramètres du mouvement sont :

- Accumuler les mesures des paramètres de mouvement d'une région au cours du temps pour affiner les estimations.
- Obtenir une bonne prédiction du mouvement nécessaire pour prédire la segmentation avant ajustement et pour initialiser l'estimation de mouvement.

### **III.6 Le filtre de Kalman**

Le filtre de Kalman, du nom de son créateur Rudolph E. Kalman, mathématicien américain né en 1930, a été présenté en 1960. Il s'agit d'un estimateur optimal de processus aléatoires.

Il permet l'estimation de variables d'état d'un système (position/vitesse, par exemple, pour des systèmes mécaniques ou encore tension/intensité, par exemple, pour des systèmes électriques...) et il fournit aussi la variance de l'erreur de l'estimation.

Son appellation « filtre » vient du fait qu'il permet, entre autre, d'estimer un signal bruité (bruit blanc), en séparant le signal de son bruit.

Le filtre de Kalman s'applique aux signaux non stationnaires, il résout le problème de filtrage linéaire et se présente sous la forme d'un ensemble d'équations différentielles ou récurrentes plus faciles à résoudre sur calculateur numérique en présence de condition initiales et d'entrées déterministe.

Ce filtre est utilisé en trajectographie pour estimer une trajectoire comme le cas de Poursuite de cibles (avions, missiles, radars...) en navigation pour localiser un engin, ou encore en balistique pour estimer des conditions initiales inconnues.

Il est également employé lorsque l'on veut coupler des informations venant de sources différentes afin de retrouver, par exemple, une trajectoire.

Le Filtre de Kalman est une approche statistique, d'assimilation de données, dont le principe est de corriger la trajectoire du modèle en combinant les observations avec l'information fournie par le modèle de façon à minimiser l'erreur entre l'état vrai et l'état filtré.

### III.7 Suivi de mouvement par filtre de Kalman

Dans le suivi du mouvement, le filtre de Kalman est une méthode de prédiction intéressant plusieurs chercheurs [55, 56,57]. Un filtre de Kalman est essentiellement un ensemble d'équations récursives, elles sont utilisées ensemble pour aider à modéliser et évaluer très précisément le mouvement d'un système dynamique linéaire. Le filtre de Kalman est appelé 'discret' parce que l'algorithme est discret dans le temps. En utilisant une multitude de différents vecteurs et de matrices, le filtre de Kalman est capable de construire une représentation interne du système et grâce à cette représentation, on peut estimer l'état futur du système. Dans le filtrage de Kalman, les équations décrivent l'état du système à temps successifs, grâce à ces équations on peut prédire l'état futur en utilisant l'état actuel sans dépendre des états antérieurs dans les calculs.

Dans le processus du suivi, le filtre de Kalman 'regarde' un objet quand il se déplace, c'est à dire il prend les informations sur l'état de l'objet à ce moment-la. Puis il utilise ces informations pour prédire où se trouve l'objet dans la prochaine trame. La position de l'objet est corrigée ensuite en considérant la prédiction et aussi l'observation.

#### III.7.1 Principe du filtre de Kalman

Le filtre de Kalman donne l'estimation d'un processus en utilisant un système à retour sous la forme de mesures (bruités). De cette manière, les équations du filtre de Kalman se regroupent en deux types d'équations :

- Les équations de prédiction (extrapolation).
- Les équations de mise à jour (correction).

Les équations d'extrapolation sont responsables de propager les estimées de l'état présent et les covariances d'erreur pour obtenir les estimées à priori de la prochaine étape.

Alors que les équations de correction sont responsables du retour, c'est à dire, introduire une nouvelle mesure avec l'estimée à priori afin d'obtenir une estimée à posteriori améliorée.

De ce fait, le filtre de Kalman ressemble finalement à un filtre prédicteur-correcteur pour la résolution des problèmes numériques (figure III.2).

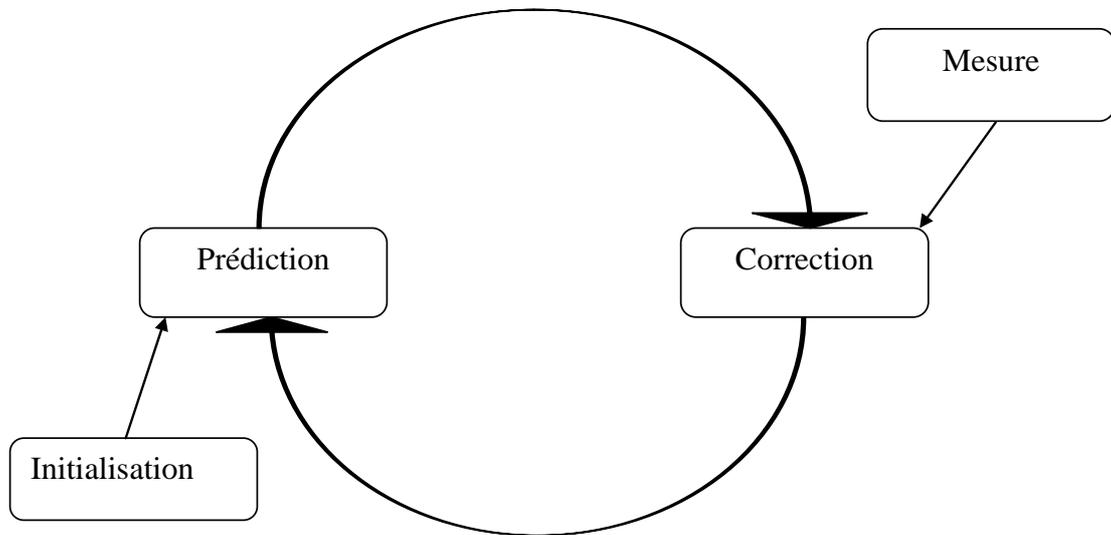


Figure III.2 le cycle du filtre de Kalman

### III.7.2 Calcul de l'estimateur de Kalman

On considère un système dynamique linéaire à temps discret par un vecteur d'équations aux différences, entaché d'un bruit blanc gaussien. L'équation dynamique est représentée par l'équation d'espace d'état suivante :

$$X(k+1) = F \cdot X(k) + V(k) \quad (\text{III.1})$$

$$Z(k) = H \cdot X(k) + W(k) \quad (\text{III.2})$$

Avec :

$X(k)$  : Vecteur d'état contenant les paramètres à estimer

$X(k+1)$  : Vecteur d'état prédit.

$F$  : matrice de transition ou d'évolution du système en absence du bruit

$V(k)$  : représente le bruit du processus supposé blanc de moyenne nulle.

$Z(k)$  : vecteur mesuré (paramètre mesurés)

$H$  : matrice de mesure

$W(k)$  : bruit de mesure blanc de moyenne nulle indépendant de  $b(i)$ .

Avec comme notation :

$$E[V(k)] = 0 \quad (\text{III.3})$$

$$E[W(k)] = 0 \quad (\text{III.4})$$

$$E[V(k) W(k)'] = 0 \quad (\text{III.5})$$

$$E[V(k) V(k)'] = Q(k) \quad (\text{III.6})$$

$$E[W(k) W(k)'] = R(k) \quad (\text{III.7})$$

Les matrices  $F$ ,  $H$ ,  $Q$ ,  $R$  sont supposées connus et peuvent varier dans le temps, donc le système peut être variant dans le temps et le bruit peut ne pas être stationnaire. L'état initial  $X(0)$  est en général connu, les deux séquences de bruit sont supposées indépendantes de l'état initial. Le problème consiste à estimer le vecteur d'état  $X(k)$  en fonction des informations disponible à l'instant  $j$  tel :

$$\hat{X}(j/k) = E[X(j) / Z^k] \quad (\text{III.8})$$

$Z^k = \{Z(i), i \leq k\}$  : est l'information disponible à l'instant  $k$ . on peut alors considérer trois cas suivants :

Si  $j=k$ , il s'agit de déterminer l'estimé à l'instant  $k$ . c'est le cas de filtrage.

Si  $j < k$ , il s'agit de déterminer l'estimé à des instants antérieur à l'instant  $j$  compte tenu des informations disponibles à l'instant  $k$ .

Si  $j > k$ , il s'agit de déterminer l'estimé à l'instant  $j$  futur, compte tenu que l'information est disponible à l'instant antérieur  $k$ .

### III.7.3 Expression des paramètres du filtre de Kalman

L'estimé optimale au sens du critère de l'erreur quadratique moyen minimale est l'espérance mathématique conditionnelle telle que :

$$\hat{X}(j/k) = E[X(j) / Z(1)...Z(k)] \quad (\text{III.9})$$

L'erreur d'estimation est définie par :

$$\tilde{X}(j/k) = X(j/k) - \hat{X}(j/k) \quad (\text{III.10})$$

La matrice de covariance de l'erreur de l'estimation conditionnelle associée est :

$$P(k/k) = E\{[X(k) - \hat{X}(k/k)].[X(k) - \hat{X}(k/k)]' / Z^k\} \quad (\text{III.11})$$

( )' : conjugué transposé.

L'objectif est de trouver une estimation récursive du vecteur d'état, obtenue par les équations d'estimation statistique :

$$\hat{X} = E[X / Z] = \bar{X} + P_{xz} \cdot P_{zz}^{-1} \cdot (Z - \bar{Z}) \quad (\text{III.12})$$

avec  $\bar{X} = \hat{X}(k+1/k)$

$$\bar{Z} = \hat{Z}(k+1/k)$$

$$P_{xz} = \text{cov}[X(k+1), Z(k+1) / Z^k]$$

$$P_{zz} = \text{cov}[Z(k+1) / Z^k]$$

$$P_{xx/z} = E[(X - \hat{X})(X - \hat{X})' / Z] = P_{xx} - P_{xz} P_{zz}^{-1} P_{zx} \quad (\text{III.13})$$

Avec  $P_{zx} = \text{cov}[Z(k+1), X(k+1) / Z^k]$

$$P_{xx} = \text{cov}[X(k+1) / Z^k]$$

L'expression de la matrice de covariance [58] :

$$P(j/k) = E[(X(j) - \hat{X}(j/k)).(X(j) - \hat{X}(j/k))'] = E[\tilde{X}(j/k).\tilde{X}(j/k)'] \quad (\text{III.14})$$

Avec  $\bar{X} = \hat{X}(k+1/k)$

$$\text{Ou } \hat{X}(k) = \hat{X}(k+1/k) + G(k) \cdot [Z - \bar{Z}] \quad (\text{III.15})$$

Le paramètre  $G(k)$  est le gain du filtre de Kalman, l'état estimé du vecteur à l'instant  $k$  est l'évaluation de son ancien terme modifié par un terme correcteur (mise à jour) qui n'est autre que l'écart entre la mesure et sa valeur prédite, l'erreur d'estimation est donnée par :

$$\tilde{X}(k/k) = X(k) - \hat{X}(k/k) \quad (\text{III.16})$$

On doit minimiser la quantité  $E[X(t) - \tilde{X}(t)']$  qui est la matrice de covariance de l'erreur d'estimation tel que :

$$P(k/k) = E[(X(k) - \hat{X}(k/k)) \cdot (X(k) - \hat{X}(k/k))'] \quad (\text{III.17})$$

Nous allons calculer les éléments de l'équation (III.13), l'équation d'estimation dynamique est obtenue par substitution à l'instant  $k+1$  pour le vecteur d'état :

$$X \rightarrow X(k+1) \quad (\text{III.18})$$

$$\bar{X} \rightarrow \hat{X}(k+1/k) = E[X(k+1)/Z^k] \quad (\text{III.19})$$

De même pour le vecteur de mesure :

$$Z \rightarrow Z(k+1) \quad (\text{III.20})$$

$$\bar{Z} \rightarrow \hat{Z}(k+1/k) = E[Z(k+1)/k] \quad (\text{III.21})$$

L'estimation à posteriori à l'instant  $k+1$  pour le vecteur d'état :

$$\hat{X} \rightarrow \hat{X}(k+1/k+1) \quad (\text{III.22})$$

La matrice de covariance de la variable estimée à l'instant  $k+1$ , en appliquant (III.14), sera :

$$P_{xx} \rightarrow P(k+1/k) = \text{cov}[X(k+1)/Z^k] = E[\tilde{X}(k+1/k).\tilde{X}(k+1/k)'/Z^k] \quad (\text{III.23})$$

Covariance de la mesure :

$$P_{zz} \rightarrow S(k+1) = \text{cov}[Z(k+1)/Z^k] = E[\tilde{Z}(k+1/k).\tilde{Z}(k+1/k)'/Z^k] \quad (\text{III.24})$$

La covariance entre la variable à estimer  $X(k+1)$  et l'observation  $Z(k+1)$ :

$$P_{xz} \rightarrow \text{cov}[X(k+1), Z(k+1)/Z^k] = E[\tilde{X}(k+1/k).\tilde{z}(k+1/k)'/Z^k] \quad (\text{III.25})$$

La mise à jour de la covariance :

$$P_{xx/z} \rightarrow P(k+1/k+1) = \text{cov}[X(k+1)/Z^{k+1}] = E[\tilde{X}(k+1/k+1).\tilde{X}(k+1/k+1)'/Z^{k+1}] \quad (\text{III.26})$$

### III.7.4 Dérivation de l'expression de la matrice de covariance

A partir de l'équation d'état et compte tenue des équations (III.1), (III.2) et (III.19) on exprime la matrice de covariance  $P_{k/k}$  [59] :

$$\begin{aligned} \hat{X}(k+1/k) &= E[X(k+1)/Z^k] = E\{[F(k).X(k)+V(k)]/Z^k\} \\ &= F(k).E[X(k)/Z^k] + E[V(k)/Z^k] \quad \text{Or } E[V(k)/Z^k] = 0 \end{aligned}$$

$$\text{D'où } \hat{X}(k+1/k) = F(k).\hat{X}(k/k) \quad (\text{III.27})$$

A partir de l'équation (III.10) l'erreur prédite sera :

$$\tilde{X}(k+1/k) = X(k+1/k) - \hat{X}(k+1/k) \quad (\text{III.28})$$

Et à partir des équations (III.1) et (III.27) l'erreur prédite sera :

$$\tilde{X}(k+1/k) = F(k).\tilde{X}(k/k) + V(k) \quad (\text{III.29})$$

La matrice de covariance s'écrit en utilisant l'équation (III.29) et l'équation (III.23) :

$$\begin{aligned} P(k+1/k) &= E[(\tilde{X}(k+1/k).\tilde{X}(k+1/k)'/Z^k)] \\ &= E[(F(k).\tilde{X}(k/k) + V(k)).(F(k).\tilde{X}(k/k) + V(k))'] \end{aligned}$$

$$\begin{aligned}
 &= F(k).E[\tilde{X}(k/k).\tilde{X}(k/k)' / Z^k].F(k)' + E(V(k).V(k)') \\
 &= F(k).P(k/k).F(k)' + Q(k)
 \end{aligned} \tag{III.30}$$

De la même manière, on trouve l'expression de la mesure prédite en utilisant l'équation (III.1) et l'équation (III.21) tel que :

$$E[Z(k+1/k)] = E[H(k).X(k+1/k) + W(k+1) / Z^k] \tag{III.31}$$

$$\hat{Z}(k+1/k) = H(k).\hat{X}(k+1/k)$$

Puisque  $w(k)$  est un bruit blanc gaussien, l'erreur de mesure prédite sera :

$$\tilde{Z}(k+1/k) = Z(k+1/k) - \hat{Z}(k+1/k) \tag{III.33}$$

$$\tilde{Z}(k+1/k) = H(k+1).\tilde{X}(k+1/k) + W(k+1)$$

Selon l'équation (III.24) la matrice de covariance de la mesure prédite est définie par :

$$S(k+1) = \text{cov}[Z(k+1/k)] = \text{cov}[\tilde{Z}(k+1/k) / Z^k] = E[\tilde{Z}(k+1/k).\tilde{Z}(k+1/k)' / Z^k] \tag{III.34}$$

En utilisant l'équation (1.34) de la mesure prédite :

$$S(k+1) = E[(H(k+1).\tilde{X}(k+1/k) + w(k+1)).(H(k+1).\tilde{X}(k+1/k) + w(k+1))']$$

$$S(k+1) = H(k+1).E[\tilde{X}(k+1/k).\tilde{X}(k+1/k)' / Z^k].H(k+1)' + E[w(k+1).w(k+1)']$$

$$S(k+1) = H(k+1).P(k+1/k).H(k+1)' + R(k+1) \tag{III.35}$$

La mesure entre le vecteur d'état et la mesure se déduit à partir de l'équation (III.25) :

$$P_{xz} \rightarrow \text{cov}[X(k+1), Z(k+1) / Z^k] = \text{cov}[\tilde{X}(k+1/k).\tilde{z}(k+1/k)' / Z^k]$$

$$E[\tilde{X}(k+1/k).\tilde{z}(k+1/k)' / Z^k] = E[\tilde{X}(k+1/k).(H(k+1).\tilde{X}(k+1/k) + W(k+1))']$$

$$\begin{aligned}
&= E[\tilde{X}(k+1/k).\tilde{X}(k+1/k)'](H(k+1) + E[\tilde{X}(k+1/k).W(k+1)]') \\
&= P(k+1/k).H(k+1)' \tag{III.36}
\end{aligned}$$

De même manière, on calcule le dernier terme de l'équation (III.13)

$$\begin{aligned}
P_{zx} &\rightarrow \text{cov}[Z(k+1), X(k+1)/Z^k] = \text{cov}[\tilde{Z}(k+1/k).\tilde{X}(k+1/k)'/Z^k] \\
E[\tilde{Z}(k+1/k).\tilde{X}(k+1/k)'/Z^k] &= E[(H(k+1).\tilde{X}(k+1/k) + W(k+1)).\tilde{X}(k+1/k)'] \\
&= H(k+1).E[\tilde{X}(k+1/k).\tilde{X}(k+1/k)'] + E[W(k+1).\tilde{X}(k+1/k)'] \\
&= H(k+1).P(k+1/k) \tag{III.37}
\end{aligned}$$

D'après l'expression (III.12) et (III.15), le gain de filtre de Kalman

$$G(k+1) = \text{cov}[X(k+1), Z(k+1)/Z^k].S(k+1)^{-1} \tag{III.38}$$

$$G(k+1) = P(k+1/k).H(k+1).S(k+1)^{-1} \tag{III.39}$$

D'où le vecteur d'état mis à jour est :

$$X(k+1/k+1) = X(k+1/k) + G(k+1).\tilde{Z}(k+1) \tag{III.40}$$

Le terme  $\tilde{Z}(k+1)$  est désigné comme l'innovation, ainsi la mise à jour de la matrice de covariance de l'équation (III.26) est effectuée en utilisant l'équation (III.13) :

$$\begin{aligned}
P(k+1/k+1) &= P(k+1/k) - p(k+1/k).H(k+1)'.S(k+1)^{-1}.H(k+1).P(k+1/k) \\
&= P(k+1/k+1) = P(k+1/k) - W(k+1/k).S(k+1).W(k+1)' \tag{III.41}
\end{aligned}$$

Le schéma bloc de la figure III.3 résume le cycle du filtre de Kalman avec les équations de prédiction et de correction

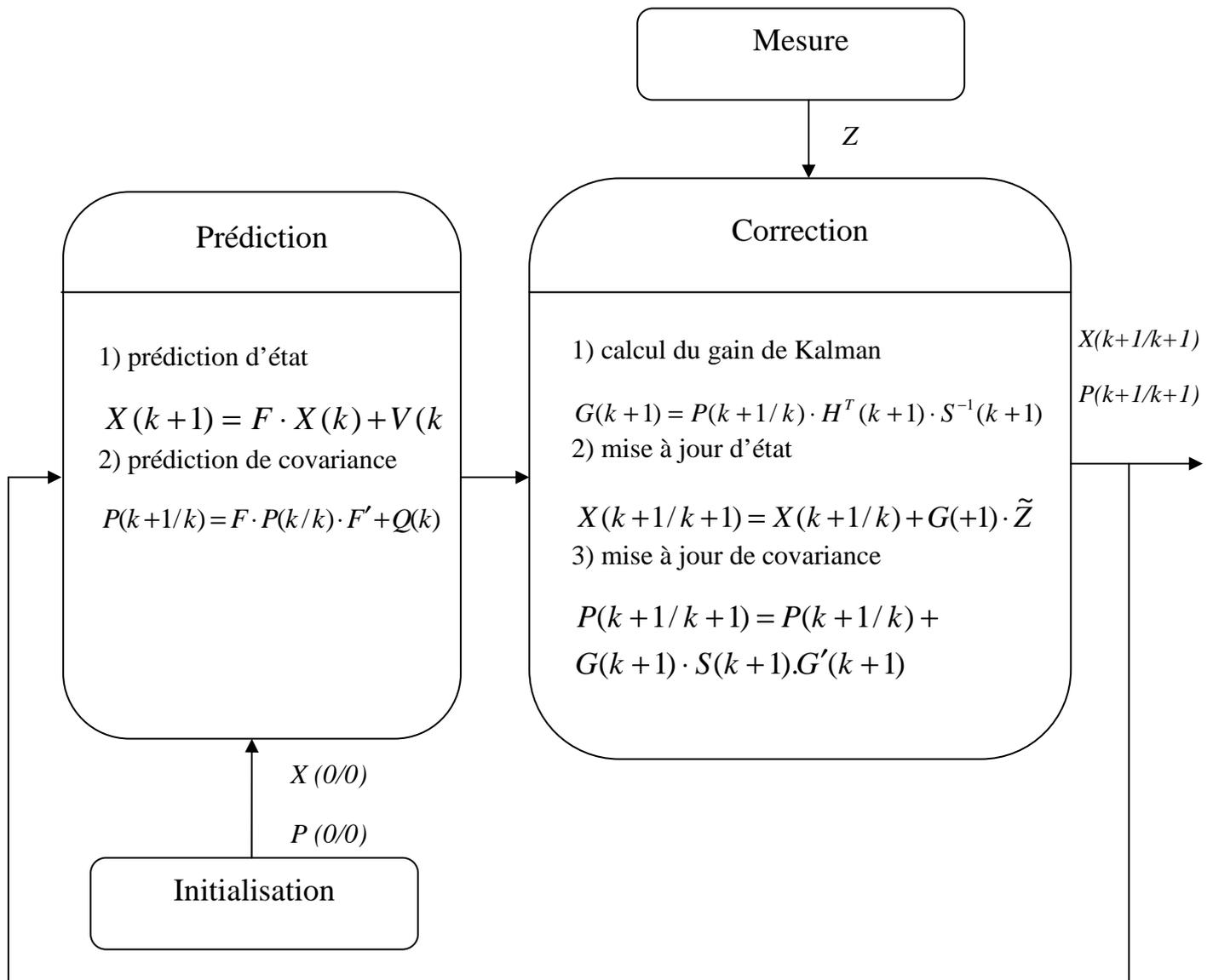


Figure III.3 le cycle du filtre de Kalman avec les équations de prédiction et de correction.

### III.8 Optimalité

L'optimalité du filtre de Kalman est définie comme le fonctionnement permettant d'obtenir une estimation qui minimise la matrice de covariance de l'erreur d'estimation.

A partir de l'équation du gain  $G(k+1)$ , nous pouvons dégager quatre paramètres déterminant la valeur :  $S(k+1)$ ,  $Q$ ,  $H$ ,  $R$ .

Premièrement, une erreur sur  $S(k+1)$  signifie que le modèle utilisé par le filtre ne correspond pas au processus physique à estimer. Par exemple, modéliser le mouvement d'une balle

tombant en chute libre par un modèle à vitesse constante. Ensuite, une mauvaise matrice de covariance  $Q$  implique une mauvaise modélisation de la variation de l'état du système par rapport à son modèle  $S(k+1)$ .

De plus une erreur sur  $H$  ou  $R$  signifie qu'on connaît mal notre système de mesure.

En général, la méconnaissance concerne plutôt  $R$ .

Donc, la sous optimalité est généralement causé par une mauvaise modélisation du processus physique observé et/ou du système de mesure utilisé. Dans ce sens l'effet sous optimalité pour les deux cas mentionnés demeure sensiblement le même. Le gain calculé ne pondère pas adéquatement la mesure et l'estimé à priori. De ce fait, la matrice  $P(k+1)$  ne représente plus comme dans le cas optimal la précision de l'estimé  $X(k+1)$  au sens des moindres carrés, le filtre peut fonctionner adéquatement, par contre l'utilisation n'a aucune mesure de la précision de l'estimer ou de la validité du résultat. Il lui faudra s'assurer du résultat par un autre moyen que par l'analyse de la matrice  $P(k+1)$ . Par exemple, dans un contexte de pistage de cible, l'unité d'association des mesures a besoin de la covariance de l'erreur en position à l'instant  $t(k)$  pour associer la prochaine mesure de l'instant  $t(k+1)$  à une piste de la base de données.

Malheureusement, dans un cas sous optimal, l'unité d'association ne dispose d'aucune information adéquate sur le processus d'estimation pour déterminer les régions de précision des pistes calculées. Il peut alors y avoir une mauvaise association.

### III.9 Application du filtre

Avant de commencer le processus récurant du filtre, il est important de spécifier le vecteur d'état estimé initial  $X(0/0)$  et sa covariance  $P(0/0)$ , de même la nature du bruit de mesure et du processus doivent être connue.

Selon le besoin, la nécessité d'utiliser le filtre de Kalman impose le choix du modèle en fonction de la nature du mouvement de l'objet. Ce modèle est décrit par un système linéaire composée d'une équation d'état et une équation de mesure. On suppose un modèle du second ordre pour un objet se déplaçant à une vitesse constante.

### III.9.1 Le modèle de mouvement

Le vecteur d'état de l'objet, il est composé de :

- La position d'objet.
- La vitesse.
- L'accélération.

- **La position :  $x, y$**

$$\left\{ \begin{array}{l} x = x_0 + \cancel{x}dt + \cancel{x} \frac{dt^2}{2} \\ y = y_0 + \cancel{y}dt + \cancel{y} \frac{dt^2}{2} \end{array} \right. \Rightarrow \left\{ \begin{array}{l} x = x_0 + \cancel{x}dt \\ y = y_0 + \cancel{y}dt \end{array} \right. \quad (\text{III.42})$$

- **La vitesse :  $\cancel{x}, \cancel{y}$**

$$\left\{ \begin{array}{l} \cancel{x} = \cancel{x}_0 + \cancel{x}dt \\ \cancel{y} = \cancel{y}_0 + \cancel{y}dt \end{array} \right. \Rightarrow \left\{ \begin{array}{l} \cancel{x} = \cancel{x}_0 = cst \\ \cancel{y} = \cancel{y}_0 = cst \end{array} \right. \quad (\text{III.43})$$

- **L'accélération :  $\cancel{x}, \cancel{y}$**

$$\left\{ \begin{array}{l} \cancel{x} = 0 \\ \cancel{y} = 0 \end{array} \right. \quad (\text{III.44})$$

### III.9.2 Le modèle de mouvement à vitesse constante

Le vecteur d'état de ce modèle ne contient que deux informations, la position( $x, y$ ) et la vitesse ( $\cancel{x}, \cancel{y}$ ) qui est supposée constante :

$$X(k) = \begin{bmatrix} x \\ y \\ \& \\ \& \end{bmatrix} \quad (\text{III.45})$$

➤ D'après l'expression de l'équation d'état :

$$X(k+1) = F \cdot X(k) + V(k)$$

Et les équations de mouvement (III.42), (III.43) la matrice de transition est :

$$F = \begin{bmatrix} 1 & 0 & dt & 0 \\ 0 & 1 & 0 & dt \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (\text{III.46})$$

➤ D'après l'expression de l'équation de mesure :

$$Z(k) = H \cdot X(k) + W(k)$$

Et le vecteur de mesure (d'observation) :

$$Z(k) = \begin{bmatrix} x \\ y \end{bmatrix} \quad (\text{III.47})$$

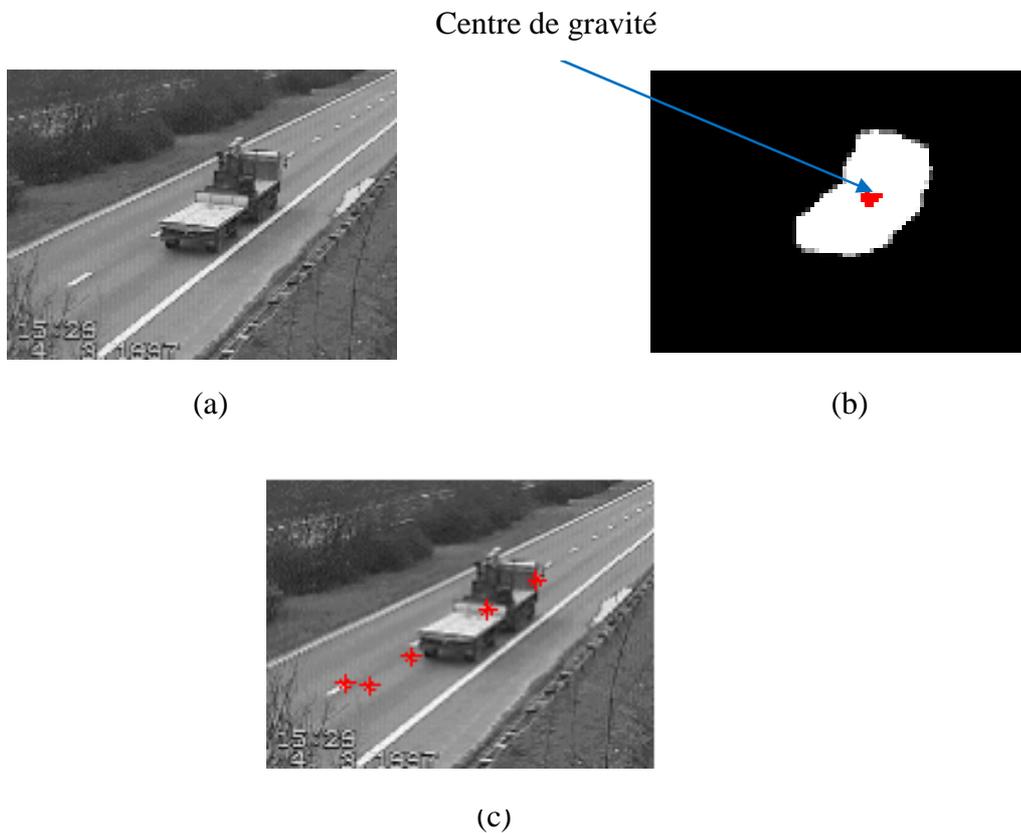
La matrice de mesure est :

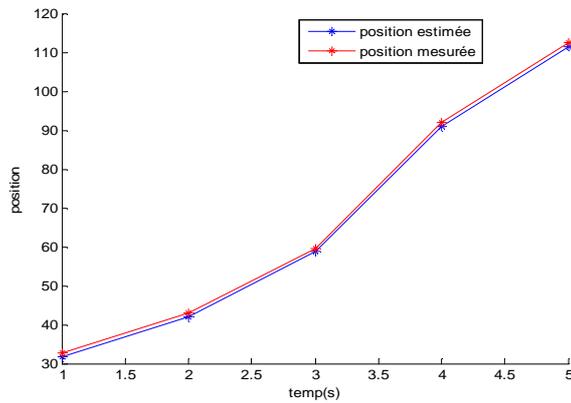
$$H = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \quad (\text{III.48})$$

### Exemple d'application du filtre de Kalman

La figure (III.4) présente un exemple d'application du filtre de Kalman sur la séquence scène routière.

- **Étape 1** : calcul du centre de gravité de l'objet détecté. Le centre de gravité étant le point d'intersection de tous les plans qui divisent le corps en deux parties de poids égaux.
- **Étape 2** : tracer la trajectoire mesurée sur 5 trames.
- **Étape 3** : appliquer le suivi par filtre de Kalman avec comme composantes du vecteur d'état les coordonnées selon l'axe x et y du centre de gravité calculé. Le résultat est un tracé conjoint des trajectoires mesurée et estimée.





(d)



(e)

Figure III.4 (a) Trame à l'instant  $t$ , (b) l'objet en mouvement détecté avec son centre de gravité, (c) trajectoire mesurée sur 5 trames, (d) représentation conjointe des trajectoires mesurée et estimée, (e) représentation conjointe des trajectoires mesurée et estimée sur l'image.

### III.10 Conclusion

Dans ce chapitre, on a présenté des algorithmes qui concernent le suivi d'objets en mouvement dans une séquence d'image. Particulièrement la méthode du filtre de Kalman a été abordée en détails.

Le filtre de Kalman est un outil performant qui permet de résoudre les problèmes d'estimation d'état instantané d'un système dynamique linéaire perturbé et noyé dans du bruit. Ceci peut être projeté au suivi d'objets en mouvement dans une scène.

L'avantage majeur de la méthode du filtre de Kalman est le fait qu'il puisse fournir à chaque itération une estimation des matrices de covariance d'erreur et de l'état. Il faut toutefois initialiser correctement ces matrices.

# *Chapitre IV*

*Implémentation  
pratique et résultats de  
simulations*

## IV.1 Introduction

Après avoir pris connaissance théorique dans les chapitres précédents des principes de base des algorithmes de détection et de suivi d'objets en mouvement dans une séquence d'image, nous allons implémenter dans ce chapitre ces algorithmes pour différentes séquences d'images. Dans une première partie nous présenterons donc les résultats de la détection de mouvement. Le suivi sera par la suite présenté dans une deuxième partie.

### IV.1 Détection de mouvement dans une séquence d'image

#### IV.2.1 Détection par la méthode de segmentation

En premier lieu, nous avons implémenté la segmentation spatiale avec ses différentes étapes : simplification d'image basée sur les segmentations morphologiques (ouverture, fermeture), approximation du gradient et algorithme de ligne de partage des eaux LPE suivi d'un fusionnement des régions.

- **Simplification d'image**

Dans la simplification d'image, on utilise les filtres morphologiques pour lisser des images bruitées, voici les résultats pratiques obtenus par l'implémentation des opérateurs d'ouverture et de fermeture.

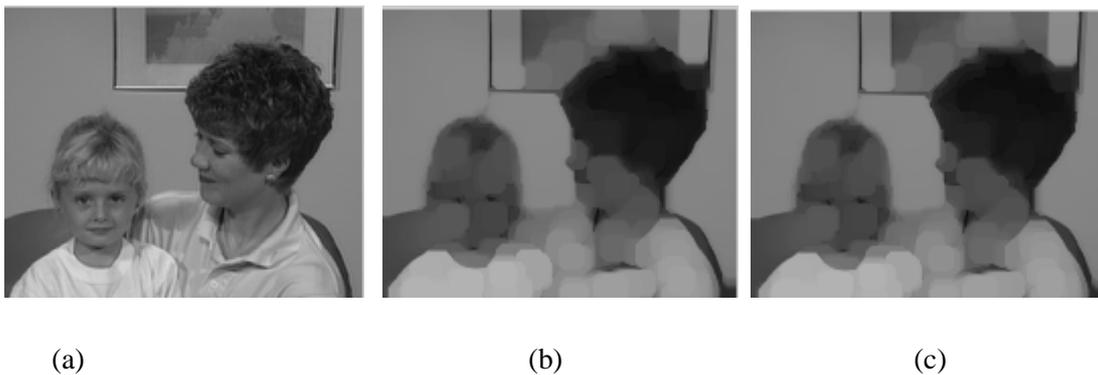


Figure IV.1. Simplification d'image (Ouverture, Fermeture). (a) 5<sup>ème</sup> trame de la séquence Mother & Daughter, (b) Ouverture de la 5<sup>ème</sup> trame de la séquence, (c) Fermeture de la 5<sup>ème</sup> trame de la séquence

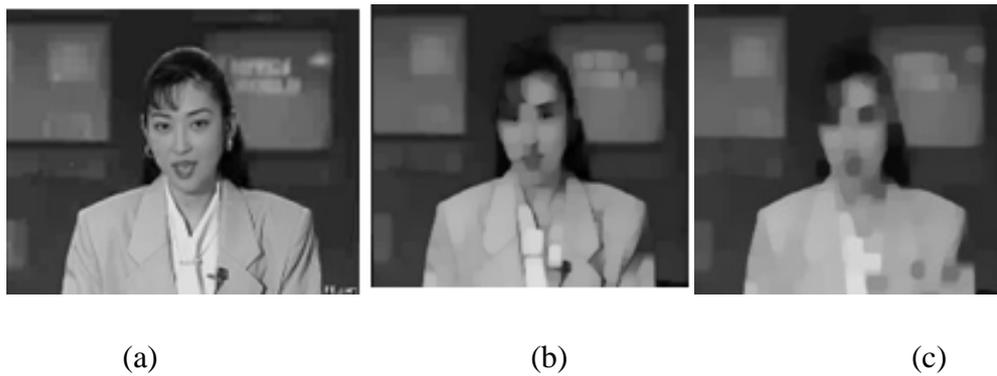


Figure IV.2. Simplification d'image (Ouverture, Fermeture). (a) 2<sup>ème</sup> trame de la séquence Akiyo, (b) Ouverture de la 2<sup>ème</sup> trame de la séquence, (c) Fermeture de la 2<sup>ème</sup> trame de la séquence

- **L'approximation de Gradient**

Le gradient est calculé par la différence entre la dilatation et l'érosion de l'image simplifiée. Les résultats obtenus sont les suivants (Figure IV.3 et IV.4).

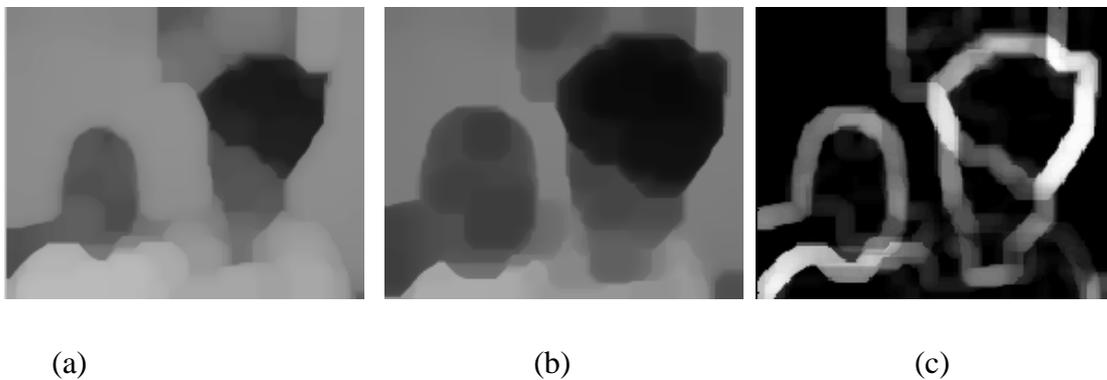


Figure IV.3. (a) La 5<sup>ème</sup> trame de la séquence Mother & Daughter dilatée, (b) La 5<sup>ème</sup> trame de la séquence érodée, (c) Image gradient de la 5<sup>ème</sup> trame de la séquence

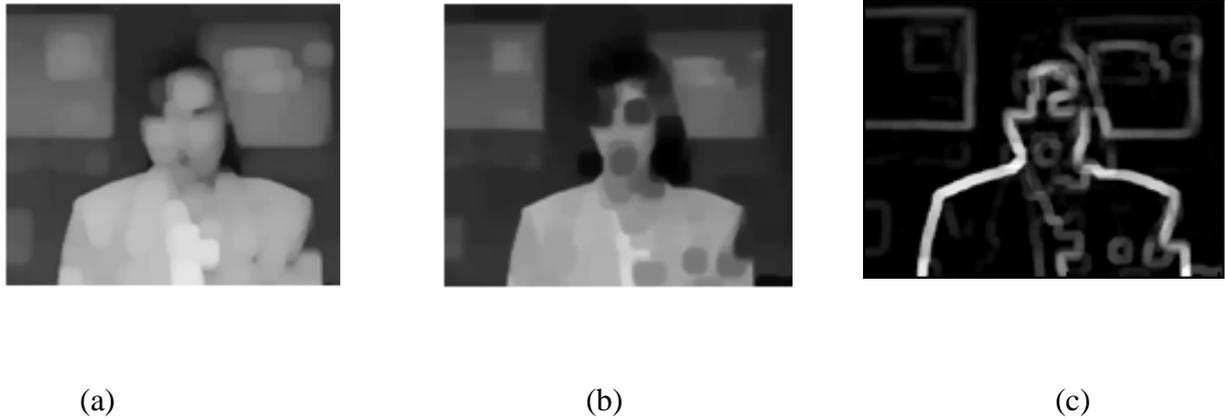


Figure IV.4. (a) La 2<sup>ème</sup> trame de la séquence Akiyo dilatée, (b) La 2<sup>ème</sup> trame de la séquence érodée, (c) Image gradient de la 2<sup>ème</sup> trame de la séquence.

- **Ligne de partage des eaux**

Comme nous l'avons mentionné dans le chapitre III, cette méthode est très sensible à la détection des frontières des objets; ce qui cause un problème de sur segmentation qu'on peut constater aisément sur la figure (IV.5)

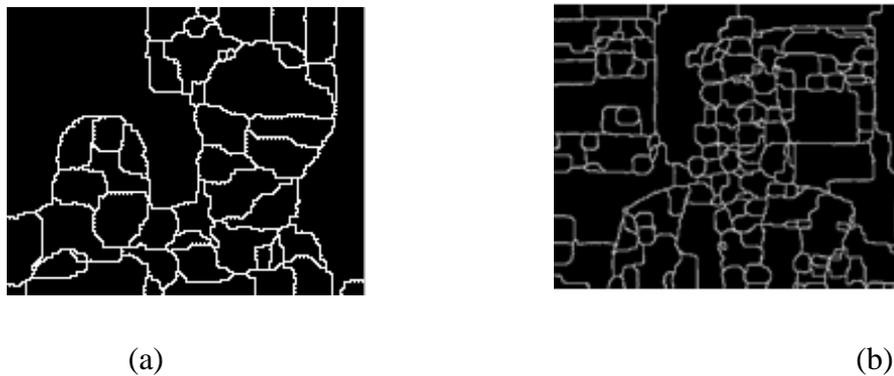


Figure IV.5. Image segmentée par LPE, (a) 5<sup>ème</sup> trame de la séquence Mother & Daughter  
(b) 2<sup>ème</sup> trame de la séquence Akiyo

On voit bien qu'il y a plusieurs régions dans un même objet. Pour éviter ce problème il faut appliquer un critère de fusion sur un graphe d'adjacence des régions.

- **Graphe d'adjacence des régions**

Le graphe d'adjacence des régions est d'abord déterminé comme illustré sur la figure IV.6.



Figure IV.6. Graphe d'adjacence des régions, (a) 5<sup>ème</sup> trame de la séquence Mother & Daughter, (b) 2<sup>ème</sup> trame de la séquence Akiyo.

- **Fusion**

L'objectif majeur de la fusion est de diminuer les sur-segmentations tout en conservant les informations sur les nouvelles zones créés.



Figure IV.7. Fusion des régions (a) 5<sup>ème</sup> trame de la séquence Mother & Daughter  
(b) 2<sup>ème</sup> trame de la séquence Akiyo

Dans le cas spatial, on peut en effet reconstituer les régions obtenues d'après leur niveau de gris. On constate une nette amélioration des résultats, en réduisant la sur-segmentation. Mais

dans le cas du mouvement (notre cas), le critère de fusion à utiliser est très délicat à définir.

#### IV.2.2 Détection par la méthode de l'image de référence

Cette méthode est efficace mais nécessite la disponibilité de l'image de référence ce qui n'est pas toujours évident.



(a)



(b)



(c)



(d)



(e)

Figure IV.8. (a) Trame de séquence scène routière, (b) image de référence, (c) image de différence, (d) image filtré par filtre médian, (e) les objets détectés après seuillage.

On peut voir sur la figure (IV.7) que les objets sont détectés correctement avec cette méthode, mais reste toujours le problème de l'obtention de l'image de référence qui requiert parfois un temps de calcul assez lent et complexe. Reste alors l'alternative d'une détection sans image de référence.

### IV.2.3 Détection sans l'image de référence

Cette méthode se base sur le calcul d'une image de différence.

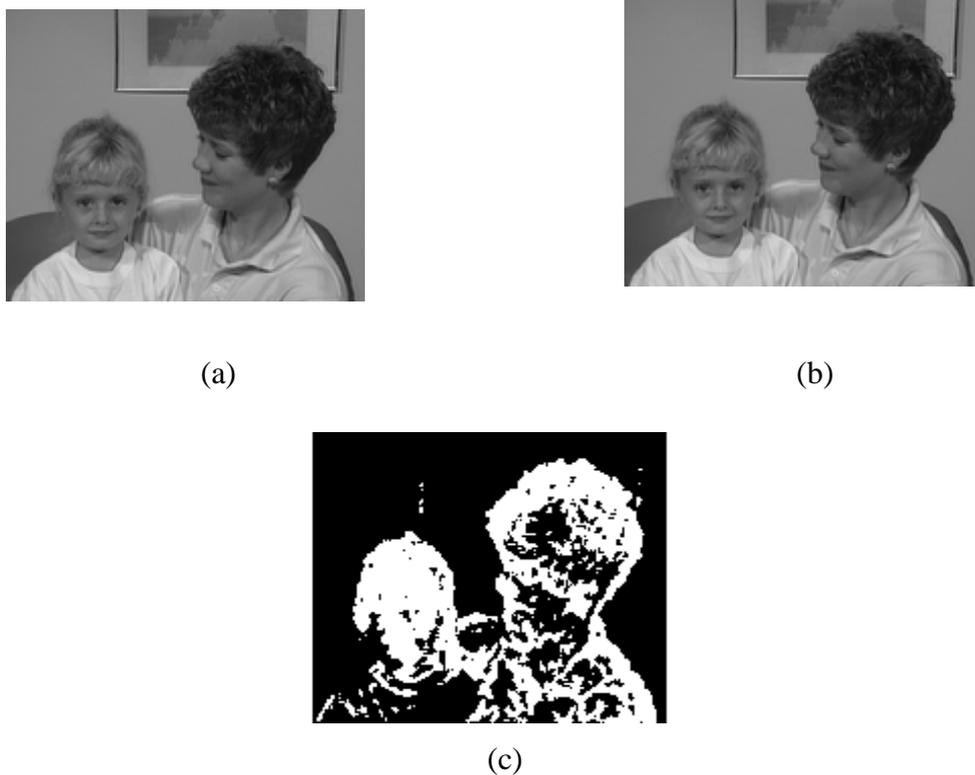


Figure IV.9. (a) Trame à l'instant  $t$ , (b) trame à l'instant  $t+1$ , (c) l'image de différence seuillée.

La figure (IV.9) présente un exemple de détection utilisant l'image de différence sur la séquence « Mother & Daughter » comportant un mouvement faible. On peut voir que les objets en mouvement sont assez bien détectés.

Un autre exemple est présenté dans la figure (IV.10).

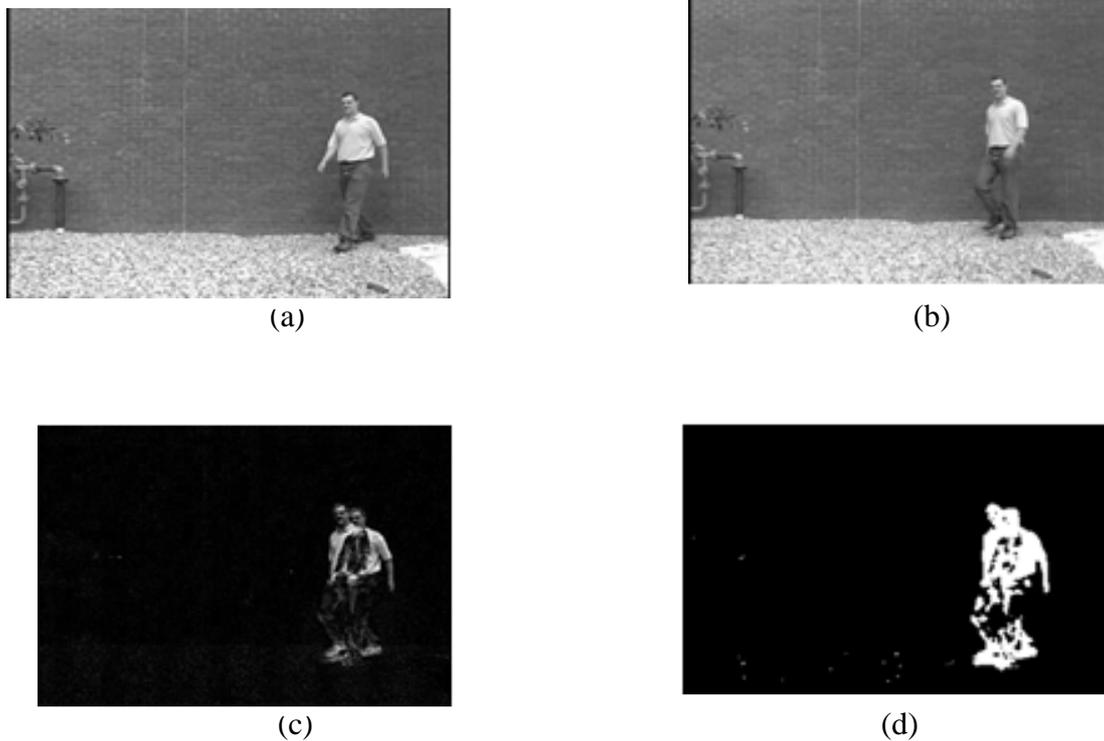


Figure IV.10. (a) 1<sup>re</sup> Trame de la séquence « chercheur », (b) 18<sup>em</sup> trame de la séquence « chercheur », (c) image de différence, (d) image de différence seuillée.

On peut constater que le résultat de la détection n'est pas satisfaisant. Ceci est dû au fait que la différence est calculée entre deux trames distantes (1<sup>re</sup> et 18<sup>em</sup>) et ce pour simuler un mouvement important. On peut donc conclure que cette méthode n'est pas efficace pour les séquences de mouvement important. Pour remédier à ce problème des méthodes se basant sur le calcul du gradient de l'image de différence ont été proposés.

#### IV.2.4 Détection des contours en mouvement

- **Proposition de HAYNES**

Le principe étant de calculer le gradient de l'image courante, le contour est détecté en faisant un produit simple (AND) entre l'image de contours courante et l'image de différence.

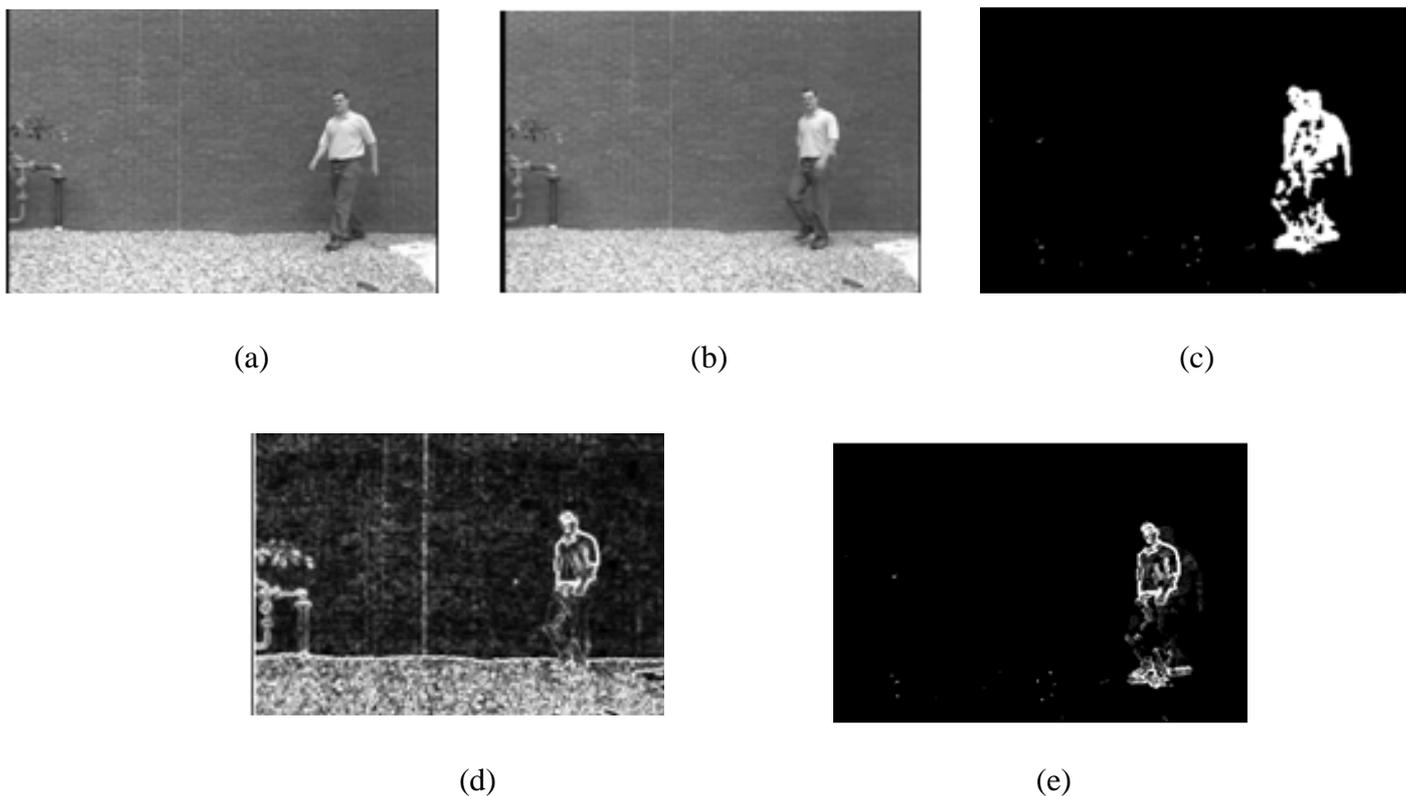
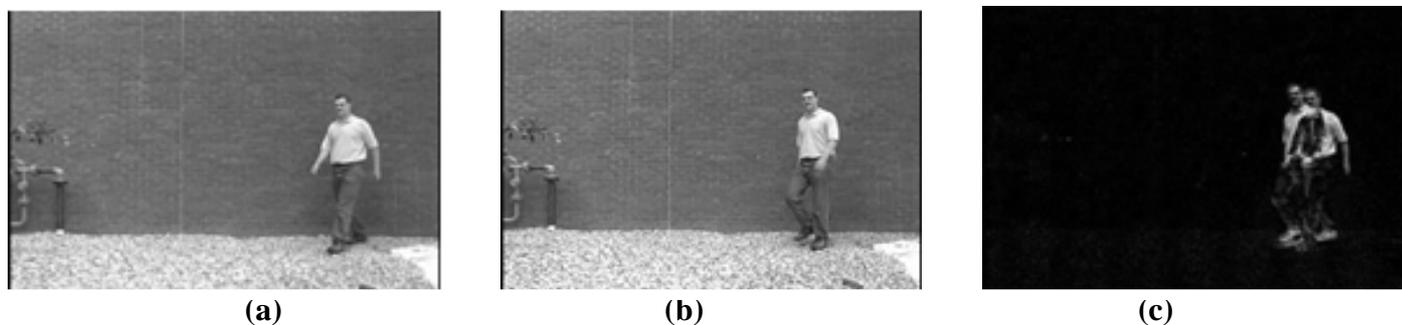


Figure IV.11. (a) Trame à l'instant  $t$ , (b) trame à l'instant  $t+1$ , (c) la différence entre les deux images, (d) gradient de l'image  $t+1$ , (e) contour détecté.

On peut voir que le résultat de la détection n'est pas satisfaisant lorsqu'un objet se déplace parallèlement à l'un de ses contours.

- **Proposition de STELMASZYK**

Dans ce cas le contour est détecté en faisant un produit simple (AND) entre l'image de contours courante et l'image de contours de l'image de différence.



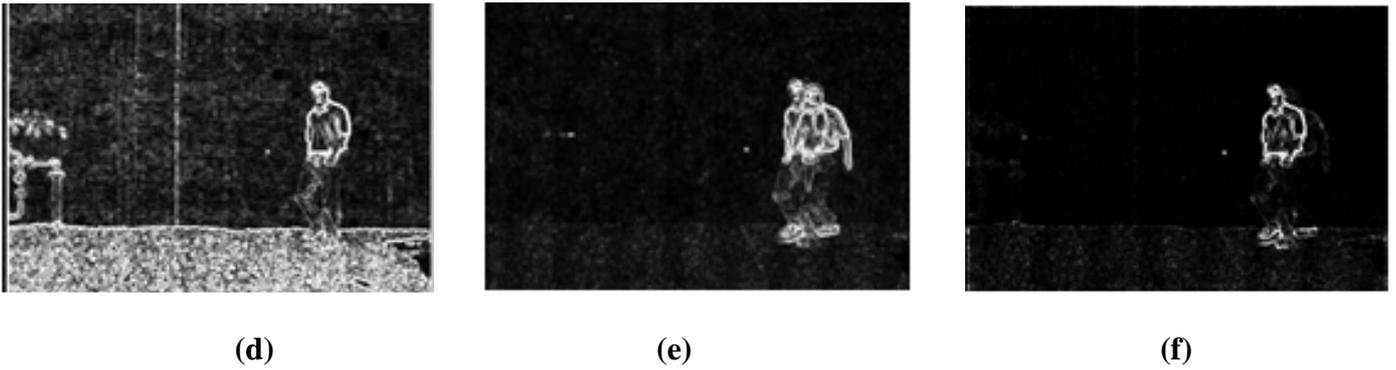


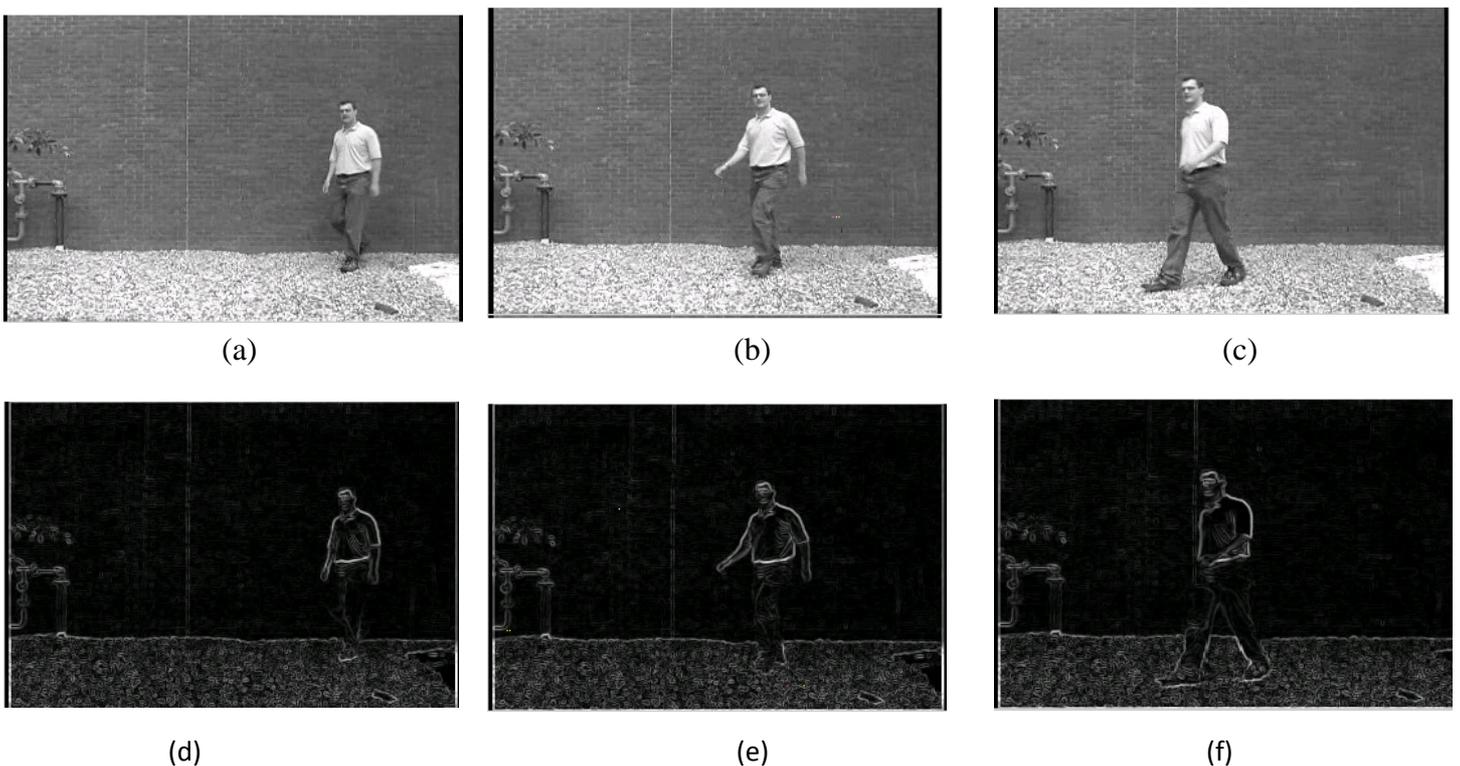
Figure IV.12. Trame à l'instant  $t$ , (b) trame à l'instant  $t+1$ , (c) la différence entre les deux images, (d) gradient de l'image  $t+1$ , (e) gradient de l'image de différence, (f) contour détecté.

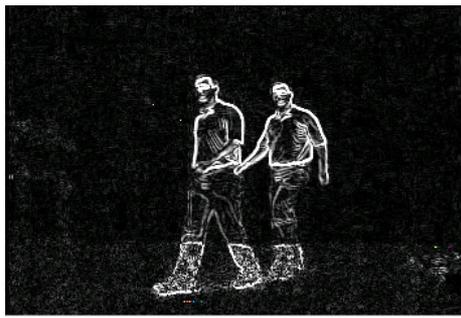
Le résultat obtenu est légèrement meilleur que celui de la méthode de HAYNES.

#### IV.2.5 Détection des contours en mouvement par l'opérateur mcm

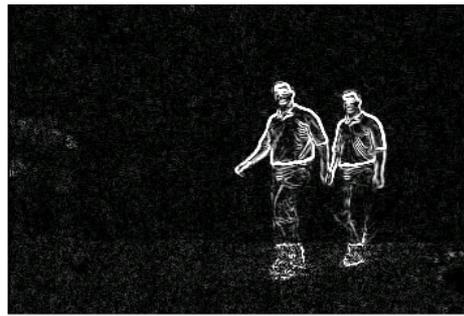
Comme c'est déjà expliqué précédemment, trois trames consécutives sont utilisées dans cette méthode. Un AND logique est opéré entre les différences des images gradients des trames ( $t-1$ ,  $t$ ) et ( $t$ ,  $t+1$ ). Finalement le contour est obtenu.

**Exemple 1 :** la séquence « chercheur » avec un mouvement rapide.





(g)



(h)



(i)

Figure IV.13. (a) 5<sup>ème</sup> Trame de la séquence « chercheur », (b) 15<sup>ème</sup> Trame de la séquence « chercheur », (c) 25<sup>ème</sup> Trame de la séquence « chercheur », (d) gradient de (a), (e) gradient de (b), (f) gradient de (c), (g) différence entre (d) et (e), (h) différence entre (e) et (f), (i) contour en mouvement.

On peut constater que le résultat est nettement meilleur que celui des méthodes précédentes.

Les figures (IV.12), (IV.13) et (IV.14) présentent d'autres exemples.

**Exemple 2 :** la séquence « chercheur » avec un mouvement lent.



(a)



(b)



(c)



(d)



(e)



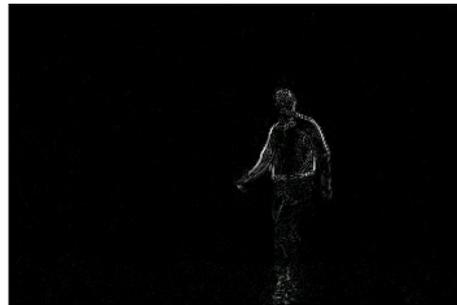
(f)



(g)



(h)



(i)

Figure IV.14. (a) 14<sup>ème</sup> Trame de la séquence « chercheur », (b) 15<sup>ème</sup> Trame de la séquence « chercheur », (c) 16<sup>ème</sup> Trame de la séquence « chercheur », (d) gradient de (a), (e) gradient de (b), (f) gradient de (c), (g) différence entre (d) et (e), (h) différence entre (e)et (f),(i) contour en mouvement.

Le résultat de la détection n'est pas satisfaisant ceci est du au fait que le mouvement est faible.

Exemple 3 : la séquence scène routière compte plusieurs objets en mouvement rapide.

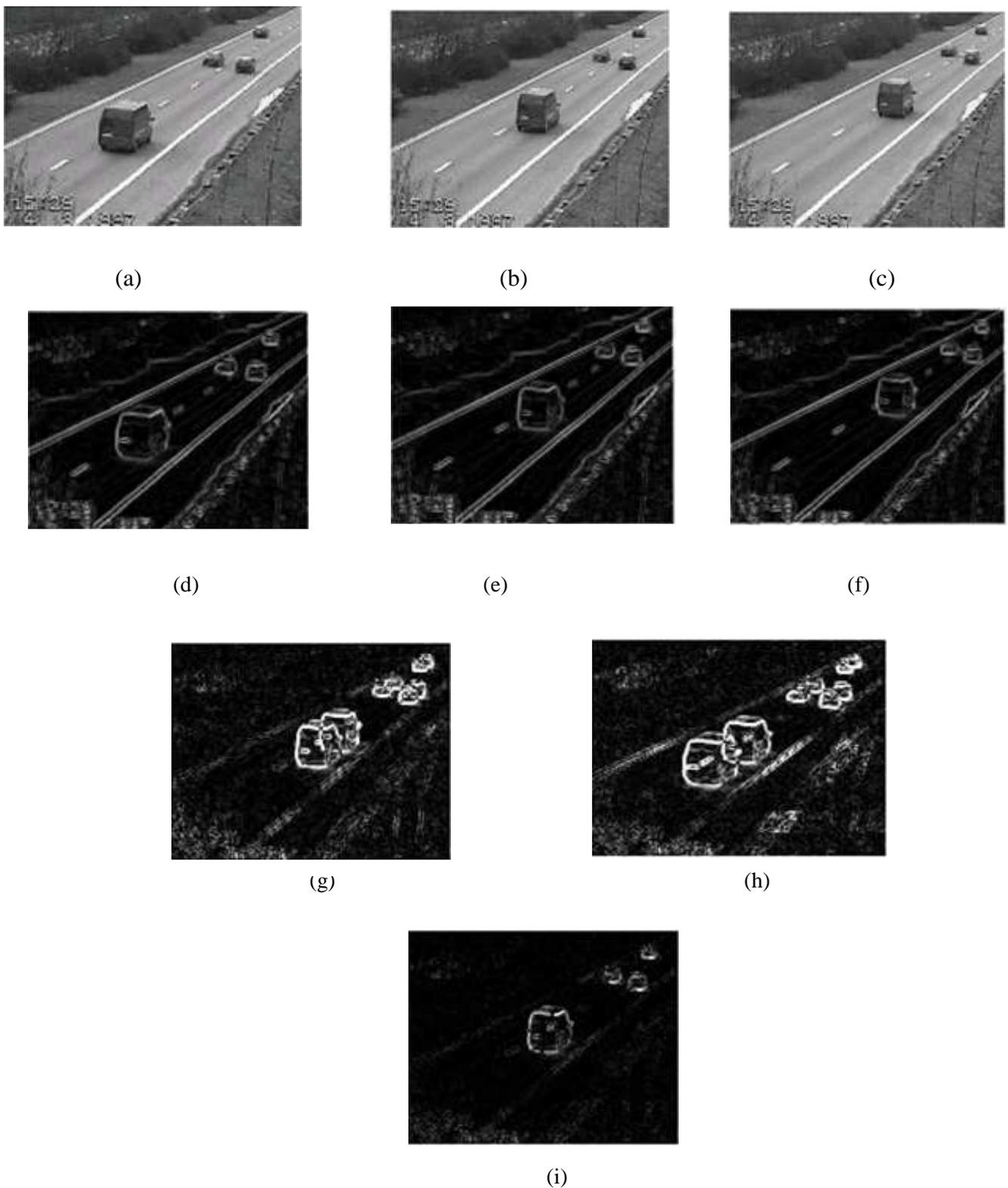


Figure IV.15. (a) Trame à l'instant  $t-1$ , (b) trame à l'instant  $t$ , (c) trame à l'instant  $t+1$ , (d) gradient de (a), (e) gradient de (b), (f) gradient de (c), (g) différence entre (d) et (e), (h) différence entre (e) et (f), (i) contour en mouvement.

On peut voir que les objets en mouvement sont assez bien détectés.

**Exemple4** : la séquence Akiyo présente un mouvement faible.

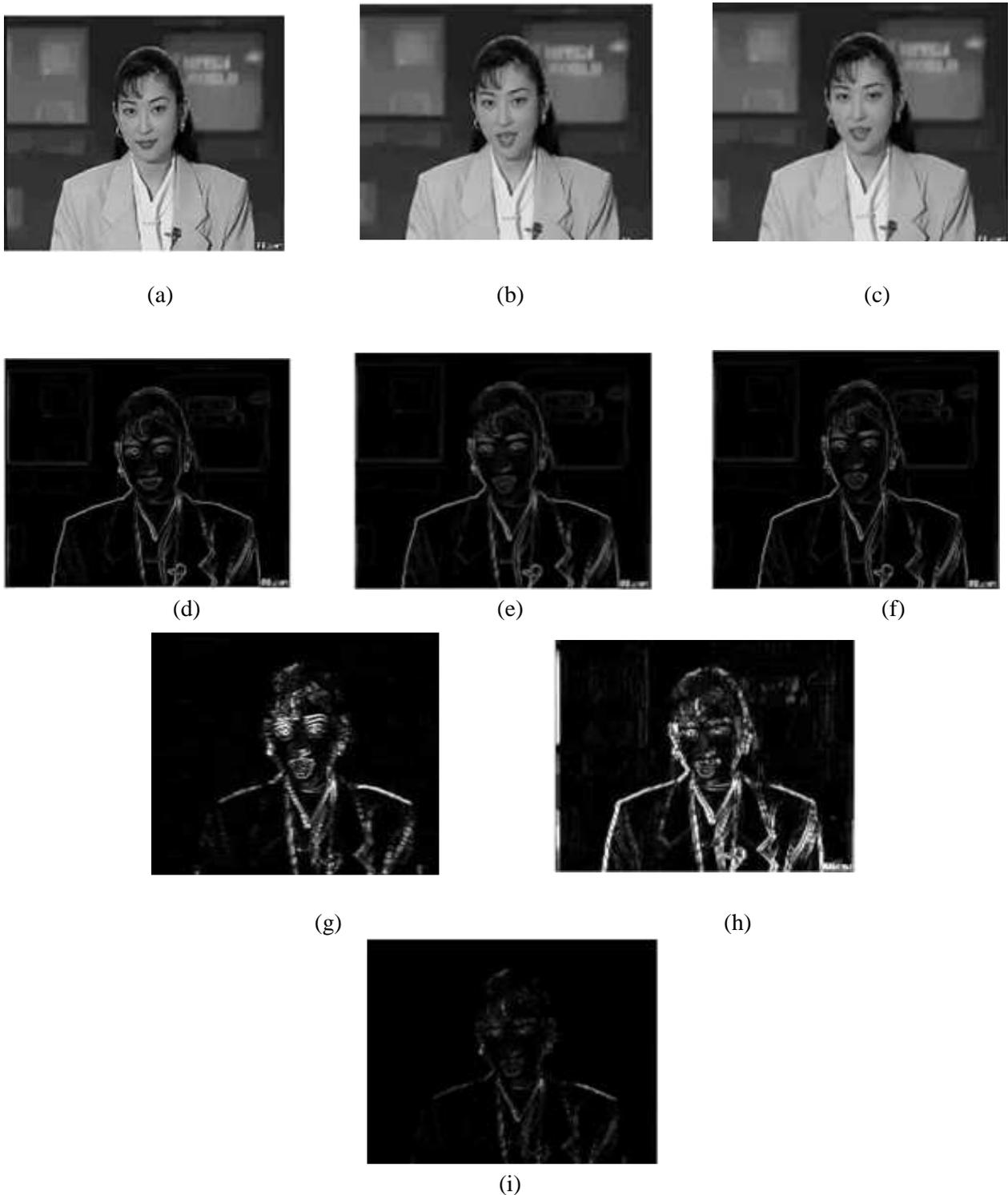


Figure IV.16. (a) Trame à l'instant  $t-1$ , (b) trame à l'instant  $t$ , (c) trame à l'instant  $t+1$ , (d) gradient de (a), (e) gradient de (b), (f) gradient de (c), (g) différence entre (d) et (e), (h) différence entre (e) et (f), (i) contour en mouvement.

Le résultat de la détection n'est pas satisfaisant ceci est du au fait que le mouvement est faible.

Compte tenu des résultats obtenus, deux cas sont envisageables pour la détection d'un objet en mouvement dans une séquence d'image :

le premier est lorsque le mouvement est lent ; dans ce cas on peut utiliser la méthode de différence entre deux images suivi par un seuillage et certaines techniques de morphologie mathématique comme l'indique l'exemple de la figure (IV.17).

Dans le deuxième cas si le mouvement est rapide, on opère une méthode pour extraire les contours des objets en mouvement Ceux-ci sont extraits à l'aide d'un opérateur de mesure des contours en mouvement (mcm) basé sur la différenciation des gradients spatiaux de trois images successives. Les contours forts de chaque image sont obtenus par le gradient. On applique ensuite les différentes techniques de la morphologie mathématique pour extraire les régions en mouvement comme l'indique la figure (IV.18)



(a)



(b)



(c)



(d)



(e)

Figure IV.17. Détection d'objet en mouvement de séquence Mother & Daughter, (a) trame à l'instant  $t$ , (b) trame à l'instant  $t+1$ , (c) différence entre la trame (a) et (b) seuillée, (d) amélioration par les filtres de la morphologie mathématique, (e) objet mobile.



(a)



(b)



(c)



(d)



(e)

Figure IV.18. Détection d'objet en mouvement de séquence scène routière, (a) image originale, (b) détection des contours en mouvement, (c) amélioration par les filtres de la morphologie mathématique, (d) masque des objets en mouvement, (e) objet mobile.

**IV.2 Suivi d’objets en mouvement par filtre de Kalman**

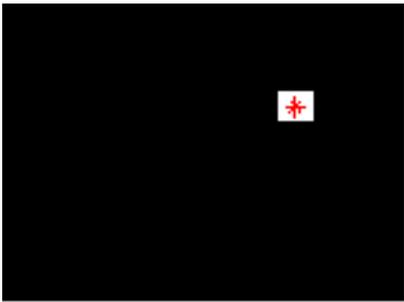
Dans cette deuxième partie, nous allons présenter les résultats des simulations de la poursuite des trajectoires d’objets en mouvement. La poursuite se fait selon les coordonnées cartésiennes sur la direction x seulement.

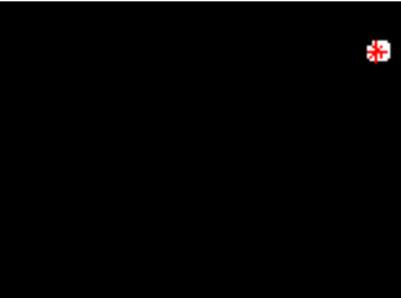
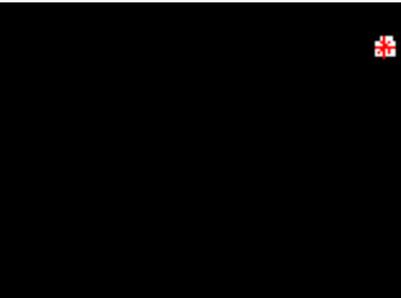
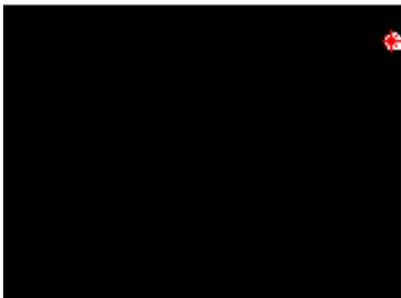
Le vecteur d’état est composé seulement de la position et de la vitesse supposée constante.

La durée de la trajectoire est de 10 trames d’images prises de trois séquences différentes.

La mesure est composée seulement de la position (centre de gravité) de l’objet mobile.

- La figure (IV.19) présente les résultats de la mesure du centre de gravité des objets mobiles (voitures) dans la séquence « scène routière ».

Image en niveau de gris 4 <sup>ème</sup> trame jusqu’à la 14 <sup>ème</sup> trame.	Objet détecté avec son centre de gravité
	
	
	

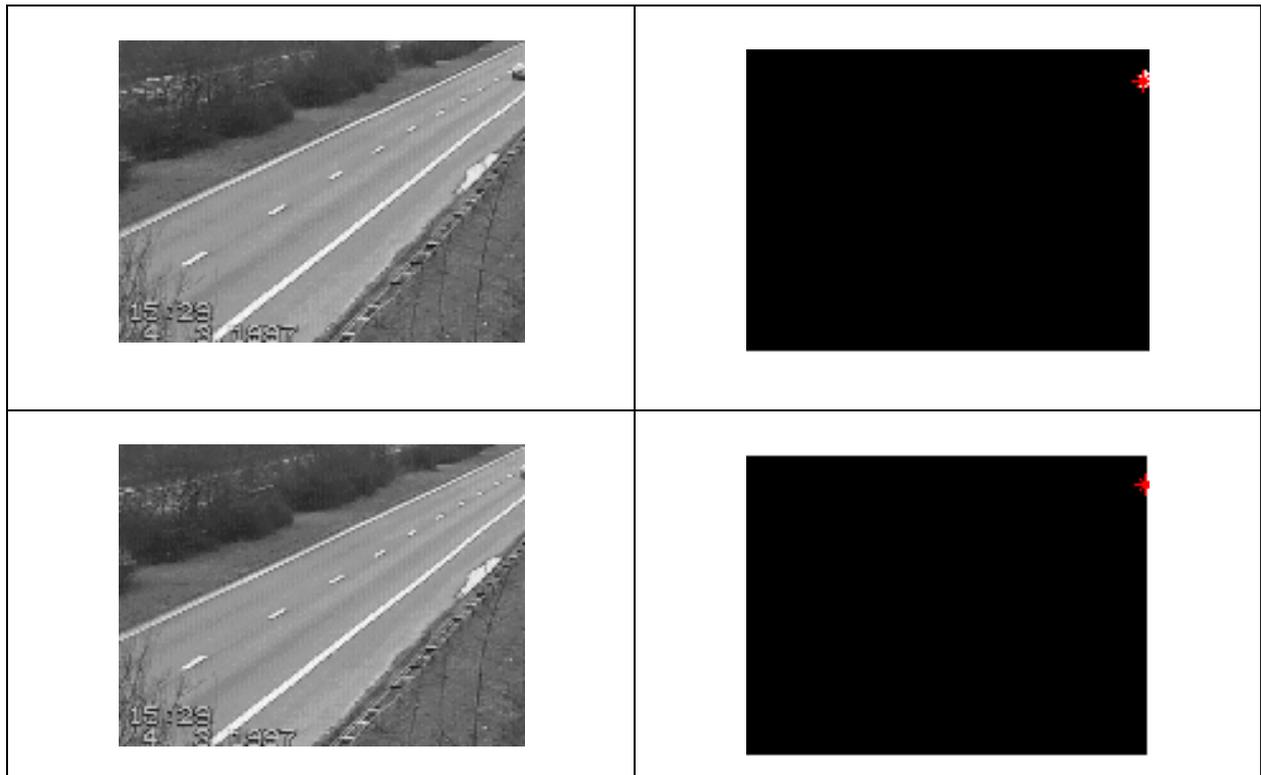


Figure IV.19. Colonne 1 : 10 échantillons d'image (scène routière), colonne 2 :  
Objet détecté avec son centre de gravité.

On peut constater que les centres de gravité des objets mobiles sont détectés correctement.

La figure IV.20 présente la trajectoire mesurée de l'objet (voiture) dans la séquence scène routière.



Figure IV.20. Trajectoire mesurée de l'objet en mouvement.

L'application du filtre de Kalman fournit la trajectoire de l'objet mobile estimée.

La figure IV.21 illustre une représentation conjointe des trajectoires des objets mobiles estimée et mesurée de la séquence « scène routière ».

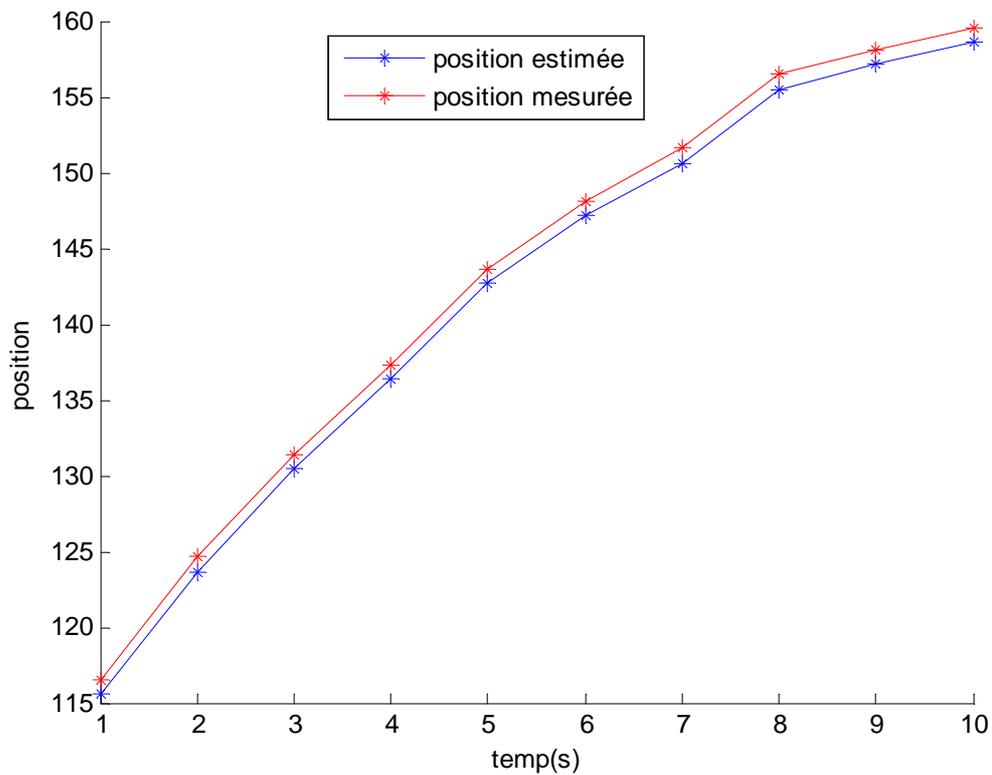


Figure IV.21. Position mesurée/estimée de l'objet mobile.

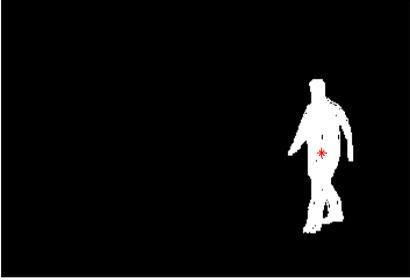
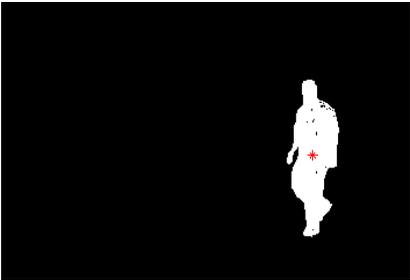
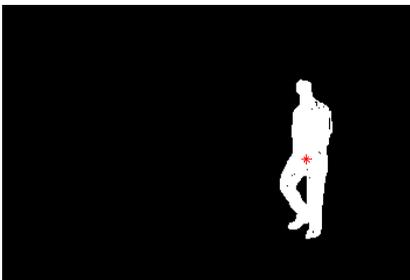
Sur la figure IV.22 les trajectoires mesurée et estimée sont conjointement illustrées sur l'image.

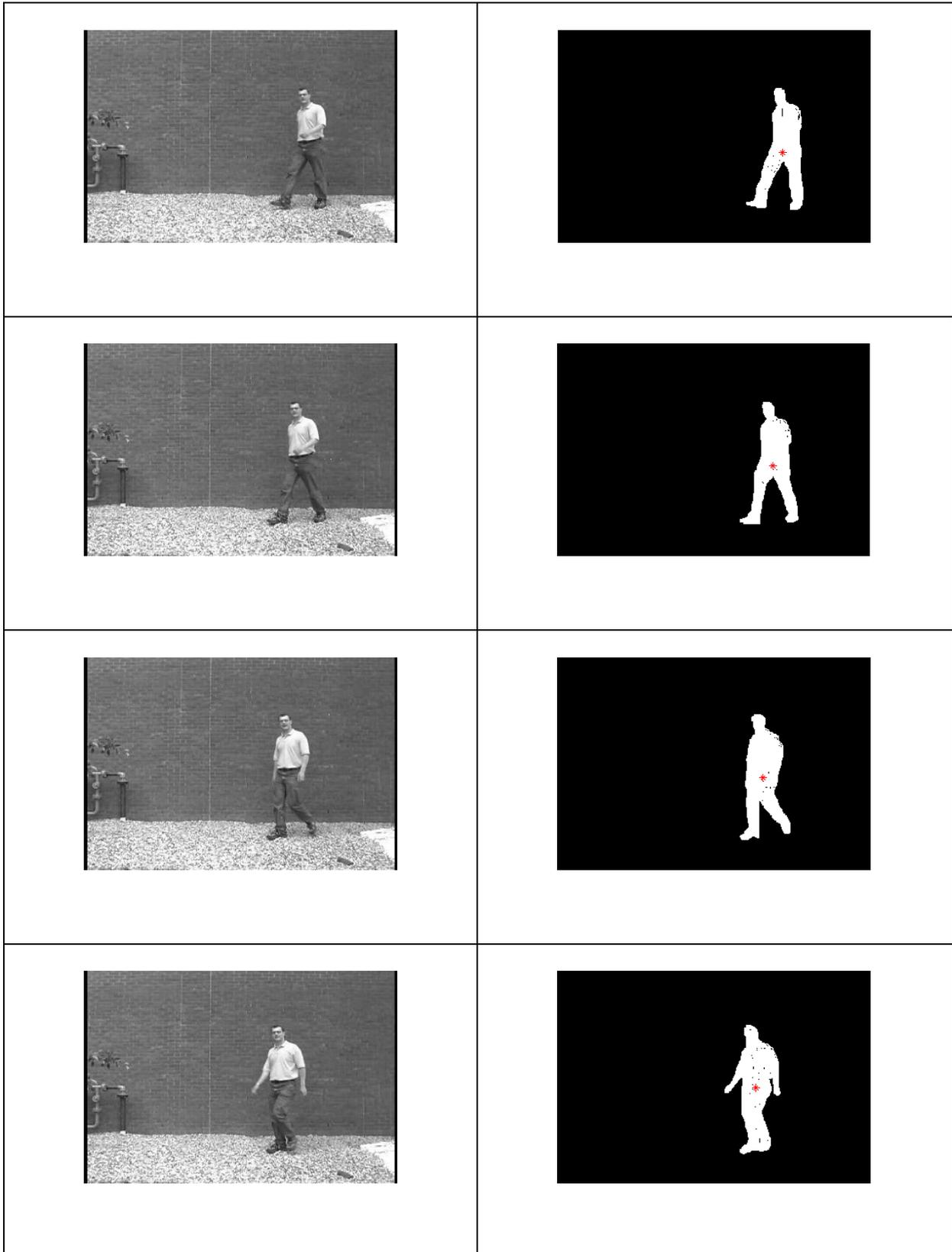


Figure IV.22. Trajectoire mesurée/estimée de l'objet en mouvement.

On peut constater sur les figures (IV.21) et (IV.22) que la différence ou l'erreur commise sur l'estimation de la trajectoire est faible.

- La figure IV.23 présente les résultats de la mesure du centre de gravité de l'objet mobile de la séquence « chercheur ».

Image au niveau de gris 1 <sup>ère</sup> trame jusqu'à la 10 <sup>ème</sup> trame	Objet détecté avec son centre de gravité
	
	
	



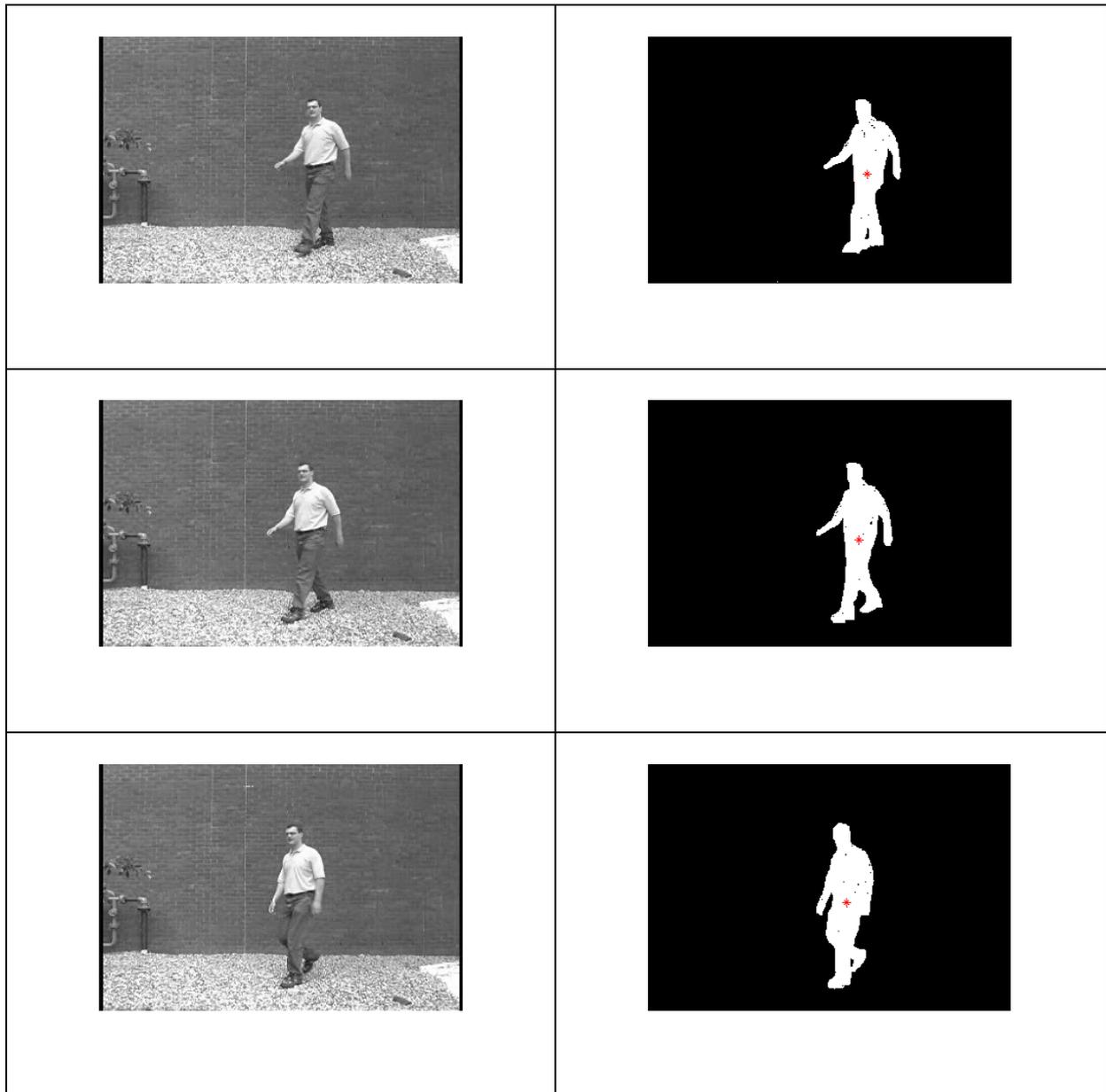


Figure IV.23. Colonne 1 : 10 échantillons d'image de la séquence « chercheur »,  
colonne 2 : Objet détecté avec son centre de gravité.

L'ensemble des centres de gravité mesurés permet de tracer la trajectoire mesurée de l'objet mobile (figure IV.24)

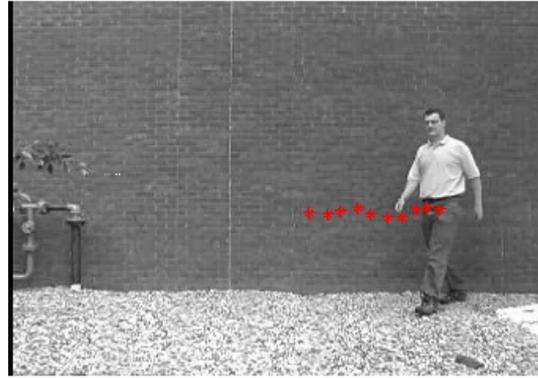


Figure IV.24. Trajectoire mesurée de l'objet en mouvement.

De même que pour la séquence précédente, l'application du filtre de Kalman fournit la trajectoire de l'objet mobile estimée.

La figure IV.25 illustre une représentation conjointe des trajectoires des objets mobiles estimée et mesurée de la séquence « chercheur ».

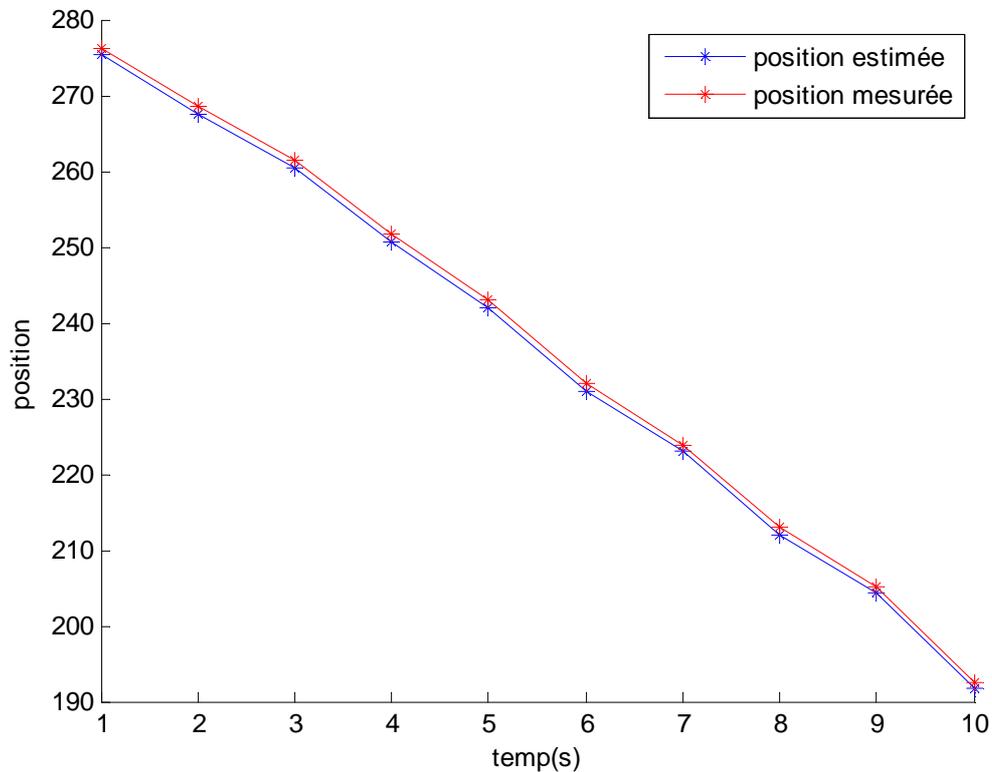


Figure IV.25. Position mesurée/estimée de l'objet mobile.

Sur la figure IV.26 les trajectoires mesurée et estimée sont conjointement illustrées sur l'image.

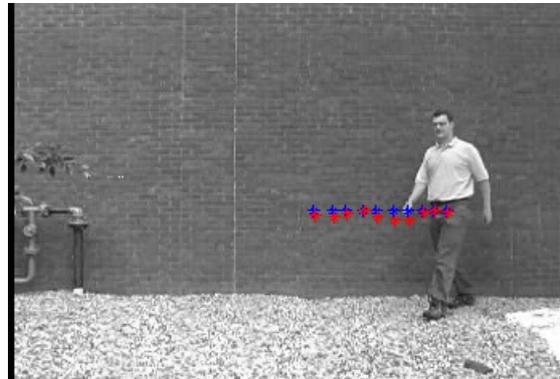
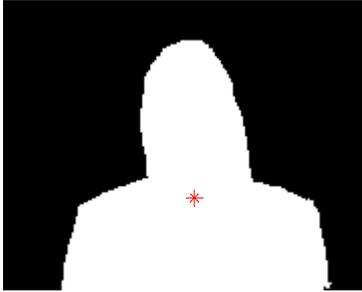


Figure IV.26. Trajectoire mesurée/estimée de l'objet en mouvement.

Dans ce cas aussi, on peut constater que l'erreur est négligeable et l'estimation de la trajectoire de mouvement est satisfaisante.

- La figure IV.27 présente les résultats de la mesure du centre de gravité de l'objet mobile de la séquence « Akiyo ».

Image en niveau de gris	Objet détecté avec centre de gravité
	

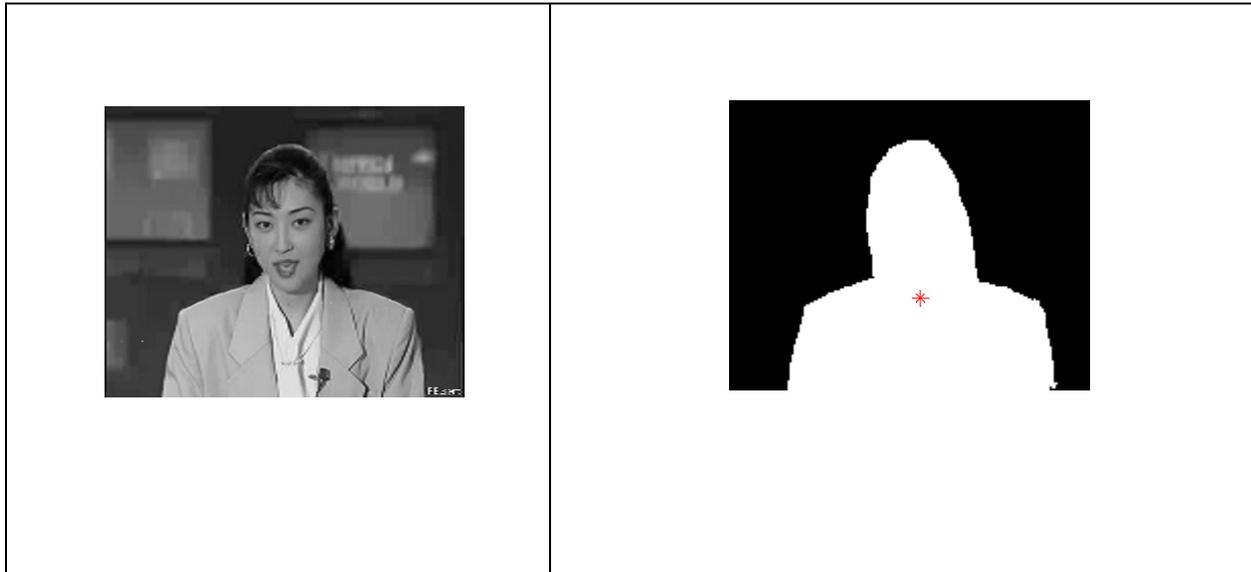


Figure IV.27. Colonne 1 : 2 échantillons d'image (Akiyo), colonne2 : Objet détecté avec son centre de gravité.

La trajectoire mesurée dans ce cas se résume en un seul point qui est le centre de gravité de l'objet mobile.



Figure IV.28. Position mesurée/estimée de l'objet

La trajectoire n'est pas significative dans ce cas parce que le centre de gravité de l'objet ne varie pas avec le temps comme l'indique la figure (IV.28). Ceci est dû au fait que le mouvement est faible contrairement aux résultats précédents (figures (IV.21) et (IV.25)), où la position estimée et la position mesurée varient avec le temps selon le déplacement considérable des objets mobiles.

#### **IV.4 Conclusion**

Dans ce chapitre, nous avons présenté les différents résultats que nous avons obtenus en intégrant les algorithmes de détection développés dans la première partie du mémoire suivis de ceux de poursuite de trajectoire de mouvement présentés dans la deuxième partie. La combinaison de ces deux types d'algorithmes permet de détecter et suivre les objets en mouvement dans une séquence d'image.

La première étape a porté donc sur la détection de l'objet en mouvement par différentes méthodes à savoir la méthode de segmentation par ligne de partage des eaux qui présente un problème de sur segmentation, la méthode de différence qui donne des résultats satisfaisants dans le cas d'un mouvement faible, et finalement la méthode de détection de contours en mouvement par l'opérateur (mcm) qui donne des bons résultats dans le cas d'un mouvement rapide.

La deuxième étape a traité le problème de suivi d'objets en mouvement basé sur l'estimation de la trajectoire d'objet mobile utilisant le filtre de Kalman. Les résultats de cette partie sont acceptable point de vue précision de l'estimation de la trajectoire du mouvement et temps de calcul requis.

# ***Conclusion générale***

# Conclusion générale

Dans ce travail, on s'est intéressé à deux des problèmes majeurs d'analyse de mouvement dans une séquence d'images; la détection et le suivi d'objets en mouvement dans des scènes simples (caméra fixe, fond statique).

Dans la première partie de ce mémoire on a donc proposé une méthode de détection d'objets mobiles, et dans une deuxième partie une méthode de suivi a été appliquée. L'implémentation des ces deux méthodes, qui peuvent être couplées, a été montrée sur des séquences diverses.

D'abord dans l'état de l'art réalisé dans le deuxième chapitre, nous avons présenté les diverses approches d'extraction des objets mobiles de l'arrière plan. Nous y avons proposé des méthodes pour chaque type de mouvement.

Dans le cas d'un mouvement lent (séquence Akiyo, Mother & Daughter, ...etc.) la méthode la plus efficace et la plus courante dans la détection d'objets en mouvement est la méthode de différence entre deux images successive dans la séquence.

Dans le cas d'un mouvement rapide les approches couramment utilisées en segmentation vidéo se limitent également à l'utilisation de deux images successives, certaines ambiguïtés sont alors impossibles à résoudre. En effet, les zones d'occultations sont difficilement attribuables à une région ou bien à un des objets vidéo. De plus, la distinction des objets est impossible lorsque les mouvements sont similaires. Dans ce contexte, nous avons introduit une méthode plus robuste à l'amplitude des déplacements des objets (mouvement rapide) permettant d'extraire les contours des objets en mouvement (mcm) correctement et ce à l'aide d'un opérateur basé sur la différenciation des gradients spatiaux de trois images successives.

En simulations, nous avons implémenté les méthodes de détection de mouvement sur différentes séquences d'images contenant différents types de mouvement. Dans le cas d'un mouvement rapide, la méthode (mcm) s'est avérée la mieux adaptée avec comme avantage

- \* aucune estimation de flot optique.
- \* aucune nécessité de connaissance à priori sur la scène (image de référence).

Par ailleurs, la méthode de différence entre deux images successive est la méthode la plus simple et la plus appropriée dans le cas d'un mouvement lent.

La deuxième partie du mémoire a porté sur le problème de suivi des objets en mouvement.

L'objectif des traitements de cette étape est d'inclure les résultats de la détection dans la continuité de la séquence d'images. Pour cela plusieurs techniques ont été abordées. Particulièrement, la méthode très populaire de résolution des problèmes de suivi qui est l'estimation par filtre de Kalman a été présentée. Le principe général du procédé de l'estimation porte sur l'utilisation de mesures courantes suivies de mises à jour. Dans ce cas le traitement se fait en deux temps. Premièrement, on génère une prédiction de ce que pourrait être le résultat courant à partir de l'ensemble des résultats précédents, puis l'on compare le résultat estimé avec les mesures réellement obtenues. Les données des étapes ultérieures ont alors un rôle de vérification

L'application du filtre de Kalman dans l'estimation des trajectoires de l'objet en mouvement à donné des résultats probants point de vue précision de l'estimation de la trajectoire du mouvement et temps de calcul requis.

### **Perspectives**

En perspectives, on propose :

- De chercher une méthode unique pour la détection et le suivi d'objets mobiles pour les deux types de mouvement (lent et rapide).
- L'application des méthodes de détection et de suivi sur les séquences à fonds dynamiques (caméras mobiles).
- L'idéal c'est d'avoir un système automatique entièrement autonome de l'opérateur humain. Malheureusement pour atteindre ce but, il faudra tout d'abord se disposer de plusieurs cameras pour avoir un maximum de prises de vue sur différents angles et d'avoir un système DSP pour le traitement parallèle. Etant donné que le filtre de Kalman présente une structure de données assez complexe, le système qui gère l'ensemble devra posséder une capacité en mémoire importante pour ne pas avoir une grande influence sur le temps de traitement.

# *Bibliographie*

## ***Bibliographie***

- [1] F. Berry, Technologies des capteurs d'images couleurs, In Imagerie Numérique Couleur GFINC, G. F, editor, EHINC'05, Lille, 2005.
- [2] N. Laouar, M. S. Laraba, Détection d'un mouvement dans une séquence vidéo par filtres morphologiques, Projet de fin d'étude, ENP, Alger, 2009.
- [3] Aek . Tahari, Détection des objets en mouvements dans une séquence d'images, Thèse de Magister, ESI (ex INI), Alger, 2002.
- [4] N. Guealia, L. Hadji, Suivi d'objet en mouvement, Projet de fin d'étude, ESI (ex INI), Alger, 2003.
- [5] C. Cafforio et F. Rocca, Methods for measuring small displacements of television images, IEEE-Trans. Information Theory, IT-22, n° 5, p. 573-579, sept. 1976.
- [6] B. Choquet et J. Santillana Rivero, Algorithmes de détection et d'estimation de mouvement des images de télévision, RT-CCETT/RTI, 007-85, mars 1985.
- [7] J. A. Leese, C. S. Novak et B. B. Clark, An automated technique for obtaining cloud motion from geosynchronous satellite data using cross-correlation, Jal of Applied Meteorology, 10, p. 118-132, fév. 1971.
- [8] A. L. Gilbert, M. K. Giles, G. M. Flachs, R. B. Rogers et Y. Hsun, A real-time video tracking system, IEEE-Trans. Pattern Analysis and Machine Intelligence, PAMI-2, n° 1, p. 47-56, janv. 1980.
- [9] J. L. Barron, D. J. Fleet, et Beauchemin, S. S, Performance of optical ow techniques, Int. Journal of Computer Vision, 12(1):43 -77,1994.
- [10] T. Yang, S. Z. Li, Q. Pan, et J. Li, Real-time and accurate segmentation of moving objects in dynamic scene, In Proc. ACM 2nd Int. Workshop on Video Surveillance & Sensor Networks, pages 136-143, New York (NY), USA, VSSN 2004.
- [11] Y. Weiss, et E. H. A. Adelson, uniéd mixture framework for motion segmentation: Incorporating spatial coherence and estimating the number of models, In Proc. IEEE Conf. on Computer Vision and Pattern Recognition (CVPR'96), pages 321-326, San Francisco (CA), USA, 1996.
- [12] J.-M. Odobez, et P. Bouthemy, Direct incremental model-based image motion segmentation for video analysis, Signal Processing, 6(2):143-155,1998.

- [13] M. Chaumont, représentation en objets vidéo pour un codage progressif et concurrentiel des séquences d'images, thèse doctorat, université de rennes 1, novembre 2000.
- [14] P. Salembier and M. Pardas, Hierarchical morphological segmentation for image sequence coding, IEEE Trans. Image Processing, vol. 3, no.5, pp. 639–651, 1994
- [15] Vincent Agnus, Segmentation spatio-temporelle de séquence d'image par des opérateurs des morphologies mathématique, thèse doctorat. Université Louis Pasteur de Strasbourg-Ecole Doctorale SPI. pp 22, 2001.
- [16] S. Castan, J. Zhao, J. Shen, Une famille de Détection de Contour Basée sur Filtre Exponentiel Optimal, 7<sup>ème</sup> Edition 1986.
- [17] Christophe Fiorio, Analyse d'image, Split and Merge, Laboratoire d'Informatique, de Robotique et Microélectronique de Montpellier LIRMM France Novembre 2005.
- [18] K. K. C. Evelet et R. Bolles, Background modeling for segmentation of video-rate stereo sequences, Dans CVPR98, pages 266–272, 1998.
- [19] W. Long et Y. Yang, Stationary background generation : An alternative to the difference of two images, Pattern Recognition, 23 :1351–1359, 1990.
- [20] P. Vannoorenberghe, Détection de mouvement par analyse de séquence d'images monoculaires, Application à l'estimation de flux de piétons en milieu urbain. PhD thesis, Laboratoire d'Analyse des Systèmes du Littoral - Université du Littoral, France, 1997.
- [21] R. Jain, Différence and accumulative difference pictures in dynamic scene analysis, Image and Vision Computing, 2(2) :99–108, may 1984.
- [22] A. Makarov, Comparison of background extraction based intrusion detection algorithms, Dans International Conference On Image Processing ICIP'96, pages 521–524, 1996.
- [23] S. Haynes, Detection of moving edges, Computer Graphics and Image Processing, 21, 1983.
- [24] P. Stelmaszyk. Analyse de scènes dynamiques par recherche des contours en mouvement. PhD thesis, Université des Sciences and Technologies de Lille, Novembre 1985.
- [25] C. Vieren, Segmentation de scènes dynamiques en temps réel. Application au traitement de séquences d'images pour la surveillance de carrefours routiers, PhD thesis, Université des Sciences and Technologies de Lille, Avril 1988.
- [26] M. Orkisz, localisation d'objets mobiles dans les scènes naturelles filmées par une caméra fixe, Traitement du Signal, 9(4) :325–346, 1992.

- [27] D.S. Zhang et G. Lu, Segmentation of Moving Objects in Image Sequence: A Review, *Circuits, Systems and Signal Processing (Special Issue on Multimedia Communication Services)*, volume 20(2), p 143-183, 2001.
- [28] M.J. Black et P. Anandan, The robust estimation of multiple motions : Parametric and piecewise-smooth flow fields, *CVIU*, volume 63, p. 75-104, Janvier 1996.
- [29] R. Deriche, Fast algorithms for low-level vision, *IEEE-PAMI*, volume 12(1), p 78-87, 1990.
- [30] Y. Ricquebourg, Analyse de mouvements articulés : mesure et suivi 2d ; application à la télésurveillance, Thèse de Doctorat de l'Université de Rennes I), Janvier 1997.
- [31] E. François, Interprétation qualitative du mouvement à partir d'une séquence d'images, Thèse de Doctorat de l'Université de Rennes I, Juin 1991.
- [32] S. Venegas-Martinez, Analyse et segmentation de séquences d'images en vue d'une reconnaissance de formes efficace, Thèse de Doctorat de l'Université René Descartes - Paris V, 2001.
- [33] M. Gelgon, Segmentation spatio-temporelle et suivi dans une séquence d'images : application à la structuration et à l'indexation de vidéo, Thèse de doctorat, Université de Rennes 1, 1998.
- [34] H. Nicolas, Hiérarchie de modèles de mouvement et méthodes d'estimation associées, application au codage de séquences d'images, Thèse de doctorat, IRISA-Université de Rennes 1, 1992.
- [35] H. Nicolas et C. Labit, Motion and illumination variation estimation using a hierarchy of models : application to image sequence coding, Publication interne 742, IRISA, 1993.
- [36] L. Bonnaudet C. Labit, Etude d'algorithmes de suivi temporel de segmentation basée mouvement pour la compression de séquences d'images, Publication interne 793, IRISA-Université de Rennes, 43 pages, Janvier 1994.
- [37] L. Bonnaud, Schémas de suivi d'objets vidéo dans une séquence animée : application à l'interpolation d'images intermédiaires, Thèse de doctorat, IRISA-Université de Rennes, 1998.
- [38] D.G. Lowe, Robust model-based motion tracking through the intégration of search and estimation, *International Journal of Computer Vision*, 8(2) :113–122, 1992.
- [39] B. Bascle, P. Bouthemy, R. Deriche, et F. Meyer, Suivi de primitives complexes sur une séquence d'images, Rapport de recherche 2428, INRIA-Sophia Antipolis, Décembre 1994. 22 pages.

- [40] T. Broida et R. Chellappa, Estimation of object motion parameters from noisy images, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-8(1) :90–99, January 1986.
- [41] M. Kass, A. Witkin, et D. Terzopoulos, Snakes : Active contour models. *International Journal of Computer Vision*, 1 :321–331, 1998.
- [42] V. Caselles, F. Catte, T. Coll, et F. Dibos. A geometric, model for active contours in image processing, *Numerische Mathematik*, 66 :1– 31, 1993.
- [43] C. Chesnaud. Techniques statistiques de segmentation par contour actif et mise en oeuvre rapide. Thèse de Doctorat de l'Université de Aix-Marseille), Février 2000.
- [44] N.K. Paragios, Geodesic active region and level set methods : Contribution and applications in artificial vision, Thèse de doctorat, Sophia Antipolis, Janvier 2000.
- [45] S. Jehan-Besson, Modèles de contours actifs basés régions pour la segmentation d'images et de vidéos, Thèse de Doctorat de l'Université de Nice - Sophia Antipolis, Janvier 2003.
- [46] F.G. Meyer et P. Bouthemy, Region-based tracking using affine motion models in long image sequences, *CVGIP : Image Understanding*, 60(2) :119–140, September 1994.
- [47] B. Bascle et R. Deriche ,Region tracking through image sequences, Rapport de recherche 2439, INRIA-Sophia Antipolis, 21 pages, Décembre 1994.
- [48] V. Garcia-Garduno et C. Labit, Suivi de segmentation spatiotemporelle pour la compression de séquences animées, Quatorzième Colloque GRETSI, Juan-les-Pins, Septembre 1993.
- [49] V. Garcia-Garduno, Une approche de compression orientée-objets par suivi de segmentation basée mouvement pour le codage de séquences d'images numériques, Thèse de doctorat, Université de Rennes 1, 1995.
- [50] A. Yilmaz, X. Li et M. Shah, Object tracking: a survey, ACM New York, NY, USA, 2006.
- [51] C.J. Veenman, M.J.T. Reinders et E. Backer, Resolving Motion correspondence for Densely Moving Points, *IEEE Transactions on Pattern*, 2001.
- [52] S. Birchfield, Elliptical Head Tracking Using Intensity Gradients and Color Histograms, *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Santa Barbara, California, pages 232-237, June 1998.
- [53] H. Schweitzer, J. Bell et F. Wu, Very fast template matching, *Computer Vision—ECCV* 2002, 2006.

- [54] H. Tao, Harpreet , S. Sawhney et R. Kumar, Object tracking with Bayesian estimation of dynamic layer representations, IEEE Trans. Pattern Analysis and Machine Intelligence (PAMI), vol. 24, no. 1, pp. 75-89, 2002.
- [55] A. Almeida, J. Almeida, and R. Araujo, Real-time tracking of multiple moving objects using particle filters and probabilistic data association, Automatika, vol. 46, no. 1-2, pp. 39-48, 2005.
- [56] E. Brookner, Tracking and Kalman Filtering Made Easy, John Wiley & Sons, 1998.
- [57] R. Raquel, Pinho, R. S. Joao Manuel, Tavares, V. Miguel. Correia, An Efficient and Robust Tracking System using Kalman Filter, VIPSI- 2006 VENICE, 2006.
- [58] Y. Bar-shalom and T.E. Fortman, Tracking and Data Association New York : Academic, 1988.
- [59] J. Thorp, Optimal tracking of Manoeuvring Targets, IEEE Trans. On Aerospace and Electronic Systems, Vol. AES-9, No. 4, pp. 512-519, 1973.

# *Résumé*

# Résumé

De nombreuses applications en vision par ordinateur nécessitent la détection et le suivi des objets en mouvement dans une séquence d'images. Le but de ce travail est de détecter et suivre les objets mobiles dans des séquences ayant un fond statique. Dans la première partie, nous proposons une nouvelle méthode pour extraire les objets en mouvement dans une séquence d'images sans aucune estimation de flot optique, ni aucun a priori sur la scène, comme une image de fond. L'objectif est d'obtenir une méthode d'extraction rapide et efficace, robuste au bruit et aux forts déplacements. On a choisi deux approches, la première est basée sur des différences entre deux d'images successives (dans le cas d'un mouvement lent) et dans le cas d'un mouvement rapide nous proposons un nouveau schéma de différenciation basé sur le gradient spatial d'images successives, permettant d'extraire les contours des objets en mouvement entre trois images successives et après on utilise des opérateurs de morphologie mathématique pour localiser l'objet en mouvement. La deuxième partie propose une méthode de suivi par filtre de KALMAN qui permette de construire les trajectoires d'objets mobiles.

## الملخص

العديد من التطبيقات في الرؤية بالكمبيوتر تتطلب كشف و تعقب الأجسام المتحركة في سلسلة من الصور الهدف من هذا العمل هو لكشف و تعقب الأجسام المتحركة في سلسلة من الصور مع خلفية ثابتة في أول الأمر علينا إيجاد طريقة جديدة لاستخراج الأجسام المتحركة بدون أي تقدير في التدفق البصري و أيضا بدون الرجوع إلى سلسلة الصور لاستخراج الصورة الخلفية الهدف هو إيجاد طريقة سهلة و فعالة لهذا لجانا الى تطبيق طريقتين الأولى عندما تكون حركة الأجسام بطيئة مثل سلسلة مقدمة الأخبار في هذه الحالة الطريقة الأكثر شيوعا هي أسلوب الفرق بين صورتين في السلسلة أما الطريقة الثانية و هي عندما تكون حركة الأجسام سريعة مثل حركة السيارات لجانا الى تطبيق طريقة استخراج حواف الأجسام المتحركة بين ثلاث صور متتالية إما الجزء الثاني من هذا العمل فنقترح طريقة مرشح كالمان التي تسمح في تعقب مسارات الأجسام المتحركة.