

# VISION PAR ORDINATEUR

Rapport de TP2

IFI-Promotion 23

---

## Compter des passagers

---



*Auteurs :*

M. OLEMBO RÉEL DEVIN

RICHMOND

M. LAMAH HENRY KPAKILÉ

*Encadrants :*

Dr. Nguyen Thi OANH

# Table des matières

<b>1</b>	<b>Introduction</b>	<b>2</b>
1.1	Spécification . . . . .	2
<b>2</b>	<b>PARTIE 1 : DETECTION DU MOUVEMENT</b>	<b>2</b>
2.1	Mode d'exécution . . . . .	2
2.2	Construire de l'image d'arrière-plan . . . . .	3
2.3	Détection des mouvements . . . . .	6
2.4	Amélioration du résultat . . . . .	8
<b>3</b>	<b>PARTIE 2 : SUIVI DU MOUVEMENT</b>	<b>9</b>
3.1	Mode d'execution . . . . .	9
3.2	Filtre Kalman . . . . .	10
3.3	Résultat . . . . .	10
3.3.1	Effec du seuil correspondance . . . . .	10
3.3.2	Suivi un objet . . . . .	11
3.3.3	Suivi plusieurs objets . . . . .	13
3.3.4	Les cas difficile de suivi . . . . .	14
<b>4</b>	<b>CONCLUSION</b>	<b>16</b>

# 1 Introduction

Dans ce TP, nous allons créer un programme qui permet de détecter les mouvements dans une vidéo, ensuite de suivre les mouvements des objets en utilisant le filtre de Kalman. Comme paramètres d'entrée du programme on aura le nom d'une vidéo ainsi que 4 autres paramètres qui seront expliqués plus bas. À la sortie du programme on aura :

- Une image de fond de la vidéo ;
- Les images binaires des mouvements détectés ;
- La trajectoire du mouvement des objets.

Ainsi, dans ce présent rapport on va présenter dans une première partie, la détection de mouvements, la méthode et les paramètres utilisés, les résultats et explications de nos expériences. Dans une deuxième partie, nous présenterons la partie qui concerne le suivi des mouvements.

## 1.1 Spécification

Notre programme permet aux utilisateurs d'entrer les paramètres. Pour lancer notre programme, il faut accéder au répertoire «TP2 VO» et taper la commande « ./TP2 ». Notre système se compose de deux fonctions principales : détection du mouvement et suivi du mouvement.

## 2 PARTIE 1 : DETECTION DU MOUVEMENT

Dans cette première partie du TP, notre travail consiste à détecter le mouvement dans les vidéos. Pour le faire, on construit d'abord une image de fond à partir de la vidéo. Ensuite, on soustrait les images de la vidéo avec l'image de fond pour obtenir les zones en mouvement.

Dans les vidéos, on constate que seuls les objets qui se déplacent ou du moins sont en mouvement. Nous expliquons notre démarche de conception plus bas.

### 2.1 Mode d'exécution

Pour la toute première option de notre application, l'utilisateur doit entrer les paramètres suivants : **le nom de l'image, le nombre de séquences, le seuil de détection.**

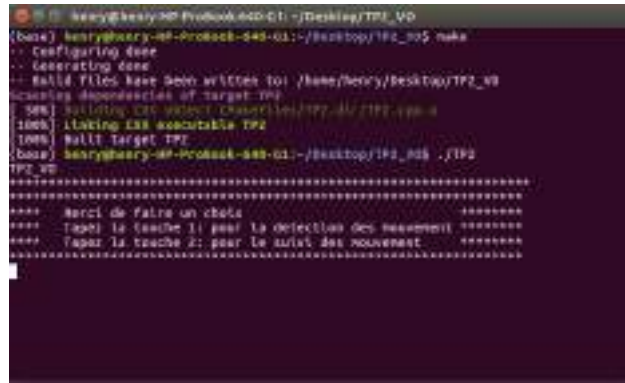


FIGURE 1 – Interface de l'application

FIGURE 2 – Fonction détection du mouvement

## 2.2 Construire de l'image d'arrière-plan

La première phase est cette phase où l'on cherche à obtenir l'image d'arrière-plan.

Pour chacune des vidéos on passe en paramètre, on procède comme ci-dessus. Alors l'idée c'est quoi ? L'idée ici dans cette première phase est d'obtenir l'image de fond de la vidéo afin de déterminer quels pixels ont bougé c'est-à-dire quels pixels appartiennent à des objets mobiles.

Notre algorithme s'inspire de la méthode de la valeur médiane. Dans le but de réduire le bruit, nous appliquons un lissage à l'image du fond et à l'image courante de la vidéo en utilisant la fonction « GaussianBlur » d'OpenCV avec la matrice kernel de la taille 5x5. Nous exécutons notre programme avec plusieurs valeurs différentes de nombre de séquences utilisées. Pour plus d'illustration vous trouverez les résultats de la construction de l'arrière plan de la vidéo **Walk1.mpg**.





1. Image dans la vidéo avec 50 séquences.
2. Image dans la vidéo avec 250 séquences.
3. Image d'arrière plan de 50 séquences.
4. Image d'arrière plan de 250 séquences

FIGURE 3 – Image d'arrière plan – vidéo Walk1.mpg

Il est visible qu'il n'y a une très grande différence entre l'arrière plan construit avec le nombre de séquence traité. Ce qui peut s'expliquer par le fait que durant les toutes premières secondes de cette vidéo, plusieurs images de l'arrière-plan vides se succèdent, donc lors de l'extraction ce sont ces valeurs des pixels de ces images qui sont choisies.

Cependant, dans la mesure où l'objet apparaît très long au début de la vidéo, avec le nombre de séquences extraites nous pouvons considérer les différences des résultats. Voir ci-dessous les résultats





1. Image dans la vidéo avec 50 séquences extraites.
2. Image dans la vidéo avec 250 séquence extraites.
3. Image d'arrière plan de 50 séquences.
4. Image d'arrière plan de 250 séquences.

FIGURE 4 – Image d'arrière plan – vidéo EnterExitCrossingPaths2cor.mpg

Quand nous choisissons le nombre de séquences extraite plus petit (le nombre de séquence est inférieur à 100), nous remarquons que Cette méthode ne fonctionne pas correctement car il considère l'objet comme le fond. Cela entraine des fautes comme dans l'image suivante. En choisissant le nombre de séquences extraite plus grande (Par exemple 250 comme nombre de sequence), nous pouvons remarquer une amélioration dans l'extraction des arrières plan. Avec le nombre de séquences extraites = 50, nous pouvons remarquer deux personnes dans l'arrière plan, ils sont considérées comme le fond. C'est la raison pour la quelle, dans tous les frames vidéo détectés, il existe des boites englobantes. Tandis que, avec le nombre de séquences extraites = 250, le fond peut contenir plus deux personnes.

Pour cette situation, il est intéressant de remarquer que le nombre de séquences extraites influence sur la qualité de résultat. Il n'est pas nécessaire de traiter toutes les séquences de la vidéo pour en extraire l'arrière-plan. Si nous ne prenons que N premières images dans la vidéo, le temps de calcul est plus rapide que celui de toutes images. Nous allons donc choisir le nombre de séquence extraites = 250 dans tous les cas suivants pour avoir un bon résultat de détection du mouvement.

## 2.3 Détection des mouvements

Dans la deuxième phase, nous utilisons l'image de fond obtenue plus haut afin de procéder normalement à la détection de mouvement. Cette image comme nous avons vue résulte de la soustraction de l'image courante et de l'image de fond dans la vidéo.

Après avoir obtenu le fond, nous avons fait la soustraction entre les images de la vidéo avec l'image du fond pour détecter le mouvement.



1. Image d'arrière plan.
2. Image de la vidéo.
3. Image de soustraction.

FIGURE 5 – Image de soustraction

Nous avons fait le seuillage de l'image soustraction. L'idée principale de cette étape est de convertir tous les pixels qui ont des valeurs inférieures au seuil en noir, à l'inverse, tous les pixels ayant des valeurs supérieures au seuil seront en blanc.



1. Image subtraction.
2. Image de seuillage.

FIGURE 6 – Image de seuillage

Comme différents resultats du seuil de détection :



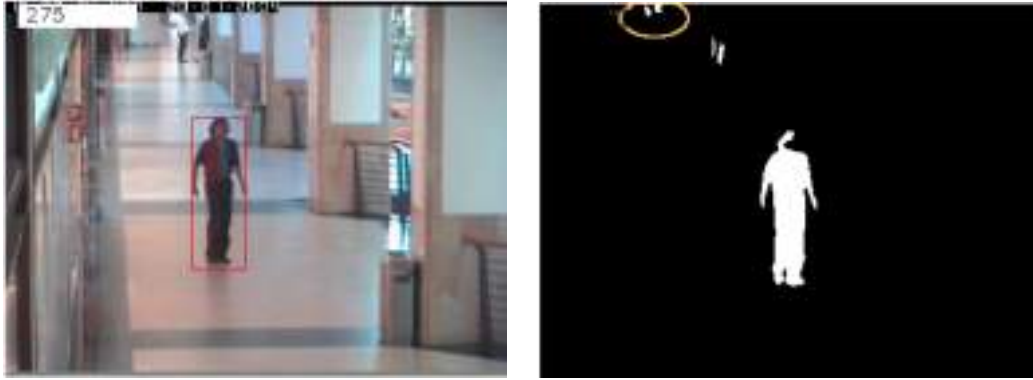
1. Image seuillage (seuil = 20).
2. Image seuillage (seuil = 50).
3. Image seuillage (seuil = 100).

FIGURE 7 – Image de seuillage avec différente valeurs de seuil

Avec un seuil près de 0 (seuil = 20), nous pouvons observer que plusieurs parties du fond sont considérées comme objet (l'ombre de l'objet, les bruits et les jambes de l'homme sont plus grandes que les résultats avec le seuil plus grand). Alors que, avec le seuil plus grand tel (seuil = 100), plusieurs parties de l'objet sont considérées comme du fond (la tête et une partie du corps de l'homme), les jambes de l'homme deviennent plus petites ainsi, les bruits semblent être perdus. Après essai de plusieurs valeurs du seuil, nous remarquons que le seuil de 50 nous donne le meilleur résultat.

À base de ce résultat, nous pouvons noter que la détection réalisée n'est pas quasi parfaite mais le résultat est satisfaisant. Une boîte englobante encadre bien le personnage en mouvement dans la vidéo originale. Une autre boîte englobante également encadre la jambe de la personne dans la vidéo. Mais la détection n'est pas toujours aussi nette comme illustrée par la figure suivante. Il y a des bruits dans l'image de seuillage (le cercle jaune).





1. Image substraction.
2. Image de seuillage.

FIGURE 8 – Image de seuillage

L'autre cas qui donne un bon résultat de la détection. Dans l'image originale de la vidéo, il y a deux boites englobantes qui encadrent deux personnes qui se déplacent. La détection est nette. Sauf le bras de l'homme à gauche qui est considéré comme le fond.

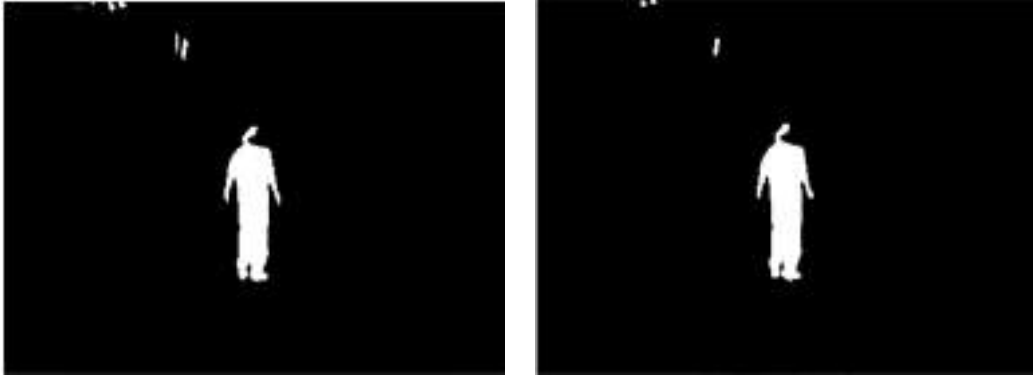


1. Image de la vidéo.
2. Image de soustraction.
3. Image de seuillage.

FIGURE 9 – Le meilleur résultat

## 2.4 Amélioration du résultat

Dans l'image de seuillage nous pouvons remarquer qu'il y a des trous . Alors les résultats doit etre ameliorer en utilisant la techinque fermeture : érosion + dilatation. En utilisant une matrice kernel de la taille 3x3. Le résultat de cette étape est affiché dans la figure suivante :



1. Image de seuillage.
2. Image améliorée.

FIGURE 10 – Image améliorée

Les objets deviennent plus complets, mais il existe des régions petites. La région de bruit devient aussi plus petite.

## 3 PARTIE 2 : SUIVI DU MOUVEMENT

Plusieurs paramètres peuvent influencer la précision d'une bonne détection lors de la détection d'un objet. Par exemple, une variation brusque de la lumière dans un programme de détection peut être interprétée comme un sujet en mouvement. Ainsi, on doit corriger ces erreurs survenues en détection pour ne pas être suivies comme des sujets.

### 3.1 Mode d'exécution

Pour la deuxième option de notre application, l'utilisateur doit entrer les paramètres suivants : **le nom de l'image, le nombre de séquences, le seuil de détection, le seuil de correspondance.**

```

henry@henry-MD-ProBook-640-G1: ~/Desktop/TP2_V0
(base) henry@henry-MD-ProBook-640-G1:~/Desktop/TP2_V0$ make
... Configuring done
... Generating done
... Build files have been written to: /home/henry/Desktop/TP2_V0
... Calculating dependencies of target TP2
... 50%: Building CXX object CMakeFiles/TP2.dir/TP2.cpp.o
100%: Linking CXX executable TP2
100%: Build Target TP2
(base) henry@henry-MD-ProBook-640-G1:~/Desktop/TP2_V0$ ./TP2
TP2_V0
*****
****  Merci de faire un choix  ****
****  Tapez la touche 1: pour la detection des mouvement *****
****  Tapez la touche 2: pour le suivi des mouvement *****
*****

```

FIGURE 11 – Fonction suivi du mouvement

## 3.2 Filtre Kalman

Pour l'implémentation de cette fonction, nous avons utilisé le filtre Kalman d'OpenCV. Pour suivre un objet, nous suivons le centre de gravité de la boîte englobante qui entoure l'objet.

Afin d'identifier dans l'image courante, un objet identifié dans l'image précédente, nous calculons pour chaque objet précédemment identifié, la distance vers les objets de l'image courante et nous choisissons l'objet de l'image courante le plus proche.

Afin de présenter les résultats de suivi des objets mouvements, nous avons créer un nouvelle image de meme taille de la vidéo. Ensuite, nous avons mis la couleur pour les points.

- Les point prédites en bleu ;
- Les point mesurées en vert ;
- Les point corrigées en rouge.

## 3.3 Résultat

### 3.3.1 Effec du seuil correspondance

Nous allons ensuit discuter sur l'influence du seuil correspondance sur les résultats finaux.



1. Image suivi – seuil correspondance = 30.
2. Image suivi – seuil correspondance = 70.
3. Image suivi – seuil correspondance = 100.

FIGURE 12 – Les résultats avec les seuil correspondance différents

Dans le but de bien étudier l'effet du seuil de correspondance sur les résultats finaux, nous avons exécuté notre programme plusieurs fois avec les différentes valeurs du seuil de correspondance. Le seuil plus petit, moins des objets petits sont suivis et aussi moins de bruits. Tandis que, avec le seuil plus grand, plus des objets petits sont suivis et aussi plus de bruits. Il n'existe pas le seuil de correspondance qui donne une bonne performance pour tous les cas. Mais, dans notre cas, le seuil de correspondance = 30 donne les résultats raisonnables pour tous les vidéos dans la base de vidéo. A présent nous présentons les résultats de nos expérimentations pour la suivi de mouvement avec ce seuil indiqué ci dessus.

### 3.3.2 Suivi un objet

Dans cette partie, nous présentons la suivi du mouvement d'un objet :

1. un seul objet en mouvement avec une trajectoire simple ;
2. un seul objet en mouvement avec trajectoire complexe.

#### (1) Un seul objet en mouvement avec une trajectoire simple



1. Image de la vidéo.
2. Image de suivi.

FIGURE 13 – Suivi de la trajectoire de l'homme dans la vidéo Walk1.mpg - objet 0

Nous pouvons remarquer sur les figures ci-dessus qui montrent des séquences d'images dont un personne est détecté et suivie dans la vidéo. Le chiffre 0 montre qu'il est l'objet numéro 0 détecté et suivi dans la présente séquence.

Dans ce cas, notre système fonctionne bien. Tous les trois trajectoire prédite, mesurée et corrigée sont bien identifié. La différence entre trois trajectoires est faible.

## (2) Un seul objet en mouvement avec trajectoire complexe



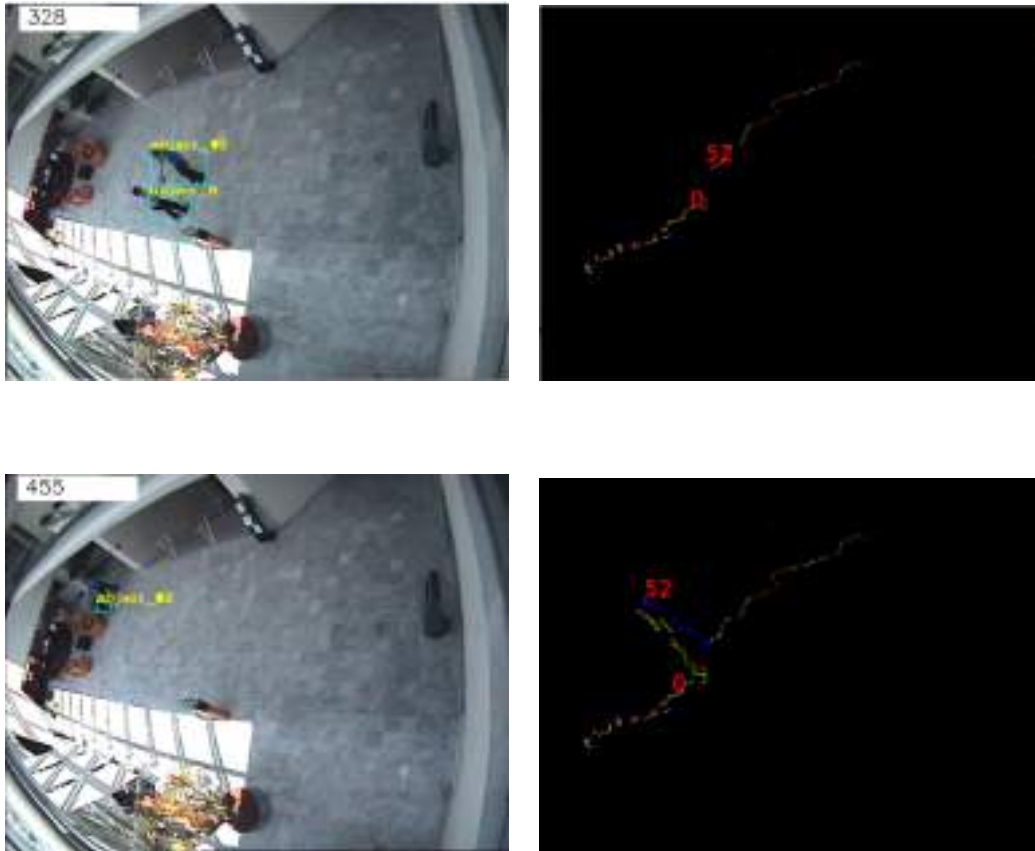
1. Image de la vidéo.
2. Image de suivi.

FIGURE 14 – Suivi de la trajectoire de la femme dans la vidéo Browse3.mpg - objet 1

Dans le second cas, nous montrons le résultat avec une autre vidéo. Dans cette vidéo, il y a une femme, au centre de la fenetre, puis elle reste pour lire un journal, enfin, elle prend d'aute trajet.

### 3.3.3 Suivi plusieurs objets

Dans cette partie, nous présentons la suivi du mouvement de plusieurs objets



1. Image de la vidéo - frame 328.
2. Image de suivi du frame 328.
3. Image de la vidéo - frame 455.
4. Image de suivi du frame 455.

FIGURE 15 – Suivi de la trajectoire de deux hommes dans la vidéo Meet WalkTogether1.mpg

Dans le cas où plusieurs objets en mouvement dans la vidéo, notre système doit créer des filtres Kalman en parallèle pour suivre ces objets en meme moment. Nous représentons le cas de deux objets en mouvement sans croisement de trajectoires dans la figure 14. La figure 14a est le frame de la vidéo où deux hommes viennent de deux cotés de la fenetre sans croisement. Dans ce cas-là, notre système fonctionne bien (voyez le figure 14b). Cependant, quand ils se déplacent ensemble, la distance entre ces deux devient plus petite, le système a donc détecté ces deux comme un personne sur une boite englobante (voyez la figure 14c).

Alors, le mouvement de deux hommes est suivi comme le mouvement d'un objet (voyez la figure 14d). Nous allons ensuite montrer le cas de où plusieurs objets en mouvement croisement. Voyez la figure 15. Dans ce cas-là, quand un homme est entré dans la porte, et l'autre homme à continuer à se déplacer vers la porte (frame 204). Ces deux homme sont détectés comme deux objets différents jusqu'à ce que deux mouvements se croisent. À cet instant, deux homme sont indentifié comme un seul (frame 209). Nous pouvons clairement remarquer que dans le cas du frame 204, notre programme identifie bien, alors que dans le cas du frame 209, la détection est moins bonne. Le cas du frame 209 est similaire à celui de la figure 14c puisqu'il y a une occlusion. Il est intéressant de remarquer que le suivi de la trajectoire des personnes n'est pas aussi bon que dans le cas précédent. Cela est dû a la difficulté à identifier les objets d'un image à une autre car la qualité de la détection n'est pas constante et il y a aussi des problèmes d'occlusion. Voyez le cercle jaune dans l'image de suivi. La partie à gauche de ce cercle montre le mouvement de l'homme entrant la porte. La partie à droite de ce cercle représente le mouvement de l'homme restant. Quand ces deux hommes se croisent dans le frame 209 de la vidéo, leurs suivis sont considérés comme le suivi d'un objet.

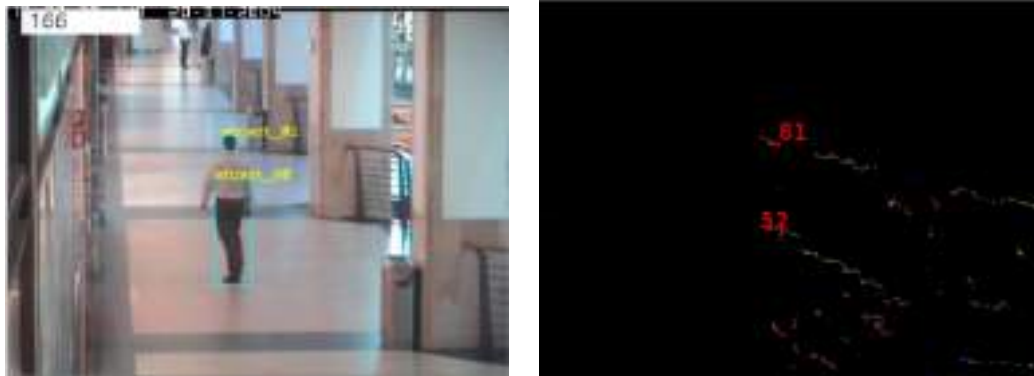


1. Image de la vidéo – frame 204.
2. Image de la vidéo – frame 209.
3. Image de suivi.

FIGURE 16 – Suivi de la trajectoire de deux homme dans la vidéo EnterExitCrossing-Paths2cor.mpg

### 3.3.4 Les cas difficile de suivi

**Le cas où les vetements des personne sont même couleur du fond**



1. Image de la vidéo du frame 166.
2. Image de suivi du frame 166.

FIGURE 17 – Le cas de suivi de la trajectoire de l’homme où les vêtements sont même couleur du fond dans la vidéo *EnterExitCrossingPaths2cor.mpg*

Dans ce cas là, on peut observer que l’homme portant une chemise qui est même couleur que celle du fond. Alors, le système détecte cet homme comme deux objets. Si vous regardez bien, la tête de l’homme est détecté comme un objets et une partie du crop de l’homme est considéré comme un autre objet.

#### **Le cas où l’objet apparaisse loin de la caméra (la taille de l’objet petite)**



1. Image de la vidéo du frame 310.
2. Image de la vidéo du frame 331.
3. Image de suivi.

FIGURE 18 – Le cas de suivi de la trajectoire de l’homme où la taille de l’objet petite dans la vidéo *FightRunAway1.mpg*

Dans ce cas là, l’homme qui porte le T-shirt rouge apparait loin de la caméra. Donc, la taille de cet homme est petit, cet homme n’est pas indentifié. Alors, son mouvement n’est pas suivi.



Il faut encore remarquer que à côté des cas où le mouvement des objets dans la vidéo est bien détecté et suivi. Il existe des cas dont le système fonctionne moins bien pour des raisons telles que des objets sont de meme couleur du fond, ou l'objet apparait loin de la caméra.

## 4 CONCLUSION

Ce rapport résume une implémentation d'un programme de détection et de suivi de mouvement réaliser en C++ et utilisant certaines fonctions existantes de OpenCV. La détection comme le suivi de mouvements dépendent de nombreux paramètres et de techniques utilisées. Un des facteurs important est la construction de l'arrière-plan, et comme on a vue en cours, on a donc utilisé la valeur médiane des pixels des images pour construire l'arrière-plan.

Cependant, elle est sensible au condition d'éclairage et le constraste entre l'objet et le fond. À base de ces résultats, on peut remarquer que : il n'est pas nécessaire d'extraire toutes les images de la vidéo pour construire le fond, mais le nombre de séquences choisit influence fortement sur les résultats ; le résultat de ce système pourrait être améliorer en appliquant les techniques tel que Gaussian Blur, fermeture.

Pour le suivi du mouvement, le Filtre Kalman est utilisé. Nous notons que la détection des objets est relativement efficace mais est perturbée par la proximité des couleurs des objets avec l'arrière plan ou par les occlusions. Le suivi quant à lui n'est pas vraiment correct que pour des objets uniques. Pour plusieurs objets les résultats sont très mitigés et cela est principalement du à l'identification des objets d'une image à l'autre et aux occlusions.