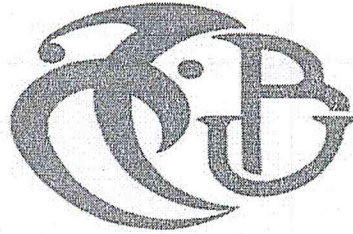


République Algérienne Démocratique et Populaire  
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique

Université Saad Dahlab Blida



Faculté des sciences

**Département d'informatique**

Mémoire Présenté par :

Kraled Kebir Nesrine & Chennouf Bahaeddine

**En vue d'obtenir le diplôme de master**

Domaine : Mathématique et informatique

**Filière :** Informatique

**Spécialité :** Informatique

**Option :** Ingénierie de logiciel

***Thème***

***Vidéosurveillance intelligente : Détection et suivi des personnes***

MA-004-341-1

Promoteur Mr. KAMECHE ABDALAH

Encadreur Mr. DJEKOUN OUALID

Promotion  
2015 / 2016

## RESUME

A l'heure actuelle, les systèmes de vision sont de plus en plus répandus et les caméras (webcam, caméra IP, etc..) se sont installées partout dans notre quotidien. Elles sont utilisées pour réaliser de la vidéosurveillance (dans les magasins, rues ou aéroports), de l'aide à la conduite (aide au guidage ou détection d'obstacles), et bien d'autres applications encore.

Dans le domaine de la vidéosurveillance, les agents de sécurité, devant un nombre important de caméras, ne peuvent pas toutes les surveiller en même temps. Ils sont contraints de ne regarder que ponctuellement chacune d'entre elles en permutant régulièrement. De plus, la répétitivité de la tâche, la faible fréquence des situations dangereuses ou anormales, rendent le travail de surveillance particulièrement difficile. Pour ces raisons, ce travail se doit d'être assisté, voire automatisé via des systèmes de vidéosurveillance ou de vidéo-assistance intelligents.

Conscient des travaux menés dans ce domaine, ce qui nous intéresse dans ce projet est d'améliorer, dans un premier temps, une plateforme distribuée permettant la visualisation et la gestion des informations issues d'un réseau de caméras IP, en temps réel. Ensuite, développer un module qui permet de détecter et suivre les personnes en temps réel, et l'intégrer dans la plateforme.

**Mots clés :** vidéosurveillance, vidéosurveillance intelligente, analyse vidéo, détection, suivi.

## **Abstract**

Nowadays, vision systems are becoming more prevalent and cameras (webcams, IP cameras, etc..) are installed is everywhere in our daily lives. They are used to perform video surveillance (in shops, streets or airports), the driving assistance (guidance assistance or obstacle detection), and many other applications.

In the field of video surveillance, the security agents, in front of a significant number of cameras, cannot monitor them all at the same time. They are constrained to look at each one only punctually while switching regularly between them. Moreover, the repetitiveness of such task, the weak frequency of the dangerous or abnormal situations, makes the monitoring particularly difficult. For these reasons, this work must be assisted, even automated via systems of video surveillance or intelligent video-assistance.

Conscious of the work carried out in this field, in this project we focus initially on improving a distributed platform allowing, in real time, visualization and information management received from an IP cameras network in real time. Then, develop a module that can detect and track people in real time, and integrate it into the platform.

**Keywords:** video surveillance, intelligente video surveillance, video processing, detection, tracking

## ملخص

في الوقت الحاضر، نظم الرؤية أصبحت أكثر انتشاراً، ويتم تثبيت الكاميرات (كاميرات IP، الخ..) في كل مكان في حياتنا اليومية. يتم استخدامها لإجراء المراقبة بالفيديو (في المحلات التجارية والشوارع أو المطارات)، ومساعدة القيادة (المساعدة التوجيه أو الكشف عن عقبة)، والعديد من التطبيقات الأخرى.

في مجال المراقبة بالفيديو، حراس الأمن، و أمام عدد كبير من الكاميرات لا يستطيعون مراقبتها كلها في نفس الوقت، إنهم مظطرون إلى مشاهدة كل كاميرا على حدى، علاوة على ذلك فإن التكرار المهمة، والتردد المنخفض من الحالات الخطرة أو غير طبيعية، جعل أعمال المراقبة صعبة للغاية. لهذه الأسباب، يجب هذا العمل يجب أن يكون مدعوماً بنظام يحول هذه العملية إلى عملية آلية، عن طريق أنظمة المراقبة أو أنظمة المراقبة الذكية

نحن مهتمون في هذا المشروع، أولاً، بتحسين منصة مراقبة وإدارة شبكة كاميرات IP. ثم وضع نموذج التي يمكن الكشف عن وتعقب الناس في الوقت الحقيقي، وإدماجه في النظام الأساسي.

**الكلمات المفتاحية:** المراقبة بالفيديو، المراقبة الذكية بالفيديو، تحليل الفيديو، كشف، تتبع

## Sommaire

<b>Introduction générale</b> .....	13
Présentation du projet .....	13
Problématique .....	13
Objectifs .....	14
Plan du mémoire .....	15
<b>Chapitre 1 : Etat de l'art de la vidéosurveillance</b> .....	16
1- Introduction .....	17
2- Les systèmes de vidéosurveillance.....	18
2.1- Définition .....	18
2.2- Types de la vidéosurveillance .....	19
2.3- Composition d'un système de vidéosurveillance .....	19
2.4- fonctions d'un système de vidéosurveillance .....	26
2.5- Objectif d'un système de vidéosurveillance .....	27
3- la vidéosurveillance IP .....	27
3.1- les composants d'un système de vidéosurveillance IP .....	28
3.2- les avantages .....	29
3.3- les inconvénients de la vidéosurveillance IP .....	31
3.4- domaine d'application .....	31
4- la vidéosurveillance intelligente .....	32
4.1- raisons pour l'utilisation de la vidéosurveillance intelligente .....	32
4.2- secteur d'application .....	33
4.3- avantages .....	35
5- conclusion .....	36
<b>Chapitre 2 : Analyse vidéo</b> .....	37

1- Introduction .....	38
Partie I : Détection de mouvement .....	39
1- Le Mouvement .....	39
1.1- Les types d'objets en mouvement .....	39
2- La vidéo .....	40
2.1- séquence d'image .....	40
2.2- image numérique .....	41
2.3- le bruit .....	41
2.4- résolution temporelle d'une vidéo .....	43
2.5- objet vidéo .....	43
3- la Détection de mouvement .....	43
3.1- les méthodes de détection de mouvement .....	44
Partie II : suivi de mouvement .....	50
1- Hiérarchique du traitement d'une séquence vidéo .....	50
2- Définition .....	51
3- Méthodes de suivi d'objet en mouvement dans une vidéo .....	51
3.1- suivi d'un point .....	52
3.2- suivi de noyau .....	54
3.3- suivi de la silhouette .....	55
4- Structure générale des algorithmes de suivi .....	56
4.1- phase de prédiction .....	56
4.2- phase d'ajustement .....	57
4.3- phase d'estimation .....	58
5- le suivi des personnes .....	59
5.1- la méthode utilisée .....	59
6- conclusion .....	61
<b>Chapitre 3 : Conception .....</b>	<b>62</b>
1- Introduction .....	63
2- Architecture de la plateforme .....	63
3- Analyse et spécification des besoins .....	64

3.1- Identification des acteurs .....	64
3.2- Etablir les fonctionnalités .....	66
3.3- diagramme de cas d'utilisation.....	67
3.4- diagramme de classe .....	73
3.5- diagrammes de séquence .....	74
4- fonctionnement de la plateforme.....	79
4.1- taches effectuées par le superviseur .....	80
4.2- taches effectuées par le surveillant .....	82
4.3- taches effectuées par l'esclave .....	82
5- conclusion .....	83
<b>Chapitre 4 : Implémentation et réalisation .....</b>	<b>84</b>
1- Introduction .....	85
2- Les outils de développement utilisé.....	85
3- Prototype matériel .....	86
4- Présentation de l'application .....	87
5- La Détection de suivi des piétons .....	94
6- Test et validation .....	95
7- Conclusion .....	98
Conclusion générale et perspectives .....	99

# Introduction générale

---

## Présentation du projet

Notre projet rentre dans le cadre du protocole de recherche proposé par l'équipe VAANIM (Vision Artificielle et Analyse d'Image) de la division Robotique du CDTA (Centre de Développement des Technologies Avancées) et dont l'intitulé est "**Système de vision pour la robotique de surveillance**".

La plateforme développée doit gérer la visualisation, en temps réel, des flux vidéo des caméras IP constituant le réseau, et ceci à la demande de l'utilisateur.

- La sauvegarde avec un minimum d'espace disque.
- Faciliter la gestion des vidéos sauvegardées pour une éventuelle enquête à postériori.
- Assurer la sécurisation à l'accès au système.
- Permettre l'intégration d'algorithmes de détection et le suivi des personnes.
- L'accès distant via le Wifi

## 1- Problématique

Dans le domaine de la vidéosurveillance, les agents de sécurité, devant un nombre important de caméras, ne peuvent pas toutes surveiller en même temps. Ils sont contraints de ne regarder que ponctuellement chacune d'entre elles en permutant régulièrement. De plus, la répétitivité de la tâche, la faible fréquence des situations dangereuses ou anormales, rendent le travail de surveillance particulièrement difficile. Pour ces raisons, ce travail se doit être assisté, voire automatisé via des systèmes de vidéosurveillance ou vidéosurveillance intelligent.

La détection et suivi d'objets temps réel est une problématique difficile qui se pose dans un grand nombre d'applications de traitement d'images comme l'interaction homme-machine,



# Introduction générale

---

la surveillance civile et militaire, la réalité virtuelle, l'analyse du mouvement humain ou encore la compression d'images. Cette difficulté est accentuée dans les environnements sans contraintes où le système de suivi devra s'adapter à la variabilité importante des objets, aux variations de luminosité, aux occlusions (partielles ou totales) ainsi qu'aux problèmes de détection de mouvement, l'analyse de vidéos se compose de 3 étapes : détection d'objets, suivi de ces objets, et analyse du mouvement de ces objets pour en déduire un comportement.

On peut citer quelques problèmes que nous allons résoudre :

- L'algorithme de détection qui existe ne donne pas des bons résultats ;
- Améliorer l'algorithme de détection pour qu'il détecte que les personnes en mouvement ;
- Problème de stockage des vidéos enregistrées ;
  - La distribution des groupes des caméras sur les surveillants ;
  - Résoudre le problème de suivi des personnes.

## 2- Objectifs :

Notre travail consiste à une contribution à l'amélioration de la plateforme de vidéosurveillance déjà développée qui présente un certain nombre de problèmes dans son exécution. Notre contribution se situe en rajoutant des fonctionnalités, et des méthodes fiables de détection et de suivi des objets ou des personnes en mouvement à partir d'un flux vidéo capturé à travers le réseau de caméras IP en cherchant à déterminer le plus efficacement et précisément possible sa position dans chaque trame et connaître sa trajectoire, La méthode proposée permet ainsi de suivre un objet tout en gérant les occultations partielles et les déformations locales des objets.

# Introduction générale

---

## 3- Plan du mémoire

**CHAPITRE I :** Le premier chapitre présente un état de l'art des systèmes de vidéosurveillance et les vidéosurveillances intelligentes de la littérature.

**CHAPITRE II :** Présente les traitements de base d'analyse vidéo utilisés dans un système de vidéosurveillance intelligente, ce chapitre est divisé en deux parties :

- **Partie 1 :** les différentes méthodes de détection d'objets et personnes en mouvement.
- **Partie 2 :** Les différents méthodes de suivi des objets en mouvement ainsi les personnes.

**CHAPITRE III :** L'étude conceptuelle et le fonctionnement de la plateforme.

**CHAPITRE IV :** Implémentation et la réalisation de l'application avec des tests réels.

**CHAPITRE I :**

**CHAPITRE I :**  
**ETAT DE L'ART DE LA**  
**VIDEOSURVEILLANCE**

ETAT DE L'ART DE LA

VIDEOSURVEILLANCE

ETAT DE L'ART DE LA

## 1- Introduction

Dans nos jours, avec l'augmentation continue de la population dans les villes, il devient difficile de surveiller le grand flux humain dans les différents endroits tel que : les aéroports, les banques et les postes polices, etc...) La plupart des responsables souhaitent accroître la sécurité en protégeant les biens et les personnes par la vidéosurveillance.

La surveillance, qui était autrefois accomplie uniquement par des agents de sécurité, a vu arriver la technologie comme une révolution, les alarmes, la vidéosurveillance permettent à ces agents d'être plus efficaces car ils visualisent directement de leur poste de sécurité tous les points clé des locaux à surveiller. Pourtant l'homme reste imparfait. En effet, qui pourrait rester devant des écrans de surveillance sans aucune faille ? C'est impossible. C'est là qu'apparaît la vidéosurveillance intelligente. Elle permet d'aider les agents dans leur travail. Le programme peut détecter une intrusion à tout moment et en avertir l'agent qui prendra les mesures nécessaires. La surveillance intelligente en est rendue bien plus efficace et moins contraignante pour les agents.

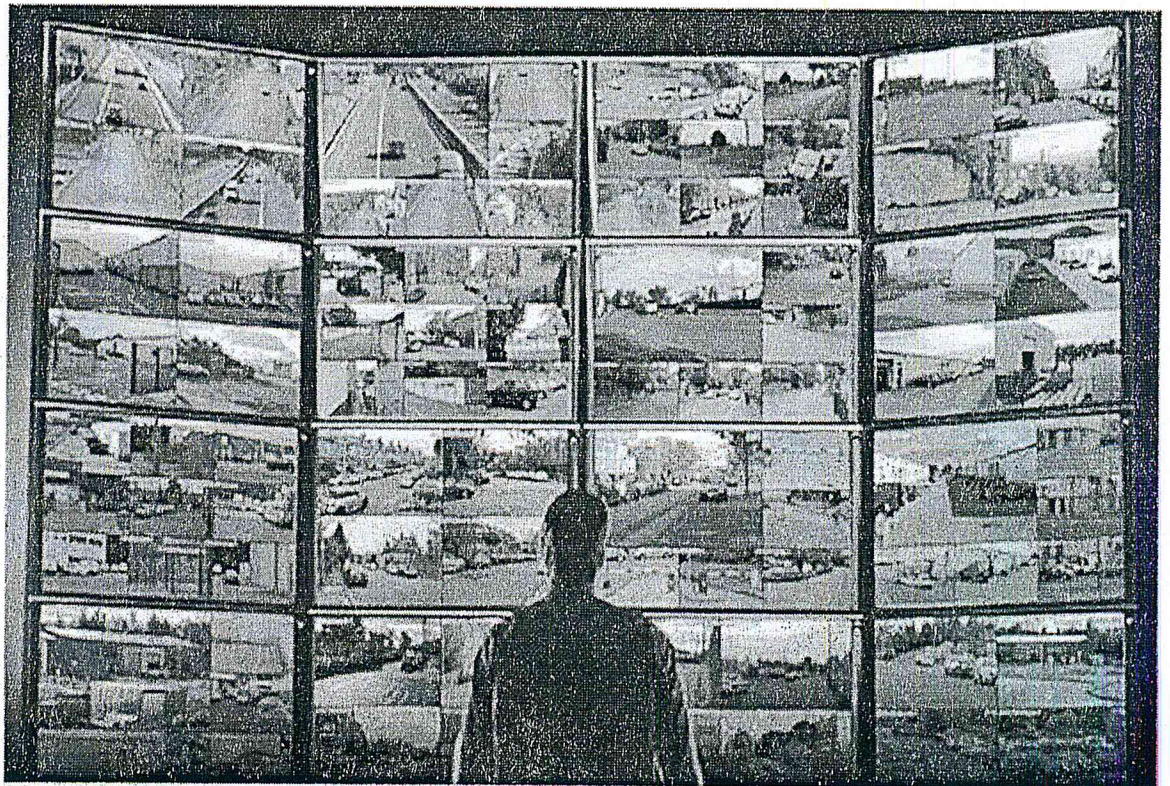
Aujourd'hui, un nombre impressionnant des vidéos sont enregistrées mais ne sont jamais ou peu visualisées ou examinées, et l'opération de traitement de tous ces enregistrements prennent un temps important. Par conséquent, des activités et des événements passent inaperçus et les comportements suspects ne sont pas détectés à temps pour prévenir les incidents. Ces inconvénients de la vidéosurveillance ont poussé au développement d'applications de vidéosurveillance intelligente, dans ce chapitre nous allons parler de la vidéosurveillance et la vidéosurveillance intelligente.

## 2-Les systèmes de vidéosurveillance

Les systèmes de vidéosurveillance sont en voie de devenir des équipements de sécurité courants au même titre que les systèmes d'alarme domestiques. Ils complètent efficacement ces derniers en offrant la possibilité de vérifier le bien fondé des alertes et peuvent également rendre de nombreux services, Il est en effet très pratique de pouvoir surveiller une ou plusieurs pièces (ou même une zone extérieure) à distance, que ce soit pour un usage privé ou professionnel [5].

### 2-1-Définition

La vidéosurveillance consiste à surveiller à distance des lieux publics ou privés, à l'aide de caméras, le plus souvent motorisées, qui transmettent les images saisies à un équipement de contrôle qui les enregistre ou les reproduit sur un écran. Elles captent les flux de personnes pour surveiller les allées et venues, prévenir les vols, agressions et fraudes, ainsi que pour gérer les incidents et mouvements de foule [1].



**Figure 1:** Centre de vidéosurveillance : difficulté de surveiller plusieurs écrans en même temps

## 2-2- Types de la vidéosurveillance

Selon son utilisation, on distingue trois catégories de vidéosurveillance [1] :

- **Active** : Surveillance d'une aire pour appuyer le travail sur place d'agents de sécurité ou lors d'intervention d'urgence.
- **Passive** : Un employé surveille un petit nombre d'écrans de télévision en s'adonnant à d'autres tâches.
- **Enregistrement** : Permet de collecter des renseignements, à des fins d'enquête et de preuve. Les enregistrements sont conservés pendant une durée déterminée, dépendant des besoins et des limites des espaces d'archivages [1].

### 2.3- Composition d'un système de vidéosurveillance

Un système de vidéosurveillance est composé de cinq types d'équipements :

#### 2.3.1- Equipement de réception

L'élément fondamental qui assure la réception dans un système de vidéosurveillance, c'est la caméra. En fonction de l'environnement, du besoin de l'utilisateur et du budget alloué à ce système, on doit choisir les caméras adéquates pour bien filmer les zones à surveiller (caméra en couleur ou noir et blanc, caméra fixe ou mobile...) [2].

##### 2.3.1.1- La caméra

La caméra est un système de prise d'images animées qui génère un signal vidéo noir et blanc ou couleur. La caméra capte la lumière pour la transformer en signal électrique. Elle se compose parfois d'un capteur d'image et d'une électronique de traitement permettant de générer le signal vidéo. [2]

##### 2.3.1.2- Type Des caméras de surveillance

Il existe souvent plusieurs types des caméras de surveillances, dans notre mémoire, on va s'intéresser sur les types suivants :

- **Fixe**

Pointée dans une direction unique, elle couvre une zone définie (une entrée, une portion de stationnement, etc.). C'est la caméra de surveillance traditionnelle. Elle constitue un excellent choix lorsqu'on désire que la présence de la caméra, ainsi que sa direction de surveillance, soient visibles. [1]



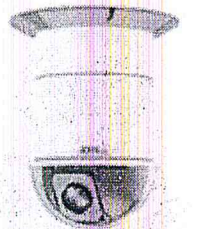
- **PTZ (Pan-Tilt-Zoom)**

Motorisée, elle peut être actionnée, manuellement ou automatiquement, dans des mouvements panoramique/inclinaison/zoom. Elle sert à suivre des objets ou des individus se déplaçant dans la scène ou à zoomer sur des régions d'intérêt (par exemple, sur une plaque d'immatriculation etc...). [1]



- **Dôme**

Recouverte d'un caisson hémisphérique, ce qui la rend discrète et, dans certains modèles, résistante au vandalisme et aux intempéries. Elle peut être fixe ou mobile. Les versions motorisées couvrent une zone très large, grâce à leur balayage horizontal de 360° et de 180° à la verticale. Bien qu'en « tour de garde », elle puisse remplacer dix caméras fixes en balayant la zone à surveiller, en observant qu'une seule direction à la fois. [1]



- **Mégapixel**

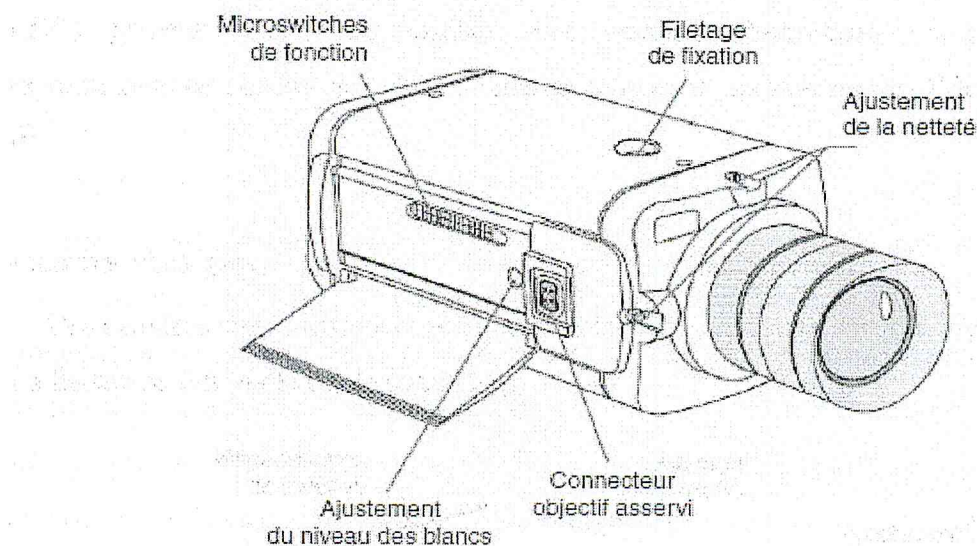
Offre une résolution plus élevée que les caméras standards, allant de 1 à 16 méga pixels. Elle permet soit de capter une image plus claire, soit de couvrir un plus large champ visuel, réduisant le nombre de caméras nécessaires pour couvrir une zone à surveiller. Lorsqu'elle est utilisée avec un grand angle d'ouverture, elle possède un espace de visualisation allant généralement de 140° à 360°, offrant la possibilité de zoomer de façon logicielle dans l'image. Elle peut ainsi devenir une alternative à la caméra PTZ mécanique qui entraîne l'usure des pièces. Sa résolution élevée contribue à l'amélioration de la performance des algorithmes de détection et de reconnaissance exigeant un haut niveau de détails, telles que la lecture de plaques d'immatriculation et la reconnaissance de visage etc... [1].

- **Infrarouge et thermique**

Sensible au rayonnement infrarouge (IR), elle est capable de produire une image de bonne qualité dans le noir pour une surveillance nocturne. De nuit, elle filme en noir et blanc, mais elle peut produire une image couleur le jour. Certaines caméras infrarouges sont équipées de leur propre source de lumière IR, allumée lorsque le niveau d'éclairage chute sous un certain seuil. Des projecteurs IR séparés (lampe ou LED) peuvent aussi être utilisés. Les caméras thermiques enregistrent le rayonnement de chaleur des objets. Elles ne requièrent aucune source d'illumination. [1]

- **Caméra analogique**

Ces caméras sont facilement reconnaissables ; elles ont une sortie de type BNC.0 La liaison se fait via le câble coaxial [2]

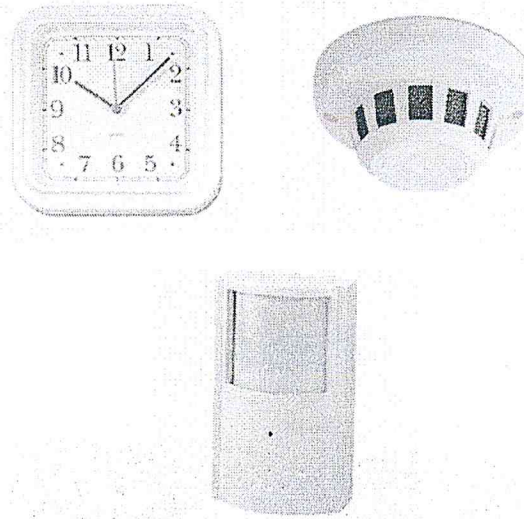


**Figure 2 : caméra analogique [2]**

- **Caméra discrète**

Les progrès des technologies font que de nos jours une caméra peut se loger dans n'importe quel accessoire, à savoir : le détecteur d'intrusion, l'horloge et la tête de détection incendie [2].





**Figure 3** : exemples des caméras discrètes [2]

- **Panoramique**

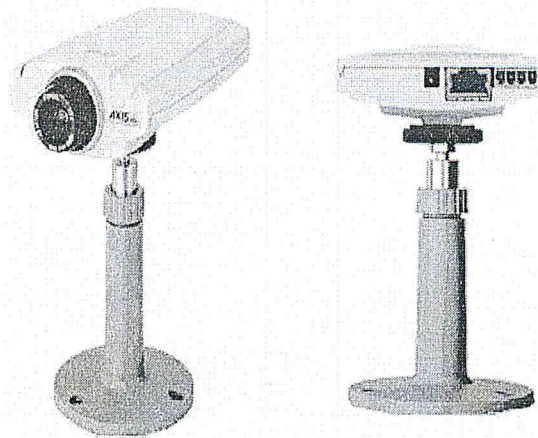
Grâce à une optique spéciale, elle offre 360° de visibilité avec une seule caméra. Elle permet un PTZ virtuel dans l'image. Les principales technologies panoramiques pour la surveillance sont le fish-eye, la lentille à miroirs et la lentille panomorphe. Toutefois, la résolution de ces caméras est souvent insuffisante pour des analyses nécessitant un niveau de détail élevé [1].

- **Caméra CCD (*Charge Coupled Device*)**

Le capteur CCD ne fait pas de distinction entre les couleurs. L'analyse des couleurs se fait *via* des filtres qui permettent de récupérer les signaux RVB avant traitement.

- **Caméra IP**

Une caméra numérique contrairement à sa consœur analogique, ne dispose pas de sortie coaxiale. En revanche elle dispose d'une liaison via une prise RJ 45 qui permet le raccordement à un réseau informatique [2].



**Figure 4:** caméra IP [2]

### **2.3.2- Equipement de transmission**

La vidéo captée par les caméras de surveillance doit être transmise aux systèmes d'enregistrement, de traitement et de visionnement. Cette transmission peut se faire par câble (câbles coaxiaux ou à fibre optique, fils de cuivre torsadés) ou à travers l'air (signaux infrarouges, transmission radioélectrique).

La vidéo filaire prédomine largement dans les systèmes de vidéosurveillance. Elle offre une plus grande bande passante et une meilleure fiabilité que les connections sans fil, à un coût inférieur. Cependant, la transmission vidéo sans fil s'impose parfois comme solution, par exemple dans le cas de surveillance de grands périmètres où l'installation de câblage s'avérerait trop coûteuse, ou lorsque les zones à surveiller sont impossibles à rejoindre par câble.

Qu'il transite par fil ou sans fil, le signal vidéo peut être analogique ou numérique. Encore aujourd'hui, la majeure partie des transmissions vidéo pour la surveillance sont analogiques. Néanmoins, les réseaux informatiques (LAN, WAN ou Internet) sont de plus en plus utilisés pour transporter la vidéo grâce au protocole TCP/IP. Les caméras IP peuvent se connecter directement sur ces réseaux, tandis que les flux vidéo émergeant de caméras analogiques doivent, au préalable, être numérisés par un encodeur, aussi appelé serveur vidéo, pour passer par les réseaux IP [1].

### 2.3.3- Equipement de compression

La vidéo numérisée représente une grande quantité de données à transmettre et à archiver. L'envoi d'une séquence vidéo peut nécessiter jusqu'à 165 mégabits de bande passante et la vidéo d'une seule caméra pour une journée peut occuper sept giga-octets d'espace disque. C'est pourquoi la vidéo de surveillance doit être compressée grâce à des codecs, algorithmes permettant de réduire la quantité de données en supprimant les redondances, par image ou entre les trames d'une séquence, ainsi que les détails imperceptibles à l'œil humain [1].

### 2.3.4- Equipement de Traitement

Les systèmes de gestion vidéo opèrent les traitements des images de vidéosurveillance, tels que la gestion des différents flux vidéo, le visionnement, l'enregistrement, l'analyse et la recherche dans les séquences enregistrées. Il existe quatre grandes catégories de systèmes de gestion vidéo. [1]

- **Enregistreur vidéo numérique (DVR) :**

Appareil qui dispose d'un disque dur interne pour l'enregistrement numérique de la vidéo et d'un logiciel intégré de traitement de la vidéo. Il n'accepte que les flux provenant de caméras analogiques, qu'il numérise. Les modèles récents permettent de visionner la vidéo à distance sur ordinateur. Encore très répandus, ils laissent toutefois peu à peu leur place aux systèmes supportant la vidéo IP de bout en bout [1].

- **Enregistreur vidéo hybride (HDVR) :**

Similaire à l'enregistreur numérique, mais accepte à la fois le branchement de caméras analogiques et IP. Il est possible de rendre hybrides plusieurs modèles d'enregistreurs vidéo numériques par l'ajout d'un logiciel [1].

- **Enregistreur numérique réseau (NVR) :**

Conçu pour les architectures réseaux IP de vidéosurveillance, il ne peut traiter que les signaux vidéo provenant de caméras IP ou d'encodeurs [1].

### **2.3.5- Equipement d'archivage :**

La période d'archivage des séquences vidéo varie selon les besoins de surveillance, allant de quelques jours à quelques années. En moyenne, les organisations conservent les preuves vidéo entre 30 et 90 jours. Le déploiement de vastes réseaux de caméras et l'utilisation de vidéosurveillance à haute résolution fait exploser les demandes pour les systèmes de stockage. Bien que le coût des supports d'enregistrement ait considérablement baissé dans les dernières années, l'archivage représente souvent une part importante des dépenses d'infrastructure en vidéosurveillance, en raison de la quantité toujours croissante de données vidéo à stocker [1].

Les solutions de stockage sont de deux types :

- **Interne** : Les disques durs intégrés aux enregistreurs vidéo numériques ou aux serveurs représentent la forme d'archivage la plus répandue. Elle peut offrir jusqu'à quatre téraoctets d'espace. Certaines caméras IP disposent même d'une carte mémoire ou d'un disque USB permettant d'enregistrer des heures, voir des jours de vidéo. Les solutions internes d'archivage conviennent bien pour les systèmes de vidéosurveillance de taille modeste, comprenant jusqu'à 50 caméras [1].

- **Rattaché** : L'archivage se fait sur des appareils externes aux enregistreurs ou serveurs vidéo. De type NAS (Network Attached Storage) ou SAN (Storage Area Network), ces systèmes offrent un espace de stockage partagé entre les différents clients du réseau. Sur un système de stockage en réseau NAS, un fichier est archivé sur un même disque dur, alors qu'avec le réseau de stockage SAN, un fichier peut être sauvegardé en fragments répartis sur plusieurs supports de stockage [1].

### 2.3.5- Equipement de visualisation

L'équipement de base qui assure la visualisation des images filmées en directe par les caméras ou enregistrées dans un système de vidéosurveillance, c'est le moniteur « écran ». Mais actuellement le client a plus de choix entre une visualisation fixe au bureau sur un téléviseur ou sur un écran d'un PC et une visualisation mobile et même à distance par internet via un PC, tablette ou Smartphone. Suivant le besoin de l'utilisateur, on doit choisir les dimensions et la technologie de ce dernier [1].

### 2.4- Fonctions d'un système de vidéosurveillance

- **Filmer** : La fonction principale d'une caméra est de filmer une scène en direct. Par exemple, il peut être intéressant de montrer aux clients d'une boutique que les lieux sont surveillés par vidéosurveillance en affichant visiblement les images sur un écran directement relié à une caméra. Outre un effet dissuasif, cela permet aux employés de la boutique de garder un œil sur une ou plusieurs scènes depuis un même endroit dans la boutique [3].

- **Enregistrer** : En fonction de la sensibilité de la scène à visualiser (ou du type de bâtiment), il peut être nécessaire d'assurer en outre l'enregistrement des images. Cela permet de revenir sur des événements qui se sont déroulés afin de mieux les interpréter (en cas de vol à l'étalage par exemple). Attention à ne pas tomber dans le piège de l'enregistrement continu : outre une saturation (relativement) rapide du disque dur, cela compliquera également la recherche d'un événement dans les enregistrements. L'enregistrement est à programmer avec pertinence, en privilégiant par exemple les accès au lieu pendant certaines plages horaires [3].

- **Visualiser et analyser** : Qu'il s'agisse d'accéder au canal vidéo d'un ou plusieurs enregistreurs ou d'accéder à leurs enregistrements, il est possible d'installer un logiciel de vidéosurveillance dédié sur un ou plusieurs postes (par ex. poste de commande, poste de police, etc.). La visualisation peut également être effectuée depuis un emplacement distant grâce à un Smartphone ou un ordinateur, pour une surveillance en toute quiétude [3].

## 2.5- Objectifs d'un système de vidéosurveillance

**Surveillance des biens :** Permettre à une zone d'être ouverte au public tout en offrant une protection contre les vols dans une boutique, un parking ou une enceinte. Le système de vidéosurveillance permet de surveiller en direct ou en différé (en cas d'enregistrement) une ou plusieurs zones du bâtiment. Enceinte, parking, zone de vente, etc [3].

**Surveillance des accès :** La surveillance des accès aux bâtiments ou aux locaux hors des heures d'ouverture permet non seulement de détecter les intrusions, mais également de retracer les incidents ou les actes de malveillance. Lors des périodes d'utilisation du bâtiment, le système de vidéosurveillance permet de gérer l'ouverture des accès en détectant la présence d'une personne à proximité d'une sortie ou en travaillant en liaison avec le contrôle des accès visiteur. Enregistrement des accès à un bâtiment [3].

**Surveillance sur événements :** Lorsqu'il est relié aux autres systèmes installés dans un bâtiment ou des locaux, le système de vidéosurveillance peut afficher à l'écran les images provenant d'une caméra associée à un événement [3].

**Intrusion :** vérification immédiate et enregistrement dès la détection de l'intrusion

**Incendie :** vérification de la présence de feu dans la zone concernée Alarme

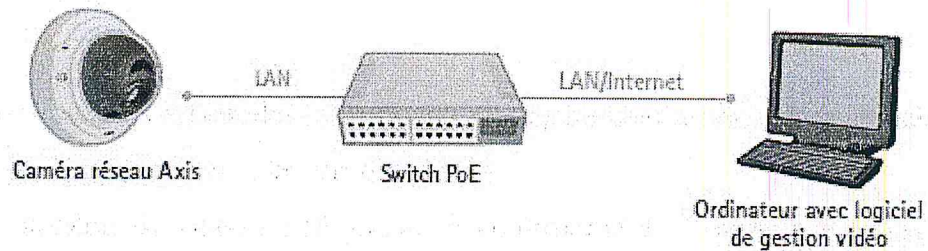
**Technique :** vérification d'un problème technique sur la machinerie [3].

## 3- la vidéosurveillance IP

La vidéo sur IP, souvent nommée vidéosurveillance IP ou IP-surveillance en raison de son application à ce secteur, utilise un réseau IP câblé ou sans fil comme principal support pour la transmission des données audio et vidéo numériques et autres. Lorsque la technologie d'alimentation par Ethernet (PoE,

Power over Ethernet) est employée, le réseau peut également servir à l'alimentation électrique des équipements vidéo sur IP [4].

Un système de vidéo sur IP permet de contrôler et d'enregistrer des images vidéo à partir de tout emplacement sur le réseau, qu'il s'agisse d'un réseau local (LAN) ou d'un réseau étendu (WAN) tel qu'Internet.



**Figure 5 :** Une caméra réseau se raccorde directement au réseau.

À la différence des systèmes analogiques, la vidéo sur IP utilise le réseau informatique plutôt qu'un système de câblage point-à-point pour transmettre les informations. Le terme vidéo sur IP englobe à la fois les sources vidéo et audio véhiculées par le système. Dans une application de vidéo sur IP, les flux d'images vidéo numériques peuvent être transférés n'importe où dans le monde via un réseau IP câblé ou sans fil, permettant une visualisation et un enregistrement vidéo en tout point du réseau [4].

### **3.1- les composants d'un système de vidéosurveillance IP**

Les principaux composants d'un système de vidéo sur IP sont la caméra réseau, l'encodeur vidéo (qui assure la connexion des caméras analogiques au réseau IP), le réseau, le serveur, l'élément de stockage ainsi qu'un logiciel de gestion vidéo. La caméra réseau et l'encodeur vidéo étant des équipements informatiques, leurs fonctionnalités sont sans comparaison avec celles d'une caméra CCTV analogique. La caméra réseau, l'encodeur vidéo et le logiciel de gestion vidéo sont les pièces maîtresses de toute solution de vidéosurveillance sur IP [4].

Le réseau, le serveur et les éléments de stockage font tous appel à des équipements informatiques standards. Le fait d'utiliser du matériel informatique courant est l'un des principaux avantages de la vidéo sur IP. Un système de vidéo sur IP comporte d'autres accessoires, comme des supports, des injecteurs intermédiaires PoE et des joysticks [4].

### **3.2- les avantages :**

Les systèmes de vidéosurveillance IP offrent beaucoup d'avantages, qu'on ne les retrouve pas dans les systèmes de vidéosurveillance classiques, parmi ces avantages, on cite :

- **Haute qualité d'image**

La qualité de l'image est essentielle aux applications de vidéosurveillance pour filmer de manière claire tout incident en cours et identifier les personnes ou les objets impliqués. Les technologies de balayage progressif et de télévision haute définition (HDTV) / méga pixels permettent à la caméra réseau de fournir une meilleure résolution que les caméras analogiques [4].

- **Accessibilité à distance**

Les caméras réseau et les encodeurs vidéo sont configurables et accessibles à distance, ce qui permet à de multiples utilisateurs habilités de visualiser à tout moment la vidéo en direct ou enregistrée depuis n'importe quel point du réseau à travers le monde [4].

- **Gestion des événements et vidéo intelligente**

La quantité d'enregistrements vidéo est bien souvent trop importante et le temps manque pour l'analyser correctement. La vidéo sur IP permet de résoudre ce problème de différentes manières. Les caméras réseau et les encodeurs vidéo, par exemple, peuvent être programmés pour enregistrer des vidéos seulement si un événement se produit, qu'il soit prévu ou déclenché. Cela réduit la quantité d'enregistrements sans intérêt. Les enregistrements vidéo peuvent aussi être



associés à des métadonnées, informations qui facilitent la recherche et l'analyse de vidéos qui présentent un intérêt [4].

- **Intégration facile et évolutive**

Les produits vidéo sur IP conformes aux normes ouvertes sont facilement intégrables à toutes sortes de systèmes de gestion vidéo. Les vidéos prises par les caméras réseau sont également facilement intégrables à d'autres systèmes existants.

Notamment ceux de points de vente (POS), de contrôle d'accès ou de gestion technique de bâtiments. En revanche, les systèmes analogiques disposent rarement d'une interface présentant la même facilité d'intégration à d'autres systèmes ou applications [4].

- **Évolutivité et flexibilité**

Il est possible d'ajouter une seule caméra à un système vidéo sur IP, alors que dans le cas d'un système analogique, cela ne peut se faire que par quatre ou 16 unités. Les systèmes sur IP permettent aussi de partager les données sur un même réseau câblé ou sans fil entre différents périphériques de vidéo sur IP et différentes applications. Un même câble sert à transporter la vidéo, le son, les commandes PTZ et E/S, l'alimentation électrique et d'autres données. Des dispositifs vidéo sur IP peuvent être ajoutés au système sans changement important ou coûteux à l'infrastructure réseau, contrairement aux systèmes analogiques. Dans un système vidéo analogique, un câble dédié, généralement coaxial, doit relier directement chaque caméra à la station de prise de vue ou d'enregistrement. Des câbles audio et PTZ (panoramique/inclinaison/zoom) séparés sont souvent nécessaires [4].

- **Communication sécurisée**

La sécurisation des dispositifs et des flux de vidéo sur IP est assurée de plusieurs manières. Elle peut se faire par authentification sur mot de passe, filtrage d'adresse IP et cryptage de données. Les caméras analogiques n'ont pas ces possibilités de cryptage ou d'authentification. N'importe qui peut avoir accès à la vidéo d'une caméra analogique ou remplacer le signal par un autre. Les systèmes

vidéo sur IP présentent aussi la flexibilité de proposer différents niveaux d'accès utilisateur [4].

### **3.3- Inconvénients de la vidéosurveillance IP**

- Exige une infrastructure réseau
- Requiert compétences en TI et réseautique
- Gestion des systèmes : sécurité ou TI ?
- Sécurité et fiabilité du réseau

### **3.4- Domaine d'application**

Le nombre d'applications dans la vidéo sur IP est pratiquement illimité. Elle est surtout utilisée dans le domaine de la sécurité et de la surveillance à distance des personnes, de lieux, de biens et d'exploitations. Elle est de plus en plus utilisée pour améliorer l'efficacité des activités au fur et à mesure du développement des applications de vidéo intelligente, on peut trouver la vidéosurveillance IP dans les domaines suivants [4] :

#### ➤ **Environnements scolaires**

Applications de sécurité, vidéosurveillance et contrôle distant des cours de récréation, des couloirs, des halls d'école et des classes, et sécurisation des bâtiments.

#### ➤ **Transports**

Contrôle distant des gares et des voies ferrées, des autoroutes et des aéroports.

#### ➤ **Secteur bancaire**

Applications courantes dans les locaux des banques, et sécurisation des appareils automatiques.

#### ➤ **Domaine public**

À des fins de surveillance, pour assurer la sécurité des lieux publics.

➤ **Commerces**

Vidéosurveillance et contrôle distant afin de faciliter et d'optimiser la gestion des magasins.

➤ **Secteur industriel**

Surveillance des processus industriels, des systèmes logistiques, des systèmes de gestion des entrepôts et de contrôle des stocks.

#### **4- la vidéosurveillance intelligente**

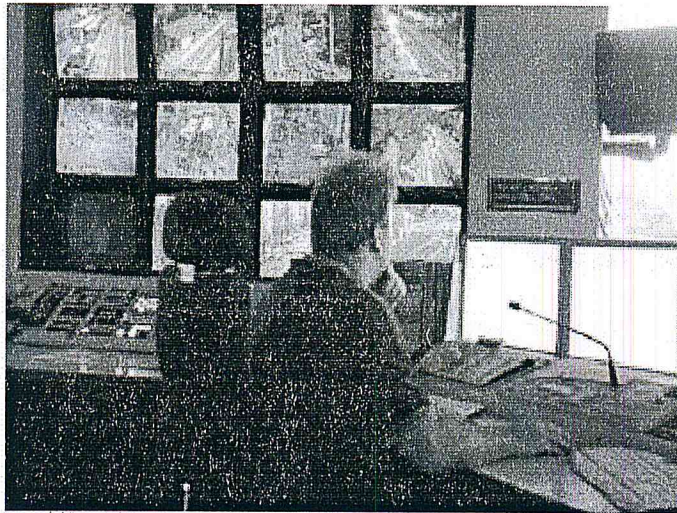
L'analytique vidéo (video analytics), aussi appelée vidéosurveillance intelligente, est une technologie qui permet, au moyen de logiciels, d'identifier automatiquement, dans des séquences vidéo, des objets, des comportements ou des attitudes spécifiques. Elle transforme la vidéo en données qui seront transmises ou archivées pour permettre au système de vidéosurveillance d'agir en conséquence. Il pourra s'agir d'actionner une caméra mobile, dans le but d'obtenir des données plus précises de la scène ou, tout simplement, d'envoyer une alerte au personnel de surveillance pour qu'il puisse prendre une décision sur l'intervention adéquate à apporter [1].

##### **4.1- Raisons pour l'utilisation de vidéosurveillance intelligente**

De nos jours, les réseaux de vidéosurveillance comptent un plus grand nombre de caméras. Pour de grandes infrastructures, par exemple un réseau de transport en commun, plus de mille caméras de surveillance peuvent être déployées. Ces installations représentent une quantité colossale de vidéo à transmettre, visionner et archiver. Il devient donc impossible pour un surveillant humain d'analyser toute cette vidéo, afin d'y déceler les comportements ou événements suspects. D'autant plus que le personnel des centres de surveillance doit gérer d'autres tâches, telles que le contrôle d'accès, l'émission des badges/clés/permis, le traitement des appels d'urgence, le suivi des alarmes d'incendie, le contrôle des communications radio [1].

D'après [5], un surveillant ne peut suivre attentivement 9 à 12 caméras plus de 15 minutes (Figure6), où la probabilité de réagir sur le fait à un événement capté par un réseau de caméras de surveillance est estimée à 1 sur 1000. Dans ce contexte la

vidéosurveillance intelligente paraît être une bonne solution pour parfaire les insuffisances des systèmes de vidéosurveillance.



**Figure 6 :** Surveillant humain contrôlant plusieurs écrans [1]

## **4.2- Secteur d'application**

### **➤ Gouvernement et sécurité publique**

Les différents paliers de gouvernement doivent assurer la sécurité de la population, ainsi que des infrastructures publiques. Au niveau national, elle sera utilisée, par exemple, pour surveiller les éléments suivants [1].

- Infrastructures sensibles
- Frontières
- Édifices et sites gouvernementaux
- Laboratoires
- Bases militaires
- Prisons

## ➤ **Éducation**

La vidéosurveillance est de plus en plus présente dans les établissements scolaires. Son utilisation vise à assurer la sécurité des enseignants et étudiants, ainsi que la protection des actifs contre le vandalisme et le vol [1].

## ➤ **Commerce de détail**

Le commerce de détail représente un marché en expansion pour la vidéosurveillance. Celle-ci y est utilisée autant pour des aspects de sécurité intérieure (magasin, entrepôt) qu'extérieure (stationnement). Même les plus petits commerces se dotent de caméras pour, minimalement, conserver des preuves vidéo en cas de vols ou d'incidents. Dans les chaînes de magasins, des systèmes de vidéosurveillance beaucoup plus sophistiqués sont déployés pour permettre la surveillance centralisée des différentes succursales [1].

## ➤ **Transports**

La sécurité et le bon fonctionnement des aéroports, des gares, des ports et des transports publics est critique pour l'économie d'un pays. Un incident de sécurité peut gravement perturber leurs opérations et se traduire en pertes significatives. Or, étant donné les flots importants de passagers qui y transitent et l'étendue de leurs infrastructures, les systèmes de transports font face à des défis de sécurité hors du commun. Les actes terroristes commis dans différents systèmes de transports dans le monde ont exacerbé ces défis [1].

## ➤ **Aéroports**

Après les événements du 11 septembre 2001, les mesures de sécurité ont connu un resserrement, particulièrement dans les grands aéroports, et de nouvelles technologies ont été déployées. Pour les aéroports, la priorité consiste à contrôler l'accès aux zones sécurisées, principalement l'accès aux avions, mais aussi à assurer la sécurité des passagers, du personnel et des biens sur l'ensemble du périmètre (pistes, stationnements, voies d'accès, etc.) [1].

## ➤ Gares et transports publics

Pour les gares et les transports publics, la vidéosurveillance vise surtout la surveillance des voies d'accès, des quais, des rails et tunnels, des stationnements et de toutes les structures auxiliaires permettant d'assurer le service. Elle s'allie à d'autres systèmes de contrôle d'accès et dispositifs de sécurité pour assurer la protection des passagers, du personnel et des infrastructures. La vidéo est visionnée en direct à partir du centre de surveillance, ou en différé pour les enquêtes suite à des incidents de sécurité [1].

## ➤ Ports

Comme pour les autres systèmes de transports, dans les ports, la vidéosurveillance intelligente est utilisée pour la protection des passagers et du personnel, pour la surveillance des allées et venues dans le périmètre et pour l'identification de véhicules dans ce périmètre. Toutefois, les ports présentent des défis particuliers pour l'analytique vidéo, parce qu'ils requièrent la surveillance des eaux avoisinantes. Les reflets sur l'eau, ainsi que le mouvement des vagues constituent des variations environnementales pouvant perturber les algorithmes analytiques, notamment pour la détection de mouvement [1].

## ➤ Milieu bancaire

Pour le milieu bancaire, la vidéosurveillance intelligente peut augmenter l'efficacité de la surveillance. Elle permet d'assurer un suivi sur l'ensemble du réseau des succursales, afin d'y détecter les individus ou comportements suspects. Elle permettra, entre autres, de retrouver toutes les séquences vidéo provenant de toutes les succursales où apparaît un certain individu, grâce à des techniques de reconnaissance faciales [1].

### 4.3- Avantages :

Les systèmes de vidéosurveillance intelligente offrent un grand nombre d'avantages, les plus importants sont [1] :

- Elle est en fonction 24 heures par jours, sept jours par semaine.
- Elle peut enclencher une alarme qui sera traitée par un opérateur humain ou commander
- Le déplacement ou zoom d'une caméra pour une surveillance plus précise de l'événement, permettant ainsi une intervention en temps réel, plutôt qu'après l'événement.
- Elle réduit la bande passante et l'espace d'archivage nécessaires en ne transmettant ou n'enregistrant que les données sur les événements pertinents.
- Elle libère le personnel de sécurité d'une surveillance continue.
- Elle permet la recherche rapide d'événements pertinents dans les séquences vidéo archivées.
- Elle permet d'identifier les objets dans une scène et de suivre leur activité.

## **5- Conclusion**

Les notions abordées dans ce chapitre survolent de façon globale les systèmes de vidéosurveillance, ses composants, ainsi que les avantages et les domaines d'application.

Par ailleurs, les systèmes de vidéosurveillance intelligente émanant de la troisième génération, sont très prometteurs et permettent beaucoup d'avancées dans différents secteurs, grâce notamment, à leurs utilisations d'algorithmes mathématiques pour la détection d'objets en mouvement et en ne gardant que les mouvements suspects.

Dans le chapitre suivant, nous allons présenter les différentes méthodes de détection et suivi des objets et des personnes en mouvement une séquence vidéo.

***CHAPITRE II :***

***ANALYSE VIDEO***



## Introduction

La détection et le suivi de mouvement dans une scène vidéo est une tâche essentielle pour l'analyse automatique des séquences vidéo pour extraire des informations depuis la vidéo dans un système de vidéosurveillance.

Beaucoup de recherche ont étaient réalisées par des chercheurs dans ce domaine, et différentes méthodes ont étaient proposées, mais ces méthodes souffrent de quelques problèmes comme le changement de la luminosité, les variations dynamiques de la scène etc.

Dans ce chapitre nous allons présenter les différentes méthodes le plus utilisés de détection et suivi d'objet en mouvement citées dans la littérature. Le but de ces méthodes est d'estimer au fil du temps les paramètres d'une cible (ou plusieurs) présentes dans le champ de vision de la caméra et initialement détectée par un moyen quelconque. Les paramètres peuvent être divers : position dans l'image, à laquelle peuvent s'ajouter la taille et l'orientation apparente, l'attitude, l'apparence, etc.

# Partie I : Détection de mouvement

## 1- Le Mouvement

Le mouvement est une des caractéristiques spécifiques à la vidéo. Il est utilisé pour caractériser les déplacements des objets dans une séquence d'images et pour caractériser les mouvements de la caméra [6].

Selon [7] il existe deux types de mouvements :

- **Le mouvement réel**

Le mouvement réel est un mouvement tridimensionnel (3D) tel qu'il est perçu par l'œil humain dans une scène réelle. Ce mouvement n'est pas toujours observable dans une séquence image, ceci est dû à l'absence de la troisième dimension [7].

- **Le mouvement apparent**

Le mouvement apparent dans une séquence d'images en 2D est perceptible grâce aux variations des intensités lumineuses. D'une façon formelle, c'est la projection des différents mouvements tridimensionnels de la scène sur le plan image [7].

### 1.1- Les types d'objet en mouvement

#### 1.1.2- Mouvement d'un objet rigide

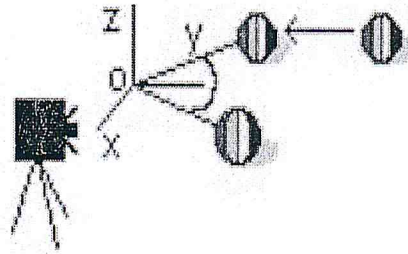
Communément, un objet rigide est un objet solide qui garde toujours sa forme. Par abus de langage, lorsqu'un tel objet se déplace, tous les points qui le composent se déplacent de façon "identique".

Mathématiquement, c'est un objet modélisé par une surface non déformable : plan, plan par morceaux, surface polynomiale, etc. Lors d'un déplacement, les paramètres du mouvement sont les mêmes pour tous les points de l'objet [8].

Le déplacement peut s'exprimer, dans l'espace 3D comme la composition :

1. d'une mise à l'échelle  $S$ , dans le cas d'un zoom effectué par la caméra ou d'un déplacement de l'objet sur l'axe des  $z$ .
2. d'une rotation  $R$  de centre  $O$ .

3. d'une translation T.



**Figure 8 :** Le repère 3-D : vue d'une caméra et d'un objet en mouvement [8]'

- **Mouvement d'un objet déformable (non rigide)**

Un objet déformable est un objet qui n'est pas rigide. Donc, lorsqu'il est en mouvement, chacun de ses points peut subir une transformation différente.

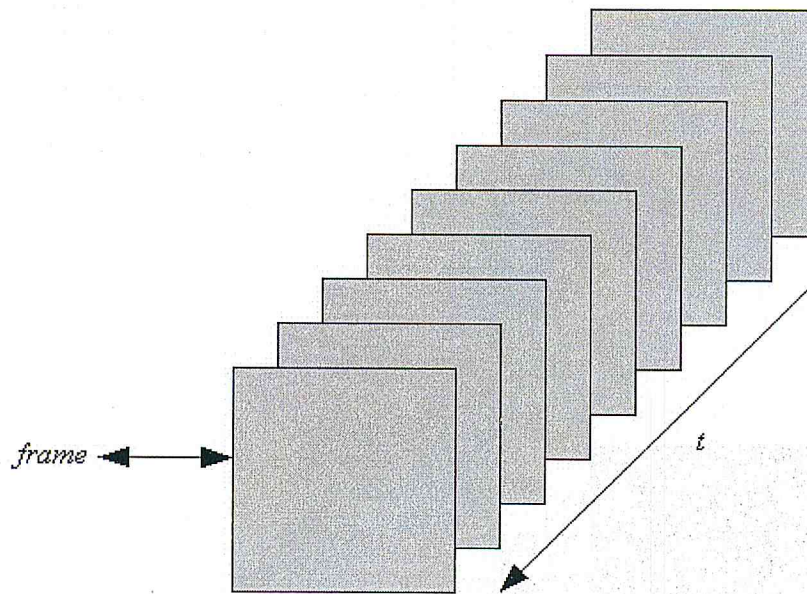
Mathématiquement, cet objet est modélisé par une surface déformable. [9]

## 2- La 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) [10].

### 2.1- Séquence d'image

Une séquence d'images est une succession d'images bidimensionnelles qui montre l'évolution temporelle d'une scène donnée. 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 [10].



**Figure 9** : exemple d'une séquence des images

## 2.2- 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 trois 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 une composante comprise au sens large entre 0 et 255, représentant la luminosité du pixel [10].

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.

## 2.3- Le bruit

En traitement d'image, toute brusque fluctuation de la valeur d'un pixel par rapport à ses voisins est considérée comme un bruit provoquant la dégradation de la

qualité d'image que ce soit de luminance ou de contraste. Ses sources les plus fréquentes sont [11] :

### **2.3.1- 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.

### **2.3.2- 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

### **2.3.3- 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 dû 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 le signal vidéo.

### **2.3.4- 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 nous admettons que chaque pixel est codé sur un octet.

## **2.4- 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 [10].

## **2.5- Objet vidéo**

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.

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 [10].

## **3- La détection de mouvement**

La détection du mouvement dans une séquence d'images consiste à distinguer les zones fixes et mobiles dans 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 [10].

### 3.1- Les méthodes de détection de mouvement

Dans cette section nous allons présenter les différentes méthodes de détection de mouvements, nous pouvons établir une classification hiérarchisée en quatre grandes familles de méthodes en fonction de la modélisation de l'arrière-plan.

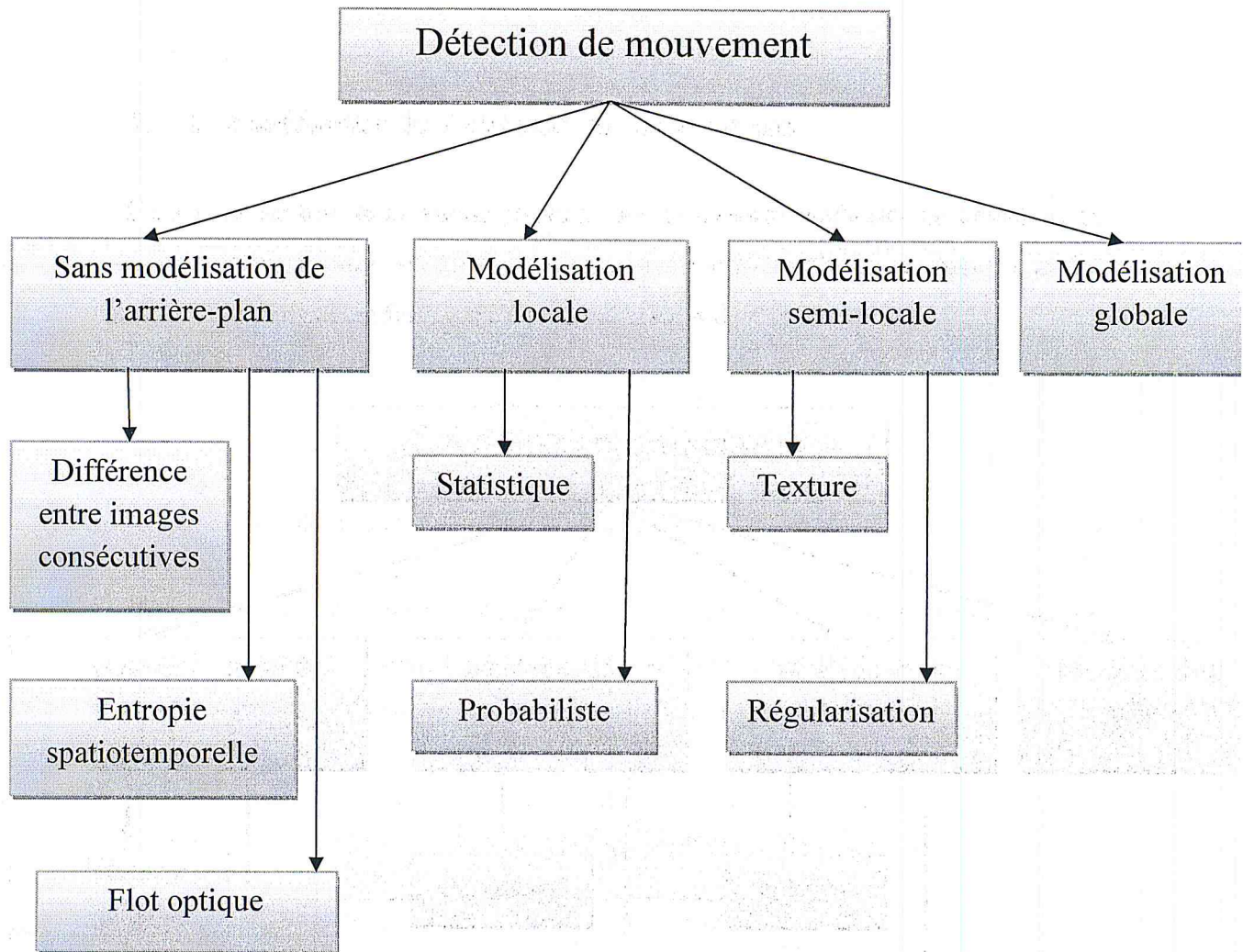


Figure 10 : Taxonomie proposée pour les méthodes de détection de mouvement [12]

#### 3.1.1- Détection sans modélisation de l'arrière-plan

Ces méthodes consistent à détecter le mouvement par le calcul en tout point de l'image d'une quantité mathématique qui est fonction de l'intensité ou de la couleur de l'ensemble des pixels et qui est censée rejeter l'importance du mouvement visible

dans la scène. La dérivée temporelle de l'intensité lumineuse, l'entropie spatiotemporelle de l'image et la norme du flot optique en sont des exemples [12].

- **Dérivé temporelle**

La manière la plus intuitive de détecter les zones en mouvement dans un champ de vision est de mesurer le changement d'apparence des pixels entre deux trames consécutives, soit la dérivée temporelle en tout point. La première utilisation de cette méthode dans l'analyse de séquence vidéo est généralement attribuée à Jain et Nagel [Jain et Nagel, 1979]. L'intensité lumineuse d'un pixel, étant un signal discret à une dimension [12].

- **Entropie spatio-temporelle**

L'entropie est une mesure issue de la thermodynamique, associée au degré de désordre d'un système. Dans notre cas, il s'agira de mesurer en chaque point, la (variabilité) de la grandeur mesurée. Concrètement, plus l'intensité lumineuse (ou la couleur, ou le gradient, etc.) aura pris de valeurs dissemblables en un point pendant un certain intervalle de temps, plus l'entropie sera élevée en ce point. Dans [Ma et Zhang, 2001], les auteurs proposent, pour calculer l'entropie, une méthode à base d'histogrammes spatio-temporels [12].

- **Flot Optique**

Le flot optique est un champ de vecteurs à deux dimensions représentant la projection sur le plan image du mouvement réel observé (tridimensionnel). De nombreuses méthodes ont été proposées, des nouveaux algorithmes sont étudiés et comparés selon des critères de précision et de la densité du champ obtenu, mais aucune mention n'est faite de la complexité algorithmique. Autres travaux permettent de combler cette lacune en mesurant les rapports précision/temps de calcul de ces méthodes. Quelle que soit la méthode choisie, le calcul du flot optique reste une opération très coûteuse en temps de calcul. Le temps-réel peut néanmoins être atteint en sous-échantillonnant les trames et en choisissant un algorithme rapide [12].



### 3.1.2- Modélisation locale de l'arrière-plan

Ces méthodes consistent à associer à tout point de l'image une valeur ou une fonction permettant de modéliser l'apparence de l'arrière-plan en ce point. Le modèle d'apparence de l'arrière-plan en un point ne dépend que des observations qui ont eu lieu en ce même point [12].

Les autres pixels de l'image n'interviennent pas. La grande majorité des méthodes présentées dans la littérature bâtissent un modèle statistique (ensemble de paramètres d'une loi, ou ensemble d'échantillons), mais il peut s'agir d'un processus stochastique, d'un filtre prédictif ou simplement d'une valeur d'intensité [12].

- **Modélisation par une image**

Le modèle d'arrière-plan le plus simple serait une image représentant la scène dépourvue d'objets. Cette méthode présente l'avantage de détecter aussi bien les mouvements lents que les mouvements rapides. Par ailleurs, même les objets momentanément immobiles sont détectés. Cependant, en environnement extérieur, les variations d'intensité lumineuse rendent rapidement obsolète un tel modèle, et il est nécessaire de mettre à jour cette image de l'arrière-plan. Par ailleurs, il n'est pas toujours possible d'obtenir d'une image de la scène totalement vide. Dans ces conditions, il est nécessaire de mettre à jour l'image de l'arrière-plan [12].

- **Modélisation Statistique**

Le problème de la modélisation de l'arrière-plan peut être exprimé d'un point de vue statistique. Il s'agit, pour chaque pixel, d'estimer la probabilité d'y observer telle ou telle couleur (ou niveau de gris) en se basant sur un modèle appris, censé représenter l'arrière-plan de la scène. Le modèle consiste en un ensemble de fonctions de densité de probabilité : une par pixel de l'image.

Les mesures dont la probabilité d'être observées est élevée correspondent des pixels qui seront étiquetés comme arrière-plan, tandis que celles dont la probabilité d'être observées est faible correspondent à des pixels qui seront étiquetés comme avant-plan [12].

- **Modélisation prédictive**

Une autre approche, assez semblable à l'approche statistique, consiste à utiliser un filtre de Wiener ou de Kalman pour prédire la prochaine valeur que l'on devrait observer en chaque point. C'est l'écart entre la prédiction et l'observation qui sera utilisé pour estimer l'amplitude du mouvement. Ces méthodes permettent de gérer les problèmes d'arrière-plans non statiques, et les perturbations intervenant à intervalles réguliers (par exemple, la bande horizontale qui apparaît lorsque l'on filme un écran à tube cathodique [12]).

### **3.1.3- Modélisation semi-local de l'arrière plan**

Ces méthodes sont très semblables à celles de la catégorie précédente, à la différence près que la modélisation de l'arrière-plan en un point dépend des observations qui ont eu lieu dans un certain voisinage de ce point, ou dans la région de l'image à laquelle il appartient [12].

- **Détection par région**

Etant donnée la nature progressive des mouvements généralement observés dans les séquences à analyser, et compte tenu des imprécisions dues aux Caméras vidéo utilisées, certains auteurs préconisent de prendre en compte l'ensemble des pixels d'un voisinage au lieu de chercher à détecter les mouvements en un point donné sans se préoccuper des pixels alentour [12].

- **Caractérisation par la texture**

La prise en compte du voisinage des points peut également être réalisée en calculant en tout point un vecteur caractérisant la texture à cet endroit, et en utilisant l'ensemble des vecteurs calculés comme espace de représentation des données [12].

### **3.1.4- Modélisation globale de l'arrière-plan.**

Ces méthodes utilisent à chaque instant l'ensemble des observations pour construire un modèle de l'ensemble de l'arrière-plan.

- **Basculement entre plusieurs modèles**

Pour prendre en compte la totalité de la scène dans le processus de segmentation entre avant-plan et arrière-plan, des auteurs décident de garder en mémoire  $k$  modèles de l'arrière-plan. La détection de mouvement au niveau local est effectuée avec chacun des modèles, et celui qui détecte le moins de pixels en mouvement est retenu pour la décision finale.

Initialement, les  $k$  modèles sont acquis en exécutant un algorithme  $k$ -means sur les images d'une séquence d'apprentissage. En cours de traitement, les modèles peuvent être mis à jour si une large majorité des pixels de l'image sont détectés comme étant en mouvement [12].

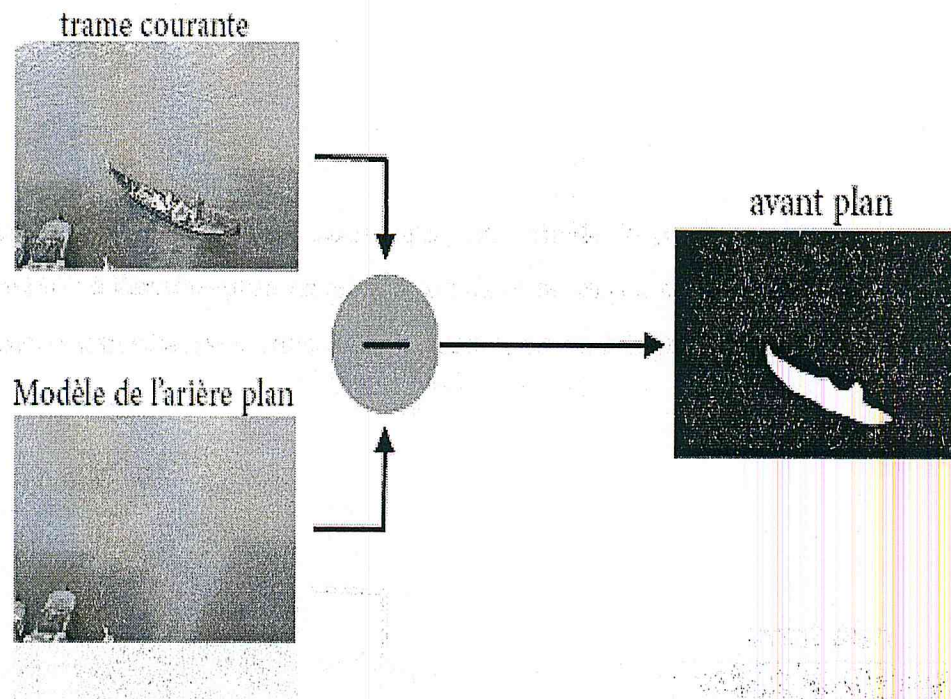
- **Espace vectoriel**

Une autre manière de prendre en compte la globalité de l'espace image pour détecter les mouvements est de considérer les pixels comme des dimensions d'un espace de représentation, et les images successives comme des individus dans cet espace. Les méthodes d'analyse de données permettent alors de considérer tous les pixels de l'image dans une approche globale pour définir de nouvelles caractéristiques que l'on pourra appliquer en tout point pour y détecter d'éventuels mouvements. La méthode des eigenbackgrounds (arrière-plans propres) constitue la première application de l'analyse de données à la détection de mouvement. Ce domaine, bien qu'en pleine expansion n'a pas à ce jour fait l'objet d'un état de l'art spécifique [12].

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

Selon Nicolas et al, 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[12].

Quand il s'agit d'un modèle statistique, on calcule la probabilité que chaque pixel appartienne à l'arrière-plan en testant la valeur observée dans le modèle, l'importance du mouvement observée varie dans le sens opposée à la probabilité calculée.

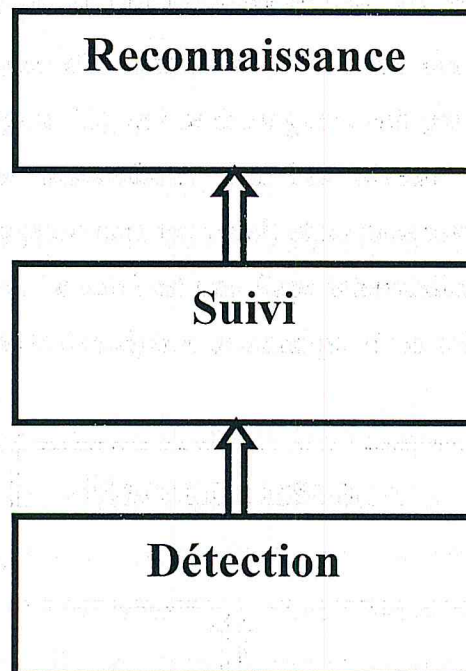


**Figure 11** : Principe de soustraction du fond

## Partie II : suivi de mouvement

### II.1- Hiérarchique du traitement d'une séquence vidéo :

Le traitement de séquence d'images pour le suivi de personnes se divise en trois niveaux hiérarchisés (figure 12), qui se distinguent soit par une approche ascendante soit par une approche descendante. Au bas niveau la détection, au niveau intermédiaire le suivi (appariement temporel) et au haut niveau la reconnaissance des actions ou des personnes. Le suivi est une étape intermédiaire entre la détection et la reconnaissance en vue de la description sémantique de la scène.



**Figure 12:** Les trois niveaux hiérarchisés du traitement d'une séquence d'images

## **2- Définition :**

Selon [13], le suivi peut être défini comme étant le problème de l'approximation de la trajectoire d'un objet dans le plan d'image se déplace autour d'une scène. Les approches du suivi des objets sont : suivi du point, suivi du noyau et la silhouette.

Le suivi d'objet en mouvement peut être aussi défini comme étant la localisation spatiotemporelle d'un objet en mouvement au cours d'une séquence vidéo, ou un objet est une zone de l'image qui peut être modélisée par des contours, silhouettes, primitives géométriques (rectangle englobant l'objet d'intérêt) ou encore par le point central. Cette localisation représente une tâche importante pour des nombreuses applications en vision par ordinateur.

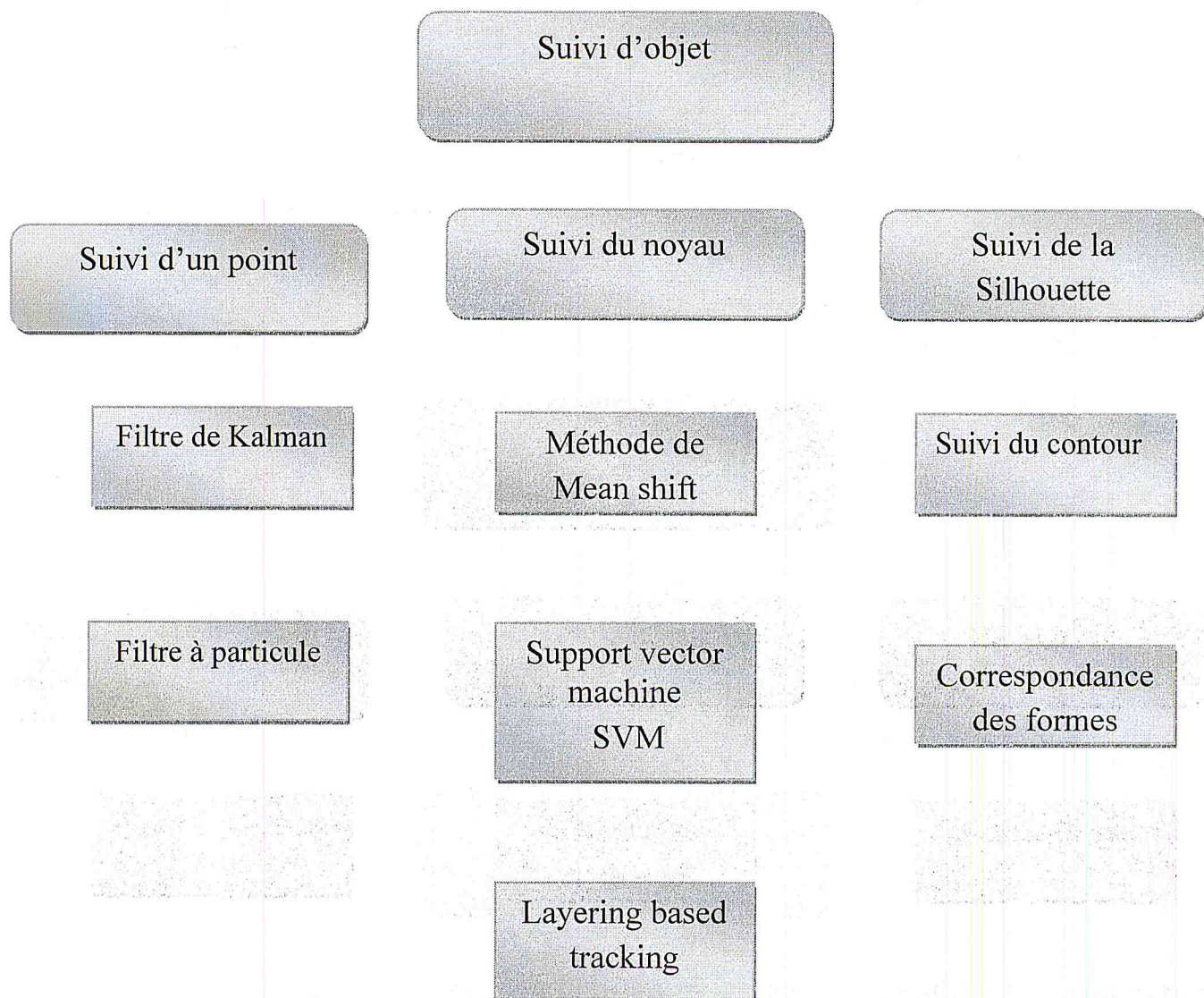
Pour résoudre ces problèmes, deux grandes approches ont été proposées dans la littérature : Les méthodes basées sur un modèle d'objet et les méthodes basées sur le mouvement (*motion-based*). Cette diversité des approches est due aux contraintes du domaine d'application, les particularités des vidéos analysées, des caractéristiques (composantes) de l'image utilisée, de la nature du mouvement estimé.

## **3- Méthodes de suivi d'objet dans une vidéo :**

De nombreuses approches de suivi d'objets ont été proposées. Outre l'algorithme lui-même, la différence entre ces méthodes réside en partie dans le choix de la représentation et de la forme des objets, des caractéristiques (composantes) de l'image utilisées, de la nature du mouvement estimé, etc.

Ce choix dépend de l'application ainsi que de la vidéo traitée. Yilmaz et al réalisent dans [14] une étude des méthodes traitant du problème du suivi d'objets. Les auteurs y établissent une classification des méthodes de suivi dont les catégories sont :

1. Suivi de points
2. Suivi de silhouettes
3. Suivi de noyau



**Figure 12:** les catégories des méthodes de suivi de mouvement dans une vidéo

### 3.1- Suivi d'un point :

Considérons qu'un objet est représenté par un ensemble de points [15]. Le problème du suivi d'objets peut être formulé comme un problème de mise en correspondance de points. La détection de points peut cependant être gênée par plusieurs facteurs tels que la présence de bruit dans les images, les occultations partielles ou totales, les erreurs de détection de points, etc.

Dans cette catégorie, on distingue 3 méthodes :

### **3.1.1- Filtre de Kalman**

Le filtre de Kalman est composé de l'ensemble des équations permettant de trouver une solution récursive à la méthode d'estimation des moindres carrés. Le filtre supporte des estimations des états passés, présents et futurs du système (l'objet suivi.)

Le filtre accepte des modèles imparfaits du système (niveau d'incertitude élevé sur les prédictions.)

L'objectif du filtre discret de Kalman est de trouver une estimation de l'état interne du système en utilisant des mesures effectuées sur le système[10].

### **3.1.2- Filtre à particules**

Filtre à particule est une approche bayésienne dans laquelle la probabilité de la configuration d'un objet (e.g sa position, son échelle), en tenant compte des observations, est représentée par un ensemble d'échantillons pondérés appelés particules, cet algorithme utilise généralement les contours, les caractéristiques de couleur. Cette méthode est également connue sous le nom d'algorithme de Condensation [16].

### **3.1.3- Suivi à multiple hypothèses (Multiple Hypothesis Tracking MHT):**

Dans l'algorithme MHT, plusieurs trames ont été observés pour des meilleurs résultats de suivi, MHT est un algorithme itératif, l'itération commence avec un ensemble des hypothèses existantes des pistes, Pour chaque hypothèse, une prédiction de la position de l'objet dans la trame suivante est effectuée. Les prédictions sont ensuite comparées en calculant une mesure de distance. MHT est capable de suivre plusieurs objets, gérer les occlusions et calculer des solutions optimales [16].



### **3.2- Suivi de noyau**

Le suivi de noyaux consiste à suivre un objet représenté par une forme géométrique basique telle qu'un rectangle ou une ellipse. Le mouvement estimé est généralement paramétrique (translation, rotation, affine, etc.).

L'approche la plus évidente pour suivre un objet repose sur l'utilisation d'un gabarit (en anglais template) [17, 18]. En effet, si l'objet à suivre est de forme connue (e.g. une voiture), il est relativement simple de trouver la partie de l'image la plus similaire au gabarit considéré. Pour ce faire, la recherche est réalisée de manière exhaustive sur tout ou partie de l'image. Les informations utilisées sont l'intensité ou la couleur. Le gradient de l'image est également utilisé pour sa robustesse aux variations d'illumination. L'inconvénient majeur de cette méthode est la lenteur de la recherche exhaustive.

#### **3.2.2- Méthode de MeanShift**

Les algorithmes mean shift sont très populaires en classification de données et segmentation d'image grâce à leurs performances et leurs faibles coûts de calcul [19][20]. Ces méthodes présentent l'avantage de fournir automatiquement le nombre de classes. Cependant, elles sont basées sur l'estimation d'une densité à noyau dont la taille doit être définie préalablement. Les méthodes mean shift usuelles utilisent un noyau de taille fixe pour tout l'espace de données. Cependant, lorsque les données varient de manière significative le long de l'espace caractéristique, des noyaux de taille variable sont préférables.

#### **3.2.3- Machine à vecteur de support (SVM)**

Le SVM [21] est une méthode de classification générale qui donne un ensemble de valeurs de formation positives et négatives. Pour le SVM, les échantillons positifs contiennent les objets images suivis, et les échantillons négatifs sont constitués de toutes les choses restantes qui ne sont pas suivis.

### **3.2.4 Layering based tracking**

C'est une autre méthode de suivi sur la base du noyau où plusieurs objets sont suivis. Chaque couche est constituée d'une représentation géométrique (ellipse), le mouvement comme la rotation, et l'apparence de la couche, selon l'intensité. La Superposition est obtenue par la première compensation du mouvement de fond de telle sorte que le mouvement de l'objet peut être estimé à partir de l'image récompensé au moyen d'un mouvement 2D paramétrique. Chaque probabilité des pixels est calculée sur la base de l'objet de mouvement et les caractéristiques de la forme [21]. Il peut être capable de suivre plusieurs images et l'occlusion totale de l'objet.

### **3.3- Suivi de la Silhouette**

Certains objets auront une forme complexe, comme les mains, les doigts, les épaules qui ne peuvent être bien définies par simple formes géométriques. Les méthodes basées sur la silhouette [14] donnent une description de la forme précise pour les objets. Le but du suivi d'objet par silhouette est de trouver la région d'objet dans chaque image au moyen d'un modèle d'objet généré par les trames précédentes. Ces méthodes sont capables de traiter avec une variété de formes d'objets, occlusion et fractionnement et fusion des objets.

#### **3.3.1- Suivi de Contour**

Les méthodes de suivi de contour [14], progressent de manière itérative un contour primaire dans la trame précédente à sa nouvelle position dans la trame actuelle. Cette progression du contour n'exige que certaine quantité de l'objet dans la superposition de la trame courante avec la région d'objet dans l'image précédente. Le suivi de contour peut être réalisé en utilisant deux approches différentes. La première approche utilise des modèles d'espace d'état pour modéliser la forme du contour et le mouvement. La deuxième approche évolue directement sur le contour en minimisant l'énergie de contour en utilisant des techniques de réduction directe, dont la descente

de gradient. L'avantage le plus significatif des algorithmes basés sur les silhouettes est leur flexibilité pour gérer une grande variété de formes d'objets.

### 3.3.2- Suivi par correspondance des formes

Ces approches examinent pour le modèle d'objet dans la trame existante. La performance de la correspondance des formes est similaire au suivi à base de modèle dans l'approche du noyau. Une autre approche de correspondance des formes détaillée dans [17], elle essaye de trouver une détection de silhouettes correspondantes dans deux trames successives correspondant.

La correspondance de silhouette, peut être considéré similaire à la correspondance des points. Des modèles d'objet sont sous la forme de fonctions de densité, limite de silhouette, les bords des objets. Capable de traiter avec un objet unique et la manipulation d'occlusion sera réalisée avec les techniques de transformée de Hough.

## 5- Structure général des algorithmes de suivi

Les algorithmes de suivi des primitives présentées précédemment de l'instant  $t$  vers un instant  $t + 1$ , sont généralement composés d'un minimum de trois étapes. Après avoir calculé une segmentation (au sens du mouvement) spatiale ou spatio-temporelle initiale, selon le type d'algorithme de suivi considéré, nous pouvons distinguer le trois étapes suivantes :

- Prédiction du mouvement des primitives de l'instant  $t$  à l'instant  $t + 1$
- Adjustment de la segmentation prédite
- Estimation du mouvement des régions suivies de  $t$  vers  $t + 1$

### 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  $t + 1$ . Quatre modèles de prédiction peuvent être envisagés [22]:

#### **4.1.1- 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.

#### **4.1.2- 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.

#### **4.1.3- 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.

#### **4.1.4- 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.

### **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 :

#### **4.2.1- 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.

### **4.2.2- Ajustement basé région**

Cette approche peut s'appuyer, par exemple, sur une modélisation markovienne des zones d'occlusion [23], ou encore sur des modèles de régions déformables [24]. 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.

### **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/objets suivies entre les images aux instants  $t$  et  $t + 1$ . Il existe trois grands types de méthodes d'estimation du mouvement [25]:

#### **4.3.1- 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.

#### **4.3.2- 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.

#### **4.3.3- 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.

## **5- Le suivi des personnes :**

Le suivi de personnes est un processus couramment employé dans diverses applications de vision artificielle, notamment en vidéosurveillance, en analyse de comportements, en sports ou en interaction homme-machine. Le corps humain est alors réduit à un modèle simple, par exemple un point, une forme ou un solide dans l'espace [26]. Les applications typiques de suivi impliquent d'abord une phase de détection où l'observation d'un humain est confirmée. Une fois ceci fait, il est possible de suivre une personne dans le temps, typiquement par des méthodes de filtrage. Détecter et localiser de façon précise un humain n'est pas toujours possible. Les environnements et éclairages sont changeants et souvent imprévisibles. Il est aussi souvent impossible d'observer certaines caractéristiques qui nous intéressent, comme la tête, les pieds ou les bras, simplement parce qu'ils sont cachés par des éléments de la scène ou par le corps de la personne elle-même. On parle alors d'occlusions visuelles.

### **5.1- La méthode utilisée**

Pour la détection et suivi des personnes dans la vidéo, on a utilisé la méthode HOG (Histogramme Of Oriented Gradient).

#### **5.1.1- Descripteur HOG :**

L'extraction des caractéristiques est une étape fondamentale vers la détection et la reconnaissance des objets, Le descripteur de HOG a montré la robustesse et la fiabilité pour représenter les caractéristiques locales de l'image, L'idée essentielle du descripteur est de capturer ou de coder l'apparence et la forme locale d'objet comme la distribution des gradients d'intensité locales ou les directions de bord.

Dalal et al ont proposé de représenter les informations de l'image en utilisant les fonctions du HOG, chaque pixel de l'image d'entrée sera calculé avec un différentiel discret du premier ordre, et le modèle de filtre est  $[-1 \ 0 \ 1]$ , ensuite, l'image est divisée en 105 blocs, et la proposition de chevauchement entre les blocs est 0.5, chaque bloc contient 4 cellules unitaires qui sont  $8 \times 8$  pixels [27].

### 5.1.2- Algorithme pour l'extraction des humains

```
initialiser win_haut,win_larg,n,BM
//win_haut : l'hauteur de la fenetre
//win_larg : la largeur de la fenetre
//n : nombre de distributions
//BM : Modèle du fond
Entrée : Image I
Sortie : Humain Détecté
1) début
2) Calculer l'hauteur h et la largeur l pour l'image
3) initialisation moyenne, le poids et la norme de déviation pour la
gaussienne.
4) pour y ∈ {1..(h-win_haut+1)} faire
    pour x ∈ {1..(w-win_larg+1)} faire
        Construire une sous image 's'
        calculer le tableau des caractéristiques pour 's'
        marquer la location [x,y]
        classifier comme une région humain/non humain
5) si région de (x,y) dans humain alors
    pour chaque pixel p(x,y) dans région
    pour i apparten {1..n} faire
        si p correspond a i alors
            if i apparten a BM alors
                p(x,y) est arriere plan
            sinon
                p(x,y) est avant plan
    fin
mettre a jour en utilisant (4) et (5)
fin
fin
fin
fin
```

**Figure 13** : algorithme de l'extraction des humains [28]

### 5.1.3- Algorithme pour l'extraction des caractéristiques

Entrée : Image I , la taille des cellules CS , nombre des ..

NOB

Sortie : tableau des caractéristiques

- 1) début
- 2) convertir l'image rgb à une image en gris
- 3) calculer l'hauteur h et la largeur l de l'image en gris
- 4) trouver les valeurs de gradient en utilisant l'équation 1 et 2
- 5) trouvé le nombre des cellules CC en utilisant la taille de cellule CS
- 6) pour j ( $\{LCC\}$ )  
calculer l'histogramme en utilisant des valeurs de gradient pour chaque j cellule j avec le nombre donné de NOB
- 7) fin

Figure 14 : Algorithme pour l'extraction des caractéristiques [28]

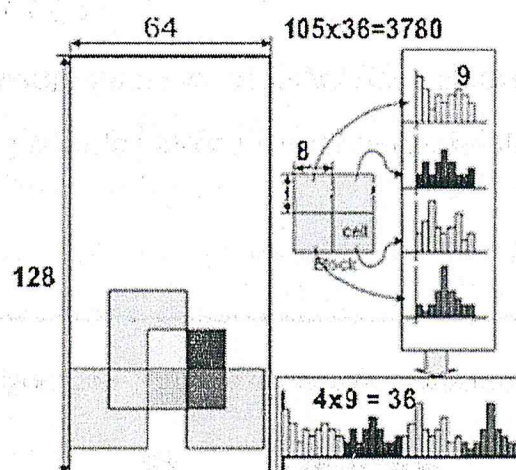


Figure 15 : Calcule du HOG



1. Afin de réduire la variance d'éclairage dans les différentes images, la normalisation de gris est effectuée de telle sorte que toutes les images ont la même plage d'intensité. L'expression est la suivante :

$$H(x, y) = \sqrt{i(x, y)}$$

2. Le même masque centré  $[-1, 0, 1]$  est utilisé pour calculer le gradient horizontal  $G_x(x, y)$  et du gradient vertical  $G_y(x, y)$  pour chaque pixel. L'amplitude  $m$  et l'orientation  $\theta$  des gradients  $G_x(x, y)$  et  $G_y(x, y)$  devrait être calculée pour chaque pixel  $(x, y)$  à l'intérieur d'une fenêtre de détection par les équations suivantes, où  $i(x, y)$  est la luminance en niveaux de gris et  $H(x, y)$  signifie la luminance normalisée avec la loi de puissance (gamma) péréquation au pixel  $(x, y)$ , les expressions sont les suivantes :

$$G_x(x, y) = H(x + 1, y) - H(x - 1, y)$$

$$G_y(x, y) = H(x, y + 1) - H(x, y - 1)$$

3. Calculer la valeur de la norme et l'orientation de chaque pixel, les expressions sont les suivantes :

$$G(x, y) = \sqrt{G_x(x, y)^2 + G_y(x, y)^2}$$

$$\alpha(x, y) = \tan^{-1}(G_y(x, y)/G_x(x, y))$$

4. Diviser l'image d'entrée dans des cellules de la même taille, et les regrouper dans des plus grands blocs, avant de calculer la fonction HOG, l'amplitude de gradient est normalisée dans le bloc
5. Après la normalisation, le bloc est calculé avec une gaussienne spatiale avec  $\sigma = 0.5 * \text{largeur de la fenêtre}$
6. Une interpolation tri linéaire est utilisée pour construire les caractéristiques HOG pour chaque cellule pour obtenir une caractéristique de bas niveau.

## 6. CONCLUSION

Ce chapitre est divisé en deux parties, la première présente les différentes méthodes de la détection des objets et des personnes en mouvement et dans la deuxième nous avons cité les différentes méthodes de suivi des objets et des personnes en mouvement, ainsi l'algorithme que nous avons utilisé dans notre travail. Chaque méthode tente de résoudre un problème bien déterminé en utilisant une modélisation d'objet adéquate qui dépend des propriétés utilisées (mouvement, histogramme de couleurs. etc.).

Dans le chapitre suivant nous allons présenter une conception détaillée de fonctionnement de la plateforme.

***CHAPITRE III :***

***CONCEPTION***

*CHAPITRE III :*

*CONCEPTION*

*CONCEPTION*

## 1- Introduction

Dans ce chapitre, nous allons aborder la conception et l'implémentation de la plateforme de visualisation basée sur les systèmes distribués et incluant les différents traitements vidéo pour la détection et le suivi des personnes.

Dans un premier temps, nous allons décrire le côté conceptuel de la plateforme en utilisant la notation UML (Unified Modeling Language), qui est un langage qui fournit des outils permettant de modéliser le système à concevoir [29].

Afin de concevoir la plateforme, nous devons tenir compte des besoins des utilisateurs. Pour ce faire, nous commencerons dans cette partie par identifier les acteurs et leurs interactions avec la plateforme, par la suite, nous allons décrire le fonctionnement interne de la plateforme ainsi que son architecture (voir figure15).

## 2- Architecture de la plateforme

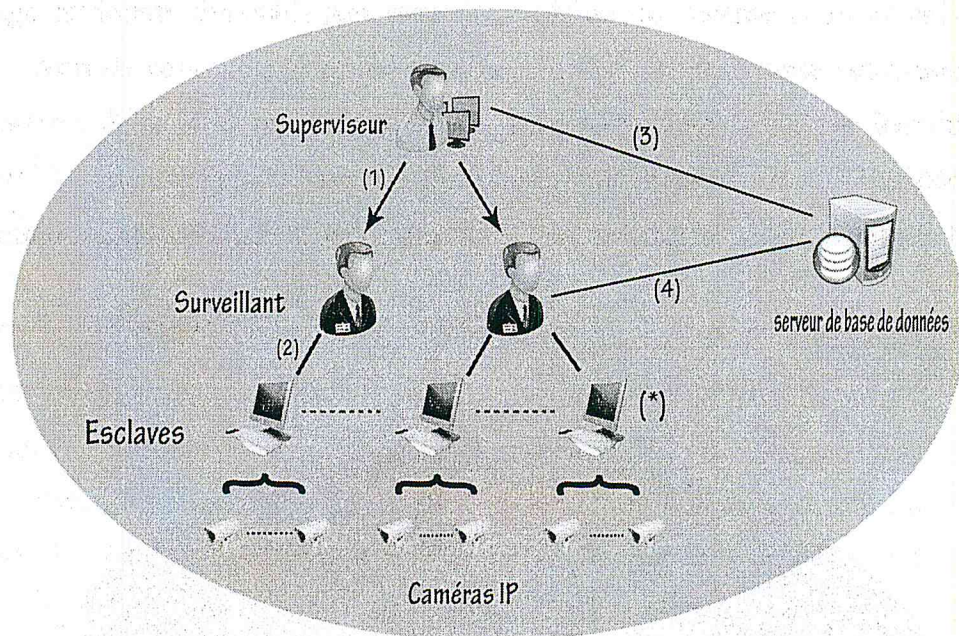


Figure 16 : Architecture de base de la plateforme.

**(1)** : Cette tâche effectuée par le superviseur :

- C'est le super utilisateur, il peut accéder à tout (tout le réseau) y compris la gestion des caméras et les groupes de caméras, il gère aussi les sauvegarde de vidéos enregistrées.

**(2)** : Cette tâche effectuée par le surveillant :

- Le surveillant a pour tâche uniquement la visualisation en temps réel des flux vidéo des caméras IP qui lui sont affectées.

**(3)** : Le superviseur peut ajouter ou supprimer, les caméras et les groupes des caméras. Il doit s'authentifier avant l'accès à la plateforme.

**(4)** : Le surveillant doit aussi s'authentifier avant l'accès à la plateforme.

**(\*)** : Les esclaves contiennent des algorithmes d'analyse vidéo de base(acquisition, détection et suivi des personnes), leur rôle c'est de faire l'acquisition, envoyer les flux vidéo au surveillant et au superviseur, exécuter les algorithmes de détection et suivi des personnes, et envoyer les résultats vers le superviseur et les surveillants.

### **3- Analyse et spécification des besoins**

#### **3.1 Identification des acteurs**

Un acteur est une entité externe qui agit sur le système (opérateur, autre système, etc.) et qui peut consulter ou modifier l'état du système. À partir des besoins et fonctionnalités de notre système, nous apercevons qu'il existe trois acteurs à savoir :

- ✓ **Superviseur** : C'est l'utilisateur qui gère la configuration du système, ainsi que la gestion des caméras IP, l'enregistrement et la visualisation.
- ✓ **Administrateur** : Sa tâche unique est de gérer les comptes des utilisateurs.
- ✓ **Surveillant** : C'est l'utilisateur qui peut seulement visualiser le flux vidéo provenant des différentes caméras IP.

Dans ce qui suit, nous allons détailler ces rôles :

### **3.1.1- La gestion des comptes :**

Chaque utilisateur de la plateforme possède un compte. Ce compte est géré par un administrateur générale qui a un seul rôle c'est la gestion des comptes de la plateforme. Il existe deux types de comptes :

- Un compte superviseur
- Un compte surveillant

### **3.1.2- la gestion des Caméras :**

Le superviseur peut à tout moment, ajouter une nouvelle caméra, modifier les informations des caméras existantes, ou bien supprimer des caméras.

### **3.1.3- la gestion des groupes de caméras :**

Les groupes sont des ensembles de caméras qui sont rassemblés ensemble dans un même endroit, par exemple groupe Direction contient des caméras qui visualisent ce qui se passe dans le bloc direction etc.

### **3.1.4- la gestion des esclaves :**

Le superviseur peut à tout moment, modifier un fichier de configuration qui contient les adresses IP des esclaves qui vont être lancés via l'interface de la configuration des esclaves.

### **3.1.5- la gestion des enregistrements**

L'enregistrement se fait en deux manières : permanent ou à la détection, c'est le superviseur qui choisit la méthode :

#### **- Enregistrement permanent :**

C'est enregistrer tout ce qui se passe en temps réel, du début de lancement de la plateforme jusqu'à son arrêt.

Par contre, le principal inconvénient c'est le problème de stockage. Les vidéos enregistrées vont prendre un très grand espace sur le disque.

#### **- Enregistrement à la détection :**

C'est enregistrer seulement les scènes où il y'a un mouvement. Le principal avantage de cette méthode c'est qu'elle ne prend pas beaucoup d'espace sur le disque par rapport à la première.

En plus de ces différents acteurs de notre plateforme, il faut souligner qu'il existe une hiérarchie entre eux, celle-ci est montrée dans la figure ci-dessous.

## SPECIFICATION

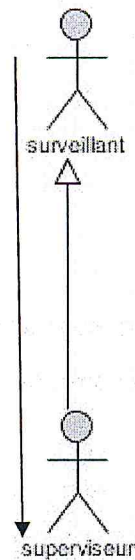


Figure 17 : Hiérarchie entre les acteurs de la plateforme.

### 3.2 Établir les fonctionnalités

Les acteurs interviennent dans les fonctions suivantes :

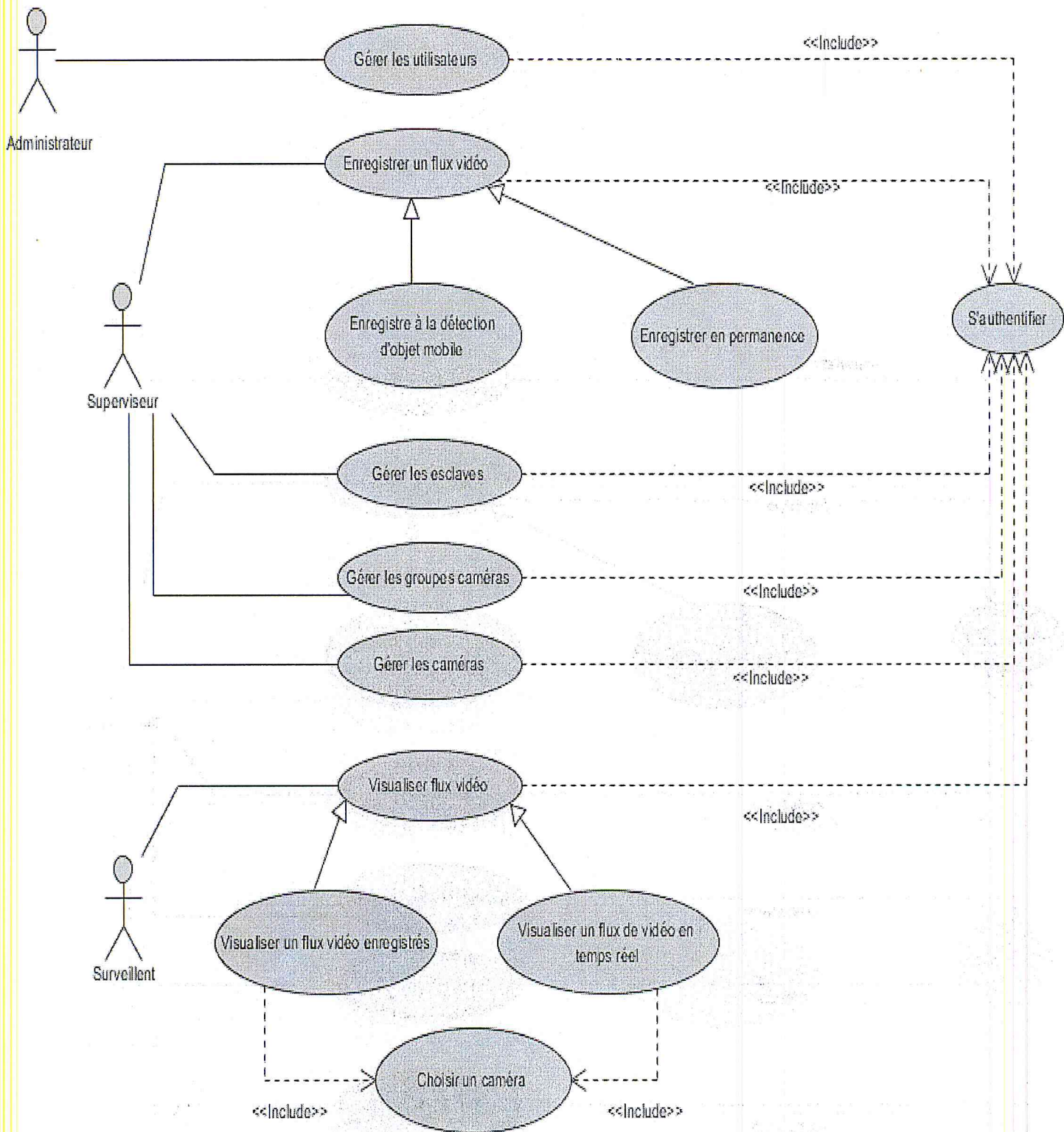
- **Gérer les utilisateurs** : Créer, modifier et supprimer un compte utilisateur.
- **Gérer les caméras** : Ajouter, modifier et supprimer une caméra.
- **Gérer les groupes de caméras** : Ajouter et supprimer un groupe de caméras.
- **Visualiser un flux vidéo** : Visualiser un flux vidéo en temps réel, Visualiser un flux vidéo enregistrés.
- **Enregistrer un flux vidéo** : Enregistrer à la détection d'objets mobiles, Enregistrer en permanence.
- **Gérer les esclaves** : Ajouter et supprimer des esclaves.

### 3.3- Diagramme de cas d'utilisation

Le diagramme de cas d'utilisation représente la structure des grandes fonctionnalités nécessaires aux utilisateurs du système, c'est le premier diagramme du modèle UML, celui où s'assure la relation entre l'utilisateur et les objets que le système met en œuvre [30].

Nous représentons dans cette section un diagramme de cas d'utilisation global pour le superviseur, l'administrateur et le surveillant (figure 17).





**Figure 18 :** Diagramme des cas d'utilisation globale de fonctionnement de la plateforme.

Nous allons maintenant décrire de façon détaillée les cas d'utilisations que nous avons identifiés dans la figure 17.

➤ **Cas d'utilisation « s'authentifier »**

- **Titre :** S'authentifier
- **Acteur :** Tous les acteurs du système
- **Objectif :** Accéder à l'espace personnel
- **Pré-condition :** 1- L'acteur doit avoir un espace personnel (il doit être déjà inscrit).
- **Scénario nominal :**
  1. Le système demande à l'utilisateur d'introduire le login et le mot de passe
  2. l'utilisateur remplit les champs qui lui sont demandés.
- **Enchainements d'erreurs :**
  - 2a. Champs requis non remplis : le système demande de remplir les champs manquants.
  - 2b. Données non valides, un message d'erreur s'affiche.
- **Post-condition :** L'utilisateur accède à son espace personnel.

➤ **Cas d'utilisation «Gérer les utilisateurs»**

- **Titre :** Gérer les utilisateurs.
- **Acteur :** Administrateur.
- **Objectif :** Ajouter/ modifier/Supprimer des utilisateurs.
- **Pré-condition :** 1- L'administrateur s'authentifie.
- **Scénario nominal :**
  - 1- Le système lui affiche la liste des utilisateurs déjà existants.
  - 2-L'administrateur demande d'ajouter un nouvel utilisateur.
  - 3-Le système lui donne la main pour la création d'un nouvel utilisateur.
  - 4- L'administrateur saisit les informations du nouvel utilisateur.
  - 5-L'administrateur enregistre les modifications.
- **Enchainements Alternatifs :**
  - 2a-L'administrateur demande de modifier un utilisateur donné :
  - 4- Le système lui donne la main pour la modification de l'utilisateur.
  - 5- L'administrateur effectue les modifications.

2b-L'administrateur demande de supprimer un utilisateur donné :

- **Enchaînements d'erreurs :**

4e-Les informations concernant le nouvel utilisateur sont erronées :

5-Le système envoie un message d'erreur.

Le cas d'utilisation se termine en échec.

- **Post-condition :** Le système met à jour les utilisateurs.

➤ **Cas d'utilisation « Gérer les caméras »**

- **Titre :** Gérer les caméras

- **Acteur :** Superviseur

- **Objectif :** Ajouter/ modifier/Supprimer des caméras

- **Pré-condition :** 1- Le superviseur s'authentifie

- **Scénario nominal :**

1-Le superviseur demande d'ajouter une nouvelle caméra.

2-Le système lui donne la main pour la création de la nouvelle caméra.

3- Le superviseur saisit les informations de la nouvelle caméra

4-Le superviseur enregistre les modifications.

- **Enchaînements Alternatifs :**

2a-Le superviseur demande de modifier une caméra donnée:

3-Le système lui donne la main pour la modification de la caméra.

4-Le superviseur effectue les modifications.

Le scénario nominal reprend au point 4.

2b-Le superviseur demande de supprimer une caméra donnée:

Le scénario nominal reprend au point 4.

- **Enchaînements d'erreurs**

3e-Les informations concernant la nouvelle caméra sont erronées :

4-Le système envoie un message d'erreur.

Le cas d'utilisation se termine en échec.

- **Post-condition :** -Le système met à jour les caméras.

➤ *Cas d'utilisation «Gérer les groupes de caméras»*

- **Titre** : Gérer les groupes de caméras
- **Acteur** : Superviseur
- **Objectif** : Ajouter/Supprimer des groupes de caméras
- **Pré-condition** : 1- Le superviseur s'authentifie
- **Scénario nominal** :
  - 1-Le superviseur demande d'ajouter un nouveau groupe de caméras.
  - 2-Le système lui donne la main pour la création d'un nouveau groupe de caméras.
  - 3- Le superviseur saisit les informations du nouveau groupe
  - 4-Le superviseur enregistre les modifications.
- **Enchaînements Alternatifs**
  - 2a-Le superviseur demande de supprimer un nouveau groupe.
  - 4-Le superviseur effectue les modifications.

Le scénario nominal reprend au point 4.
- **Enchaînements d'erreurs**
  - 3e-Les informations concernant le nouveau groupe sont erronées :
  - 4-Le système envoie un message d'erreur.

Le cas d'utilisation se termine en échec.
- **Post-condition**

-Le système met à jour les groupes de caméras.

➤ *Cas d'utilisation «Visualiser un flux vidéo en temps réel»*

- **Titre** : Visualiser un flux vidéo en temps réel
- **Acteur** : Superviseur, surveillant
- **Objectif** : voir le flux vidéo des caméras
- **Pré-condition** 1- L'acteur s'authentifie
- **Scénario nominal**
  1. Le système demande à l'acteur de sélectionner les caméras qu'il veut visualiser
  2. L'acteur sélectionne les caméras
  3. L'acteur valide sa sélection
- **Post-condition**

Le système affiche à l'acteur le flux vidéo des caméras sélectionnées

➤ *Cas d'utilisation «Visualiser un flux vidéo enregistré»*

- **Titre :** Visualiser un flux vidéo enregistré
- **Acteur :** Superviseur
- **Objectif :** Visionner un enregistrement
- **Pré-condition**
  - 1- Le superviseur s'authentifie
- **Scénario nominal**
  1. Le système demande au superviseur de sélectionner la caméra qu'il veut visionner ainsi que la date d'enregistrement
  2. Le superviseur sélectionne la caméra et la date
  3. Le système valide sa sélection
- **Post-condition**

Le système affiche au superviseur l'enregistrement souhaité

➤ *Cas d'utilisation « Gérer les esclaves»*

- **Titre :** Gérer les esclaves
- **Acteur :** Superviseur
- **Objectif :** Ajouter/Supprimer des esclaves
- **Pré-condition**
  - 1- Le superviseur s'authentifie
- **Scénario nominal**
  - 1-Le superviseur demande d'ajouter un nouvel esclave.
  - 2-Le système lui donne la main pour la création d'un nouvel esclave.
  - 3- Le superviseur saisit les informations du nouvel esclave
  - 4-Le superviseur enregistre les modifications.
- **Enchainements Alternatifs**
  - 2a- Le superviseur demande de supprimer un nouvel esclave:
  - 4-Le superviseur effectue les modifications.

Le scénario nominal reprend au point 4.
- **Post-condition**

-Le système met à jour les esclaves.

### 3.4- Diagramme de classe :

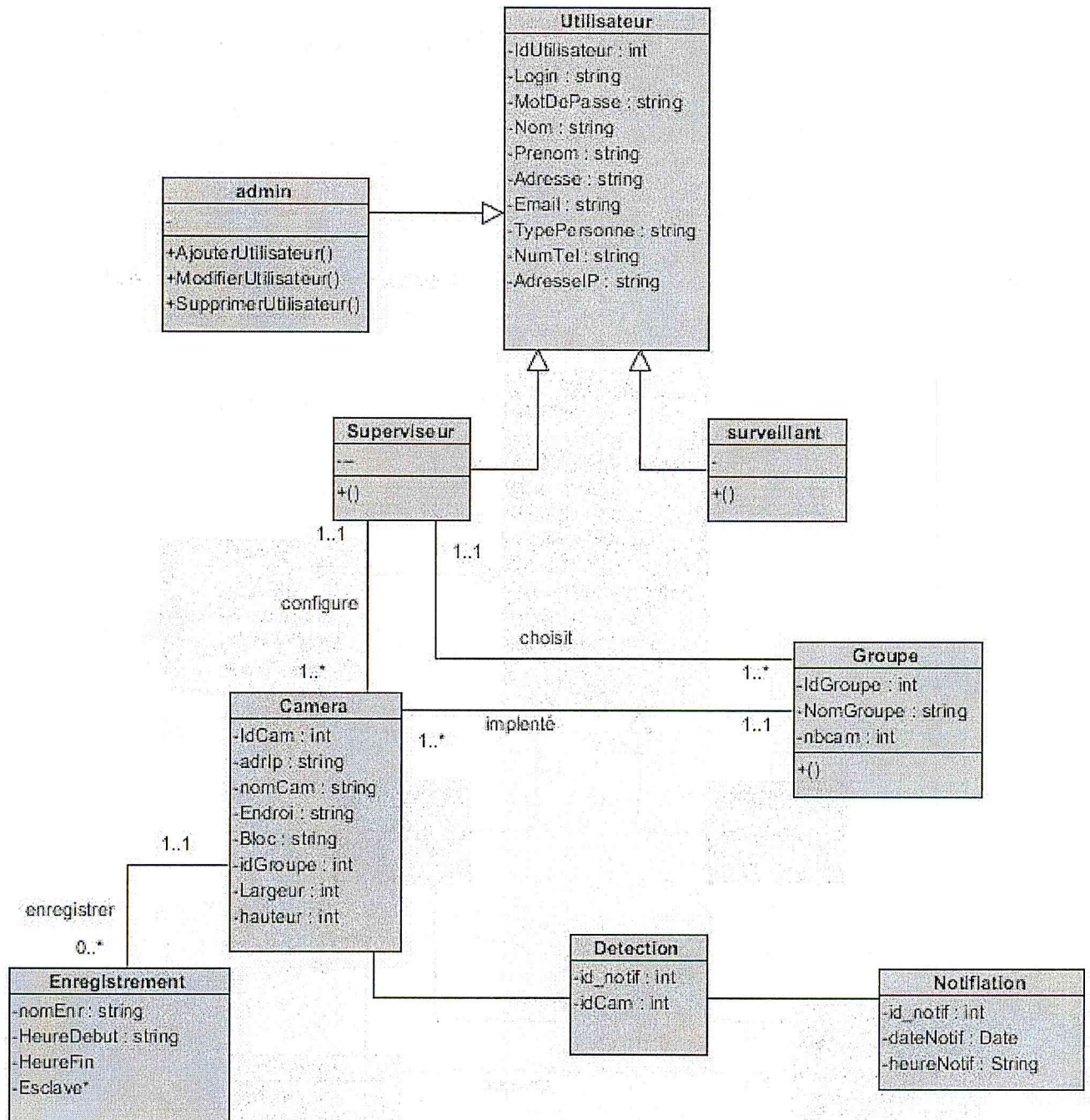


Figure 19 : Diagramme de classe de la plateforme

### 3.5- Diagrammes de séquence :

Les diagrammes de séquences sont une représentation séquentielle du déroulement des traitements et des interactions entre les éléments du système et/ou de ses acteurs.

Nous présentons ci-après les scénarios possibles dans la plateforme.

#### 3.5.1- S'authentifier :

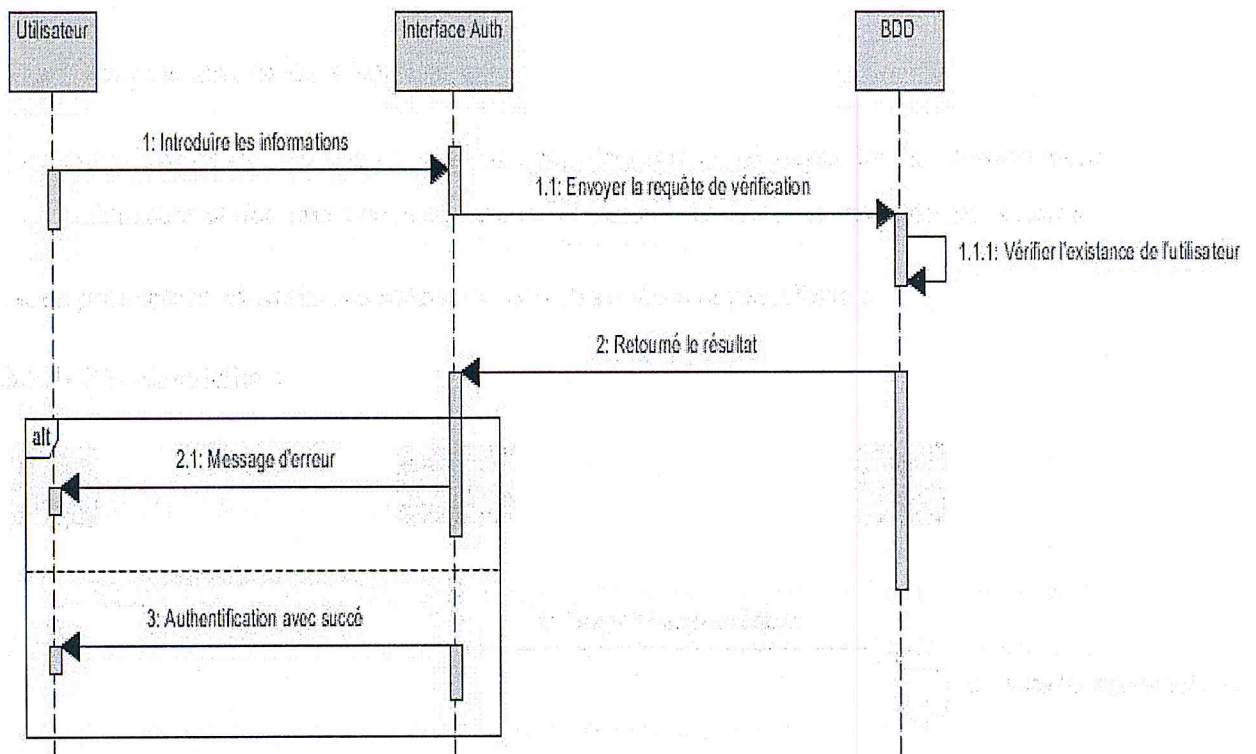


Figure 20 : Diagramme de séquence de l'authentification.

### 3.5.2- Gérer les utilisateurs :

Nous présentons dans la figure 20, le scénario qui se déroule lorsque l'administrateur veut ajouter un nouvel utilisateur.

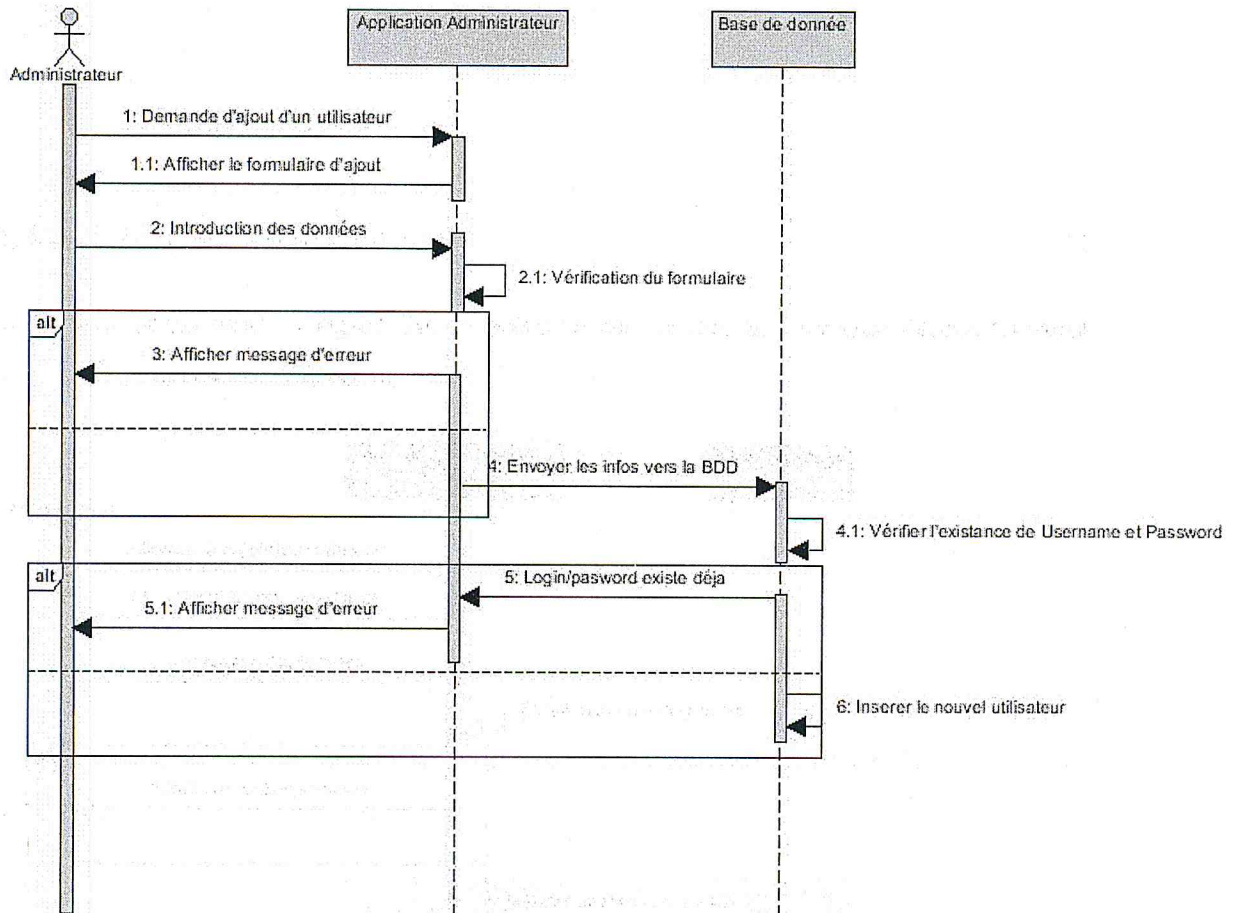


Figure 21 : Diagramme de séquence de gérer les utilisateurs.



### 3.5.3- Visualiser un flux vidéo enregistré :

Il est possible qu'un utilisateur (Superviseur) demande de visualiser un enregistrement. Le scénario se déroule comme illustré dans la figure 21

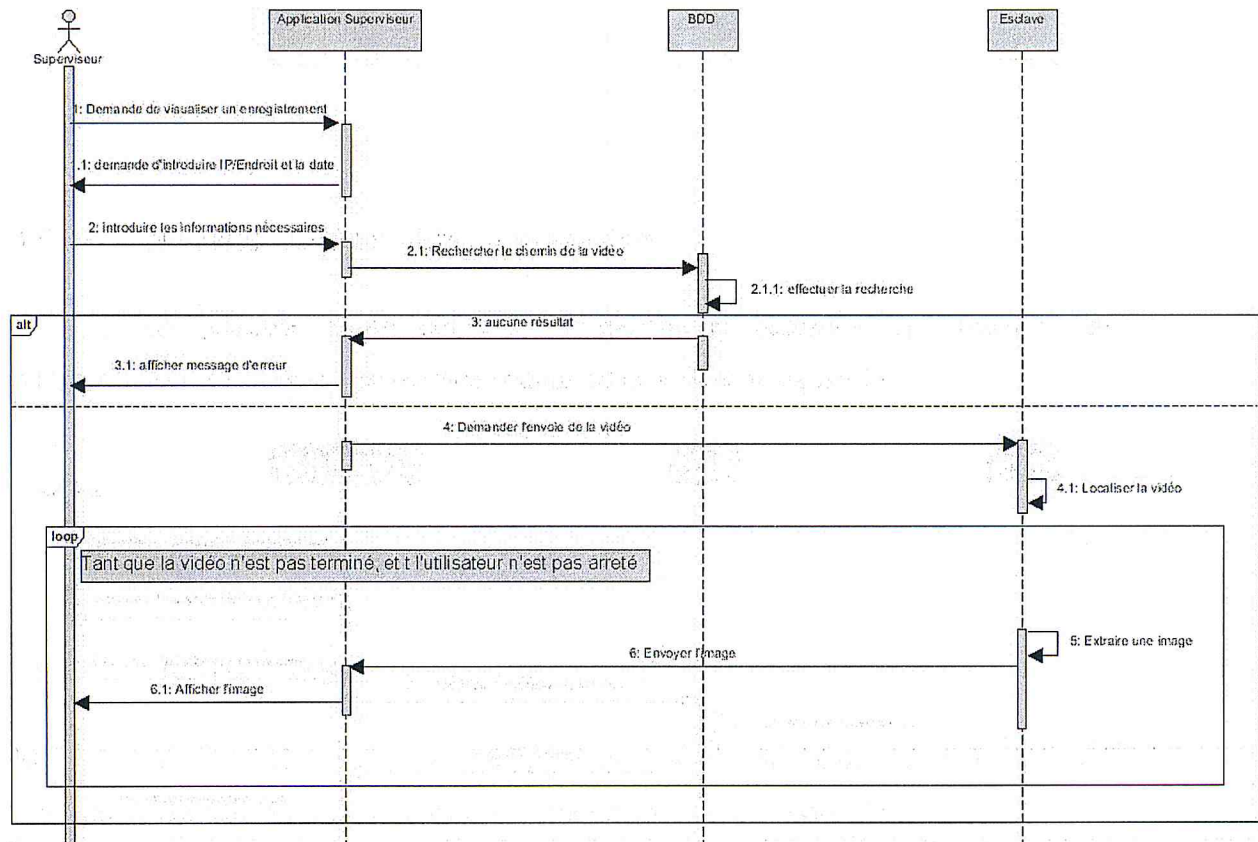
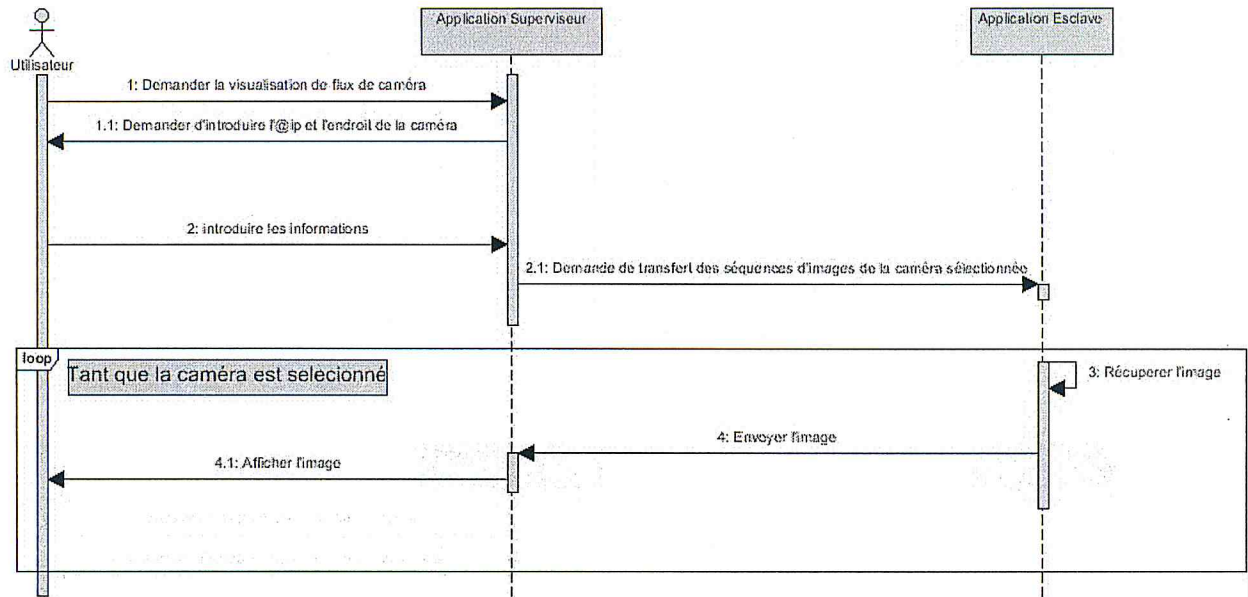


Figure 22 : Diagramme de séquence de visualiser un flux vidéo enregistré.

### 3.5.4- Visualiser un flux vidéo en temps réel

Il est possible qu'un utilisateur quel que soit son grade demande de visualiser un flux vidéo en temps réel. Le scénario se déroule comme illustré dans la figure 22



**Figure 23:** Diagramme de séquence de visualiser un flux vidéo en temps réel.

### 3.5.5- Gérer les caméras

Nous présentons ci-dessous, le scénario qui se déroule lorsqu'on veut ajouter, modifier une caméra. Le diagramme de séquence correspondant est montré dans la figure 23.

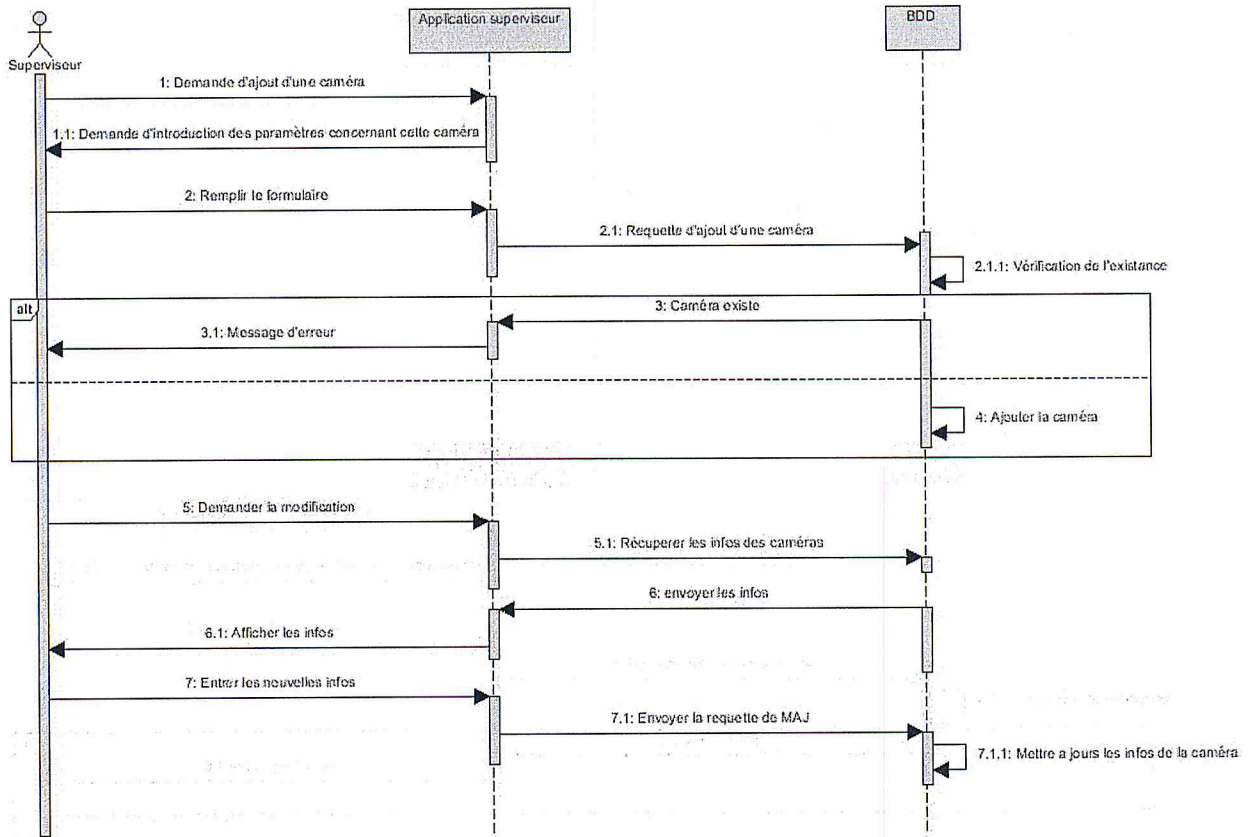


Figure 24: Diagramme de séquence de "gérer les caméras".

### 3.5.6- Gérer les groupes de caméras :

Nous présentons ci-dessous, le scénario qui se déroule lorsque le superviseur veut ajouter un groupe de caméras. Le diagramme de séquence correspondant est montré dans la figure 24

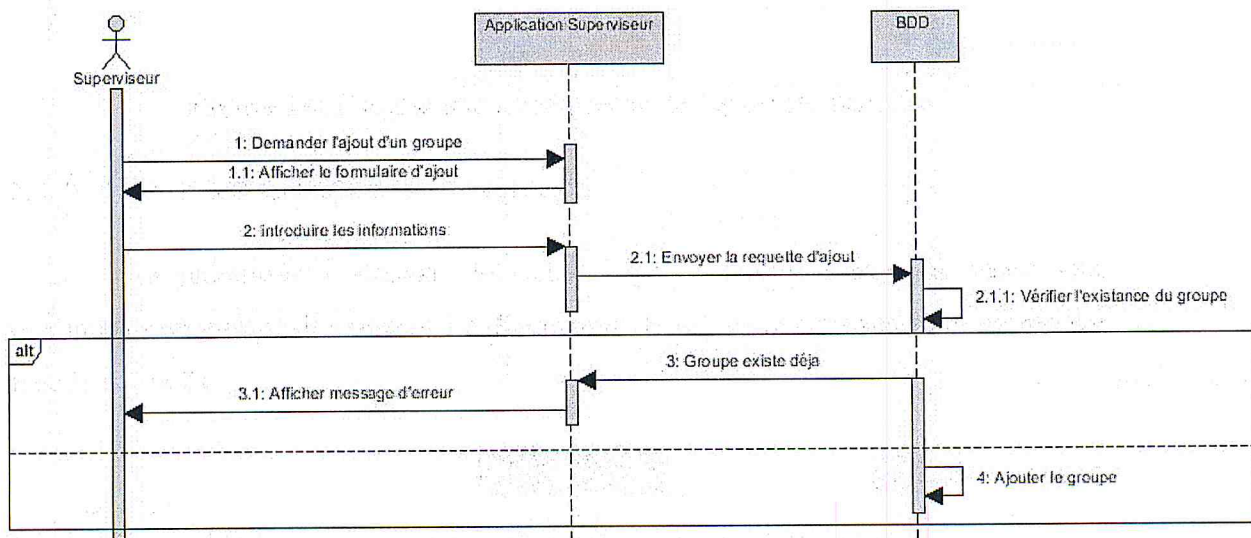


Figure 25 : Diagramme de séquence de "gérer les groupes des caméras".

### 3.5.7- Gérer les esclaves

Nous présentons ci-dessous, le scénario qui se déroule lorsque le superviseur veut ajouter un esclave. Le diagramme de séquence correspondant est montré dans la figure.

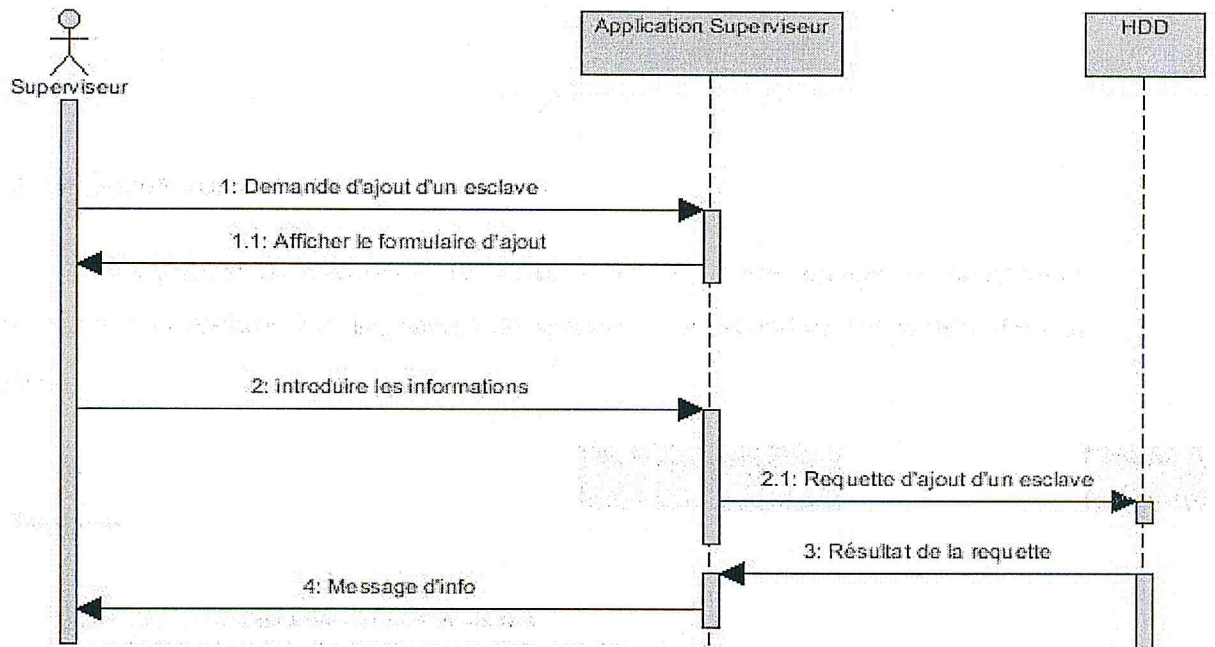
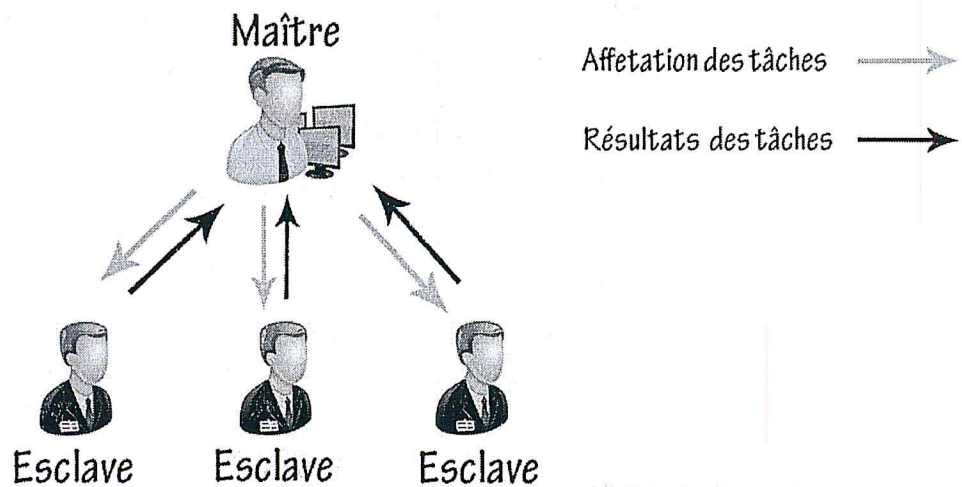


Figure 26 : Diagramme de séquence de "gérer les esclaves".

## 4- Fonctionnement de la plateforme :

Vu le nombre important de caméras IP et la quantité colossale d'informations vidéo générées, une puissance de calcul est nécessaire pour effectuer le traitement du flux vidéo. De ce fait, nous avons opté pour une architecture distribuée pour répartir les données et ainsi les traitements sur plusieurs sites (ordinateurs), en utilisant le paradigme maître-esclaves. Dans ce paradigme, le maître distribue une partie des tâches sur un ensemble d'esclaves (voir figure 26)



**Figure 27:**Paradigme Maître / Esclaves.

Dans la plateforme plusieurs esclaves sont utilisés (des machines autonomes) pour effectuer les traitements, plusieurs maîtres (surveillants) pour la visualisation et un maître (superviseur) pour la visualisation, la gestion et la coordination entre les différentes entités de la plateforme (surveillants, esclaves). Les esclaves, les surveillants et le superviseur communiquent par l'intermédiaire d'un réseau en utilisant les services web.

#### **4.1- Tâches effectuées par le superviseur**

Dans cette section on va présenter les différentes tâches effectuées par le superviseur :

##### **4.1.1- Gestion de la configuration**

Avant le premier démarrage de la plateforme, l'utilisateur (superviseur) doit effectuer les opérations suivantes :

- Création des groupes des caméras dans la base de données.
- Ajout des caméras dans la base de données, en associant à chaque caméra l'un des groupes précédemment créés.
- Ajout des adresses IP des esclaves dans un fichier de configuration.

Notons qu'à tout moment l'utilisateur (superviseur) peut modifier les opérations précédemment citées en respectant la règle suivante : le nombre d'esclaves doit être supérieur ou égal au nombre de groupe.

Au démarrage de la plateforme, le superviseur effectue la répartition des caméras entre les esclaves, de la manière suivante :

- Répartir les esclaves entre les groupes
- Récupérer les caméras associées à chaque groupe pour les répartir entre les esclaves

Au moment du démarrage, tous les esclaves dont les adresses IP se trouvent dans le fichier de configuration doivent être connectés à la plateforme pour le bon fonctionnement de cette dernière.

Au moment de la connexion d'un utilisateur (surveillant), le superviseur lui affecte un groupe ainsi que toutes les caméras appartenant à ce groupe. Notons que le nombre de surveillants connecté à la plateforme doit être inférieur ou égal au nombre de groupe.

Afin de permettre la communication entre les surveillants et les esclaves, le superviseur doit informer chaque surveillant des adresses des esclaves qui lui appartiennent et chaque esclave de l'adresse du surveillant à qui il appartient.

#### **4.1.2- Gestion de pannes**

Dès le lancement de la plateforme, le superviseur communique continuellement avec les esclaves et les surveillants. Dans le cas où l'un des surveillants répond plus, le superviseur affiche un message indiquant à l'utilisateur (superviseur) que l'un de ces derniers a cessé de fonctionner, dans le cas de l'esclave, il récupère les caméras affectées à l'esclave et les répartit sur les autres esclaves fonctionnels.

#### **4.1.3- Gestion d'alarme**

À chaque fois qu'un événement est détecté par un esclave, ce dernier envoie une notification au superviseur et au surveillant. À l'issue de cet événement, une alarme sera déclenchée par la génération d'un son spécifique (un bip sonore).

#### **4.1.4- Visualisation en temps réel**

Ce module permet la visualisation en temps réel des images issues des caméras IP et envoyées par l'esclave. À tout moment l'utilisateur (superviseur) peut choisir une ou plusieurs caméras à visualiser.

#### **4.1.5- Visualisation des enregistrements**

Dans le cas où l'utilisateur (superviseur) souhaite visualiser une vidéo enregistrée, il spécifie l'adresse ou l'endroit de la caméra ainsi que la date d'enregistrement, le superviseur se charge de localiser le chemin de cette vidéo et l'esclave à qui elle appartient, ensuite il envoie une demande à cet esclave pour lui renvoyer la vidéo sous forme d'une séquence d'images.

### **4.2- Tâches effectuées par le surveillant**

Dans cette section on va présenter les différentes tâches effectuées par le superviseur :

#### **4.2.1 Visualisation en temps réel**

Ce module permet la visualisation en temps réel des images issues des caméras IP et envoyées par l'esclave. À tout moment l'utilisateur (surveillant) peut choisir une ou plusieurs caméras à visualiser.

### **4.3- Tâches effectuées par l'esclave**

#### **4.3.1- Acquisition**

Le but du module acquisition est d'extraire les images provenant du flux vidéo d'une caméra IP par le biais d'une requête HTTP.

#### **4.3.2- Enregistrement**

Le but de l'enregistrement est de sauvegarder les vidéos pour une éventuelle utilisation future. L'enregistrement est automatique en cas de détection de mouvement, de plus l'utilisateur peut choisir entre un enregistrement permanent ou à la détection. Les vidéos sont enregistrées en format AVI (Audio Video Interleave).

### **4.3.3-Détection de mouvement**

L'une des tâches d'un système de vidéosurveillance intelligent est la détection de mouvement. Notre solution intègre un module intelligent qui permet de détecter les personnes en mouvements en temps réel et d'envoyer des notifications au superviseur et aux surveillants.

### **4.3.4-Suivi du mouvement**

Après la détection des objets en mouvement, nous avons le suivi des personnes en mouvement dans notre cas le suivi devra être avec un rectangle englobant et chaque personne possédera sa propre couleur, en traçant aussi une trajectoire simple de son mouvement.

## **5- Conclusion**

Dans ce chapitre, on a présenté l'étude conceptuelle et fonctionnelle de la plateforme, on a détaillé les fonctions de chaque acteur dans la plateforme, la hiérarchie entre les utilisateurs ainsi que l'interaction entre eux, dans le chapitre suivant, on va aborder la partie implémentation, tests et validation des algorithmes de détection et suivi des personnes ainsi que les améliorations proposées dans la plateforme



***CHAPITRE IV :***

***IMPLEMENTATION ET  
REALISATION***

## 1- Introduction

Dans le chapitre précédent, nous avons présenté la conception de la plateforme en décrivant son architecture et ses fonctionnalités. Nous allons dans cette partie, aborder l'implémentation réelle de la partie qui nous concerne dans la plateforme, ainsi que son fonctionnement à travers une série de tests expérimentaux.

Nous commençons par exposer les outils utilisés pour le développement, ensuite nous présenterons les différentes interfaces qui constituent notre application et nous terminerons par quelques tests réels.

## 2- Les outils de développement utilisés

Au cours de l'implémentation de notre projet, nous avons utilisé l'environnement de développement Microsoft Visual studio 2010. Nous avons implémenté les interfaces sous le langage C#, nous avons utilisés SQL Server 2008 comme SGBD (Système de gestion de base de données), quant à l'environnement d'implémentation des services nous avons opté pour WCF (Windows Communication Foundation) et d'autres bibliothèques externes de développement comme EMGUCV et krypton toolkit.

### - Microsoft Visual Studio 2010

Un environnement de développement intégré riche pour créer des applications époustouflantes pour Windows, Android et iOS, ainsi que des applications Web et des services de cloud modernes [sw1].

### - Le langage C#

Le C# est un langage de programmation orienté objet, commercialisé par Microsoft depuis 2002<sup>1</sup> et destiné à développer sur la plateforme Microsoft .NET.

Il est dérivé du C++ et très proche du Java dont il reprend la syntaxe générale ainsi que les concepts, y ajoutant des notions telles que la surcharge des opérateurs, les indexeurs et les délégués. Il est utilisé notamment pour développer des applications web sur la plateforme ASP.NET [sw2].

## - **Système de gestion de base de données : SQL Server 2008 R2**

SQL Server est un système de gestion de base de données développé et commercialisé par la société Microsoft. Il est considéré parmi les leaders mondiaux des SGBD. [sw3].

## - **WCF (Windows Communication Foundation)**

Windows Communication Foundation (WCF) est le modèle de programmation unifié de Microsoft permettant de générer des applications orientées service. Il permet aux développeurs de générer des solutions transactionnelles sécurisées et fiables qui s'intègrent à plusieurs plateformes et interagissent avec les investissements existants. [sw4].

## - **EMGU CV**

EMGU CV est un WRAPPER de la bibliothèque OPEN CV en C#. On y retrouve la quasi-totalité des fonctionnalités d'OPEN CV permettant d'effectuer toutes les tâches basiques d'analyse et de traitement d'image [sw5].

## - **Crypton toolkit**

Krypton est un addons pour Visual Studio, qui permet de créer des interfaces très riches, telle que la barre d'outils Office et pleines d'autres fonctionnalités [sw6].

## **3- Prototype matériel**

### • **Ordinateurs**

Ordinateurs portables et /ou de bureau dotés d'un système d'exploitation contenant le Framework 4.0 et équipés d'une carte Ethernet.

### • **Caméras IP**

Le réseau de caméras que nous avons utilisé pour développer notre application est composé de plusieurs caméras IP fixes de marque Axis, type M1011. La caméra IP Axis M1011 offre une qualité d'image vidéo supérieure pour sa

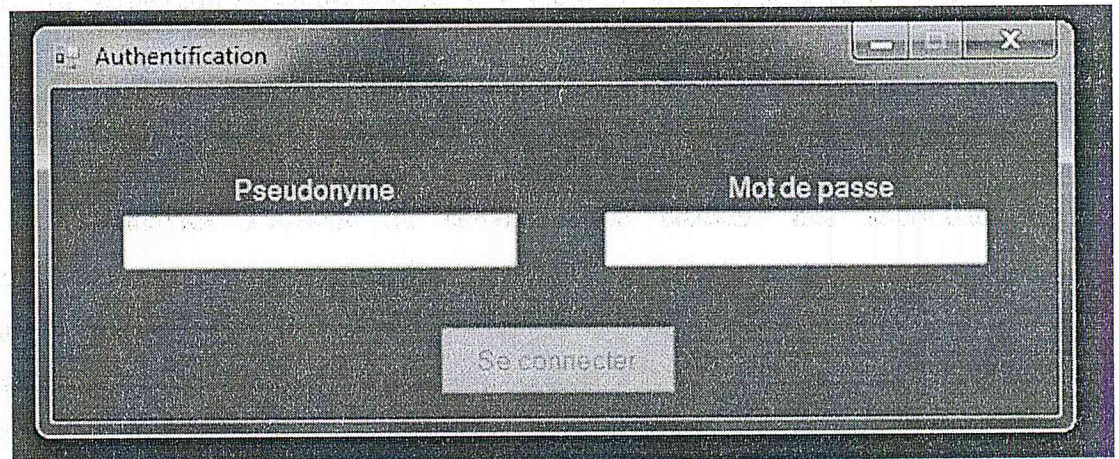
catégorie, pouvant atteindre 30 images par secondes en résolution VGA (Video Graphics Array), ainsi qu'un décodage des flux vidéo multiples.

#### **4- Présentation de l'application**

Dans ce qui suit, nous allons présenter les principales fonctionnalités de l'application. La plateforme interagit avec trois types d'acteur, à savoir : l'administrateur, le superviseur et les surveillants. Chacun des acteurs dispose d'une interface lui permettant de communiquer et d'interagir avec la plateforme. Afin de montrer les différentes interactions, nous allons présenter chacune des interfaces.

##### **4.1-Interfaces d'authentification**

Cette interface s'affiche au lancement de chacune des applications (administrateur, superviseur et surveillant) et exige à l'utilisateur de renseigner les informations nécessaires afin de lui permettre de se connecter à la plateforme. Les informations sont le pseudonyme et le mot de passe.



**Figure 27 :** Interface d'authentification.

##### **4.2- Interface administrateur**

La seule tâche de l'administrateur consiste à gérer les utilisateurs (créer, modifier et supprimer les comptes utilisateurs).

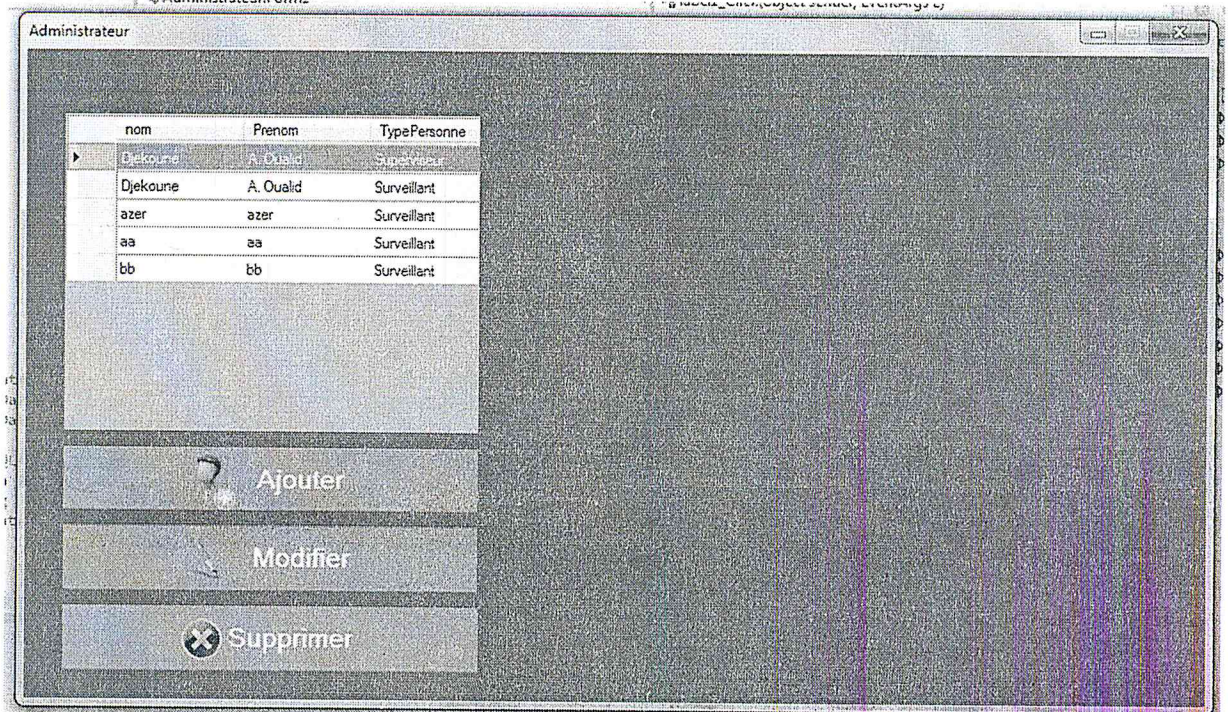


Figure 28 : Fenêtre principale d'administrateur

### 4.3- Interface superviseur

Dans cette partie, nous allons présenter la partie du superviseur

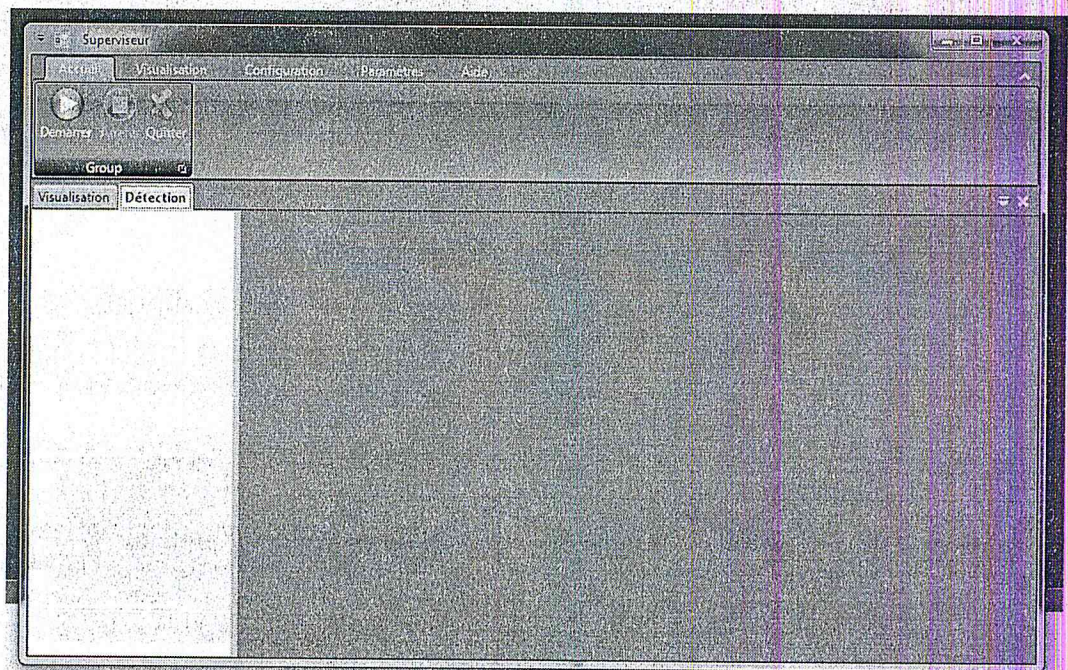
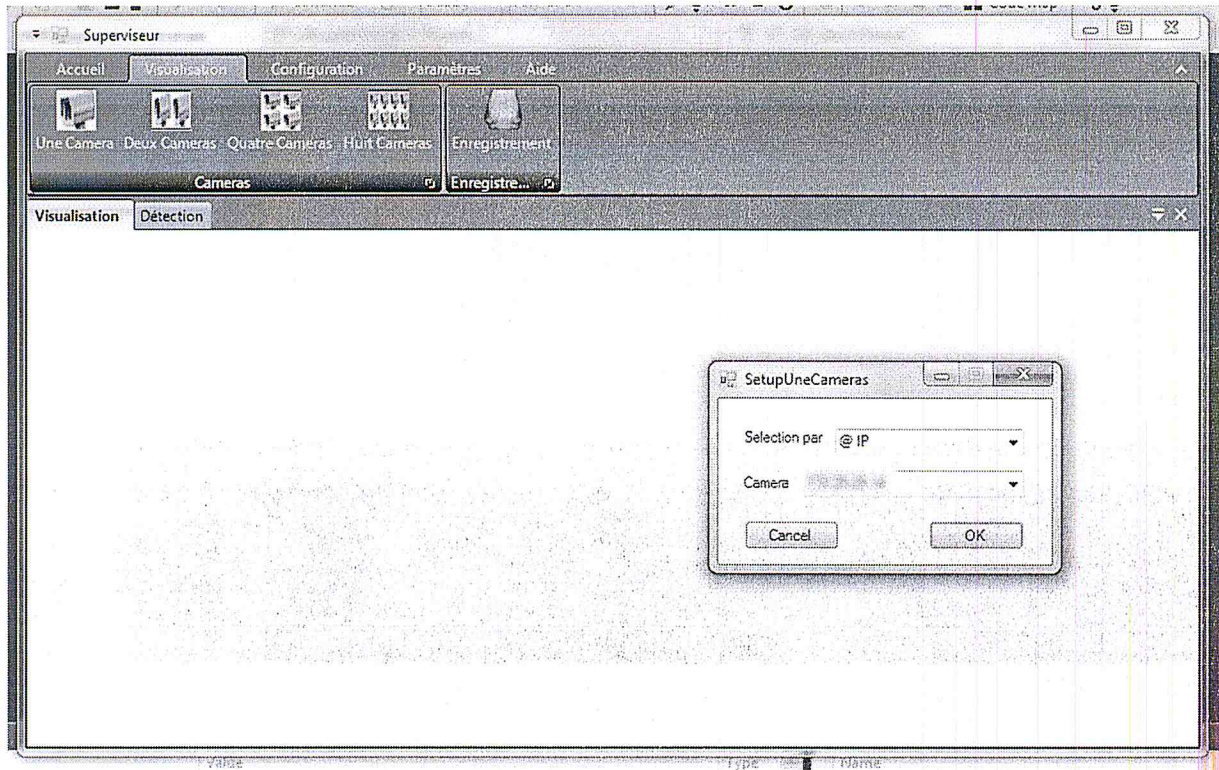


Figure 29 : Fenêtre principale du superviseur

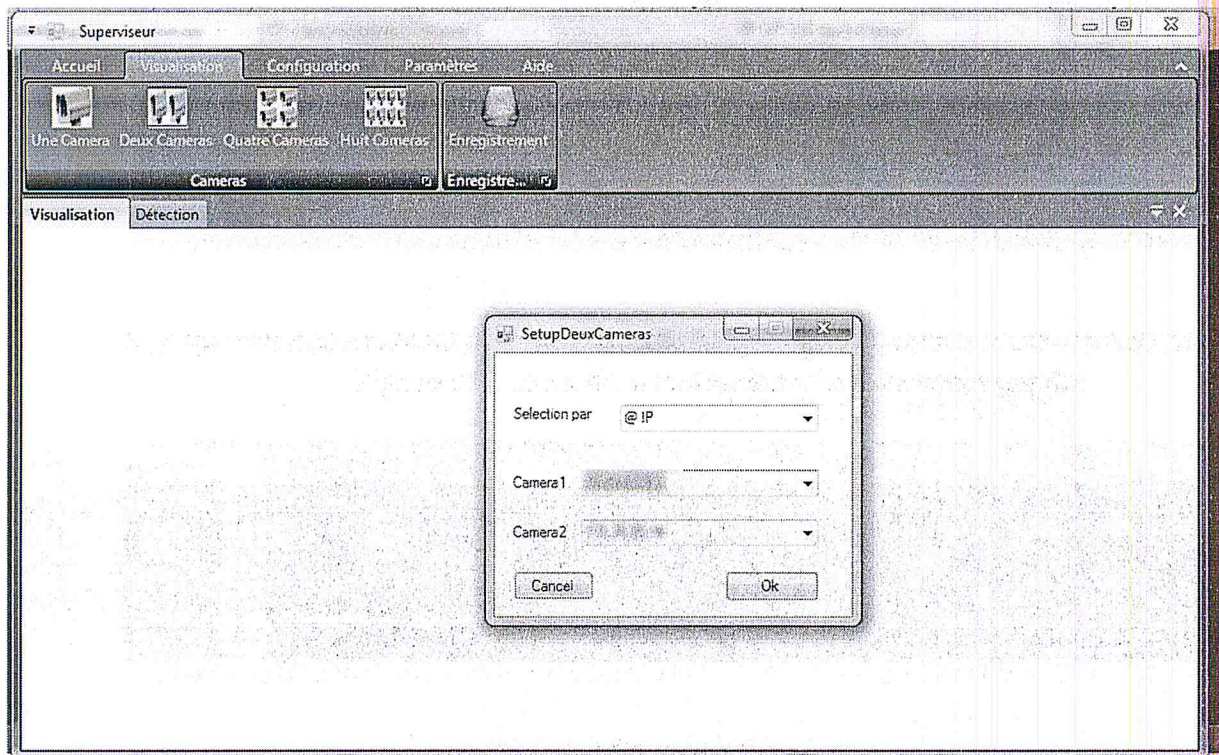
L'interface du superviseur est constituée de trois onglets à savoir : un menu, une page réservée à la visualisation en temps réel des flux vidéo provenant des caméras IP et la visualisation différée des enregistrements vidéo, et une page pour la visualisation des mouvements détectés en temps réel sur les différentes caméras IP ainsi qu'une zone de notification contenant la date, l'heure et l'endroit où le mouvement est détecté.

Les éléments les plus importants dans le menu sont : l'onglet « Visualisation » et l'onglet « Configuration » :

- 1- L'onglet « visualisation » permet de choisir le nombre de caméras à visualiser en temps réel. L'utilisateur peut choisir entre une, deux, quatre ou huit caméras. Après avoir inséré les adresses IP\emplacement des caméras, les flux vidéo provenant de ces dernières sont affichés dans la page réservée à cet effet. L'onglet « visualisation » permet aussi à l'utilisateur de choisir de visionner un enregistrement vidéo en spécifiant la date d'enregistrement et l'adresse de la caméra dont il provient.



**Figure 30 : choix de visualisation avec une seule caméra**



**Figure 31 : choix de visualisation avec 2 caméras**