

**République algérienne démocratique et populaire**  
**Ministère de l'enseignement supérieur et de la recherche**  
**scientifique**

---

**Université du 20 Août -1955-Skikda**

**Faculté des sciences**

**Département d'informatique**



---

Mémoire de fin d'études en vue de l'obtention du diplôme de  
Master professionnel - Option : systèmes d'information  
avancée et application (SIAA)

THEME :

**Conception et réalisation**  
**d'une caméra de**  
**surveillance intelligente :**  
**Détection et**  
**reconnaissance d'objet**

Réalisé par :

GUENNAD Rayane

Encadré par :

LAYADI Saïd

Session : juillet 2023

# *Remerciement*

Au terme de ce travail, nous tenons à exprimer notre gratitude et notre reconnaissance à tous ceux qui ont contribué de près ou de loin à l'élaboration de ce mémoire, en tout premier lieu, notre encadreuse : **Mr. LAYADI Saïd** pour son aide, sa disponibilité, son suivi permanent et ses nombreux conseils qui nous ont permis de mener à bien ce travail. Nous ne manquerons pas également de montrer notre gratitude à l'ensemble du personnel du département informatique notamment le chef de département **Mr. ZEGHIDA Djamel** et tous les professeurs qui nous ont suivi durant nos études.



\*\*\*\*\*

## *Dédicace :*

*Tout d'abord, je tiens à remercier DIEU De m'avoir donné la force et le courage de mener à bien ce modeste travail.*

*Je tiens à dédier cet humble travail :*

*Aux deux êtres qui me sont très cher, eh oui, ma mère Noura et mon père ALI, vous qui m'avez élevé et éduqué... vous qui avez toujours été là pour moi et n'avez jamais cessé de croire en moi, aucun mot ni aucune langue ne pourrait exprimer ma profonde gratitude à votre égard.*

*A toute la famille Guennad et Chaib, à mon frère Mohammed lamine sans oublier leur femme Chaima, à mes sœurs Asma et soraya, à mes deux nièces ALAA AL-Rhmane et Mohamed Djawed.  
à Khaled le mari de ma chère sœur.*

*A tous mes amis et à tous ceux que j'ai rencontrés...*

*A Tous mes amis d'enfance et du long parcours scolaire et universitaire.*

*Tous ceux qui m'aiment et que j'aime.*



*Rayane*

## Résumé :

La vision par ordinateur a connu des avancées significatives dans les domaines de la détection et de la reconnaissance d'objets. Plusieurs approches existent pour résoudre ces problèmes, notamment la détection basée sur des caractéristiques, les régions et le Deep Learning. Cependant, il y a encore des défis à relever dans ce domaine, tels que les variations de l'éclairage, les changements de point de vue et les occlusions. Les avancées récentes dans l'apprentissage profond ont ouvert la voie à de nouvelles approches pour surmonter ces limites.

Dans ce projet, on a créé un système de détection et de reconnaissance d'objets en temps réel en utilisant des techniques telles que les réseaux de neurones convolutifs, les algorithmes de détection en cascade et les méthodes basées sur des caractéristiques. Le système a été conçu en utilisant UML et implémenté en Python. Ce système représente une contribution significative dans le domaine de la surveillance automatisée et peut servir de base pour des améliorations futures.

## ملخص

شهدت رؤية الحاسوب تقدماً كبيراً في مجالات كشف وتعرف الكائنات. هناك العديد من النهج المتاحة لحل هذه المشكلات، بما في ذلك الكشف المعتمد على الميزات والمناطق وتعلم العمق. ومع ذلك، لا تزال هناك تحديات في هذا المجال، مثل تغيرات الإضاءة وتغييرات زاوية النظر والتغطية. فقد فتح التقدم الأخير في التعلم العميق الباب أمام نهج جديدة للتغلب على هذه القيود.

في هذا المشروع، قمنا بتطوير نظام للكشف والتعرف على الكائنات في الوقت الحقيقي باستخدام تقنيات مثل الشبكات العصبية التابعة للتصوير وخوارزميات الكشف التسلسلي والأساليب المعتمدة على الميزات. تم تصميم النظام باستخدام لغة نمذجة موحدة وتنفيذها باستخدام لغة برمجة متعددة النماذج قابلة للتفسير. يمثل هذا النظام مساهمة مهمة في مجال المراقبة التلقائية ويمكن أن يكون أساساً لتحسينات مستقبلية.

## Summary:

Computer vision has made significant advancements in the fields of object detection and recognition. Several approaches exist to solve these problems, including feature-based detection, region-based methods, and deep learning. However, there are still challenges to overcome in this field, such as lighting variations, changes in viewpoint, and occlusions. Recent advances in deep learning have paved the way for new approaches to overcome these limitations.

In this project, we created a real-time object detection and recognition system using techniques such as convolutional neural networks, cascade detection algorithms, and feature-based methods. The system was designed using UML and implemented in Python. This system represents a significant contribution in the field of automated surveillance and can serve as a basis for future improvements.

# Table de matière

**Remerciement**

**Dédicace**

**Résumé**

**Liste des figures**

**Liste des tableaux**

**Abréviations**

**Introduction générale .....1**

**Chapitre 01 Généralité sur la vidéo surveillance .**

Introduction.....	4
1.vidéosurveillan.....	4
1.1. Définition .....	4
1.2. Objectifs.....	7
1.3. L'évolution de la vidéosurveillance .....	8
1.4. Les types de vidéosurveillance.....	12
1.5. Domaines d'application de vidéosurveillance.....	15
1.6. Ses buts.....	16
1.7. Les différents types de caméra de vidéosurveillance.....	16
1.8. Les Avantages du système de la vidéo-surveillance.....	17
1.9. Les problèmes liés à la vidéosurveillance.....	18
Conclusion.....	18

**Chapitre02 la représentation de l'image et de la vidéo.**

Introduction.....	20
2.1.image.....	21
2.1.1. Définition de l'image.....	21
2.1.2. Naissance de l'image numérique.....	21
2.1.3. Représentation d'une image numérique.....	22
2.1.4. Les types d'image numérique.....	22
2.1.4.1.Image binaire (noir ou blanc).....	22
2.1.4.2.Image de niveaux de gris.....	22
2.1.4.3.Image couleur (RVB).....	23
2.1.4.3.1.Codage d'une image en couleur 8 bits.....	24
2.1.4.3.2.Codage d'une image en couleurs 24 bits.....	24
2.1.5. Les caractéristiques d'image.....	25
2.1.5.1. Pixel.....	25
2.1.5.2. voisinage.....	25
2.1.5.3.Poids de l'image.....	26
2.1.5.4.Bruit.....	26
2.1.5.5.Transparence.....	26

2.1.5.6.Luminance.....	26
2.1.5.7.Contraste.....	27
2.1.5.8.Histogramme.....	27
2.1.6. Qualité de l'image numérique .....	28
2.1.7. Les différents formats d'images.....	28
2.1.7.1. Format vectorielle.....	28
2.1.7.2.Format matricielle.....	29
2.1.8.Frame.....	30
2.1.9.Objet.....	30
2.1.9.1. les objets contexte .....	30
2.1.9.2.Les objets mobiles .....	30
2.1.10.La détection d'objet .....	30
2.2.Vidéo.....	31
2.2.1. Définition.....	31
2.2.2. Les fondamentaux physiques et techniques .....	31
2.2.2.1. La vision .....	31
2.2.2.2. La couleur .....	31
2.2.2.3. le signal.....	32
2.2.2.4. Affichage vidéo.....	33
2.2.2.5. Le son .....	34
2.2.3. Types de vidéo.....	34
2.2.3.1. La vidéo analogique.....	34
2.2.3.2. La vidéo numérique .....	35
i) La fréquence d'échantillonnage.....	35
ii) La quantification de l'échantillonnage.....	35
iii)Les Avantages de la vidéo numérique.....	36
2.2.4. Composition des fichiers vidéo.....	36
2.2.4.1. Le conteneur .....	36
2.2.4.2. Le(s) contenu(s).....	36
2.2.5. Les formats de fichiers vidéo.....	37
2.2.5.1. les codecs vidéo .....	37
2.2.5.2. Les Conteneurs vidéo.....	38
2.2.6. Hiérarchie des données vidéo.....	38
2.2.6.1. Séquence vidéo (Video Sequence).....	38
2.2.6.2. Groupe images (Group of Pictures) .....	39
2.2.6.3. Image (Picture).....	39
2.2.6.4. Tranche (Slice).....	39
2.2.6.5. Macroblocs (Macroblocks).....	39

2.2.6.6. Blocs (Block).....	39
2.2.7. Représentation d'une séquence vidéo.....	40
2.2.7.1. La résolution en luminance.....	40
2.2.7.2. La résolution spatiale.....	40
2.2.7.3. La résolution temporelle.....	40
2.2.8. Les paramètres clés d'une vidéo.....	40
Conclusion.....	41

### **Chapitre 03 détection et reconnaissance d'objet.**

Introduction.....	43
3.1. Définition d'objet.....	44
3.2. Caractéristiques d'un objet .....	44
3.3. La détection d'objet.....	45
3.4. La reconnaissance d'objet.....	45
3.5. Le principe de la détection d'objets.....	45
3.6. Classification d'objet.....	46
3.6.1. classification basée sur la texture .....	47
3.6.2. classification basée sur la forme .....	47
3.6.3. classification basée sur la couleur.....	47
3.6.4. classification basée sur le mouvement.....	47
3.7. Quelques Domaines d'application .....	49
3.8. technique de reconnaissance d'objet .....	50
3.9. Méthode de Détection d'objets en mouvement.....	51
3.9.1. La différence temporelle des images (méthode de mouve.....	51
3.9.2. double de différence temporelle et caractère de contour.....	52
3.9.3. La soustraction de l'image de fond (méthode de différence).....	54
3.9.4. Méthode de détection des objets selon la Couleur.....	55
Conclusion.....	57

### **Chapitre 04 Analyse , conception et implémentation**

Introduction.....	59
4.1 Environnement du travail.....	59
4.1.1. Environnement matériel ou le hardware.....	59
4.1.2. Environnement immatériel ou le software.....	60
4.2. Environnement de développement.....	60
4.3. Langages de ... programmation.....	60
4.3.1. Python.....	60
4.3.2. Open CV (Open Computer Visio).....	61
4.4. Diagramme de sequence.....	62
4.5. Présentation de l'interface du système.....	62
Conclusion.....	67

<b>Conclusion général...</b> .....	<b>69</b>
------------------------------------	-----------

**Bibliographies** .....71

## Liste des figures

<b>Figure 1.1</b> : Caméra de surveillance.....	5
<b>Figure 1.2</b> : Caméra de surveillanc.....	5
<b>Figure 1.3</b> : Câbles de liaison.....	5
<b>Figure 1.4</b> : Alimentation électrique.....	6
<b>Figure 1.5</b> : Un enregistreur pour sauvegarder les images.....	6
<b>Figure 1.6</b> : Un logiciel pour accéder aux images.....	6
<b>Figure 1.7</b> : Un écran pour visualiser les images.....	6
<b>Figure 1.8</b> : Modem pour connexion internet.....	6
<b>Figure 1.9</b> : Schéma de principe de vidéosurveillance.....	7
<b>Figure 1.10</b> : Système de vidéosurveillance analogique.....	8
<b>Figure 1.11</b> : Système de vidéosurveillance analogique avec enregistreur numérique.....	9
<b>Figure 1.12</b> : Système de vidéosurveillance analogique avec enregistreur numérique réseau.....	10
<b>Figure 1.13</b> : Systèmes de vidéo sur IP avec serveurs vidéo.....	10
<b>Figure 1.14</b> : Système tout IP.....	11
<b>Figure 1.15</b> : Exemple d'installation numérique.....	13
<b>Figure 1.16</b> : Exemple d'installation IP.....	15
<b>Figure 1.17</b> : les types de caméra de surveillance.....	17
<b>Figure 1.18</b> : Vigiles de surveillance.....	18
<b>Figure 2.1</b> : Image binaire.....	22
<b>Figure 2.2</b> : Image en niveaux de gris (Nuance de 256 gris Exemple de photo possible en 8bpp).....	23
<b>Figure 2.3</b> : Image en couleur.....	24

<b>Figure 2.4</b> : pixel dans une image numérique.....	25
<b>Figure 2.5</b> : Voisinage à 4.....	26
<b>Figure 2.6</b> : Exemple d'image en niveaux de gris.....	27
<b>Figure 2.7</b> : Représentation d'un flux de vidéo.....	31
<b>Figure 2.8</b> : codage de couleur en RVB.....	32
<b>Figure 2.9</b> : Schéma de principe d'un Capture Tri-CCD (chargedcoupledevice).....	32
<b>Figure 2.10</b> : Hiérarchie des données dans le flux vidéo.....	39
<b>Figure 3.1</b> : Plusieurs objets détectés dans la scèn.....	45
<b>Figure 3.2</b> : Principe de la détection d'objet.....	46
<b>Figure 3.3</b> : Classification d'objet.....	46
<b>Figure 3.4</b> : Exemples des domaines d'applications.....	49
<b>Figure 3.5</b> : Un exemple de la différence temporelle. ....	51
<b>Figure 3.6</b> : Détection de mouvement par différence temporel.....	54
<b>Figure 3.7</b> : Exemple soustraction du fond.....	54
<b>Figure 3.8</b> : Exemple1, pixels de couleur rouge reconnus dans les particules et reconnaissance de l'objet.....	56
<b>Figure 3.9</b> : Exemple 2, reconnaissance d'un objet à l'aide de la couleur après modification d'un paramètre.....	56
<b>Figure 4.1</b> : Exemple d'un diagramme de séquence.....	61
<b>Figure 4.2</b> : diagramme de séquence.....	62
<b>Figure 4.3</b> : interface d'authentification.....	63
<b>Figure 4.4</b> : interface d'erreur lors de la saisie du mot de passe et nom d'utilisateur.....	64
<b>Figure 4.5</b> : interface de succès lors de la saisie du mot de passe et nom d'utilisateur.....	65
<b>Figure 4.6</b> : la détection de l'objet en temps réel.....	66
<b>Figure 4.7</b> : la détection de l'objet en temps réel.....	67

## **Liste des tableaux :**

<b>Tableau 3.1 :</b> Étude comparative des méthodes de classification des objets.....	48
---	----

# **Abbreviations**

**DVR3:** Digital Video Recorder 3

**CD-ROM:** Compact Disc Read-Only Memory

**IP:** Internet Protocol

**PDA:** Personal Digital Assistant.

**GSM:** Global System for Mobile

**LAN :** Local Area Network

**CPL :** Courant Porteur en Ligne

**VPN:** Virtual Private Network

**H.264:** AVC (Advanced Video Coding)

**MP3 :** MPEG-1/2 Audio Layer 3

**Codec :** Encodeur/Décodeur

**RAM:** Random Access Memory

**RPM:** Revolutions Per Minute

**MP4:** MPEG-4Part 14

**UML:** Unified Modeling Language

**IA :** Intelligence Artificielle

*Introduction*

*Générale.*

# Introduction Générale

La détection et la reconnaissance d'objets sont des domaines clés de la vision par ordinateur, qui ont connu des avancées significatives ces dernières années. Dans cet article, nous allons explorer les différentes approches qui existent pour résoudre ces problèmes. Parmi ces approches, nous pouvons citer la détection basée sur des caractéristiques, qui utilise des descripteurs de bas niveau pour identifier les objets, basée sur les régions, qui recherche les régions d'intérêt dans l'image pour identifier les objets, et basée sur le Deep Learning, qui utilise des réseaux de neurones pour apprendre à identifier les objets.

Cependant, il y a encore des défis et des limitations à relever dans ce domaine. Par exemple, la détection et la reconnaissance d'objets peuvent être affectées par les variations de l'éclairage, les changements de point de vue et les occlusions. Il est donc important de développer des méthodes qui soient robustes à ces variations.

Enfin, nous allons examiner les perspectives d'avenir pour la recherche dans ce domaine en constante évolution. Les avancées récentes dans l'apprentissage profond ont ouvert la voie à de nouvelles approches, telles que l'apprentissage par renforcement et l'apprentissage actif, qui pourraient permettre de surmonter les limites actuelles de la détection et de la reconnaissance d'objets. En outre, les progrès en matière de traitement des données en temps réel et de l'utilisation des technologies de réalité augmentée ouvrent des perspectives prometteuses pour des applications encore plus sophistiquées dans le domaine de la détection et de la reconnaissance d'objets.

Dans ce mémoire nous aurons quatre chapitres,

**Le premier chapitre :** intitulé GENERALITES SUR Le VIDEO SURVEILLANCE nous parlerons de c'est quoi la vidéosurveillance, ses buts, ses défauts, les différents types de vidéosurveillance et les domaines liés à cette dernière.

**Dans le deuxième chapitre :** intitulé LA REPRESENTATION DE L'IMAGE ET DE Le VIDEO nous parlerons en détail de l'image numérique et la vidéo numérique leurs caractéristiques, leurs différentes formes et leurs paramètres.

**Le troisième chapitre** intitulé LA DETECTION ET LA RECONNAISSANCE D'OBJET c'est plus technique car on va spécifier nos besoins et ceux du système, nous connaissons aussi les

différentes méthodes de détection de mouvement et de reconnaissance leurs points forts et leurs points faibles.

**Pour le dernier chapitre** intitulé ANALYSE, CONCEPTION ET IMPLIMENTATION nous mettront en place un diagramme de séquence qui va nous aider à réaliser ce système, aussi nous présenterons les différents outils utilisés pour aboutir notre travail des différents Framework utilisés et nous présenterons l'interface du système de reconnaissance tel qu'il implémenter dans nos machines.

*Chapitre 1 :*

*Généralités sur la*

*vidéo surveillance.*

## Chapitre 01 :

# Généralités sur la vidéo surveillance

## Sommaire

<a href="#">Chapitre 01 :</a> .....	4
<a href="#">Introduction</a> .....	4
<a href="#">1. La vidéosurveillance</a> .....	4
<a href="#">1.1. Définition</a> .....	4
<a href="#">1.2. Objectifs</a> .....	7
<a href="#">1.3. L'évolution de la vidéosurveillance :</a> .....	8
<a href="#">1.4. Les types de vidéosurveillance</a> .....	12
<a href="#">1.5. Domaines d'application de vidéosurveillance</a> .....	15
<a href="#">1.6. Ses buts</a> .....	16
<a href="#">1.7. Les différents types de caméra de vidéosurveillance</a> .....	16
<a href="#">1.8. Les Avantages du système de la vidéo-surveillance</a> .....	17
<a href="#">1.9. Les problèmes liés à la vidéosurveillance</a> .....	18
<a href="#">Conclusion</a> .....	18

## Introduction

La vidéosurveillance est, le plus souvent, implantée dans le cadre d'un programme de prévention de la criminalité ou de renforcement de la sécurité publique. Dans ce chapitre, nous allons présenter des généralités sur les systèmes de vidéosurveillance.

### 1. La vidéosurveillance

#### 1.1. Définition

La vidéo surveillance est un système de caméras permettant de surveiller à distance un espace privé ou public. Des images sont enregistrées avec ce système et sont par la suite visionnées et sauvegardées. [S5] Les systèmes de vidéosurveillance sont composés de différents types de matériel en fonction des besoins de son utilisateur (les caméras de surveillance, l'écran de vidéosurveillance, l'alimentation des caméras de vidéosurveillance, les enregistreurs de vidéosurveillance, Câble de vidéosurveillance ou liaison sans fil, .. etc). [S1]



**Figure 1.1** Caméras de surveillance. [1]

Bien que ses premières utilisations remontent aux années 1950, la surveillance au moyen de systèmes de télévision en circuit fermé (TVCF) s'est vraiment développée à partir des années 1970, principalement au Royaume-Uni, pour lutter contre les activités terroristes[1]. À l'heure actuelle, le Royaume-Uni constitue d'ailleurs la société « la plus surveillée» avec un nombre estimé à plus de quatre millions de caméras de surveillance déployées sur son territoire. L'implantation de la vidéosurveillance s'est intensifiée au cours des années 1990, mais elle a connu un développement fulgurant suite aux attentats de septembre 2001 aux États-Unis, et puis ceux de Londres en 2005 [1].

### **✚ C'est quoi un système de vidéosurveillance ?**

Un Système de vidéosurveillance est un ensemble : [S33]

1. De caméras pour filmer les zones à surveiller



**Figure 1.2** : Caméra de surveillance. [S33]

2. De câbles ou des liaisons sans fil pour relier les caméras au réseau ou à l'enregistreur.



**Figure 1.3** : Câbles de liaison. [S33]

3. D'alimentations au niveau de chaque caméra ou déportées grâce aux câbles qui véhiculent le courant.



**Figure 1.4** : Alimentation électrique. [S33]

4. Un enregistreur pour sauvegarder les images.



**Figure 1.5** : Un enregistreur pour sauvegarder les images. [S33]

5. Un logiciel pour accéder aux images.



**Figure 1.6** : Un logiciel pour accéder aux images. [S33]

6. Un écran pour visualiser les images de l'enregistreur.

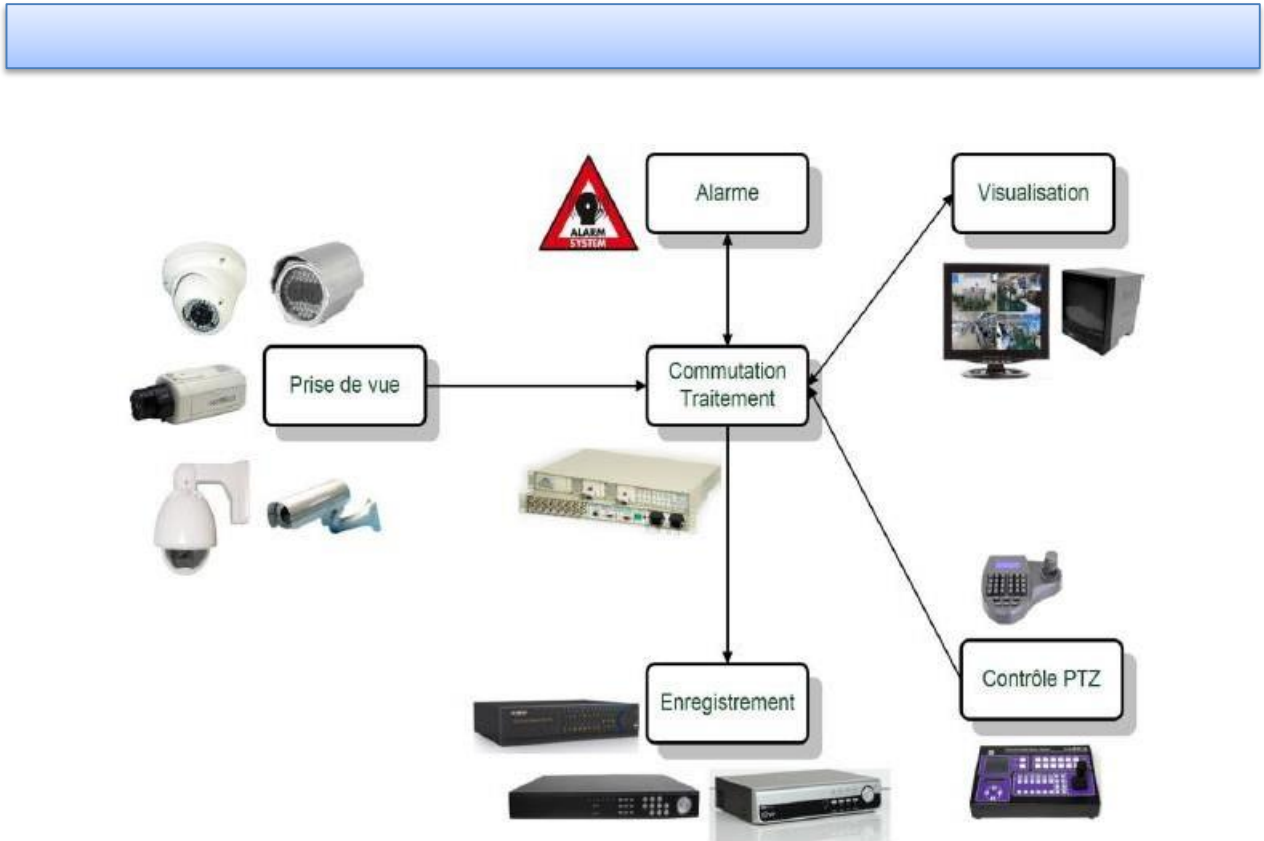


**Figure 1.7** : Un écran pour visualiser les images. [S33]

7. Une connexion internet pour accéder aux caméras à distance. [S33]



**Figure 1.8** : Modem pour connexion internet. [S33]



**Figure 1.9** : Schéma de principe de vidéosurveillance. [2]

Le schéma ci-dessus illustre les diverses fonctions qui peuvent être effectuées par un système de vidéosurveillance, qu'elles soient assurées par un seul ou plusieurs dispositifs. Il est possible qu'un DVR3, par exemple, soit en mesure de remplir à la fois les fonctions de commutation et d'enregistrement. Cependant, il convient de noter que ce schéma présente un modèle "exhaustif", ce qui signifie que certaines installations peuvent ne pas impliquer autant de fonctions [2]

## 1.2.Objectifs

L'objectif général d'un système de vidéo surveillance est de contribuer à la sécurité de biens et/ou de personnes. [3]

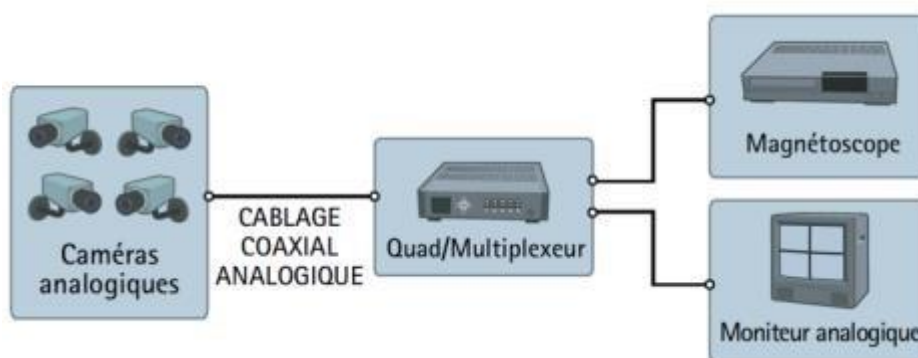
Cette contribution peut se focaliser sur diverses composantes :

- Prévention de la criminalité.
- Sécurité routière.
- Sécurité industrielle.
- Sûreté.
- Collecte de donnée.

### 1.3.L'évolution de la vidéosurveillance :

- **Système de Vidéosurveillance analogique avec magnétoscope :**

Dans un système de vidéosurveillance analogique on utilise des caméras vidéos analogiques avec des sorties coaxiales. Des magnétoscopes sont reliés à cette sortie pour enregistrer la vidéo et un moniteur pour visionner les images tel que ce schéma nous le décrit : [4]



**Figure 1.10 :** Système de vidéosurveillance analogique. [4]

Autrefois les images n'étaient pas numérisées, elles étaient sous forme de signaux électriques et leurs transmissions se faisaient avec des câbles coaxiaux comme celles de la télévision [4]

L'un des principaux avantages du système analogique c'est qu'il est simple à installer, et il ne requière pas de compétence informatique, ajoutant aussi que le coût d'installation d'un tel système est largement inférieur au système actuel. Tout de même il présente des inconvénients importants qui sont [5] :

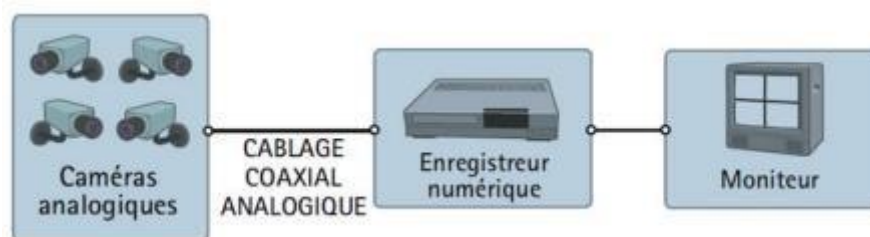
- La qualité de la vidéo est inférieure à celle des systèmes numériques.
- Il faut changer les cassettes fréquemment (trois jours ou plus).
- Nécessite un nettoyage et un entretien régulier des magnétoscopes.
- La qualité de la vidéo enregistrée se détériore avec le temps.
- Ne permet pas le visionnement à distance, comme sur les réseaux numériques.
- Ce sont des systèmes propriétaires [5].

- **Système de Vidéosurveillance analogique avec enregistreur numérique :**

Un système de vidéosurveillance analogique utilisant un enregistreur numérique (DVR) est un système analogique permettant l'enregistrement numérique des images. Avec un enregistreur numérique, l'enregistrement vidéo ne se fait plus sur bandes magnétiques mais sur des disques durs où les séquences sont numérisées et compressées de manière à emmagasiner chaque jour un maximum d'images. Les Premiers enregistreurs numériques disposaient d'un espace disque limité. La durée des enregistrements était donc assez restreinte, à moins de réduire la fréquence d'images. Grâce aux progrès récents dans ce domaine, l'espace disque ne pose plus Réellement problème. La plupart des enregistreurs numériques disposent en outre de Plusieurs entrées vidéo (en général 4, 9 ou 16), ce qui leur permet d'intégrer d'emblée Les fonctionnalités du quad ou des multiplexeurs. [S2]

Les enregistreurs numériques présentent les avantages suivants :

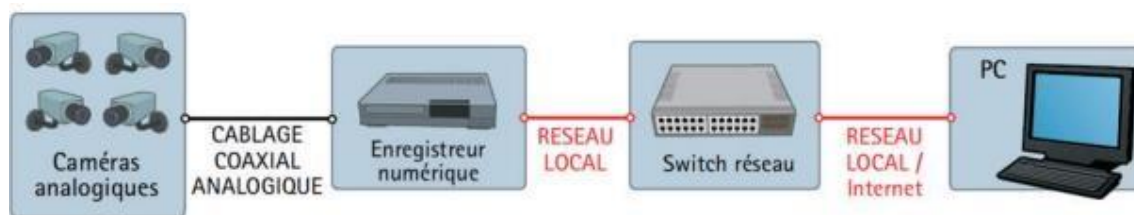
- Pas besoin de changer de cassette
- Qualité constante de l'image [S2]



**Figure 1.11 :** Système de vidéosurveillance analogique avec enregistreur numérique. [S2]

- **Système de vidéosurveillance analogique avec enregistreur numérique réseau :**

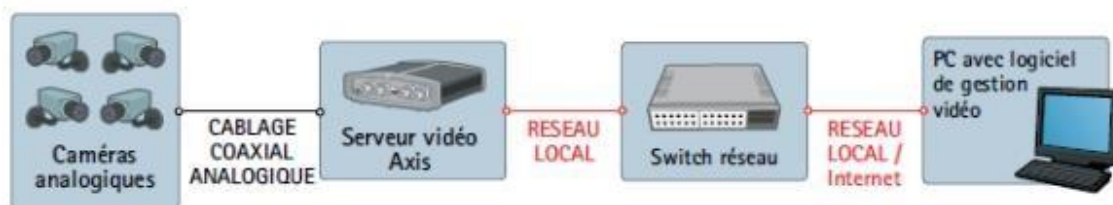
Le fonctionnement de ce système est similaire à celui décrit précédemment, à une différence près : les signaux électriques sont convertis en une séquence binaire (signal numérique) une fois qu'ils arrivent à l'enregistreur. Cette technologie innovante permet d'enregistrer les images sur des supports numériques tels que des CD-ROM ou des disques durs, augmentant considérablement la capacité de stockage. Certains enregistreurs numériques sont équipés d'un port RG45, ce qui leur permet de transmettre les images directement sur un réseau informatique pour être visualisées sur un ordinateur [4].



**Figure 1.12 :** Système de vidéosurveillance analogique avec enregistreur numérique réseau. [4].

- **Systèmes de vidéo sur IP avec serveurs vidéo :**

Un système de vidéo sur IP associé à un serveur vidéo comprend un serveur vidéo, un commutateur réseau et un PC équipé d'outils de gestion vidéo. La caméra analogique est branchée sur le serveur vidéo, lequel assure la numérisation et la compression des séquences vidéo. De son côté, le serveur vidéo est connecté sur le réseau qui transporte la vidéo vers un PC ou serveur via un commutateur réseau. La vidéo est alors enregistrée sur le disque dur du PC. Il s'agit alors d'un véritable système de vidéo sur IP [S2]



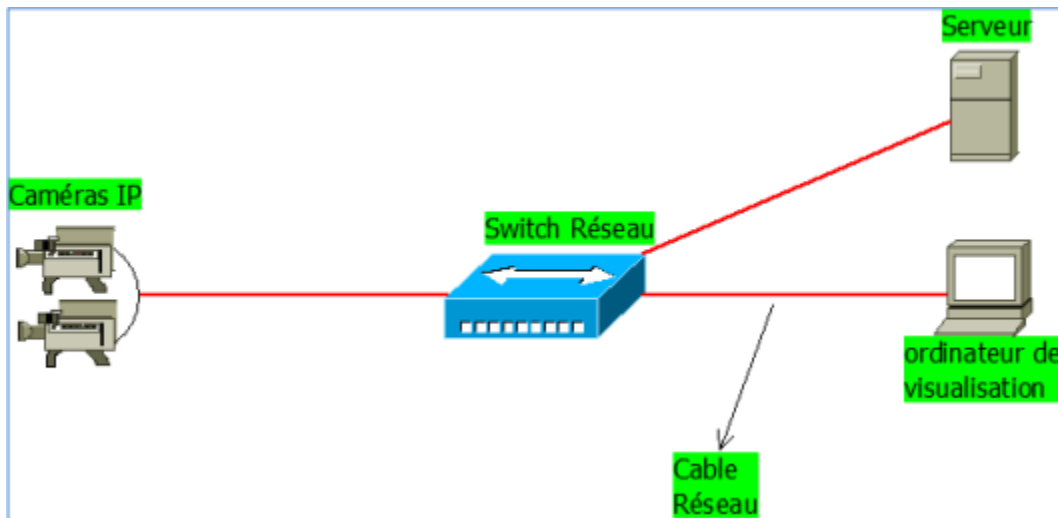
**Figure 1.13 :** Systèmes de vidéo sur IP avec serveurs vidéo. [S2]

Ce système utilisant les serveurs vidéo présentent les avantages suivants :

- Recours à un réseau standard et à un serveur informatique standard l'enregistrement et le traitement vidéo
- Système capable d'évoluer d'une caméra à la fois
- Possibilité d'enregistrement hors-site
- Système évolutif pouvant être élargi par ajout de nouvelles caméras réseau [S2].

- **Système Tout numérique et IP :**

On dit d'un système de vidéosurveillance est complètement IP (Figure1.14), lorsque tous ces composants sont numériques et que toutes les transmissions sont effectuées suivant le protocole IP, on peut trouver aussi des caméras analogiques mais qui devons être reliées à des encodeurs qui se chargent de convertir le signal analogique en signal numérique [5].



**Figure 1.14 :** Système tout IP. [5]

Le système de vidéosurveillance reposant sur les cameras réseau présente les avantages suivants :

- ✓ Caméras haute résolution (mégapixels)
- ✓ Qualité constante de l'image
- ✓ Fonction d'alimentation par câble Ethernet (Power over Ethernet) et réseau sans fil
- ✓ Fonctions panoramique/inclinaison/zoom, audio, entrée et sorties numériques sur IP
- ✓ Grandes flexibilité et évolutivité. [5]

## 1.4. Les types de vidéosurveillance

Il existe deux types principaux de mode de vidéosurveillance : la vidéosurveillance IP et la vidéosurveillance analogique, plus traditionnelle.

### a. Vidéosurveillance analogique : la plus répandue

Le système de la vidéo surveillance analogique est équipé de caméras analogiques dont le seul rôle est de capturer les images et les envoyer à un enregistreur à durée limitée (type magnétoscope) via un signal analogique. Ce système est composé d'une ou plusieurs caméras, d'un moniteur (ou téléviseur), d'un enregistreur et d'un câblage (le transfert d'images se fait via un câble dit coaxial) [3].

#### Avantages de vidéosurveillance analogique :

Ce type de vidéosurveillance conserve une position dominante sur le marché, ce qui présente plusieurs avantages significatifs [3] :

- Sa qualité d'image supérieure, sa fluidité et sa facilité d'utilisation sont ses principaux atouts.
- La diversité des caméras analogiques est également un avantage, avec différentes tailles, formes et applications disponibles.
- Les caméras analogiques sont généralement moins coûteuses que les caméras IP.

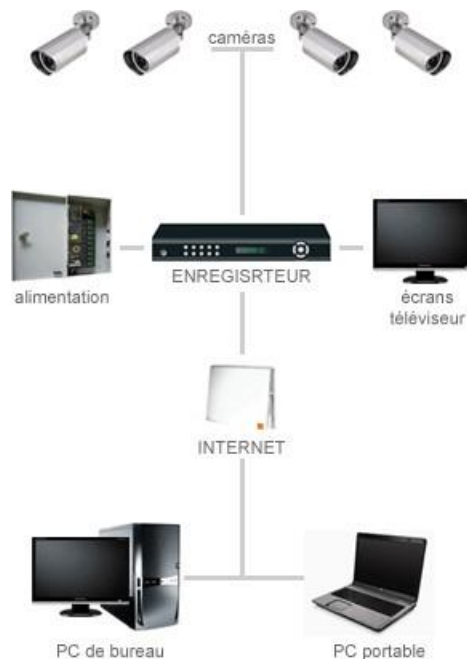
#### Inconvénients de vidéosurveillance analogique :

En comparaison avec la vidéosurveillance en réseau, le système analogique présente des limitations importantes en termes de fonctionnalités [3] :

- La durée de stockage limitée (par exemple, via des cassettes).
- Le format peu flexible des images.
- Le manque d'accès en temps réel via Internet (sauf si les caméras sont connectées à un réseau) et l'Internet mobile (PDA, GSM...).
- L'évolutivité limitée : il est difficile d'ajouter des caméras supplémentaires dans le temps en raison de problèmes de longueur de câble ou de travaux à réaliser.
- L'absence de gestion à distance : installation, maintenance, etc
- Dépasser par le temps. [39]

## Coût de la vidéosurveillance analogique :

Ce qui est le plus cher dans la vidéosurveillance analogique, c'est l'installation car celle-ci nécessite souvent des travaux lourds (passage de câbles, installation des caméras...). Ce système nécessite également la présence à plein temps d'un agent de sécurité pour surveiller les images diffusées sur le moniteur et pour effectuer les rondes. Les caméras analogiques sont souvent moins chères que les caméras numériques ou IP. [6]



*Figure 1.15* : Exemple d'installation analogique. [S33]

### b. Vidéosurveillance numérique : fonctionnement en réseau

La vidéosurveillance IP (protocole Internet) est venue compléter la vidéosurveillance analogique, qui reste encore la norme dans la plupart des installations. Elle utilise les mêmes composants que la vidéosurveillance analogique (caméras, moniteur, enregistreur, câbles), mais elle est connectée à un réseau informatique :

Les caméras IP sont reliées à un réseau IP (intranet, internet, LAN, CPL ou VPN) et connectées à un serveur central de vidéosurveillance. Elles capturent les images et les transmettent sur le réseau.

Le serveur de vidéosurveillance est le cœur du système et est équipé d'un logiciel de vidéosurveillance. Il récupère les images des caméras et les stocke sur un disque dur.

### **Avantages de la vidéosurveillance IP :**

La vidéosurveillance IP (Internet Protocole) offre de nombreux avantages par rapport à la vidéosurveillance analogique [3] :

- ✓ Elle peut exploiter toutes les fonctionnalités d'Internet.
- ✓ Grâce à la technologie réseau, elle peut s'intégrer facilement avec d'autres technologies IP telles que la visioconférence ou les systèmes de sécurité.
- ✓ Elle peut être installée sur un réseau informatique existant, ce qui réduit les coûts d'installation.
- ✓ Elle est compatible avec tous les types de câblage, y compris le Wifi.
- ✓ Le système est entièrement piloté par un ordinateur, éliminant le besoin de moniteur ou d'enregistreur dédié.
- ✓ Les caméras peuvent être facilement ajoutées ou déplacées en fonction des besoins.
- ✓ Les images peuvent être visualisées en temps réel et ajustées selon les besoins.
- ✓ Certaines caméras peuvent être commandées à distance.
- ✓ Le système peut être maintenu à distance par le service informatique, sans nécessité de contrat de maintenance.
- ✓ Les images peuvent être sécurisées par un mot de passe, et le système peut être intégré avec une alarme ou d'autres systèmes de sécurité.
- ✓ Il est possible d'ajouter des fonctionnalités intelligentes au système

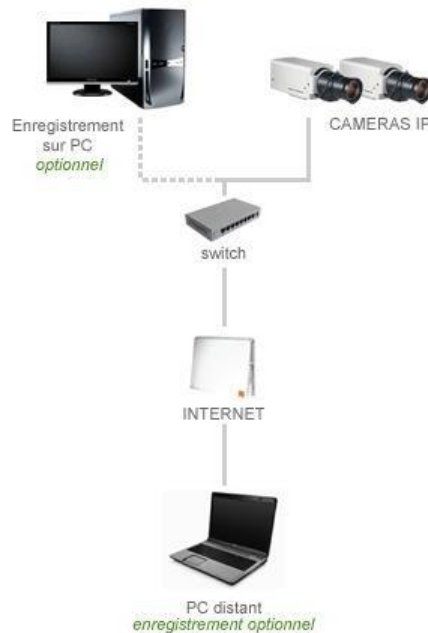
### **Inconvénients de la vidéosurveillance IP :**

- ✓ Choix de caméras plus limité en termes de taille et de forme par rapport à l'installation analogique.
- ✓ Installation plus complexe en raison des nombreux réglages nécessaires et de la nécessité de faire appel à un informaticien compétent.
- ✓ Les images capturées peuvent être de moins bonne qualité que celles d'une installation analogique, en particulier si le réseau n'est pas correctement configuré.
- ✓ L'installation du réseau peut être un aspect important à prendre en compte lors de la mise en place de la vidéosurveillance IP.

### **Coût de la vidéosurveillance IP : pas toujours économique**

Les systèmes de vidéosurveillance IP sont réputés moins chers que la vidéosurveillance analogique s'ils sont installés sur un réseau informatique déjà existant. Il n'est pas nécessaire de remplacer tout le matériel pour accéder à des fonctionnalités plus récentes : il suffit de changer de logiciel de vidéosurveillance (c'est le cœur du système). En

revanche, ils sont beaucoup moins économiques s'il faut installer le réseau. Ils nécessitent en outre les services permanents d'un informaticien ou d'un service informatique. [7]



**Figure 1.16** : Exemple d'installation IP. [S33]

### **Passer de l'analogique au numérique : pas difficile**

Si vous disposez d'un système de vidéosurveillance analogique et que vous souhaitez passer à un système numérique, il n'est pas nécessaire de remplacer tout votre matériel. Il existe des options pour profiter des avantages de la vidéosurveillance numérique, telles que l'utilisation d'une carte d'acquisition vidéo à installer sur votre ordinateur ou d'un enregistreur numérique (DVR), qui permettent de convertir les signaux analogiques en signaux numériques et de stocker les images numérisées sur un ordinateur. Grâce à ces enregistreurs numériques, vous pouvez facilement mettre en place un système hybride de vidéosurveillance, qui associe des caméras analogiques et des caméras numériques ou IP via un réseau unique.

### **1.5. Domaines d'application de vidéosurveillance**

La vidéo-surveillance, est un champ de plusieurs d'applications [PPTD, 2015] :

- L'industrie : Salle de contrôle, Zones de stockage.
- Le transport privé ou collectif.
- Le commerce et la distribution.

- Les administrations et les services publics : les banques.
- Les lieux publics : Cabinets médicaux et cliniques.
- L'enseignement : Les écoles et les établissements d'enseignement.
- Les loisirs : parc d'attraction, cinéma, les clubs.

## **1.6. Ses buts**

Il existe diverses raisons qui justifient l'installation de systèmes de vidéosurveillance. Outre la sécurité publique et la protection des biens, d'autres éléments jouent un rôle essentiel. En Angleterre, par exemple, les attentats de juillet 2005 ont contribué à augmenter le nombre de caméras. Bien que cette menace ait toujours été présente, elle n'avait jamais réellement créé un sentiment d'insécurité. Cependant, les attentats du 11 septembre 2001 ont changé la donne, faisant prendre conscience aux gens que personne n'était à l'abri.

Néanmoins, l'installation de la vidéosurveillance ne peut être uniquement justifiée par l'insécurité croissante ou la protection des biens. Il existe d'autres raisons moins connues du grand public. Par exemple, la vidéosurveillance améliore la gestion des incidents et permet une intervention plus rapide et efficace. Elle peut être utile dans la prévention du suicide ou lors d'accidents sur la voie publique, contribuant indirectement à maintenir les primes d'assurance à un niveau raisonnable. De plus, la surveillance des axes routiers permet d'informer les automobilistes en temps réel sur les conditions de circulation.

Les enregistrements fournis par les caméras de surveillance ont également permis de résoudre certaines affaires criminelles. Par exemple, après les attentats du métro de Londres en juillet 2005, les enregistrements ont été utilisés pour identifier les poseurs de bombes, bien qu'ils n'aient pas été considérés comme indispensables. Cependant, les études indépendantes menées en France et à l'étranger n'ont pas pu démontrer de manière concluante si la vidéosurveillance prévient ou réduit les crimes. Le gouvernement britannique a lui-même estimé qu'il était difficile d'évaluer les effets bénéfiques, bien que Scotland Yard ait affirmé en 2008 que les 500 000 caméras de vidéosurveillance à Londres n'avaient permis d'élucider que 3 % des vols dans la rue. [S3]

## **1.7. Les différents types de caméra de vidéosurveillance**

Il existe différents types de caméras de surveillance pour une installation de vidéosurveillance à domicile, dans une entreprise ou un commerce. Elles se distinguent par leurs fonctionnalités mais aussi par leur niveau de gamme. [S4]



*Figure 1.17* : les types de caméra de surveillance.

## 1.8. Les Avantages du système de la vidéo-surveillance

- Il agit comme un effet dissuasif sur la criminalité. Voyant un appareil photo ou un avertissement d'affichage des caméras, peut être déroutant à un soi-criminel.
- Un autre avantage d'avoir un système de vidéo-surveillance est la capacité à appréhender un suspect quand un crime se produit.
- L'amélioration de la productivité en milieu de travail.
- Lot de cas de harcèlement au travail ne sont pas signalés. Installation du système de surveillance vidéo aide à tout suivi par les deux enregistrements vidéo et audio. [DM, 2012]

## 1.9. Les problèmes liés à la vidéosurveillance

Le plus important des problèmes est lié à l'atteinte de la vie privée puisque nous sommes constamment surveillés. Les familles ne se sentent pas à l'aise en sachant qu'une caméra les surveille en permanence. Cela pose une restriction des libertés.

L'autre problème est celui de la mise en place de caméra qui est très coûteuse puisque si l'on souhaite un système sophistiqué de vidéo surveillance cela représentera un investissement considérable et à cela s'ajoutera la maintenance du système qui devra être régulière et coutera beaucoup trop chère à long terme. [S5]

Sans oublier les vigiles qui travaillent derrière les écrans de surveillance comme le montre la figure suivante :



*Figure 1.18* : Vigiles de surveillance.

De plus, on ne peut pas surveiller toutes les rues d'une ville puisque pour le faire, il faudrait un nombre important de caméras. Enfin, cela enlève de l'emploi des vigiles dans les rues. [S5]

### Conclusion

Dans ce chapitre ont à vue c'est quoi un système de surveillance, c'est quoi sont objectif et les différents type de vidéosurveillance et pour finir les problèmes liés à la vidéosurveillance.

*Chapitre 2 :*

*La représentation de l'image et*

*de la vidéo*

## Chapitre 02

# La représentation de l'image et de la vidéo

## Sommaire

<a href="#">Chapitre 02</a>	20
<a href="#">Introduction</a>	20
<a href="#">2.1. L'image</a>	21
<a href="#">2.1.1. Définition de l'image</a>	21
<a href="#">2.1.2. Naissance de l'image numérique</a>	21
<a href="#">2.1.3. Représentation d'une image numérique</a>	22
<a href="#">2.1.4. Les types d'image numérique</a>	22
<a href="#">2.1.5. Les caractéristiques d'image</a>	26
<a href="#">2.1.6. Qualité de l'image numérique</a>	28
<a href="#">2.1.7. Les différents formats d'images</a>	28
<a href="#">2.1.8. Frame</a>	30
<a href="#">2.1.9. Objet</a>	30
<a href="#">2.1.10. La détection d'objet</a>	30
<a href="#">2.2. Vidéo</a>	31
<a href="#">2.2.1. Définition</a>	31
<a href="#">2.2.2. Les fondamentaux physiques et techniques</a>	31
<a href="#">2.2.3. Types de vidéo</a>	34
<a href="#">2.2.4. Composition des fichiers vidéo</a>	36
<a href="#">2.2.5. Les formats de fichiers vidéo</a>	37
<a href="#">2.2.6. Hiérarchie des données vidéo</a>	38
<a href="#">2.2.7. Représentation d'une séquence vidéo</a>	40
<a href="#">2.2.8. Les paramètres clés d'une vidéo</a>	40
<a href="#">Conclusion</a>	41

### Introduction

De nos jours, l'image est un pilier de notre société, elle est utilisée dans différents milieux comme : la presse, le tourisme, la communication, le commercial....

Dans cette chapitre là ; nous allons présentons une vue globale sur l'image et la vidéo.

## 2.1.L'image

### 2.1.1. Définition de l'image

Une image est définie comme étant la représentation exacte ou analogique d'une scène par la peinture, la sculpture, le dessin, photographie, film... Elle représente un ensemble structuré d'informations qui après affichage sur un support (écran, toile, papier, ...) ont une signification pour l'œil humain. Si  $x$  et  $y$  sont les coordonnées spatiales d'un point de l'image,  $I(x, y)$  est une fonction de l'intensité lumineuse et de la couleur [8] Autrement dit ; l'image numérique est l'image dont la surface est divisée en élément de taille fixe appelée cellules ou pixels. Une image peut être représentée sous quatre types.

### 2.1.2. Naissance de l'image numérique

Les premières applications des images numériques remontent aux années 1920 et ne concernent que les premiers pas en transfert d'images et en correction- amélioration d'images.

- ❖ Transmission d'images de journaux par câble sous-marin entre Londres et New York en 1921 (une image transmise en trois heures) par codage au départ et décodage à l'arrivée par une imprimante spécialisée.

Chaque niveau de gris étant codé par un caractère particulier (déclenchant une impression plus ou moins dense en surface encrée).

1921 : 5 niveaux de gris.

1929 : 15 niveaux de gris

- ❖ En 1929 premier brevet déposé sur la reconnaissance optique de caractères.
- ❖ Les potentialités du traitement numérique des images digitales pour le transfert et l'amélioration des images sont apparues avec le développement des grands ordinateurs et surtout avec les nécessités des programmes de recherche spatiale :

Jet Propulsion Laboratory en 1964 : Images de la lune obtenues par RANGER7 avec la correction numérique des distorsions de la caméra par ordinateur.

- ❖ Puis l'explosion des applications dans tous les domaines..., mais toujours pas de méthodologie universelle pour automatiser par ordinateur la vision humaine... d'où de
- ❖ nombreuses recherches dans ce domaine en pleine expansion [9]

### 2.1.3. Représentation d'une image numérique

La représentation d'une image numérique peut être effectuée selon plusieurs caractéristiques, telles que la couleur ou les coefficients de Fourier ou d'ondelette. En informatique, le pixel est l'élément visuel le plus petit pouvant être affiché à l'écran. Ainsi, une image est représentée sous forme d'un tableau de pixels de différentes couleurs ou niveaux de gris. [10]

### 2.1.4. Les types d'image numérique

#### 2.1.4.1. Image binaire (noir ou blanc)

Image binaire contient avec seulement 2 valeurs possibles : chaque pixel a valeur que 0 pour le noir ou 1 pour le blanc [S6]

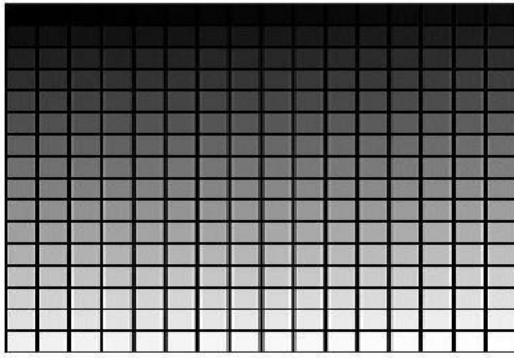
L'expression de la valeur d'image binaire  $N_g = 2$  et la relation sur les niveaux de gris devient :  $p(i,j) = 0$  ou  $p(i,j) = 1$ .

1	1	1	1	1	1	1	1	1	1
1	1	1	0	0	0	0	1	1	1
1	1	0	1	1	1	1	0	1	1
1	0	1	1	1	1	1	1	0	1
1	0	1	0	1	1	0	1	0	1
1	0	1	1	1	1	1	1	0	1
1	0	1	0	1	1	0	1	0	1
1	0	1	1	0	0	1	1	0	1
1	1	0	1	1	1	1	0	1	1
1	1	1	0	0	0	0	1	1	1
1	1	1	1	1	1	1	1	1	1

Figure 2.1 : Image binaire.

#### 2.1.4.2. Image de niveaux de gris

Une image de niveaux de gris autorise un dégradé de gris entre le noir et le blanc. En général, on code le niveau de gris sur un octet (8 bits) soit 256 nuances de dégradé. Le niveau de gris est codé sur un bit (BinarydigIT).L'expression de la valeur du niveau de gris avec  $N_g=256$  devient :  $p(i,j) \in [0, 255]$  .[11]



**Figure 2.2 :** Image en niveaux de gris (Nuance de 256 gris Exemple de photo possible en8bpp).

[18]

### ➤ Codage d'une image à niveaux de gris

Si on code chaque pixel sur 2 bits on aura 4 possibilités (noir, gris foncé, gris clair, blanc). L'image codée sera très peu nuancée.

En général on code chaque pixel sur 8 bits = 1 octet. On a alors 256 possibilités (on dit 256 niveaux de gris).

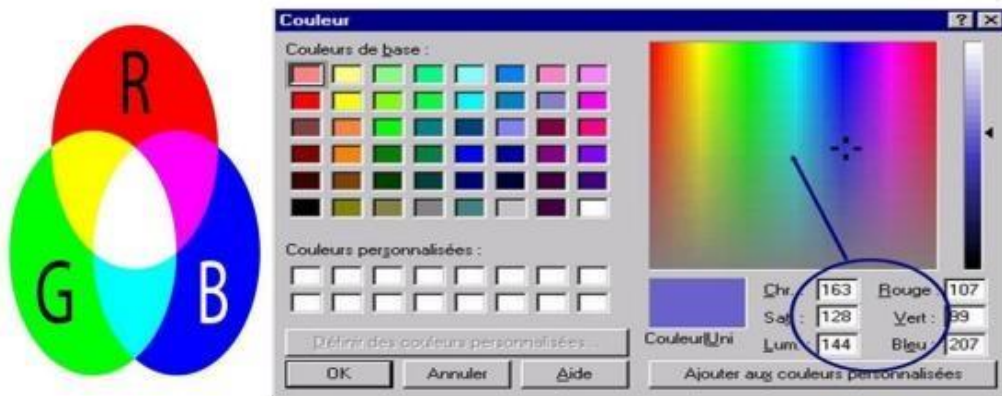
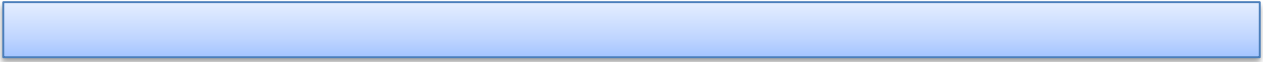
L'image de 10 000 pixels codée occupe alors 10 000 octets en mémoire.

Exemple d'image en 72 pixels par pouce (environ 30 pixels par cm), codée en 256niveaux de gris.

Cette image de 303 x 303 pixels occupe 303 x 303 = 91809 octets puisque chaque pixel occupe 1 octet en mémoire. [S7]

### 2.1.4.3. Image couleur (RVB)

Dans image en couleurs chaque pixel possède une couleur décrite par la quantité de ces 3 composants : rouge(R), vert(V) et bleu (B), chacune de ces couleurs est codée sur l'intervalle [0, 255].) La combinaison de ces trois couleurs donne un point lumineux (un pixel) d'une certaine couleur. Donc Le système RVB est une des façons de décrire une couleur en informatique. Par exemple : Le triplet {255, 255, 255} donnera du blanc, {255, 0, 0} un rouge pur, {100, 100, 100} un gris, etc. Le premier nombre donne la composante rouge, le deuxième la composante verte et le dernier la composante bleu [S6]



*Figure 2.3 : Image en couleurs. [S6]*

#### **2.1.4.3.1. Codage d'une image en couleur 8 bits**

Dans ce cas on attache une palette de 256 couleurs à l'image.

Ces 256 couleurs sont choisies parmi les 16 millions de couleurs de la palette RVB. Pour chaque image le programme recherche les 256 couleurs les plus pertinentes.

Chaque code (de 0 à 255) désigne une couleur.

L'image occupe 3 fois moins de place en mémoire qu'avec un codage 24 bits. L'image est moins nuancée : sa qualité est bonne mais moindre. [S7]

#### **2.1.4.3.2. Codage d'une image en couleurs 24 bits**

Il existe plusieurs modes de codage de la couleur. Le plus utilisé est le codage Rouge, Vert, Bleu (RVB). Chaque couleur est codée sur 1 octet = 8 bits. Chaque pixel sur 3 octets c'est à dire 24 bits : le rouge de 0 à 255, le vert de 0 à 255, le Bleu de 0 à 255. Le principe repose sur la synthèse additive des couleurs : on peut obtenir une couleur quelconque par addition de ces 3 couleurs primaires en proportions convenables.

On obtient ainsi  $256 \times 256 \times 256 = 16777216$  (plus de 16 millions de couleurs différentes). [S7]

## 2.1.5. Les caractéristiques d'image

### 2.1.5.1. Pixel

Le mot pixel provient d'une abréviation de l'expression britannique « PICture Element ». Il est le plus petit élément constitutif d'une image numérique [S6]

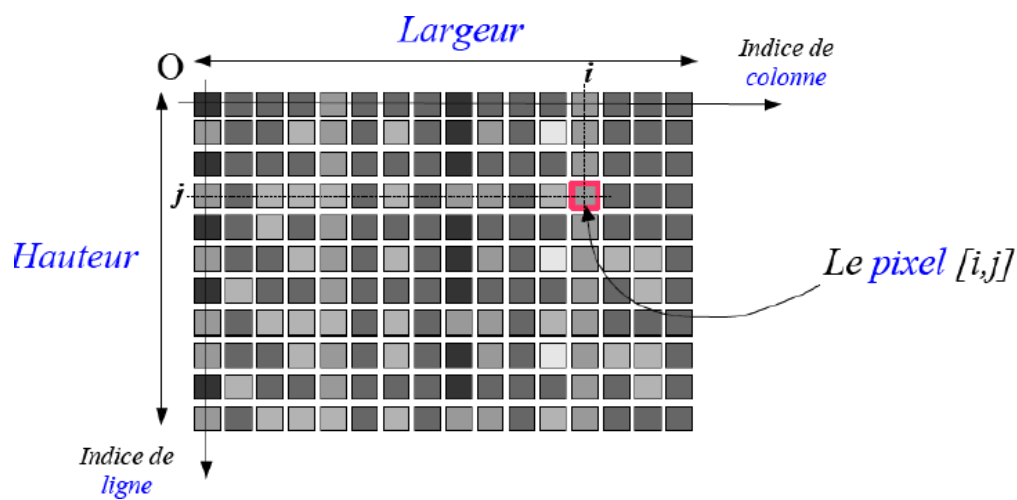
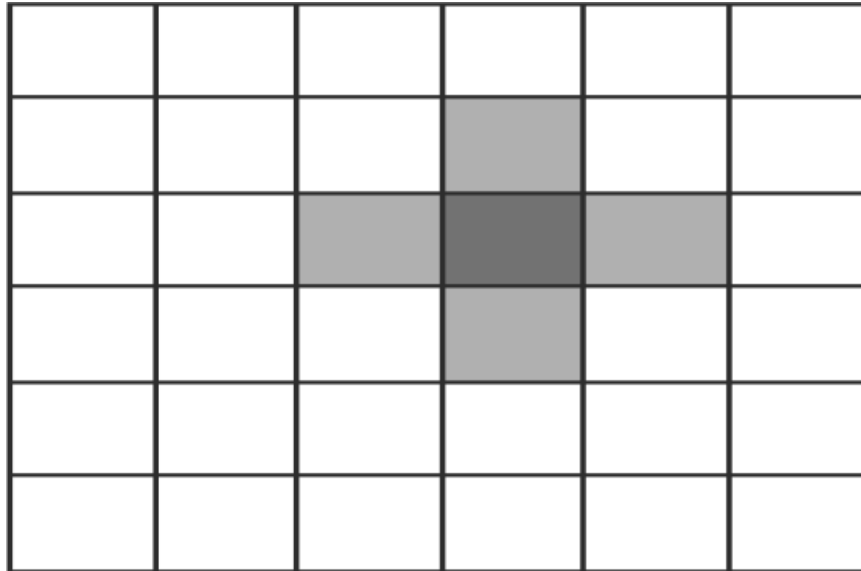


Figure 2.4 : pixel dans une image numérique. [14]

### 2.1.5.2. Voisinage

C'est une notion fondamentale en traitement d'image, c'est l'ensemble des pixels qui se situent autour d'un pixel. [8]



**Figure 2.5 :** Voisinage à 4. [8]

### 2.1.5.3. Poids de l'image

C'est la taille de l'image. Etant donné que cette dernière est représentée sous forme d'une matrice dont les valeurs représentent l'intensité (pixels), le nombre de colonne (W) multiplié par le nombre de ligne (H) donne le nombre total de pixels dans l'image.

Pour une image de 640x480 en couleur :

- ❖ Nombre de pixel =  $640 \times 480 = 307200$
- ❖ Poids de chaque pixel = 3 octets
- ❖ Le poids de l'image =  $307200 \times 3 = 921600$  octets = 900 Ko. [S8]

### 2.1.5.4. Bruit

Un bruit (parasite) dans une image est considéré comme un phénomène de brusque variation de l'intensité d'un pixel par rapport à ses voisins, il provient de l'éclairage des dispositifs optiques et électroniques du capteur. [12]

### 2.1.5.5. Transparence

La transparence est une caractéristique définissant le niveau d'opacité des éléments de l'image, c'est la possibilité de voir à travers l'image des éléments graphiques situés derrière celle-ci. [S9]

### 2.1.5.6. Luminance

C'est le degré de luminosité des points de l'image. Elle est définie aussi comme étant le quotient de l'intensité lumineuse d'une surface par l'aire apparente de cette surface, le mot luminance est

substitué au mot brillance, qui correspond à l'éclat d'un objet. Une bonne luminance se caractérise par :

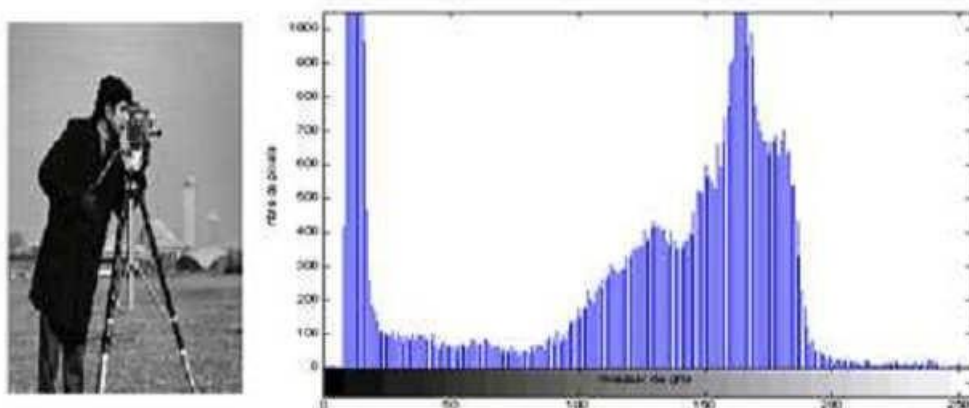
- ✓ Des images lumineuses (brillantes) ;
- ✓ Un bon contraste : il faut éviter les images où la gamme de contraste tend vers le blanc ou le noir ; ces images entraînent des pertes de détails dans les zones sombres ou lumineuses.
- ✓ L'absence de parasites [13]

### 2.1.5.7. Contraste

Le contraste, une caractéristique inhérente d'une image, se réfère à la distinction et à la quantification des différences entre les zones claires et sombres d'une image, permettant de distinguer les teintes claires des teintes foncées. En photographie, le contraste est défini comme la disparité entre la densité la plus élevée et la plus faible présente dans une image. Maîtriser le contraste constitue un élément essentiel de la pratique photographique. Le contraste final d'une image dépend simultanément du sujet capturé, de la nature même de la scène et du traitement appliqué sur le négatif et le positif. [S10]

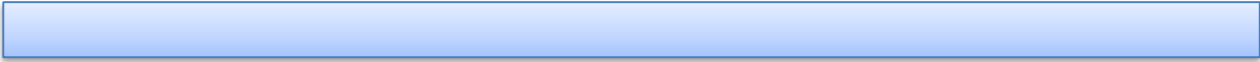
### 2.1.5.8. Histogramme

Un histogramme est un graphique statistique permettant de représenter la distribution des intensités des pixels d'une image. Il fournit diverses informations comme les statistiques d'ordre (moyenne, variance,...), l'entropie, et peut permettre d'isoler des objets. [S11]



*Figure 2.6* : Exemple d'image en niveaux de gris. [S12]

Pour les images en couleur plusieurs histogrammes sont nécessaires.



Par exemple pour une image codée en RVB :

- un histogramme représentant la distribution de la luminance
- trois histogrammes représentant respectivement la distribution des valeurs respectives des composantes rouges, bleues et vertes. [S13]

### **2.1.6. Qualité de l'image numérique**

Elle dépend, d'une part, de la qualité des images d'origine et, d'autre part, des moyens mis en œuvre pour convertir un signal analogique en signal numérique. Elle dépend aussi de :

- La qualité des périphériques de numérisation de l'image, du nombre de niveaux de gris ou de couleurs enregistrées, etc.
- La qualité de l'affichage à l'écran : définition de l'écran, nombre de teintes disponibles. [14]

Les critères d'appréciation de la qualité d'une image, tels que cités succinctement ci-dessus, dépendent largement de la structure même de l'image réaliste ou conceptuelle et de son mode de représentation (matricielle ou vectorielle). [15]

### **2.1.7. Les différents formats d'images**

On peut classer les images en deux formats :

#### **2.1.7.1. Format vectorielle**

Dans une image vectorielle les données sont représentées par des formes géométriques simples qui sont décrites d'un point de vue mathématique.

Par exemple, un cercle est décrit par une information du type (cercle, position du centre, rayon). Ces images sont essentiellement utilisées pour réaliser des schémas ou des plans [S14]

- **Avantage :**

- Les fichiers sont légers et occupent peu de place en mémoire [35]
- Adapté aux objets mathématiques et leur extension (voiture, avion...) [35]

- **Inconvénients :**

- Inutilisables pour des images complexes, des photographies [35]
- Un fichier vectoriel est plus fragile qu'un fichier matriciel dont l'en-tête surtout doit être intacte. La moindre dégradation de l'information est souvent irréparable [36]

### 2.1.7.2. Format matricielle

Une image matricielle est formée d'un tableau de points ou pixels. Plus la densité des points sont élevée, plus le nombre d'informations est grand et plus la résolution de l'image est élevée. Corrélativement la place occupée en mémoire et la durée de traitement seront d'autant plus grandes.

Les images vues sur un écran de télévision ou une photographie sont des images matricielles [S14].

On obtient également des images matricielles à l'aide d'un appareil photo numérique, d'une caméra vidéo numérique ou d'un scanner. Parmi ces formats on peut citer :

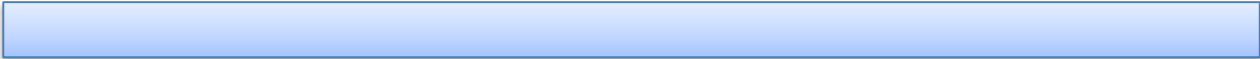
- **BMP (BitMap)** : Le format BMP est le format par défaut du logiciel Windows. C'est un format matriciel. Les images ne sont pas compressées. Son logiciel d'origine.
- **Le format EPS (Encapsulated PostScript)** : matriciel n'est pas très différent du EPS vectoriel. En fait seules les données contenues dans le fichier sont différentes. Ainsi un logiciel de retouche de photos tel que Photoshop permet l'importation, la modification et l'exportation de fichiers en format EPS.
- **GIF (Graphical Interchange Format)** : Le format GIF est un format qui a ouvert la voie à l'image sur le World Wide Web. C'est un format de compression qui n'accepte que les images en couleurs indexés codé sur 8 bits, C'est un format qui perd beaucoup de son marché suite à une bataille juridique concernant les droits d'utilisation sur Internet.
- **JPEG (Joint Photographique Experts Group)** : Les images JPEG sont des Images de 24 bits. C'est-à-dire qu'elles peuvent afficher un spectre de 16 millions de couleurs. C'est la meilleure qualité d'images disponible. [S15]

#### • **Avantage :**

- Simplicité de stockage en mémoire, puisqu'il suffit de coder la succession des valeurs de la matrice. [37]
- Adaptable aux images complexes [35]

#### • **Inconvénients :**

- La création d'une image à la main est difficile et il est conseillé de passer par un périphérique de numérisation : scanner, appareil photo-numérique... mais les retouches sont délicates [36]

- 
- Les images bitmap sont « lourdes » : les fichiers, lorsque l'on traite des images en haute définition, ont des tailles qui varient entre 10 et 30 Mo par image. Elles sont donc encombrantes, difficiles à faire passer sur le réseau, etc. [35] [38]

### **2.1.8. Frame**

Image composants la vidéo, les photogrammes [S16]

### **2.1.9. Objet**

Les objets physiques sont divisés en deux types : les objets de contexte et les objets mobiles.

#### **2.1.9.1. Les objets de contexte :**

Les objets physiques sont généralement immobiles, tels que les murs qui restent en place. Cependant, certains objets contextuels comme les chaises ou les portes peuvent être en mouvement, mais leurs déplacements peuvent être anticipés en se basant sur les informations environnantes. [16]

#### **2.1.9.2. Les objets mobiles :**

Sont des objets physiques qui peuvent être perçus dans les scènes par leurs mouvements [16]. Il est cependant difficile de prédire leurs mouvements. P. ex. les personnes, les véhicules.

Les objets dans les vidéos sont des objets mobiles. Ils sont habituellement détectés et suivis dans plusieurs frames.

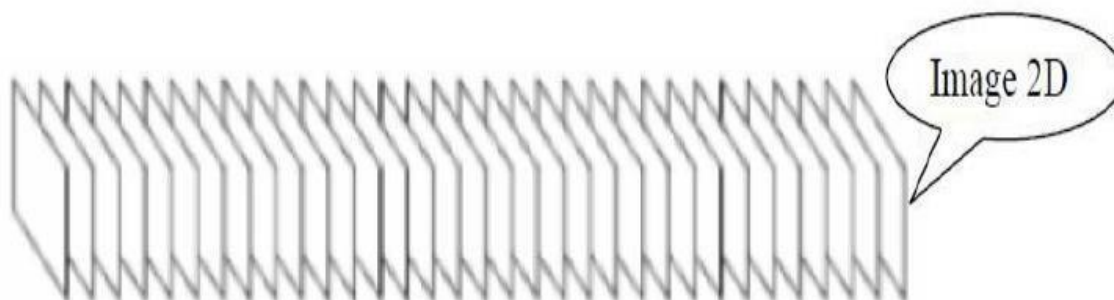
### **2.1.10. La détection d'objet :**

La détection d'objet englobe l'identification des éléments et la perception des formes dans le contexte d'une séquence vidéo. Un dispositif de détection d'objet est indispensable pour le suivi, que ce soit à chaque image ou lors de l'apparition initiale de l'objet dans la vidéo. [17]

## 2.2. Vidéo

### 2.2.1. Définition

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). [18]



*Figure 2.7* : Représentation d'un flux de vidéo.

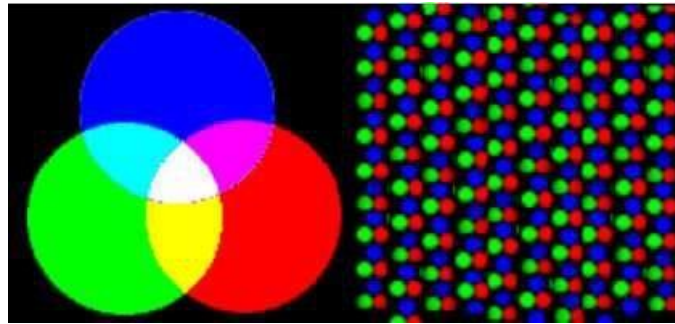
### 2.2.2. Les fondamentaux physiques et techniques

#### 2.2.2.1. La vision

Lorsque la lumière pénètre dans l'œil, elle se divise en deux éléments essentiels : l'intensité lumineuse, mesurée par la luminance (Y), et la couleur, mesurée par la chrominance (C). Les bâtonnets, présents dans l'œil, captent et réagissent principalement à l'intensité de la lumière, tandis que les cônes se chargent de détecter les différentes nuances de couleur. Ces informations sont ensuite transmises au cerveau qui les synthétise pour créer une image cohérente. Il est intéressant de noter que l'œil humain accorde une plus grande sensibilité à l'intensité lumineuse (Y) qu'à la couleur (C). [S17]

#### 2.2.2.2. La couleur

La vidéo utilise une approche connue sous le nom de synthèse additive de la couleur. Dans ce système colorimétrique, le rouge, le vert et le bleu (RVB ou RGB en anglais) sont considérés comme les trois couleurs primaires. En mélangeant ces trois couleurs, il est possible de reproduire l'ensemble du spectre visible par l'œil humain. Ainsi, l'écran de visualisation vidéo est composé d'une série de triades rouge-vert-bleu. Lorsque toutes ces triades sont activées simultanément, elles créent l'image affichée à l'écran [S17]

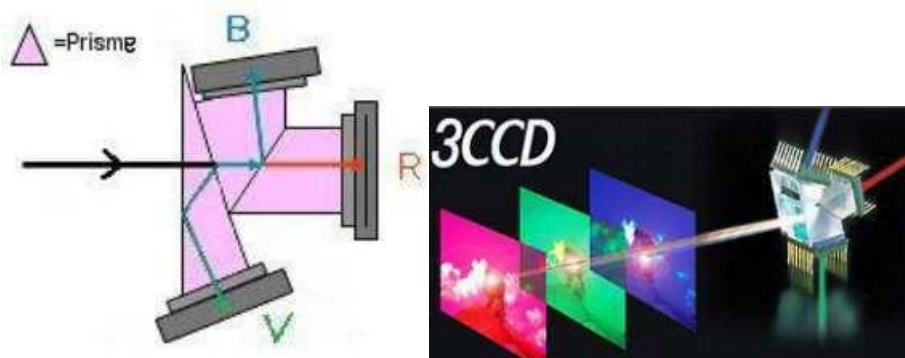


*Figure 2.8* : codage de couleur en RVB. [S18]

### 2.2.2.3. Le signal

En vidéo, c'est la caméra qui transforme l'information lumineuse (photons) en signal électrique (électrons). En vidéo analogique, l'intensité de ce signal électrique varie de façon continue. Le processus de transformation de l'information lumineuse en signal électrique est le suivant :

1. L'élément optique de la caméra, l'objectif, sépare la lumière en trois composants le rouge, le vert, le bleu. Cette opération est réalisée en faisant passer le flux lumineux par une succession de filtres dichroïques réfléchissant certaines couleurs et en laissant passer d'autres.
2. En vidéo professionnelle, les trois images sont projetée sur trois capteurs photo sensibles distincts formés chacun de centaines de milliers de points généralement entre 400 000 et 700 000. Ces capteurs sont nommés CCD (charged coupled device) ou dispositifs à transfert de charge. Les caméras domestiques ne sont généralement équipées que d'un seul capteur CCD. [S17]



*Figure 2.9* : Schéma de principe d'un Capture Tri-CCD (chargedcoupledevice). [S19]

3. Pour chacun des points de chacun des capteurs, l'énergie lumineuse sera transformée en énergie électrique. Ainsi, à la sortie des capteurs, trois signaux électriques d'intensité variable, un signal pour chacune des trois composantes. La lumière blanche est formée par la somme des trois composantes RVB. Toutefois, elle n'est pas composée des trois couleurs primaires en quantité égale. Les propositions sont les suivantes :

$$100\% Y = 29,9\%(R) + 58,7\%(V) + 11,4\%(B)$$

4. Une information supplémentaire devra être ajoutée à ces signaux. Comme nous l'avons dit, une valeur doit être établie pour chacun des points des capteurs. C'est cette valeur pour chacun des points de chacun des pixels qui permettra de recomposer l'image sur un écran de visualisation, chacun des points de la surface d'affichage pouvant alors prendre la même valeur qu'au moment de la capture par la caméra. La restitution de l'information pour chacun des pixels se fait par balayage ligne par ligne de la surface de l'écran de reproduction, les valeurs enregistrées pour chacun des pixels au moment de la capture étant transférées aux pixels de la surface de reproduction. Pour synchroniser ces balayages, un signal de fin de ligne est ajouté pour chacune des lignes. Un signal de fin d'image est également inséré. [S17]

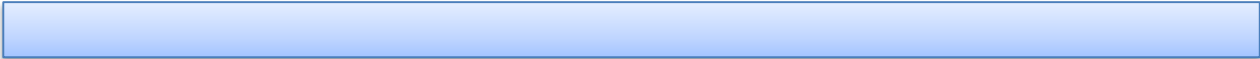
#### ➤ Les caractéristiques d'un signal vidéo

Chaque fichier vidéo a des attributs qui décrivent ce qui constitue le signal vidéo. Ces caractéristiques comprennent :

- **Châssis de taille** : Il s'agit de la dimension de pixel de l'image
- **Le ratio d'aspect** : C'est le rapport entre la largeur et la hauteur
- **Vitesse de défilement** : C'est la vitesse à laquelle les images sont capturées et destinées à la lecture.
- **Débit** : Le taux de débit ou de données est la quantité de données utilisées pour décrire la partie audio ou vidéo du fichier. Il est généralement mesuré en unités par seconde et peut être en kilo-octets, méga-octets ou giga-octets par seconde. En général, plus la vitesse de transmission, meilleure est la qualité.
- **Le taux d'échantillonnage audio** : C'est à quelle fréquence le signal audio est échantillonné lors de la conversion d'une source analogique à un fichier numérique. [19]

#### 2.2.2.4. Affichage vidéo

La technique de composition d'image par balayage présente une particularité : contrairement au cinéma, où l'image s'affiche instantanément, dans le cas de la vidéo, elle est affichée ligne par ligne,



en commençant par le coin supérieur gauche et se terminant par le coin inférieur droit. Afin

d'assurer une uniformité de luminosité lors de cet affichage, c'est-à-dire évité que l'intensité lumineuse des pixels en haut de l'écran ne diminue pendant que ceux du bas sont activés, une méthode spécifique a été développée : l'entrelacement.

Cette technique consiste à effectuer un premier balayage des lignes impaires, suivi d'un deuxième balayage des lignes paires. Ainsi, il faut deux passages pour reconstituer une image complète. La fréquence d'entrelacement est basée sur la fréquence du réseau électrique. Par exemple, en Amérique du Nord, où la fréquence du réseau est de 60 Hz, le système de balayage fonctionne à 60 demi-images, ce qui équivaut à 30 images par seconde. Plus précisément, le taux de rafraîchissement est de 29,97 images par seconde.

Aux États-Unis, la norme télévisuelle est le NTSC (National Television Standard Committee).

Une image NTSC est composée de 525 lignes, dont seules 486 sont utilisées pour afficher l'image. En Europe, les deux principales normes sont le PAL (Phase Alternation by Line) et le SECAM (Séquentiel Couleur à Mémoire). Une image PAL ou SECAM est constituée de 625 lignes, dont 576 sont utilisées pour l'image. Étant donné que la fréquence des réseaux électriques européens est de 50 Hz, le système de balayage est de 50 demi-images, ce qui équivaut à 25 images par seconde [S17]

### **2.2.2.5. Le son**

Une attention particulière sera accordée à la gestion du son. Le traitement audio sera séparé de la vidéo, garantissant ainsi des signaux distincts. De plus, lors de l'enregistrement sur bande vidéo, le son sera enregistré sur des pistes séparées. [S17]

## **2.2.3. Types de vidéo**

Le signal vidéo est le signal qui permet de transporter une séquence d'images de la source à un dispositif d'affichage sous forme électrique. Selon la façon dont les signaux sont traités on peut distinguer les deux modes : [20]

### **2.2.3.1. La vidéo analogique**

Un signal analogique est défini comme un signal électrique dont l'intensité varie de manière continue dans le temps. Malheureusement, la qualité du signal obtenu est souvent altérée par le bruit ajouté lors du traitement, ce qui peut réduire sa qualité globale

### **2.2.3.2. La vidéo numérique**

Le signal numérique est caractérisé par une série de valeurs qui représentent l'information à transmettre, telles que les valeurs 0 et 1. Contrairement au signal analogique, qui est un signal électrique continu, le signal numérique offre l'avantage de faciliter la distinction entre l'information et le bruit, ce qui améliore la qualité du signal final.

#### **i) La fréquence d'échantillonnage**

Le théorème de Shannon, également connu sous le nom de théorème d'échantillonnage, énonce qu'un signal dont le spectre est à support borné ne subit aucune altération lorsqu'il est échantillonné, à condition que la fréquence d'échantillonnage soit au moins deux fois supérieure à la plus haute fréquence présente dans le signal. Ainsi, il est possible de reconstituer le signal original en utilisant un filtre passe-bas idéal avec une fréquence de coupure égale à la moitié de la fréquence d'échantillonnage.

Dans le contexte de la numérisation d'un signal vidéo, qui peut avoir une bande passante pouvant atteindre 6 MHz, une fréquence d'échantillonnage de 13,5 MHz serait utilisée. Cela signifie que le signal vidéo serait mesuré 13 500 000 fois par seconde. Étant donné que l'œil est plus sensible à la luminance qu'à la chrominance, il est courant d'utiliser une fréquence d'échantillonnage moins élevée pour la chrominance. Ainsi, un échantillonnage à 13,5 MHz est utilisé pour la luminance (Y), tandis qu'un échantillonnage à 6,75 MHz est utilisé pour la chrominance (C, U et V). En conséquence, cet échantillonnage correspond à une profondeur de couleur de 16 bits par pixel. [S17][S20]

#### **ii) La quantification de l'échantillonnage**

Pour coder les informations relatives à un seul pixel d'une image, un certain nombre de bits sera attribué à chaque mesure. Connaissant la capacité de l'œil humain à distinguer jusqu'à 256 niveaux différents de luminance, un codage à 8 bits est suffisant pour reproduire toutes les variations perceptibles ( $8 = 256$  combinaisons possibles). De même, chaque composante couleur de l'image sera quantifiée sur 8 bits (8 bits pour le Rouge, 8 bits pour le Vert, 8 bits pour le Bleu), ce qui donne un total de 24 bits (3 octets) utilisés ( $2^{24} = 16\,777\,216$  combinaisons possibles). [S17][21]

### iii) Les Avantages de la vidéo numérique

- Le passage au numérique a pour fonction principal de tout ramener à l'outil informatique.
- Simplification de la chaîne de production (workflow) d'une vidéo (utilisation uniquement d'ordinateurs et de logiciels particuliers).
- Simplicité de conversion de vidéo format.
- Plusieurs applications possibles : télévisions, vidéo-conférence, cinéma, ....
- Opération de manipulation et de traitement d'image plus facile et plus intuitive.
- Simplicité de requête de recherche de vidéo sur les différentes bases de données. [S21]

#### 2.2.4. Composition des fichiers vidéo

Un fichier vidéo se compose généralement de 2 éléments :

##### 2.2.4.1. Le conteneur

Il correspond généralement au format du fichier. Son rôle est de rassembler et d'organiser dans un fichier, différents types de données (flux audio, vidéos, sous-titres, métadonnées).

Exemple de conteneurs :

- ✓ Audio Video Interleave (.AVI), Quicktime Movie (.MOV), Real Media (.RM), MP4
- ✓ FLV (Flash Video), MKV (Matroska), le WMV (Windows Media Video), Chaque conteneur possède ses spécificités en termes de nombre de pistes acceptées pour la vidéo et l'audio, de codecs reconnus. [S22]

##### 2.2.4.2. Le(s) contenu(s)

Ils se composent essentiellement de flux audio et/ou vidéo. Ceux-ci sont généralement compressés à l'aide d'un codec (algorithme de compression/décompression) comme le Divx, le H264, le mp3.

**Exemple :** un flux vidéo au format Divx peut être « encapsulé » dans un conteneur AVI ou QuickTime. [S22]

## 2.2.5. Les formats de fichiers vidéo

Une manière alternative de décrire l'ordre et la structure des images est la suivante : Les images dans un flux vidéo, qui peuvent être accompagnées de sons sous forme de flux audio, sont extrêmement volumineuses. Afin de les stocker (sur disque dur ou sur des supports d'enregistrement tels que des CD ou des DVD) ou/et de les transmettre (et donc de les adapter au débit des réseaux), il est impératif de les compresser (coder) à l'aide d'un codec. Une fois encodés, les flux vidéo (et éventuellement les flux audio associés) sont généralement encapsulés dans des fichiers conteneurs. Ces fichiers conteneurs permettent notamment une lecture simultanée des différents éléments. [S23]

### 2.2.5.1. Les codecs vidéo

Il existe de nombreux codecs permettant de compresser et de décompresser une vidéo dans son conteneur.

**1. MJPEG (Motion JPEG) :** Le codec vidéo MJPEG compresse la vidéo image par image, en utilisant la technologie JPEG appliquée à l'image fixe, et réunit ces images en mouvement et le son dans un même format de fichier.

Le MJPEG est le codec le plus utilisé pour les captures vidéo des ensembles cartes d'acquisition et logiciels d'édition vidéo. La conservation d'une bonne qualité d'image produit toutefois de gros fichiers.

Le format MJPEG est un format non normalisé. Les solutions M-JPEG, au début des années 90, ont été développées sans concertations par des fabricants, conduisant à MPEG des solutions propriétaires et à des fichiers très souvent incompatibles entre eux.

**2. MPEG (Moving Picture Experts Group) :** Les formats MPEG sont des formats de compression avec pertes pour les séquences vidéo.

Le groupe MPEG (Moving Picture Experts Group), est un groupe d'experts créé en 1988 et chargé du développement de normes internationales pour la compression, la décompression, le traitement et le codage de la vidéo, de l'audio et de leur combinaison, de façon à satisfaire un large panel d'applications. Les formats produits par MPEG sont ouverts, mais non libres

**3. DivX (Digital Vidéo Express)**

- **Extension : .avi**
- Codec vidéo propriétaire et fermé proposé par DivX Inc., conçu à partir de MPEG-4 part 2, ce dernier ayant été modifié afin d'y ajouter la possibilité de compresser le son au format MP3. Cela permet ainsi d'obtenir des vidéos compressées très peu volumineuses avec une

perte de qualité raisonnable.

Ainsi le format DivX permet de stocker un film complet de plusieurs heures sur un CD-ROM de 650 ou 700 Mo. [S24]

### **2.2.5.2. Les Conteneurs vidéo**

Les conteneurs les plus courants sont les suivants :

#### **1. AVI (Audio Video Interleave)**

- Extension : **.avi**
- Format propriétaire et ouvert.
- Le format AVI, développé par Microsoft, très répandu et lisible sur tous les lecteurs vidéo, c'est le format d'en capsulage le plus populaire. Dans un fichier AVI, chaque piste audio et/ou vidéo peut théoriquement être compressée par n'importe quel codec.

#### **2. RealMedia**

- **Extension : .rm, .ram, .rpm**
- Format développé par Real Networks.
- Format propriétaire et fermé.
- Lisible uniquement avec Realplayer.
- Conteneur de flux audio (notamment : RealAudio) et vidéo (notamment : RealVideo)

#### **3. MP4**

- **Extension : .mp4, .mp4a, .mp4v, .m4P**
- Conteneur officiel pour la norme MPEG-4. [S24]

### **2.2.6. Hiérarchie des données vidéo**

La hiérarchie des données vidéo est présentée dans la *Figure 2.10*

#### **2.2.6.1. Séquence vidéo (Video Sequence)**

Qui commence par un en-tête de séquence, contient un ou plusieurs groupes d'images et s'achève par un code de fin de séquence.

### 2.2.6.2. Groupe images (Group of Pictures)

Qui regroupe un en-tête et une série d'une ou plusieurs images facilitant l'accès direct.

### 2.2.6.3. Image (Picture)

L'unité fondamentale pour le codage des séquences vidéo est une image, qui se compose de trois matrices rectangulaires représentant la luminance (Y) et la chrominance (Cb et Cr). Au lieu d'utiliser la représentation RGB (Rouge, Vert, Bleu), on préfère utiliser la représentation YCbCr. Cette préférence est due à la sensibilité de l'œil humain, qui est plus attiré par la luminosité que par la chrominance. Par conséquent, il n'est pas nécessaire de stocker autant d'informations dans les matrices Cb et Cr que dans la matrice Y. En comparaison, la représentation RGB aurait nécessité une matrice deux fois plus grande que la matrices-Y.

### 2.2.6.4. Tranche (Slice)

Dans cette méthode de compression, les données sont organisées en macroblocs qui sont disposés de gauche à droite, puis de haut en bas. Chaque macrobloc contient une partie de l'information nécessaire pour reconstituer l'image ou la vidéo. En cas d'erreur dans une tranche de données, le décodeur peut simplement passer à la tranche suivante pour obtenir les informations manquantes. Cette approche permet de traiter les erreurs de manière plus efficace, car plus il y a de tranches, meilleure est la capacité à récupérer les données perdues. Cependant, l'inconvénient est que cela nécessite plus d'espace pour stocker les multiples tranches de données.

### 2.2.6.5. Macroblocs (Macroblocks)

C'est une matrice rectangulaire de dimension 2 et constituée de blocs.

### 2.2.6.6. Blocs (Block)

C'est un ensemble des valeurs de luminance et chrominance de 8 lignes de 8 pixels. [S25]

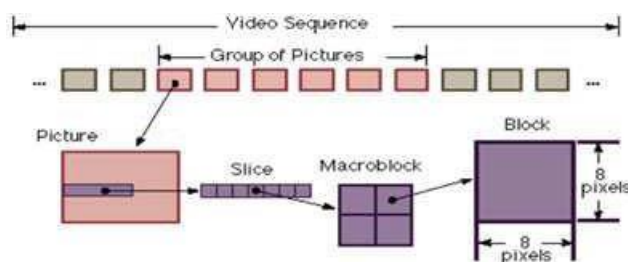


Figure 2.10 : Hiérarchie des données dans le flux vidéo. [S25]

## 2.2.7. Représentation d'une séquence vidéo

Une séquence vidéo brute est une suite d'images fixes, qui peut être caractérisée par trois principaux paramètres :

### 2.2.7.1. La résolution en luminance

Le nombre de nuances ou de couleurs possibles pour un pixel peut être déterminé selon deux formats couramment utilisés. Pour les niveaux de gris, la plupart des images utilisent généralement un codage sur 8 bits, ce qui signifie qu'il y a 256 nuances de gris. En revanche, pour les images en couleurs, on utilise souvent un codage sur 24 bits, ce qui permet de représenter jusqu'à 16,7 millions de couleurs distinctes.

### 2.2.7.2. La résolution spatiale

Définit le nombre de lignes et de colonnes de la matrice de pixels.

### 2.2.7.3. La résolution temporelle

Est le nombre d'images par seconde. La valeur de ces trois paramètres détermine l'espace mémoire nécessaire pour stocker chaque image de la séquence. Cet espace mémoire est caractérisé par le débit, qui est le coût de stockage pour une seconde (capacité mémoire nécessaire pour stocker une seconde de vidéo). [22]

## 2.2.8. Les paramètres clés d'une vidéo

Le stockage et la diffusion d'une vidéo exigent un espace volumineux et un taux de transfert plus élevé. Le contrôle de qualité, et la taille d'une séquence vidéo est déterminé par deux paramètres clés,

- 1. le nombre d'images par seconde** : le nombre d'images du système visuel humain exigé en général 25 ou 30 images par seconde.
- 2. La résolution** : Ce terme fait référence à la capacité limitée de transmettre des informations visuelles dans une image. En d'autres termes, il s'agit du nombre de pixels pouvant être affichés par un dispositif d'affichage. Parvenir à un compromis entre ces paramètres et les limitations imposées par la technologie est essentiel pour obtenir une qualité vidéo optimale.



## **Conclusion**

Dans ce chapitre ont à vue c'est quoi une image, la représentation d'une image numérique et les différents types d'image numérique ont à vue la détection de l'objet en première partie, en deuxième partie ont à vue c'est quoi une vidéo, ses fondamentaux technique et physique, la composition d'un fichier vidéo ont passant par les formats de fichier vidéo, hiérarchie des données vidéo et pour finir la représentation d'une séquence vidéo et les paramètres clés d'une vidéo.

*Chapitre 3 :*

*Détection et*

*reconnaissance d'objet.*

## Chapitre 03

# Détection et reconnaissance d'objet

## Sommaire

<a href="#">Chapitre 03</a> .....	43
<a href="#">Introduction</a> .....	43
<a href="#">3.1. Définition d'objet</a> .....	44
<a href="#">3.2. Caractéristiques d'un objet</a> .....	44
<a href="#">3.3. La détection d'objet</a> .....	45
<a href="#">3.4. La reconnaissance d'objet</a> .....	45
<a href="#">3.5. Le principe de la détection d'objets</a> .....	45
<a href="#">3.6. Classification d'objet</a> .....	46
<a href="#">3.7. Quelques Domaines d'application</a> .....	49
<a href="#">3.8. Technique de reconnaissance d'objet</a> .....	50
<a href="#">3.9. Méthode de Détection d'objets en mouvement</a> .....	51
<a href="#">Conclusion</a> .....	57

### Introduction

La détection et la reconnaissance d'objets sont des domaines de recherche en constante évolution dans le domaine de la vision par ordinateur. La capacité à identifier automatiquement les objets dans une image ou une vidéo est une tâche cruciale pour de nombreuses applications telles que la surveillance de sécurité, la navigation autonome, la réalité augmentée et bien d'autres. La détection d'objets consiste à localiser les instances d'un objet particulier dans une image ou une vidéo, tandis que la reconnaissance d'objets consiste à identifier l'objet en question en fonction de sa catégorie ou de son identité. Ces deux tâches sont intimement liées et sont souvent réalisées ensemble. Dans ce chapitre, nous allons explorer les différentes approches de détection et de reconnaissance d'objets, ainsi que les techniques les plus récentes et les plus performantes pour résoudre ces problèmes. Nous examinerons également les défis et les limitations de ces méthodes, ainsi que les perspectives d'avenir pour la recherche dans ce domaine en constante évolution.

### 3.1. Définition d'objet

Dans le domaine de la vision par ordinateur, la notion d'objet constitue un élément clé, puisque notre étude se concentre toujours sur la notion d'objet. L'objet désigne une zone de l'image qui est caractérisé par sa texture, sa forme, sa couleur, son orientation du gradient ou encore son mouvement et qui représente. D'une autre manière on peut dire que les objets dans les séquences d'images représentent les pixels qui appartiennent à l'avant plan ou le fond de la scène.

- **Les objets de contexte**

Les objets physiques sont généralement immobiles, tels que les murs qui restent en place. Cependant, certains objets contextuels comme les chaises ou les portes peuvent être en mouvement, mais leurs déplacements peuvent être anticipés en se basant sur les informations environnantes. [32]

- **Les objets mobiles**

Sont des objets physiques qui peuvent être perçus dans les scènes par leurs mouvements [32]. Il est cependant difficile de prédire leurs mouvements P. ex. les personnes, les véhicules.

Les objets dans les vidéos sont des objets mobiles. Ils sont habituellement détectés et suivis dans plusieurs frames.

### 3.2. Caractéristiques d'un objet

Dans le domaine du mouvement d'un objet dans une séquence, un objet peut avoir différentes caractéristiques, il peut être : [23]

- **Constant ou variable** : La forme, le mouvement, les couleurs ou les textures d'un objet varient ou non au cours du temps. Ces considérations jouent en faveur de l'adoption ou non d'un modèle de l'objet pour en améliorer le suivi.
- **Rigide ou non rigide** : pour la forme nous parlerons d'objet rigide (voiture) ou non-rigide (corps humain) c.à.d. la distance entre les points de l'objet est constante ou variable au cours du temps.
- **Unique ou multiple** : Dans le cadre d'un suivi labellisé où chaque objet porte une étiquette et n'en change pas durant l'opération de son suivi, une distinction entre chaque objet est nécessaire c.à.d. que chaque objet doit être unique. L'unicité repose, là aussi, sur un modèle. Cependant celui-ci n'est pas nécessairement connu au préalable. Cependant une connaissance

- des critères potentiellement discriminants permettent une amélioration de la charge de calcul. La description de chaque objet selon des critères pertinents assure une labellisation fiable.

### 3.3.La détection d'objet

La détection d'objets est une tâche importante de la vision par ordinateur qui consiste à localiser et à identifier les objets dans une image ou une vidéo en utilisant des algorithmes de traitement d'image et d'apprentissage automatique. Cette tâche est cruciale dans de nombreux domaines tels que la surveillance, la reconnaissance de visages, la conduite autonome et la réalité augmentée. Les méthodes de détection d'objets les plus performantes sont basées sur les réseaux de neurones convolutifs, qui sont des modèles d'apprentissage profond largement utilisés dans la vision par ordinateur. [S26]



Figure 3.1 : Plusieurs objets détectés dans la scène.

### 3.4.La reconnaissance d'objet

La reconnaissance d'objets en informatique est le processus qui consiste à classifier les objets dans une image ou une vidéo en catégories prédéfinies. Cette tâche utilise des techniques d'apprentissage automatique, comme les réseaux de neurones, pour extraire des caractéristiques de l'image et effectuer la classification. Les résultats obtenus peuvent être utilisés dans divers domaines tels que la recherche d'images, la surveillance de sécurité et la réalité augmentée. [S26]

### 3.5.Le principe de la détection d'objets

Lorsqu'il s'agit de détecter des objets dans une image, le processus repose sur une approche spécifique. Tout d'abord, on effectue une recherche au sein de l'image afin d'identifier les zones susceptibles de contenir des objets. Une fois ces régions détectées, elles sont extraites

individuellement et soumises à un modèle de classification d'images pour être étiquetées. Les régions de l'image originale qui obtiennent de bons résultats lors de la classification sont conservées, tandis que les autres sont écartées. En conséquence, une méthode de détection d'objets efficace nécessite à la fois un algorithme fiable pour la détection des régions et un solide algorithme de classification d'images. [S27]



Figure 3.2 : Principe de la détection d'objet. [52]

### 3.6. Classification d'objet

Les régions de mouvement extrait peuvent être de différents types d'objet tels que les humains, les véhicules, les oiseaux, les nuages flottants, les arbres balançant et d'autres objets en mouvement. C'est pourquoi les caractéristiques de forme des régions en mouvement sont utilisées [24]. Selon la littérature les approches de classification des objets sont les suivants :

:

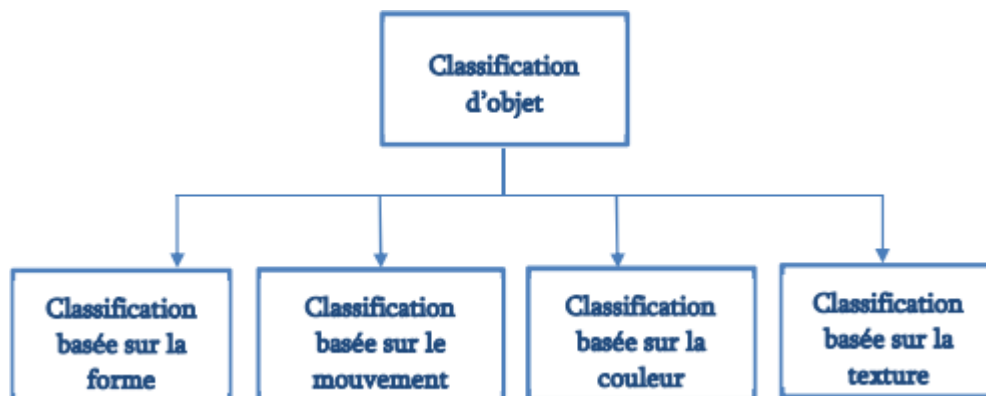


Figure 3.3 : Classification d'objet. [24]

### **3.6.1. Classification basée sur la texture**

La technique de classification basée sur la texture [25] compte les occurrences d'orientation du gradient dans des parties localisées d'une image et calculé sur une grille dense de cellules uniformément espacés et utilise une normalisation de chevauchement de contraste local pour une meilleure précision.

### **3.6.2. Classification basée sur la forme**

Il existe différentes descriptions pour les informations de forme de mouvement des régions telles que les représentations de points et des boites qui sont disponible pour classer les objets en mouvement. Les caractéristiques d'entrée sont un mélange de paramètres d'objet basé sur l'image et la scène. La classification est effectuée sur chaque objet chaque frame [24].

### **3.6.3. Classification basée sur la couleur**

Contrairement à d'autres aspects de l'image tels que la forme, la couleur reste relativement constante malgré les changements de point de vue, et elle est facile à acquérir. Bien qu'elle ne soit pas toujours suffisante pour détecter et suivre des objets, la faible charge de calcul des algorithmes disponibles en fait une caractéristique souhaitable. Pour la détection et le suivi en temps réel de véhicules ou de piétons, une technique basée sur l'histogramme de couleur est utilisée. Selon Sankari et al. [26], un modèle de mélange de Gaussiennes est créé pour décrire la distribution des couleurs à l'intérieur de la séquence d'images et segmenter l'image en arrière-plan et en objets. Les objets en situation d'occlusion sont traités à l'aide d'un tampon d'occlusion.

### **3.6.4. Classification basée sur le mouvement**

Le mouvement d'objet non rigide offre une caractéristique périodique, ce qui permet son utilisation comme un signal puissant pour la classification des objets en mouvement. De plus, le flux optique constitue également un outil précieux dans la classification des objets. L'analyse du flux résiduel permet d'examiner la rigidité et la périodicité des entités en mouvement. On peut prévoir que les objets rigides présenteront peu de flux résiduel, tandis qu'un objet en mouvement non rigide, tel qu'un être humain, présentera un flux résiduel moyen plus élevé, voire même une composante périodique affichée. [24]

Le tableau suivant montre quelques points de comparaison entre les méthodes de classification des objets :

**Tableau 3.1 : Étude comparative des méthodes de classification des objets [26].**

<b>Méthodes</b>	<b>Précision</b>	<b>Temps de calcul</b>	<b>Commentaires</b>
<b>basée sur la forme</b>	<b>Modéré</b>	<b>Faible</b>	-Peut être appliqué avec des modèles appropriés  -Ne fonctionne pas bien dans les situations dynamiques et elle est incapable de déterminer les mouvements internes
<b>basée sur le mouvement</b>	<b>Modéré</b>	<b>Elevé</b>	-Ne nécessite pas des modèles de motifs prédéfinis  -Lutte pour identifier un humain qui n'est pas en mouvement
<b>basée sur la texture</b>	<b>élevé</b>	<b>Elevé</b>	-Fournit une meilleure qualité -Besoin de temps de calcul supplémentaire
<b>Basée sur la couleur</b>	<b>élevé</b>	<b>Elevé</b>	- Cout de calcul faible des algorithmes - Pas toujours approprié en raison de faible précision

### 3.7. Quelques Domaines d'application

Dans ce titre, nous allons ici explorer l'impact que cette technique de vision par ordinateur peut avoir dans différents secteurs. Plus précisément, nous allons examiner comment la détection et la reconnaissance d'objets peut être utilisée dans les domaines suivants : [S26]



**Figure 3.4** : Exemples des domaines d'applications.

**1. Vision par ordinateur** : La détection et la reconnaissance d'objets sont largement utilisées dans la vision par ordinateur, notamment pour des tâches telles que la détection de piétons, la détection de véhicules, la reconnaissance faciale, la reconnaissance de gestes, etc.

**2. Automobile autonome** : Les voitures autonomes utilisent la détection et la reconnaissance d'objets pour détecter et suivre les véhicules, les piétons, les panneaux de signalisation, les feux de circulation, les obstacles, etc., afin de naviguer en toute sécurité sur les routes.

**3. Médecine** : Dans le domaine médical, la détection et la reconnaissance d'objets sont utilisées pour l'analyse d'images médicales, la détection de tumeurs, la reconnaissance de caractéristiques anatomiques, la planification chirurgicale, etc.

**4. Surveillance et sécurité** : Dans le domaine de la surveillance et de la sécurité, la détection et la reconnaissance d'objets sont utilisées pour détecter les intrusions, surveiller les comportements suspects, suivre les objets en mouvement, identifier les personnes, etc.

### 3.8. Technique de reconnaissance d'objet

Diverses approches de reconnaissance d'objets sont possibles. Récemment, les techniques proposées par le Machine Learning et le Deep Learning remportent un succès croissant pour la résolution des problèmes posés par la reconnaissance d'objets.

**1. Le Deep Learning** : est aujourd'hui couramment utilisé pour la reconnaissance d'objets.

Des modèles de Deep Learning, tels que des réseaux de neurones à convolution, ou CNN (Convolutional Neural Network), sont utilisés pour l'apprentissage automatique des caractéristiques intrinsèques d'un objet afin de rendre possible son identification.

Deux approches sont possibles pour mettre en œuvre de la reconnaissance d'objets à l'aide du Deep Learning :

- **Entraînement d'un modèle en partant de zéro** : pour entraîner un réseau profond en partant de zéro, il vous faut réunir un très gros jeu de données labellisées et concevoir une architecture de réseau qui apprendra les caractéristiques voulues et bâtira le modèle. Les résultats peuvent être impressionnants, mais cette approche nécessite une grande quantité de données d'apprentissage, et vous devez configurer les couches et les pondérations dans le CNN.
- **Utilisation d'un modèle de Deep Learning pré-entraîné** : La plupart des applications de Deep Learning utilisent l'approche dite de l'apprentissage par transfert, qui consiste à perfectionner un modèle pré-entraîné. Vous démarrez avec un réseau existant, tel qu'AlexNet ou GoogLeNet, qu'il vous faut enrichir avec de nouvelles données contenant des classes auparavant inconnues du réseau. Cette méthode, moins fastidieuse, est aussi susceptible de fournir plus rapidement des résultats dans la mesure où le modèle a déjà été entraîné sur des milliers voire des millions d'images

**2. Le Machine Learning**, également fréquemment utilisé pour la reconnaissance d'objets, propose des approches différentes de celles du Deep Learning. Parmi les techniques de Machine Learning couramment mises en œuvre, citons à titre d'exemple les suivantes

L'extraction de caractéristiques HOG (Histogram of Oriented Gradients) avec un modèle de Machine Learning SVM (Support Vector Machine)

Les représentations par sac de mots, avec des algorithmes SURF (Speeded Up Robust Features) et MSER (Maximally Stable Extremal Regions), par exemple

L'algorithme de Viola-Jones, qui peut être utilisé pour la reconnaissance de divers d'objets, y compris les visages et les bustes. [S32]

### 3.9.Méthode de Détection d'objets en mouvement

Pour pouvoir suivre les mouvements dans la vidéo, la première étape est de détecter les mouvements. Cette étape joue un rôle très important dans l'analyse vidéo.

Dans cette partie, on va étudier quatre méthodes pour détecter le mouvement :

#### 3.9.1. La différence temporelle des images (méthode de mouvement)

La différence temporelle détecte la région de mouvement grâce à la différence de pixel par pixel de deux trames consécutives dans un flux vidéo [27] [28].

Cette méthode adapte le changement de la scène. Mais elle est moins d'efficacité parce que dans une durée de temps  $\Delta t$ , peut-être, on détecte seulement une partie d'objet, par exemple : la main, la tête etc. Et le vide (la région où l'objet s'est déplacé l'autre lieu)est aussi détecté.

Dans ce cas-là, c'est très difficile à extraire des propriétés de mouvement (la taille, la position, la vitesse etc.) et à suivre le mouvement. [29]



**Figure 3.5** : Un exemple de la différence temporelle. [29]

(a) une scène simple avec deux objets, (b) les régions rouges sont la différence entre deux images consécutives.

Voici l'idée principale de cette méthode : Soit  $I_t$  l'image à l'instant  $t$  et  $I_{t-1}$  l'image à l'instant  $t-1$ . L'objet du mouvement se compose des pixels qui satisfont l'équation suivante :

$$\max (| I_t(x, y) * c - I_{t-1}(x, y) * c |, c=(R,G,B)) \geq \text{seuil}$$

Comme ci-dessus a déjà expliqué, on peut améliorer cette méthode de la façon suivante au lieu de soustraire l'image à l'instant t-1, on soustrait la moyenne de N images dernières [28] it  $I_m$  la moyenne de N dernières images à l'instant t. L'objet du mouvement se compose des pixels qui satisfont l'équation suivante :

$$\max (| I_t(x, y) * c - I_m(x, y) * c |, c=(R,G,B)) \geq \text{seuil}$$

La moyenne de N images à l'instant t +1 est mise à jours :

$$I_{t+1}(x, y) * c = \alpha I_t(x, y) * c + (1-\alpha) I_m(x, y) * c, c=(R, G, B).$$

Où  $\alpha \in (0,1)$  est une constante et est décidée par la pratique.

➤ **Avantage :**

- Adapte le changement de la scène.
- Détecte la région de mouvement.

➤ **Inconvénients :**

- Ne permet pas de détecter le mouvement dans les zones uniformes intérieures à l'objet
- Ne fonctionne pas dans plusieurs cas, pour différentes raisons, telles que : la présence de bruit du capteur et les changements de luminosité de la scène qui modifient les intensités des pixels.

### 3.9.2. Double de différence temporelle et caractère de contour

Dans cette méthode, on utilise aussi la différence de pixel par pixel des trames consécutives dans un flux vidéo comme la 1<sup>ère</sup> méthode. Mais on va utiliser trois trames consécutives. [30]

Cette façon nous donne le résultat meilleur que celui de la 1<sup>ère</sup> méthode, tandis qu'elle adapte aussi le changement de la scène.

Voici l'idée principale de cette méthode : Soit I l'image à l'instant t, I l'image à l'instant t-1 et I l'image à l'instant t-2.

L'objet du mouvement se compose les pixels qui satisfont l'équation suivante :

$$I_1(x, y) = \max (| I_t(x, y) * c - I_{t-1}(x, y) * c |, c=(R,G,B)) \geq \text{seuil}$$

$$I_2(x, y) = \max (| I_{t-1}(x, y) * c - I_{t-2}(x, y) * c |, c=(R,G,B)) \geq$$

$$\text{seuil} \quad I_{\text{résultat}}(x, y) = I_1(x, y) \cup I_2(x, y).$$

Nous extrayons la région des objets en mouvement par la double méthode de différence. Dans le cas où l'objet se déplace lentement ou il n'y a qu'une partie d'objet qui se déplace, nous ne pouvons pas obtenir complètement la forme d'objet. Peut-être, un objet sera divisé en plusieurs régions.

Nous utilisons donc le caractère de contour pour combiner ces régions.

Après avoir masqué l'image à l'instant t-1 avec la région mouvante obtenue à partir de la double méthode de différence, le contour est calculée dans la région où le mouvement s'est produit.

Le contour est calculé à partir de l'image F qui est l'image à l'instant t-1 masquée avec R

$$G = \begin{bmatrix} F_x^2 & F_x F_y \\ F_x F_y & F_y^2 \end{bmatrix}$$

$$= I_1(x, y) \cup I_2(x, y).$$

Le contour est représenté par :

En utilisant un caractère de gradient avec l'intensité de l'image masquée F. En chaque pixel que le mouvement est détecté par l'intermédiaire de la double méthode de différence dans l'armature courante, le coefficient est calculé. C'est-à-dire, les pixels commandés par le caractère maximal de contour sont choisis en tant que points de caractère de contour parce que les valeurs propres minimum plus grandes sont les caractères plus forts de contour.

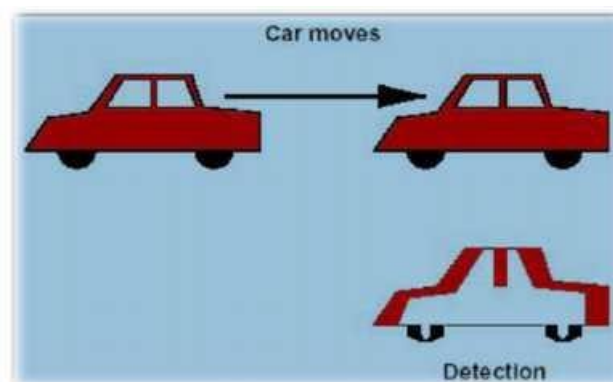
Avec cette façon, à l'instant t, on va obtenir des objets mouvant à l'instant t-1. C'est-à-dire, on est en retard de 0,5 seconde. Autre chose, si l'objet ne se déplace pas pendant quelques secondes, cette manière ne peut pas détecter le mouvement. Dans ce cas, nous utilisons le contour d'objet avec le résultat de dernière étape pour détecter le mouvement. [31]

### ➤ Avantages

- Le résultat de cette méthode est meilleur par rapport la méthode précédente.
- Détecte la région de mouvement.
- Utilise trois trames.

### ➤ Inconvénient

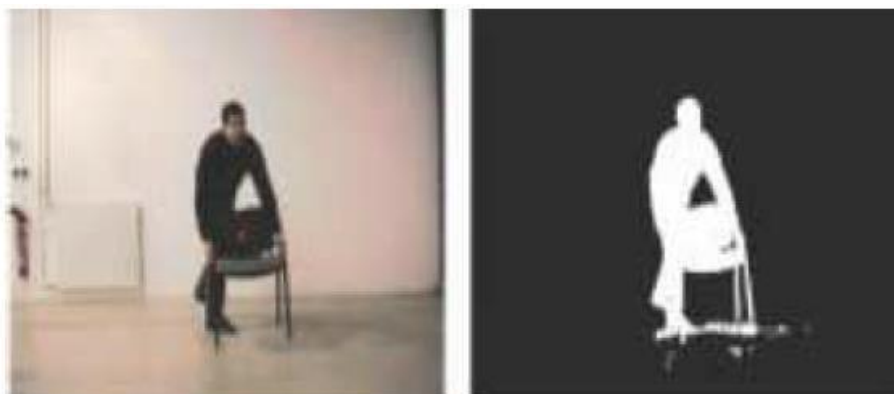
- Si l'objet ne se déplace pas pendant quelques secondes, cette méthode ne peut pas détecter le mouvement.



*Figure 3.6* : Détection de mouvement par différence temporel. [32]

### 3.9.3. La soustraction de l'image de fond (méthode de différence)

La troisième méthode, on utilise une image de fond. Cette méthode est très populaire et elle est utilisée par plusieurs applications. Cette façon détecte la région de mouvement en soustrayant pixel par pixel l'image courante de l'image de fond.



*Figure 3.7* : Exemple soustraction du fond. [33]

Soit  $I_t$  l'image à l'instant  $t$ .  $B_t$  est l'image de fond à l'instant  $t$ . L'objet de mouvement se compose des pixels qui satisfont l'équation suivante:

$$\max (| I_t(x, y) * c - B_t(x, y) * c |, c = (R, G, B)) > = S(x, y)$$

Où  $S_t(x, y)$  est le seuil de la position  $(x, y)$  à l'instant  $t$ . [31]

#### ➤ **Création d'image de fond**

Il existe trois manières pour créer l'image de fond :

- Le calcul de la moyenne des  $N$  premières images.
- Choisir l'image de fond qui a peu de changement ou qu'il n'y a pas de mouvement pendant une durée longue.
- Choisir la première image dans le flux vidéo. [S28]

#### ➤ **Les Avantages**

- Cette méthode est très populaire.
- Utilisée par plusieurs applications.
- Détecte l'objet complètement.

#### ➤ **Inconvénient**

- Cette technique se limite aux caméras en position fixe.

### **3.9.4. Méthode de détection des objets selon la Couleur**

La quatrième méthode est basée sur la couleur d'objets détectés.

Lecture pixel par pixel de la valeur de celui-ci en rouge vert et bleu. Puis test sur le pixel, suivant la couleur à reconnaître le rouge, le vert et le bleu doivent être supérieur ou inférieur au seuil impose. S'ils respectent ces conditions alors on incrémente. [34]

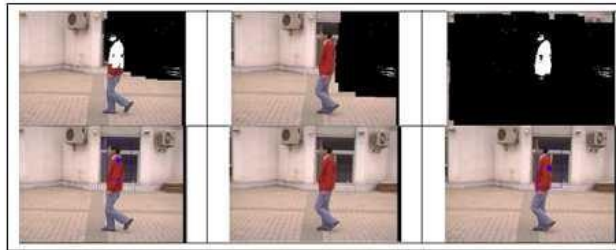
#### ➤ **Test**

Condition de supériorité ou d'infériorité sur chaque couleur.

#### ➤ **Paramètres**

Les paramètres principaux sont les différents seuils des couleurs (RGB). Dans l'exemple utilise, les conditions sont  $R > 120$ ,  $G < 100$ ,  $B < 100$ . [S16]

## ➤ Résultats

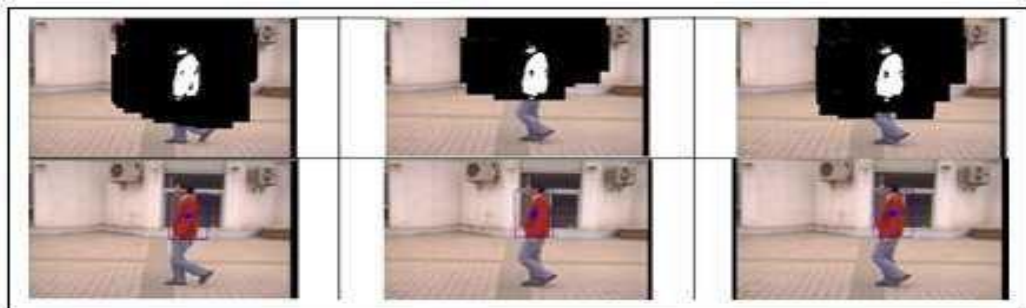


**Figure 3.8** : Exemple1, pixels de couleur rouge reconnus dans les particules  
etreconnaissance de l'objet. [S16]

Sur les images ci-dessus, une erreur apparait, l'objet n'est pas reconnu sur le frame du milieu, cette erreur est due à la reconnaissance dans une particule.

En ajoutant un paramètre, cette erreur peut être évitée, ce paramètre est un seuil applique à la moyenne de la particule, en effet si la moyenne est trop faible alors elle est considérée comme nulle.

Le même exemple repris avec l'ajout de ce seuil donne le résultat suivant. [S16]



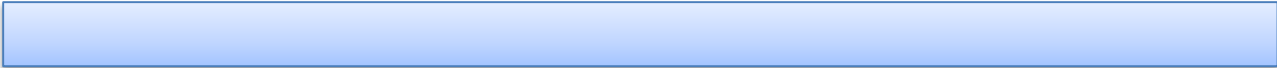
**Figure 3.9** : Exemple 2, reconnaissance d'un objet à l'aide de la couleur  
aprèsmodification d'un paramètre. [S16]

## ➤ Avantage

- C'est une méthode assez simple à mettre en œuvre, il suffit de reconnaître une couleur qui ressort par rapport au reste de l'image.

## ➤ Inconvénients

- Il existe beaucoup de nuance de couleur (2563) il faut alors avoir une grande base de données puis choisir la couleur ou définir à chaque fois la couleur souhaitée ce qui sous-entend que ces valeurs soit connues.

- 
- Si un autre objet possède la même couleur on ne peut pas les différencier, exceptes'il y a une grande différence de taille entre les deux objets et que l'objet à reconnaître est le plus grand.
  - Si notre objet change de couleur au cours de la vidéo, la couleur à reconnaître devra également être modifiée.

## **Conclusion**

En conclusion, ce chapitre a exploré les différentes approches de détection et de reconnaissance d'objets utilisées. Nous avons vu que la détection d'objets implique la localisation et la classification d'objets dans une image, tandis que la reconnaissance d'objets se concentre principalement sur la classification d'objets dans une image ou une vidéo. Les méthodes les plus performantes pour la détection et la reconnaissance d'objets sont souvent basées sur l'apprentissage en profondeur qui est capables d'extraire des caractéristiques complexes des images pour effectuer la classification.

*Chapitre 4 :*

*Analyse, conception et*

*implémentation.*

## Chapitre 04

# Analyse, conception et Implémentation

## Sommaire

<a href="#">Chapitre 04</a> .....	59
<a href="#">Introduction</a> .....	59
<a href="#">4.1 Environnement du travail</a> .....	59
<a href="#">4.2 Environnement de développement</a> .....	60
<a href="#">4.3 Langages de programmation</a> .....	60
<a href="#">4.4 Diagramme de séquence</a> .....	61
<a href="#">4.5 Présentation de l'interface du système</a> .....	62
<a href="#">Conclusion</a> .....	67

### Introduction

Dans ce chapitre, nous allons présenter l'étape de réalisation. Tout d'abord, par la représentation de l'environnement matériel et logiciel utilisé pour développer notre système et la discussion du choix du langage de programmation, le diagramme séquence. Ensuite, nous présenterons l'interface du système et quelques résultats expérimentaux.

### 4.1 Environnement du travail

Dans cette section, nous présentons les environnements matériels et logiciels de notre travail.

#### 4.1.1. Environnement matériel ou le hardware

- CPU : Intel(R) Core(TM) i3
- RAM : 4GB Memory DDR3 1666Hz
- ROM : 500 GB HDD
- Webcam intégrée 720P

#### 4.1.2. Environnement immatériel ou le software

- Système d'exploitation : Windows 7 Home
- Logiciel : PyCharm Community Edition 2020.3.5 x64
- La bibliothèque : OPENCV version 4.5.2 et Tkinter

#### 4.2. Environnement de développement

**PyCharm** : est un environnement de développement intégré utilisé pour programmer en Python. Il permet l'analyse de code et contient un débogueur graphique. Il permet également la gestion des tests unitaires, l'intégration de logiciel de gestion de versions, et supporte le développement web avec Django. Développé par l'entreprise tchèque JetBrains, c'est un logiciel multiplateforme qui fonctionne sous Windows, Mac OS X et Linux. Il est décliné en édition professionnelle, diffusé sous licence propriétaire, et en édition communautaire diffusé sous licence Apache. [S31]



#### 4.3. Langages de programmation

##### 4.3.1. Python

Python est un langage de programmation interprété multi-paradigme. Il favorise la programmation impérative structurée, et orientée objet. Il est doté d'un typage dynamique fort, d'une gestion automatique de la mémoire par ramasse-miettes et d'un système de gestion d'exceptions il est ainsi similaire à Perl, Ruby, Scheme, Smalltalk et Tcl.



Le langage Python est placé sous une licence libre proche de la licence BSD et fonctionne sur la plupart des plates-formes informatiques, des supercalculateurs aux ordinateurs centraux, de Windows à Unix en passant par Linux et MacOS, avec Java ou encore .NET. Il est conçu pour optimiser la productivité des programmeurs en offrant des outils de haut niveau et une syntaxe simple à utiliser. Il est également apprécié par les pédagogues qui y trouvent un langage où la syntaxe, clairement séparée des mécanismes de bas niveau, permet une initiation plus aisée aux concepts de base de la programmation. [S30]

Parmi les bibliothèques de python les plus importantes dans notre projet que nous avons utilisé :



#### 4.3.2. OpenCV (Open Computer Vision)

Est une bibliothèque graphique libre, initialement développée par Intel, spécialisée dans le traitement d'images en temps réel. La société de robotique Willow Garage et la société ItSeez se sont succédé au support de cette bibliothèque. Depuis 2016 et le rachat d'ItSeez par Intel, le support est de nouveau assuré par Intel. Cette bibliothèque est distribuée sous licence BSD. NVidia a annoncé en septembre 2010 qu'il développerait des fonctions utilisant CUDA pour OpenCV7. [S29]

#### 4.4. Diagramme de séquence

Un diagramme de séquence est un outil de modélisation visuelle qui représente la séquence d'actions et de messages échangés entre les différents objets ou acteurs impliqués dans un scénario donné. Il est utilisé dans le domaine de l'ingénierie logicielle pour concevoir et décrire le comportement d'un système en termes de séquences d'interactions entre ses différents composants. [S26]

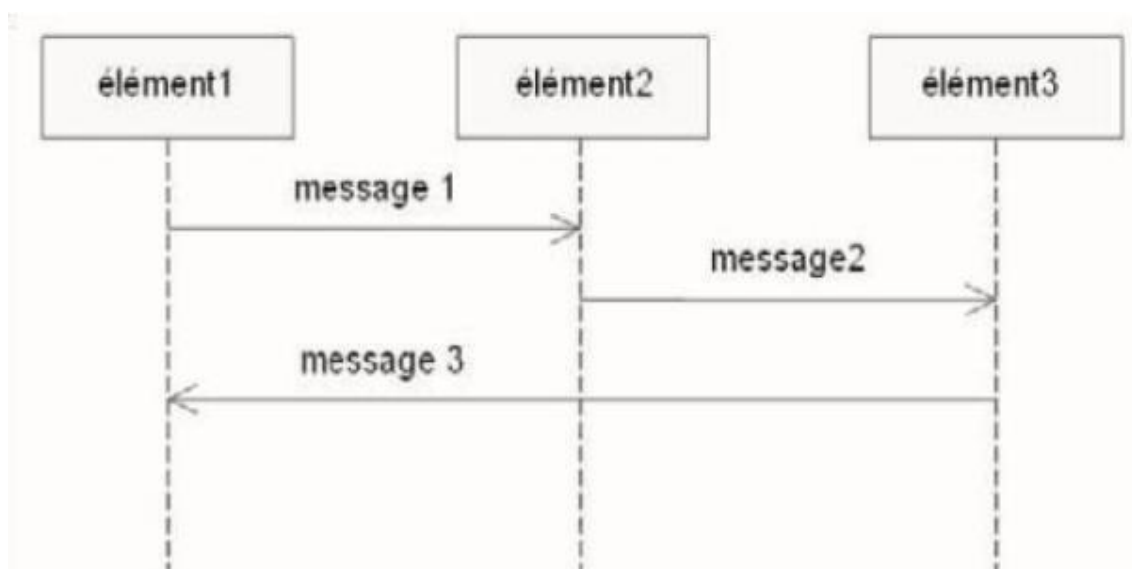
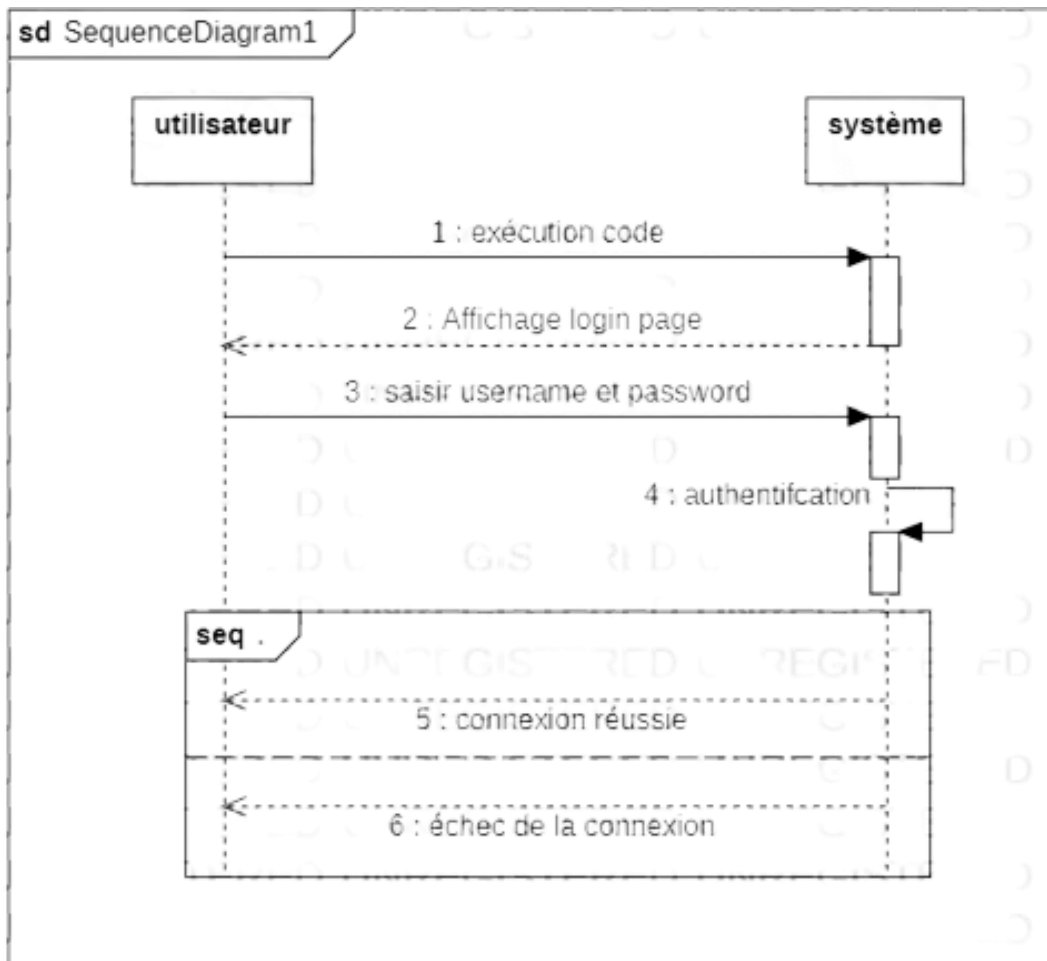


Figure 4.1 : Exemple d'un diagramme de séquence.

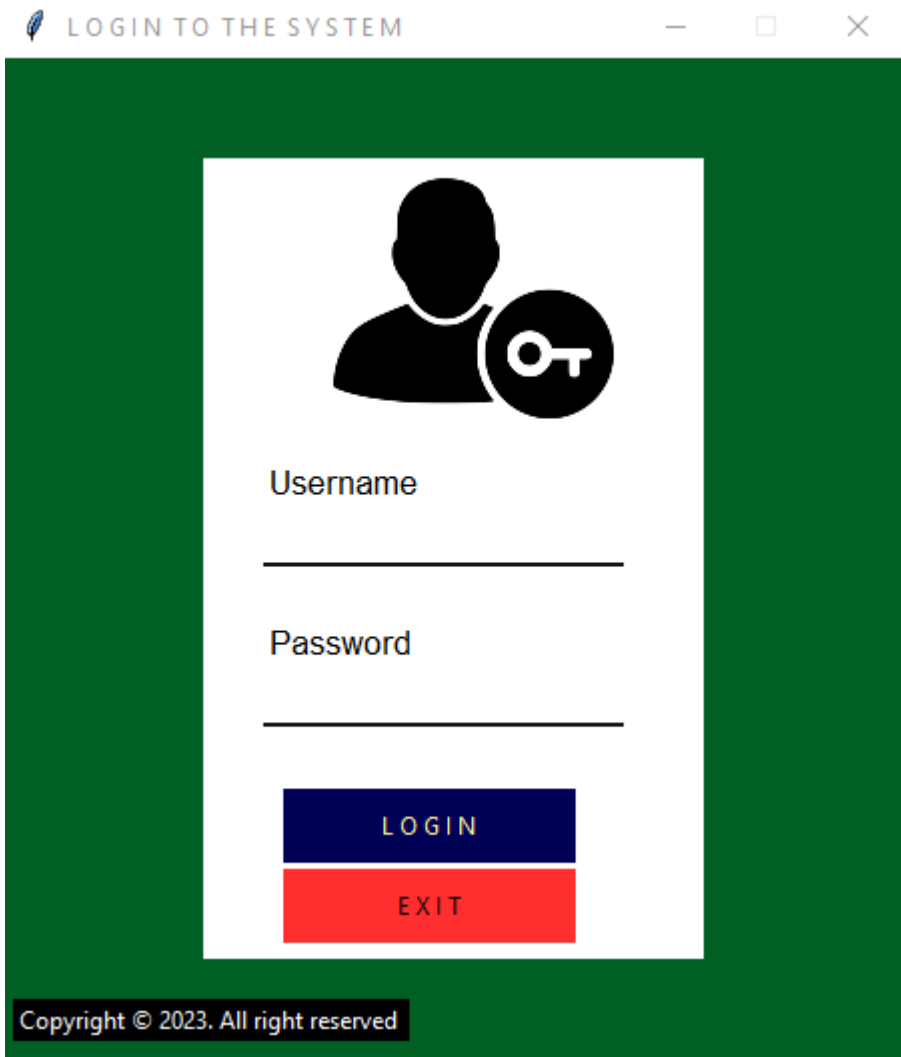
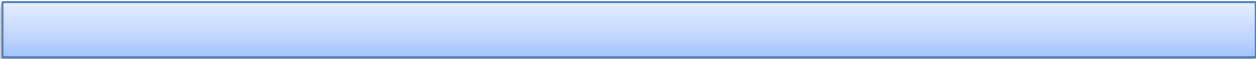
Voici le diagramme de séquence pour mon domaine :



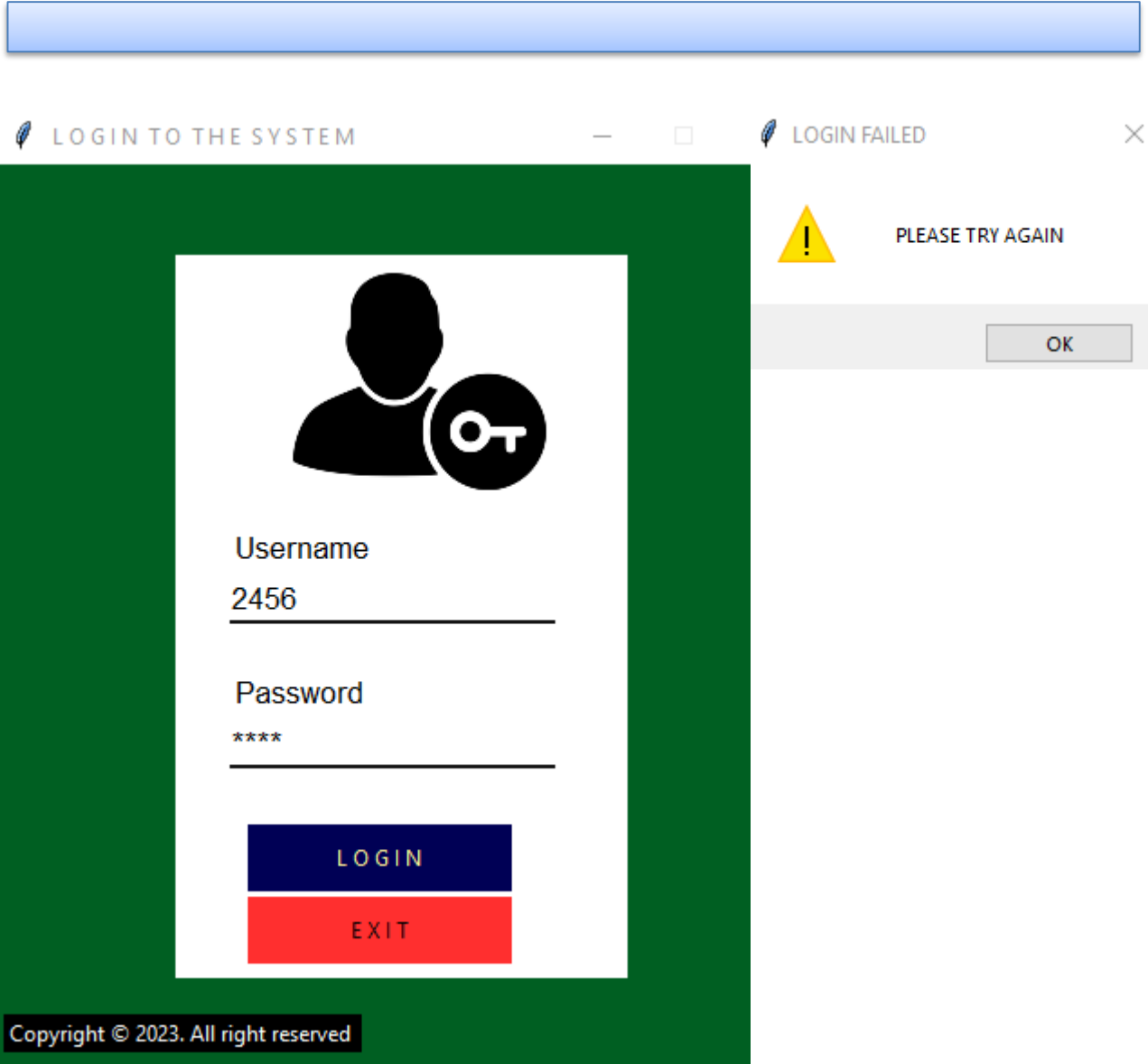
*Figure 4.2* : diagramme de séquence.

#### 4.5. Présentation de l'interface du système

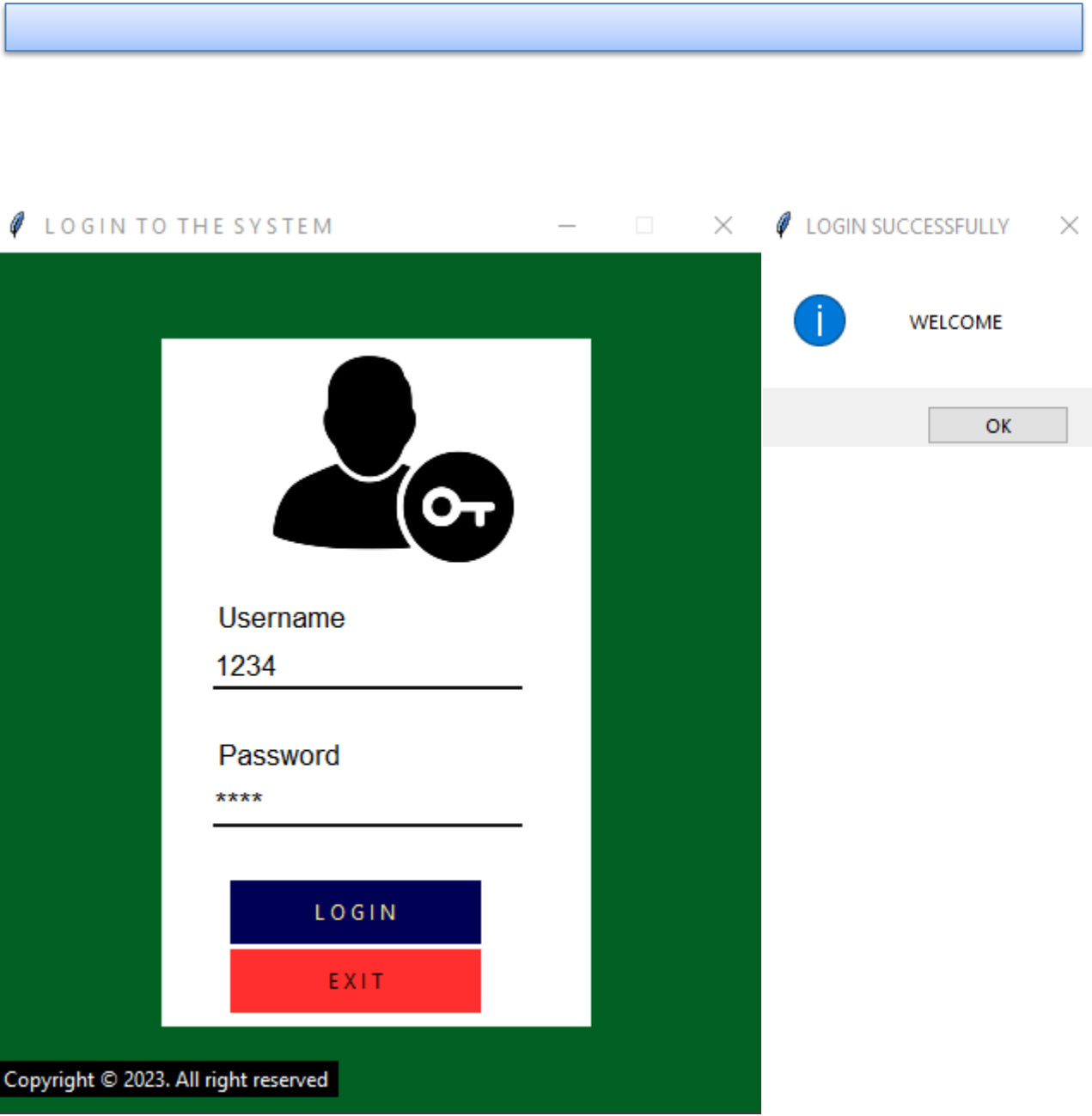
Dans cette section nous allons présenter quelques captures d'écran de l'interface de notre système, chacune de ces captures va montrer une fonctionnalité de l'application. Pour protéger l'application il faut d'abord entrer un nom d'utilisateur et le mot de passe.



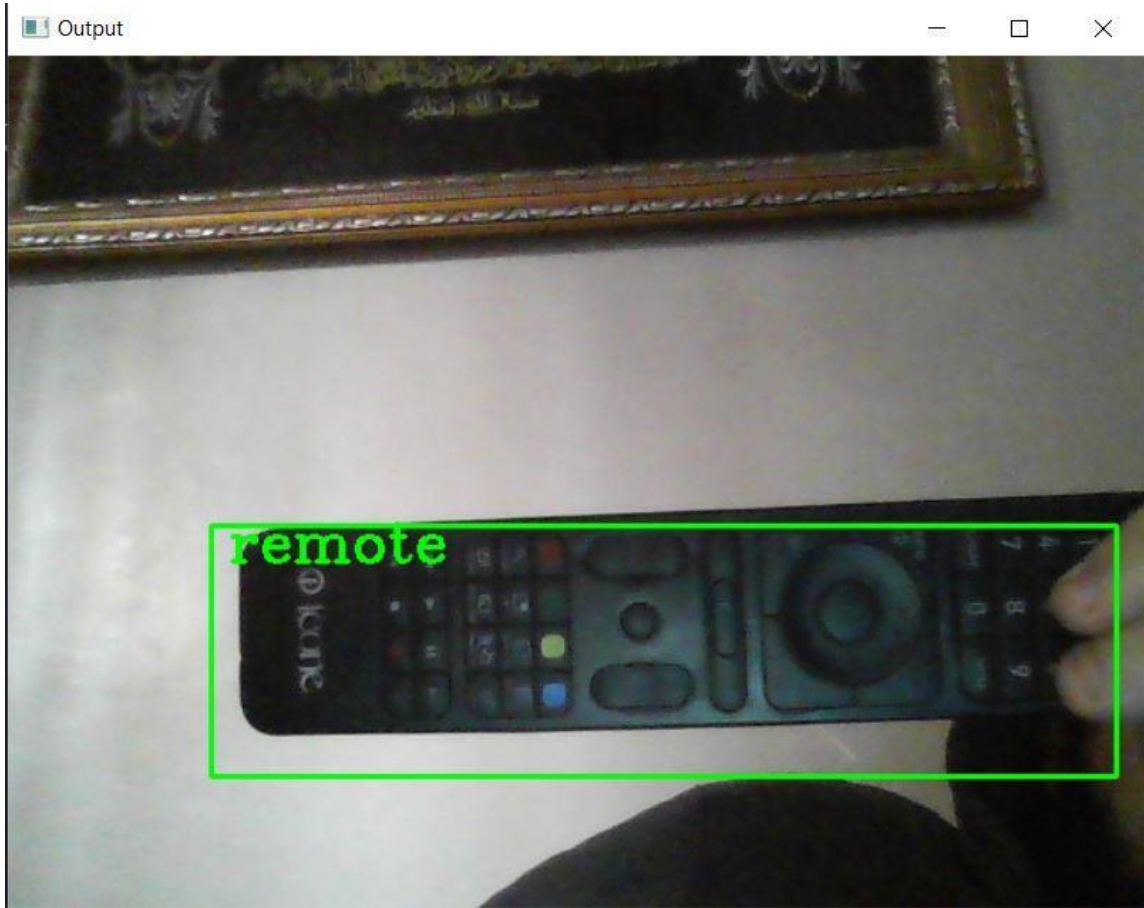
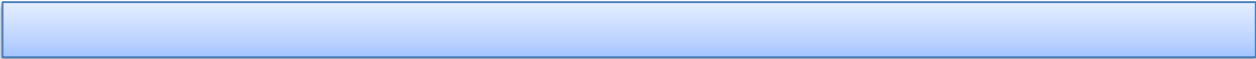
*Figure 4.3* : interface d'authentification.



*Figure 4.4* : interface d'erreur lors de la saisie du mot de passe et nom d'utilisateur.



*Figure 4.5* : interface de succès lors de la saisie du mot de passe et nom d'utilisateur.



*Figure 4.6* : la détection et la reconnaissance d'objets en temps réel.



*Figure 4.7* : la détection et la reconnaissance d'objets en temps réel

## **Conclusion**

Dans ce chapitre nous avons présenté l'étape de la réalisation de notre application ainsi que sa mise en œuvre, les interfaces principales et ses fonctionnalités. Durant l'implémentation des différentes fonctionnalités et des différentes fenêtres, nous avons pris soin de développer une application efficace et capable de détecter des mouvements en temps réel d'un côté, et faire la comparaison entre des images existantes d'un autre côté. Le choix de Python comme langage de programmation nous a permis d'atteindre notre objectif.

# *Conclusion Générale.*

## **Conclusion Générale**

Dans ce projet de fin d'étude, nous avons créé un système de détection et de reconnaissance d'objets qui permet de détecter la présence d'objets spécifiques dans une zone surveillée en temps réel. Nous avons d'abord étudié les différentes approches pour la détection et la reconnaissance d'objets en utilisant des techniques telles que les réseaux de neurones convolutifs, les algorithmes de détection en cascade, et les méthodes basées sur des caractéristiques. Ensuite, nous avons spécifié les besoins fonctionnels et non fonctionnels du système et avons établi une analyse détaillée en utilisant UML pour concevoir l'architecture de notre système. Nous avons également utilisé Python pour implémenter notre système, en tirant parti de ses bibliothèques riches et de sa capacité à gérer l'IA.

Notre système de détection et de reconnaissance d'objets représente une contribution significative dans le domaine de la surveillance automatisée en offrant une solution efficace pour détecter les objets spécifiques en temps réel.

Nous espérons que notre travail pourra servir de base pour des améliorations futures dans ce domaine en constante évolution. En conclusion, ce projet nous a permis d'approfondir nos connaissances dans la détection et la reconnaissance d'objets ainsi que dans l'utilisation de Python pour développer des systèmes intelligents.

# *Bibliographies.*

## Bibliographies

- [1] Valérie Gouaillier, Aude-Emmanuelle Fleurant « La vidéosurveillance intelligente : promesses et défis », Avril, 2009
- [2] Les différents matériels de vidéo surveillance
- [3] RAKOTONARIVO Laza Manampisoa CONCEPTION D'UN SYSTEME INTELLIGENT POUR LA VIDEO SURVEILLANCE UNIVERSITE D'ANTANANARIVO ECOLE SUPERIEURE POLYTECHNIQUE 2019 / 2020
- [4] : Armand Renaud Ngaamou Nana de l'Institut supérieur des technologies et du design industriel du Cameroun, mémoire « Etude et mise en place d'un système de vidéosurveillance cas de l'immeuble Folep à Bali », le 02 septembre 2011, URL : <http://www.memoireonline.com/01/13/6765/Etude-et-mise-en-place-d-un-systeme-devideosurveillance-Cas-de-l-immeuble-Folepe--Bali.html>
- [5] Amar SADAT Youghourta TALEB Etude et mise en place d'un système de vidéosurveillance mobile UNIVERSITE MOULOUD MAMMERI DE TIZI-OUZOU 25/09/2016
- [6] .Boutadara, « Proposition d'une approche intelligente pour la reconnaissance d'actions humaines à partir d'image de vidéosurveillance », Université Ahmed Draia – Adrar, 2016
- [7]Introduction\_reseau\_local\_prive, Support de formation.pdf
- [PPTD, 2015] PPTD, "Aide-mémoire pour l'installation d'une vidéosurveillance ", 2015/domaine d'application de la vidéo
- [DM, 2012] Djamila Mokhtari, Université de Montréal Département d'Informatique et de Recherche Opérationnelle Faculté des Arts et des Sciences, " Détection des chutes
- [8] J-L. Baril "Outils Mathématique pour l'informatique" Université de Bourgogne Labo.le2i,UMR-CNRS 5158. 2007.
- [9]Jacky Desachy, « Analyse d'image » notes de cours - version 1.3, Université des Antilles et de la Guyane Pointe à Pitre, pp 5-10. janvier 2001
- [10]Guesdon VINCENT,"Détection efficace de contours d'images" ,université du Québec en Outaouais, Octobre 2004.
- [11] Y. Kabir " Segmentation d'images de films de radiographie dédiée au contrôle non destructif (CND) ", Thèse de magister, université de Saad Dahleb, Blida,1999
- [12]: Chérif TAOUCHE, " Implémentation d'un Environnement Parallèle pour la Compression d'Images à l'aide des Fractales", Université Mentouri Constantine, 2005.
- [13]M .T.Chikh , amélioration des images par un modèle de réseau de neurones ccccc (comparaison avec les filtres de base), université abou-bakr belkaid tlemcen.2011

- [14]K. Aounallah, les approches de segmentation d'image par coopération régions-contours, 2010
- [15]M Sandeli, traitement d'images par des approches bio-inspirées application à la segmentation d'images, université constantine 2. 2014
- [16] : Thi-Lan le, "Indexation et recherche de vidéo pour la vidéo surveillance", université de NICE-SOPHIA ANTIPOLIS, 3 février 2009.
- [17] :Djamila Mokhtari, " Détection des chutes par calcul homographique ", Université de Montréal Département d'Informatique et de Recherche Opérationnelle Faculté des Arts et des Sciences, Août 2012
- [18] : MEDJAHED Fatiha, " Détection et Suivi d'Objets en Mouvement Dans Une Séquence d'Images ", Université Mohamed Boudiaf Oran, 2011.
- [19]: David Salomon, "Data-compression-the-complete-reference ", 2007.
- [20] Touati R. , "Reconnaissance des actions humaines à partir d'une séquence vidéo", Mémoire pour l'obtention du grade de Maîtres sciences (M.Sc.) en informatique, Université de Montréal, Janvier 2014
- [21] : Ahmed Ben Atitallah, " Etude et Implantation d'Algorithmes de Compression d'Images dans un Environnement Mixte Matériel et Logiciel", thèse de doctorat en électronique, Université Bordeaux 1, Juillet 2007.
- [22] : Nicolas DUMOULIN, "Compression de séquences vidéo choix des images de références ", Université de Rennes1-IFSIC, Juin 2003.
- [23] H.YEDJOUR. Détection de contours et suivi d'objet dans une séquence d'images par les réseaux de neurones impulsionnels Université Mohamed Boudiaf (USTO) ORAN Thèse Magister 2010.P13,14,P16,17,18,19,20,21,22 ;23,P24.
- [24]BharatiD. PatilRupaliS . Rakibe. Background subtraction algorithm based human motion detection. International Journal of Scientific and Research Publications, May 2013.
- [24]DarshakGThakoreHiteshAPatel.Movingobjecttrackingusingkal-manlter. International Journal of Computer Science and Mobile Com-puting, page 326332, April2013.
- [25] Jae-YeongLee Wonpil Yu.Visual tracking by partition-based histogram back projection And maximum support criteria. In International Conference on Robotics and Biomimetic2011. [26] C. MeenaM.Sankari.Estimation of dynamic background and object detection in noisy visual surveillance. International Journal of Advanced Computer Science and Applications, pages 7783, 2011.
- [27] : Cao Tien Dung, " La vidéos surveillance ", Institut de la Francophonie pour l'Informatique, Janvier 2007.
- [28] : HORAUD, Radu, MONGA, Olivier, "Vision par ordinateur", 1995

- [29] : Centre de recherche informatique de Montréal, "La vidéosurveillance intelligente", Avril 2009.
- [30]: Sung Wook Seol, Jee Hye Jang, Hyo Sung Kim, Chul Hun Lee, and Ki Gon Nam "Detection and Tracking System of Moving Objects Using Double Difference based Motion Estimation", 2003
- [31] : Centre de recherche informatique de Montréal, "La vidéosurveillance intelligente", Avril 2009.
- [32] : R. Collins, A. Lipton, T. Kanade, H. Fujiyoshi, D. Duggins, Y. Tsin, D. Tolliver, N. Enomoto, O. Hasegawa, P. Burt et L. Wixson, "A System for Video Surveillance and Monitoring", Robotics Institute, May 2000.
- [33] : Thierry Chateau, Antoine Vacavant, " Real Time tracking of human gestures using color image processing", Université Blaise Pasca France, Juin 2004
- [34] : Taoufik EL Kabir, Amine Boukouta, " Reconstruction du mouvement humain par système de vision ", INRIA, Christine Azevedo, 2013
- [35]D. Zeroual, implémentation d'un environnement parallèle pour la compression d'image à l'aide des fractales, Thèse de Doctorat, Batna, 2006.
- [36]L. Bouanzi, Les formats de numération des images fixes, DESS en informatique Documentaire, Rapport de stage, Ecole Nationale Supérieure des Sciences de l'Information et des Bibliothèques (ENSSIB), Université Claude Bernard Lyon 1, 1999.
- [37] S. Katzenbeisser, F. Petitcolas, Information Hiding Techniques for steganography and Digital Watermarking, Artech House, London, 2000
- [38]J. Luc, Les images numériques, Généralités, Cours TERI, Diplôme d'Université Assistant à l'Usage des Technologies de l'Information et de la Communication (DU AUTIC) DEUST, Université de Bretagne Occidentale, 2002/2003
- [39] N.Boutadara, « Proposition d'une approche intelligente pour la reconnaissance d'actions humaines à partir d'image de vidéosurveillance », Université Ahmed Draia – Adrar, 2016

## **SITE INTERNET**

- [s1] <http://www.inter-assistance.com/videosurveillance-guide.html>
- [S2]<https://www.technologuepro.com/cours-videosurveillance/Chapitre-3-evolution-des-systemes-de-videosurveillance.html>

- [S3] <http://videosurveillance.comprendrechoisir.com/comprendre/videosurveillance-ip>
- [S4] <http://www.techno-science.net/definition/11698.html>
- [s5] <http://tpemultimedia.e-monsite.com/pages/qu-est-ce-que-la-videosurveillance.html>
- [s6] <https://iast.univsetif.dz/documents/Cours/CoursInformatiqueL1GAT21.pdf>
- [S7]: [www.acnice.fr/ia06/iencagnes/file/tice/tuto/Image\\_numerique.pdf](http://www.acnice.fr/ia06/iencagnes/file/tice/tuto/Image_numerique.pdf)
- [S8]: [www.crdp.ac-grenoble.fr/image/general/general.htm](http://www.crdp.ac-grenoble.fr/image/general/general.htm)
- [S9]: [webcache.googleusercontent.com/search?q=cache:http://obligement.free.fr/articles/codagedesimages.php](http://webcache.googleusercontent.com/search?q=cache:http://obligement.free.fr/articles/codagedesimages.php).
- [S10]: [www.mediatheque.mc/dotAsset/1bc69e0a-28a1-4399-8193-6a831697f818.pdf](http://www.mediatheque.mc/dotAsset/1bc69e0a-28a1-4399-8193-6a831697f818.pdf)
- [S11]: [http://www.foad-mooc.auf.org/IMG/pdf/D226\\_Chapitre-5.pdf](http://www.foad-mooc.auf.org/IMG/pdf/D226_Chapitre-5.pdf)
- [S12]: [www.lagis.univ-lille1.fr/formations/sii/06-07/projets/cmucam-rapport.pdf](http://www.lagis.univ-lille1.fr/formations/sii/06-07/projets/cmucam-rapport.pdf) [S13]: <http://S2-e-monsite.com/2010/03/16/85124207image-numerique.pdf>
- [S14]: [www.st-tremeur.fr/courstechno/imagenum/www.crdp.ac-grenoble.fr/image/general/img\\_num.htm](http://www.st-tremeur.fr/courstechno/imagenum/www.crdp.ac-grenoble.fr/image/general/img_num.htm).
- [S15]: <http://tecfa.unige.ch/tecfa/teaching/staf13/fiches-mm/formatfichier.htm>
- [S16]: <http://www.vision.uji.es/~montoliu/docs/pfm/AnneClaireMaheo.pdf>
- [S17]: <http://gillesboulet.ca/textes/VideoNumerique.pdf>.
- [S18]: <http://www.nymphomath.ch/info/images/images.pdf>
- [S19]: <http://www.tdechardin.org/web/page.php?id=1350>.
- [S20]: <http://www.techno-science.net/?onglet=glossaire&definition=7376>
- [S21] <http://www.techno-science.net/?onglet=glossaire&definition=7376>
- [S22]: [icar.univlyon2.fr/projets/corinte/documents/Acquisition\\_et\\_compression.ppt](http://icar.univlyon2.fr/projets/corinte/documents/Acquisition_et_compression.ppt)
- [S23]: [http://lucbor.fr/formats\\_videos\\_et\\_codecs.pdf](http://lucbor.fr/formats_videos_et_codecs.pdf)
- [S24]: [https://climat63.files.wordpress.com/2012/04/format\\_vid3a9o.pdf](https://climat63.files.wordpress.com/2012/04/format_vid3a9o.pdf)
- [S25]: <http://benjlaiel.e-monsite.com/medias/files/video-numerique.pdf>
- [s26] <https://openai.com/blog/chatgpt>



[S27] [https://www.fritz.ai/object-detection/visité le 20/02/2022](https://www.fritz.ai/object-detection/visité%20le%2020/02/2022)

[S28]:<http://docslide.fr/download/link/27714afb8a63911930a39ea51e575c0e-pfe-oussama>

[s29] <https://fr.wikipedia.org/wiki/OpenCV>

[S30] : <https://www.techno-science.net/glossaire-definition/Python-langage.html>

[S31] <https://fr.wikipedia.org/wiki/PyCharm>

[S32]<https://fr.mathworks.com/solutions/image-video-processing/object-recognition.html>

[S33][http://www.gcya-holding.com/pages/all/alltech/pages/video\\_surveillance.php](http://www.gcya-holding.com/pages/all/alltech/pages/video_surveillance.php)