

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

UNIVERSITE MOULOUD MAMMARI DE TIZI-OUZOU



FACULTE DU GENIE ELECTRIQUE ET D'INFORMATIQUE
DEPARTEMENT D'INFORMATIQUE

Mémoire de Fin d'Etudes De MASTER ACADEMIQUE

Domaine : **Mathématiques et Informatique**

Filière : **Informatique**

Spécialité : **Conduite de Projet Informatique**

Présenté par

Prénom : Nouria NOM : Ouaras

Prénom : Fazia NOM : Tabani

Thème

Annotation Vidéo Basé Sur MapReduce Dans le Domaine Vidéosurveillance

Mémoire soutenu publiquement le ...13.../...07.../2016..... Devant le jury composé de :

Président : M MHEND KERBICHE

Encadreur : M AHMED DIB

Co-Encadreur : M Prénom NOM

Examineur : M Loukman Chebouba

Examineur : M Bourkache G

Remerciements

Nous tenons à exprimer nos sincères remerciements à notre promoteur Mr DIB pour l'intérêt et l'aide qu'il nous a donné.

Nos remerciements vont également aux membres du jury pour l'honneur qu'ils nous font en acceptant d'examiner et de juger notre travail.

Nous remercions aussi tous ceux, et celles qui ont contribué de près ou de loin pour l'accomplissement de ce modeste travail.

Dédicaces

Je dédie ce modeste travail à :

A mes chers parents qui m'ont beaucoup aidé durant ma vie et surtout dans mes études. je les remercie chaleureusement pour leur compréhension et leur soutien.

A mes très chers frères et mes sœurs.

A mes beaux frères et belles sœurs

A mes adorable nièces : Melissa, Amel, Saïd, Lydia, Yassine, Tiziri , Yasmine.

A mon mon mari Madjid ainsi sa famille

A mon binôme Nouria ainsi sa famille.

A tous mes amis(es), surtout dyhia, koka, lynda, naima, monia.

A toutes ma famille.

A tous ceux qui m'ont aidé de près ou de loin à la réalisation de ce travail.

Tabani Fazia.

Dédicaces

Je dédie de fond de mon cœur ce

Petit travail à :

- ❖ *A mon cher père **HACENE** et ma chère mère **ZHOR** ;*
- ❖ *A mon Fiancé **HOCINE** ;*
- ❖ *A mes chers frères : **IDIR** et sa petite famille (**KENZA** et **LILICHE,Lina**), **MOURAD** et sa fiancée **Hamida**, et **SAMIR** ;*
- ❖ *Mes quatre sœurs : **ZAKIA** et son marie **HACENE**, **SOUHILA** et son marie **FERHAT**, **Sonia** et son marie **youcef**, **LYNDA** et son marie **Arzki**, et ma Chère tante **FEROUDJA** et ses deux filles **OUARDIA** et **ZHOR** ;*
- ❖ *A toute la famille **OUARAS** de plus petit au plus grand.*
- ❖ *A tous mes Amis (**Monia**, **Henia**, **Naima**, **Fazia**, **Zohra**, **Zanida**) pour aide et encouragement.*

A tous ceux qui m'ont aidé de près ou de loin à la réalisation de ce travail.

Nouria.

Sommaire

Introduction générale

Chapitre I : Cloud computing et Hadoop

Introduction	1
I. Cloud computing	1
I.1. Historique	1
I.2. Définition	2
I.3. Architecture de Cloud computing	3
I.3.1 .Iaas : Infrastructure as a service	3
I.3.2.Paas: Platform as a service	3
I.3.3.SaaS: Software as a service	3
I.4. Modèles de service associés	5
I.5.Caractéristique d'un Cloud	5
I.6. Modèle de déploiement de Cloud computing	5
I.6.1 : Clouds publics	6
I.6.2.Clouds privés	6
I.6.3.Clouds communautaires	6
I.6.4.Clouds hybrides	6
I.7. Quelques avantages et inconvénients d'un Cloud	6
I.7.1. Avantages	7
I.7.2.Inconvénients	7
II. Hadoop	8
II.1 Historique	8
II.2. Présentation d'Hadoop	8
II.3.Typologie d'un cluster Hadoop	9
II.4. HDFS	11
II.5. Le composant MapReduce	15
III. MapReduce.....	16
III.1 : Définition	16
III.2. Principe de MapReduce	17
III.3 : Le Modèle de Programmation MapReduce	17

Conclusion	18
-------------------------	----

Chapitre II : Vidéo_surveillance

Introduction	19
I. Définition	19
II. Les différents types de caméra de surveillance	20
III.les différents types de systèmes	23
III .1 Système sur réseaux IP	23
III.2 Kit de vidéosurveillance	24
III.3 Réseau « classique » de vidéosurveillance	25
III.4 Système « hybride » de vidéo surveillance	25
VI. Technologie	26
VI.1 Architecture des systèmes de vidéosurveillance	26
VI.1.1 Composantes d'un système de vidéosurveillance	20
VI .1.1.1 Acquisition	27
VI.1.1.2. Transmission	29
VI.1.1.3 Compression	29
VI.1.1.4 Traitement	30
VI.1.1.5 Archivage	30
VI.1.1.6 Affichage 31	
V .Evolution des systèmes de vidéosurveillance	32
V .1 Systèmes de vidéosurveillance classique CCTV.....	32
V.2 Systèmes de vidéosurveillance analogique avec magnétoscopes traditionnels	32
V.2.1 Multiplexeurs	33
V.3. Systèmes de vidéosurveillance analogique avec enregistreurs numériques	33
V.4 Systèmes de vidéosurveillance analogique avec enregistreurs numériques réseau	34
V.5Systèmesde vidéo sur IP avec serveurs vidéo	35
V.5.1 Serveurs vidéo	36
V.5.2 Décodeur vidéo	36
3.6 Systèmes de vidéo sur IP avec cameras réseau	37
V.7 Réseau sans fil	38
IV .Domaines d'applications	40

VI- Le traitement d'image.....	40
VI.1 Définition d'une vidéo	40
VI.2- Définition d'une image	41
VI.3. Caractéristiques d'une image	41
VI.4.Les caractéristiques locales et globales.....	42
1) Les caractéristiques globales.....	42
2) Les caractéristiques locales	44
VI.5.Technique de détection de caractéristiques locales et caractéristiques globales.....	44
VI.6.Descripteurs de caractéristiques d'image	46
Conclusion	48

Chapitre III : Annotation de vidéo

I. Introduction.....	49
II. Présentation d'annotation de vidéo.....	49
III. L'annotation vidéo basée sur les images clé	50
IV. Etat de l'art sur les techniques et les méthodes d'annotation vidéo.....	51
IV.1.1 Méthodes basées sur l'échantillonnage :.....	52
IV.1.2 Méthodes basées sur les plans :.....	52
IV.2 L'annotation vidéo dynamique :.....	53
V. Traitement vidéo de la base de données sur hadoop :.....	56
V.1.MapReduce pour traitement la base de données de la vidéo :.....	56
V.2.Traitement MapReduce de base de données vidéo:.....	57
VI. Etat de l'art sur l'annotation de Vidéo à grand échelle:.....	58
Conclusion :.....	59

Chapitre VI : Annotation de vidéo

I. Introduction :.....	64
II. Architecture de Système :.....	65
III. La méthode d'annotation utilisée	66
IV. Algorithme d'annotation avec MapReduce :.....	67
V. La plateforme utilisé:	69
1. Machine Virtuelle :.....	70

2. Virtual Box :.....	70
2.1 Présentation :.....	70
2.2 Fonctionnement global de VirtualBox :.....	70
2.3 Créations d'une machine virtuelle ubuntu :.....	71
4. OpenImage :.....	71
VI. Réalisation avec des captures d'écran :	72
Conclusion :	81

Liste des Figures

Chapitre I :

Figure I. 1: Illustration d'un Cloud Computing	2
Figure I. 2 : représente l'Architecture d'un Cloud computing.....	3
Figure I.3: Modèle de déploiement de cloud computing	5
Figure II 1 Architecture d'un Cluster Hadoop	10
Figure II. 2.: Schéma HDFS	12
Figure II 3. : Réplication des données à l'aide du HDFS.....	13
Figure II.4 Hadoop Distributed File System (HDFS)	14
Figure II.5 : Fonctionnement de MapReduce dans Hadoop.....	16
Figure III.1 : les étapes Map et Reduce.....	17

Chapitre II :

Figure I.1 : la vidéo de surveillance	19
Figure.III.2 : Vidéo de surveillance sur réseau IP	24
FigureIII.3 : Kit de vidéo surveillance	24
Figure III.4 : système de vidéo surveillance analogique	25
Figure IV.5 : schéma de principe	27
Figure V.6 : Vidéo surveillance classique CCTV	32
Figure V.7 : Vidéosurveillance analogique avec magnétoscopes traditionnels	33
Figure.8 : Vidéosurveillance analogique avec enregistreurs numérique.....	34
Figure V .9 : Vidéosurveillance analogique avec enregistreurs numérique réseau.....	34
Figure V .10 : Vidéo sur IP avec serveur vidéo	35
Figure V 11 : Description du système	36
Figure V.12 : Décodeur vidéo	37
Figure V 13 : système de vidéo sur IP avec cameras réseau.....	38
Figure VI 14 : deux découpages différents de l'image	39
Figure VI.16 : Dérivées de gaussienne, jusqu'à l'ordre	40

Chapitre III :

Figure I .1 : Keyframe ont basé sur annotation vidéo	52
Figure IV.2 :La structure hiérarchique d'une séquence Vidéo	54
Figure V.3 : Flux de traitement d'image utilisant MapReduce.....	61
Figure V.4 : Flux de traitement séquence d'image vidéo qui utilise Map et HDFS	62
Figure V.1: L'interface de Hibernate	76
Figure V.2 :L'interface EasyPHP.....	77
Figure V.3 :L'interface PhpMyAdmin.....	77
Figure V.4 :L'interface de Netbeans	78
Figure V.5 :L'interface de JEE	79
Figure V.6 : Interface page d'accueil	82
Figure V.7 : Interface authentification	82
Figure V.8 : Interface d'un produit	83
Figure V.5 : Interface ajouter un produit.....	83

Introduction générale

En quelques années, le volume des données exploité par les entreprises a considérablement augmenté. Emanant de sources diverses (transaction, systèmes d'information automatisés, réseaux sociaux), elles sont souvent susceptibles de croître très rapidement. Lorsqu'on parle de manipulation de gros volume de données, on pense généralement à des problématiques sur le volume des données et sur leur rapidité de traitement qu'on trouve généralement dans la vidéo surveillance avec ce nombre de déploiements de leur systèmes qui augmente régulièrement et constitués par de plus en plus de caméras, les opérations de vidéo surveillance ont des difficultés grandissantes d'exploitation dues aux grandes quantités de données vidéo stockées chaque jour.

Dans une solution efficace qui est l'annotation de vidéo à pour objectif de résumer les tâches d'analyse vidéo basée sur le contenu en utilisant le populaire Hadoop Apache et modèle de programmation MapReduce .

Pour ce faire, et pour mener à bien notre travail nous avons opté pour la démarche articulée autour des chapitres suivants :

Le premier chapitre consacré à la présentation des concepts du cloud computing et framework Hadoop, la description de ses principaux composants HDFS et sa propre version de MapReduce.

Le deuxième chapitre présentation de vidéo surveillance et ces systèmes, nous présentons les concepts généraux sur le traitement d'image, dans lequel nous allons faire un aperçu sur l'image, ses caractéristiques.

Le troisième chapitre, que nous intitulerons, nous verrons en quoi consistent l'annotation vidéo proprement dite, les techniques et les différentes méthodes d'annotation existantes, et pour finir, nous verrons quelques travaux apparentés au notre.

Le quatrième chapitre, la présentation des différents outils de développement utilisés lors de la mise en œuvre effective de l'application, ainsi que la description de son fonctionnement et l'analyse des résultats obtenus.

Nous terminons notre mémoire avec une conclusion générale.

Introduction :

Indéniablement, la technologie de l'internet se développe de manière exponentielle depuis sa création. Actuellement, une nouvelle « tendance » à fait son apparition dans le monde des IT (Technologies de l'information et de la communication), il s'agit du Cloud Computing. Cette technologie, s'appuie sur le WEB 2.0, offre des occasions aux sociétés de réduire les coûts d'exploitation des logiciels par leurs utilisations directement en ligne. Divers fournisseurs comme Google, Amazon, IBM offrent une vaste gamme de services de cloud computing. [1] Les services Cloud sont optimisés pour les applications d'analytique et de Technique Computing utilisant des gros volumes de données, mais également pour toute une gamme de distributions Hadoop, qui font désormais partie intégrante du nouveau IBM High Performance Hadoop Service. Apache Hadoop (High-availability distributed Object-oriented Platform) est un système distribué qui répond à ces problématiques. D'une part, il propose un système de fichier distribué HDFS (Hadoop Distributed File System) pour assurer le stockage et l'intégrité des données en dupliquant plusieurs copies d'un même bloc à travers des dizaines, des centaines, voire des milliers de machines différentes, ce qui amène un cluster Hadoop à être configuré sans requérir un système RAID. D'autre part, Hadoop fournit un système d'analyse de données appelé MapReduce pour réaliser des traitements sur des gros volumes de données grâce à sa répartition efficace du travail sur différents nœuds de calcul.

Dans ce chapitre nous allons présenter la notion fondamentale du cloud computing, Hadoop et MapReduce.

I. Cloud computing :

I.1. Historique :

Il n'y a pas de date-clé à laquelle nous puissions dire que le Cloud computing est né. La notion de Cloud fait référence à un nuage, tel que l'on a l'habitude de l'utiliser dans des schémas techniques lorsque l'on veut représenter internet. Un réseau comme internet est constitué d'une multitude de systèmes fournissant des services et des informations. Le Cloud computing est dans cette ligne : un ensemble de service et de données consommables.

Cette notion de consommation a été proposée en 1961, lors d'une conférence au MIT (Massachusetts Institute of Technology), par John McCarthy [1.23] aussi connu comme l'un des pionniers de l'intelligence artificielle (dont il proposa le nom en 1955) et pour avoir inventé LISP en 1958.

Lors de ce discours, John McCarthy suggéra que la technologie informatique partagée (« time-sharing ») pouvait construire un bel avenir dans lequel la puissance de calcul et même les applications spécifiques pouvaient être vendues comme un service public. C'est donc depuis presque 50 ans que l'idée d'une informatique à la demande est présente dans les esprits même si les technologies n'étaient jusqu'alors pas au rendez-vous pour pouvoir concrétiser cette idée. Avec les différents progrès technologiques réalisés durant ces 50 dernières années, tant sur le plan matériel, logiciel et conceptuel, aux avancées des mécanismes de sécurité, à l'élaboration de réseaux complexes mais standardisés comme internet, et à l'expérience dans l'édition et la gestion de logiciels services, infrastructures et stockage de données, il est maintenant possible d'entrer dans l'ère du Cloud computing, telle que rêvait par John McCarthy en 1961.

I.2. Définition :

Le Cloud Computing permet la mise à disposition de ressources informatiques, délocalisées sur Le réseau, sous la forme de services. Le National Institute of Standards and Technology en a donné Une définition générale qui reprend ces principes de base : « L'informatique en nuage est un modèle Permettant d'établir un accès par le réseau à la fois commode et à la demande, à un bassin partagé de Ressources informatiques configurables (par exemple : réseaux, serveurs, stockage, applications et Services) qui peuvent être rapidement mobilisées et mises à disposition avec un minimum d'effort De gestion ou d'interaction avec le fournisseur de services [1]. ».

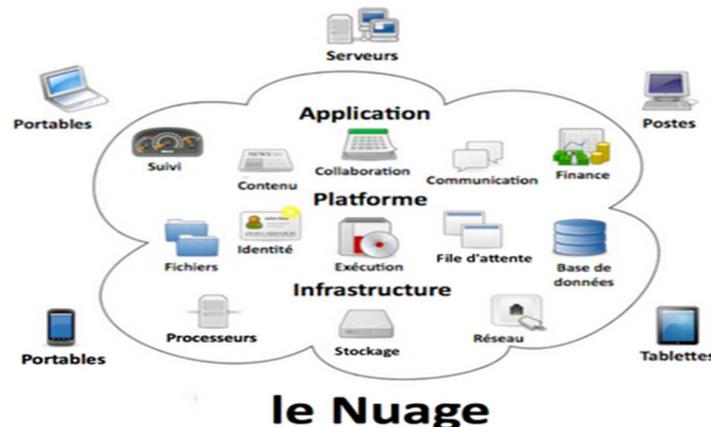


Figure I.1: Illustration d'un Cloud Computing

I.3. Architecture de Cloud computing :

Le Cloud computing propose trois modèles principaux :

- L'infrastructure (IaaS : Infrastructure as a service)
- La plate-forme (PaaS : Platform as a service)
- L'application (SaaS : Software as a service)



Figure I.2 : représente l'Architecture d'un Cloud computing

I.3.1 .IaaS : Infrastructure as a service :

L'infrastructure en tant que service ou, en anglais Infrastructure as a service (IaaS) est le premier modèle de Cloud on désigne une infrastructure matérielle, louée à la demande : stockage, machines virtuelles OS, etc.

I.3.2.PaaS: Platform as a service:

La plate-forme en tant qu'un service désigne la mise à disposition d'un environnement de développement et d'exploitation de logiciels sur Internet.ces plates-formes sont donc directement utilisable par des éditeurs qui proposeront leurs logiciels en mode Cloud.

I.3.3.SaaS: Software as a service:

Le logiciel en tant qu'un service désigne la mise à disposition par Internet d'applications informatiques (logiciels) comme un service dans le cadre d'un abonnement, c'est en quelque sorte la partie visible du Cloud computing pour l'utilisateur final, qui n'a plus besoin d'installer l'application sur son poste, et qui accède à son compte par le web, Le SaaS, souvent associé au « Cloud computing » peut être vu comme

un modèle économique de consommation des applications : celles-ci sont consommées et payées à la demande (par utilisateur et par minute d'utilisation par exemple) et non plus acquises par l'achat de licences.

Principaux acteurs :

- Google(Apps)
- Zoho(Apps)
- Salesforce .com, Microsoft, IBM, etc.

L'Infrastructure en tant que service offre une base matérielle (hardware) aux plateformes en tant que service. Ces infrastructures sont mises en place et gérés par des administrateurs réseau avec un bon niveau d'expertise. Elles sont le plus souvent constituées d'équipements réseaux et de serveurs la plus part du temps entièrement virtualisés.

La plateforme en tant que service est un ensemble de composants reposant sur l'infrastructure offerte par la couche IaaS. Elle permet aux développeurs d'applications d'avoir une plateforme de travail adaptable, distribuée et virtualisée dans laquelle ils n'ont plus besoin d'installer l'architecture sous-jacente (réseaux, matériels, serveur, système d'exploitation)

L'application en tant que service est une application souple et déployée dans une plateforme en tant que service. C'est une application souple qui est accessible uniquement à travers un réseau et qui est le plus souvent facturée à l'utilisateur final.

I.4. Modèles de service associés :

A côté de ces modèles de services, on trouve une multitude de modèles de service construits mécaniquement sur l'expression « as a service ». Voici par ordre alphabétique :

- ✓ **BPaaS** : il s'agit du concept de Business Process as a service (BPaaS) qui consiste à externaliser une procédure d'entreprise suffisamment industrialisée pour s'adresser directement aux managers d'une organisation, sans nécessiter l'aide de professionnels de l'informatique.
- ✓ **DaaS** : Le Desktop as a Service (aussi appelé en français « bureau en tant que service », « Bureau virtuel » ou « bureau virtuel hébergé ») est l'externalisation d'une Virtual Desktop Infrastructure auprès d'un fournisseur de service. Généralement, le Desktop as a Service est proposé avec un abonnement payant.
- ✓ **Network as a Service (NaaS)** : le Network as a service correspond à la fourniture de service réseaux, suivant le concept de Software Defined Networking(SDN).

- ✓ **STaaS** : SStorage as a Service correspond au stockage de fichiers chez des prestataires externes ,qui les hébergent pour le compte de leurs clients .Des services grand public ,tels que SugarSync et Box .net ,proposent ce type de stockage ,le plus souvent à des fins de sauvegarde ou de partage de fichiers. Voici d'autres exemples : Amazon Simple Storage Service, Dropbox, Google Drive, Ubuntu One.

I.5.Caractéristique d'un Cloud :

- ✓ **Libre service à la demande** : le consommateur récupère des ressources de calcul et de stockage à la demande (machine virtuelle).
- ✓ **Elasticité** : Croissance ou décroissance dynamique du nombre de ressources de nuage en fonction de la demande et des besoins donnant une vue de ressources infinies, avec ce dynamisme dès qu'une ressource se libère, on peut la louer à n'importe quel moment à un autre client.
- ✓ **Large accès au réseau** : Des services sont mis tous au long du réseau et peuvent être accéder par des mécanismes standard qui favorisent l'utilisation et l'accès par de plateformes hétérogènes des clients (ex : les PC portable, mobile, PDA, etc.).
- ✓ **Une mise en commun des ressources** : les ressources de calcul sont mises à disposition des clients sur un modèle multi-locataires, avec une attribution dynamique des ressources physiques et virtuelles en fonction de la demande .le client n'a généralement aucun contrôle ou connaissance sur l'emplacement exact des ressources fournies (ressources cachées).Toutefois ,le client peut imposer de spécifier l'emplacement à un niveau plus haut d'abstraction (par exemple le pays ,l'état ou le Data Center).
- ✓ **Service mesuré et facturation à l'usage** : la facturation est calculée en fonction de la durée et de la quantité de ressources utilisées.

I.6. Modèle de déploiement de Cloud computing :

Les nuages de calcul peuvent être classés en différentes modèle de déploiement .On observe les clouds publics, privés, communautaires, hybrides comme le montre la figure



Figure I.3: Modèle de déploiement de cloud computing

I.6.1 : Clouds publics :

Le principe est d'héberger des applications, en général des applications Web, sur un environnement partagé avec un nombre illimité d'utilisateurs. La mise en place de ce type de Cloud est gérée par des entreprises tierces (exemple Amazon, Google, etc.). Les fournisseurs du Cloud publique, les plus connus sont Google et Amazon. Ce modèle est caractérisé par :

- Mise en place de lourds investissements pour le fournisseur de services.
- Offre un maximum de flexibilité.
- N'est pas sécurisé.

I.6.2. Clouds privés :

L'infrastructure Cloud fonctionne pour une organisation unique .Elle peut être gérée par l'organisation elle-même (Cloud Privé interne) ou par un tiers (Cloud Privé externe). Dans ce dernier cas l'infrastructure est entièrement dédiée à l'entreprise et accessible via réseaux sécurisés.

I.6.3. Clouds communautaires :

L'infrastructure est partagée par plusieurs organisations qui ont des intérêts communs (par exemple les exigences de sécurité, de conformité...). Comme le Cloud privé, il peut être géré par les organisations elles-mêmes ou par un tiers.

I.6.4. Clouds hybrides :

L'infrastructure se compose de deux nuages ou plus (privé, communautaire ou public), qui restent des entités uniques, mais qui sont liées par une technologie normalisée ou propriétaire, permettant la portabilité des données ou des applications.

I.7. Quelques avantages et inconvénients d'un Cloud :

I.7.1. Avantages :

- ✓ Un démarrage rapide: Le cloud computing permet de tester le business plan rapidement, à coûts réduits et avec facilité.
- ✓ L'agilité pour l'entreprise: Résolution des problèmes de gestion informatique simplement sans avoir à vous engager à long terme.

- ✓ Un développement plus rapide des produits: Réduisons le temps de recherche pour les développeurs sur le paramétrage des applications.
- ✓ Pas de dépenses de capital: Plus besoin des locaux pour élargir vos infrastructures informatiques

I.7.2. Inconvénients :

- ✓ La bande passante peut faire exploser votre budget: La bande passante qui serait nécessaire pour mettre cela dans le Cloud est gigantesque, et les coûts seraient tellement importants qu'il est plus avantageux d'acheter le stockage nous-mêmes plutôt que de payer quelqu'un d'autre pour s'en charger.
- ✓ Les performances des applications peuvent être amoindries: Un Cloud public n'améliorera définitivement pas les performances des applications.
- ✓ La fiabilité du Cloud: Un grand risque lorsqu'on met une application qui donne des avantages compétitifs ou qui contient des informations clients dans le Cloud.
- ✓ Taille de l'entreprise:

Si votre entreprise est grande alors vos ressources sont grandes, ce qui inclut une grande Consommation du cloud. Vous trouverez peut-être Plus d'intérêt à mettre au point votre propre Cloud Plutôt que d'en utiliser un externalisé. Les gains sont bien plus importants quand on passe d'une Petite consommation de ressources à une consommation plus importante.

Cloud Computing regroupe plusieurs nouvelles technologies et les clouds privés sont largement utilisés dans les entreprises modernes. Une mise en œuvre bien connue de privé nuages est Apache Hadoop, qui est un open-source Cadre de logiciel pour traiter un volume important de données sur un cluster. Généralement, un Cloud privé services utilisateurs multiples. Chaque utilisateur peut avoir des priorités différentes, les types de tâches et les données tailles. En règle générale, de telles variations sont significatives de l'utilisateur, ce qui conduit à un défi pour la planification des tâches.

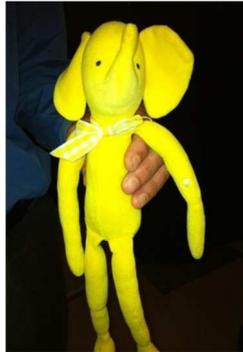
II. Hadoop :

II.1 Historique :

2003/2004: publication par Google de deux whitepapers, le premier sur GFS (un système de fichier distribué) et le second sur le paradigme Map/Reduce pour le calcul distribué. En 2004 c'est le premier développement version du framework qui deviendra Hadoop par Doug Cutting (archive.org). Doug Cutting (désormais chez Yahoo) développe en 2006 une première version exploitable de Apache Hadoop pour l'amélioration de l'indexation du moteur de recherche. Encore primitif (au maximum quelques machines, etc.). En 2008 développement très abouti, Hadoop utilisé chez Yahoo dans plusieurs départements.

En **2011**: Hadoop désormais utilisé par de nombreuses autres entreprises et des universités, et le cluster Yahoo comporte 42000 machines et des centaines de peta-octets d'espace de stockage.

Le nom lui-même n'est pas un acronyme: il s'agit du nom d'un éléphant en peluche du fils de l'auteur originel (Doug Cuttin) de Hadoop.



II.2. Présentation d'Hadoop :

Hadoop est un ensemble de logiciels et d'outils qui permettent de créer des applications distribuées. C'est une plate-forme logicielle open-source, écrite en Java et fondée sur le modèle MapReduce de Google et les systèmes de fichiers distribués (HDFS). Elle permet de prendre en charge les applications distribuées en analysant de très grands ensembles de données.

Hadoop est utilisé particulièrement dans l'indexation et le tri de grands ensembles de données, le data mining, l'analyse de logs, et le traitement d'images. Le succès de Google lui est en partie imputable. En 2001, alors qu'il n'en est encore qu'à ses balbutiements sur le marché des moteurs de recherche, le futur géant développe ce qui inspira les composants phares d'Hadoop: MapReduce, Google BigTable et Google BigFiles (futur Google File System). Ces deux points forment l'écosystème Hadoop, écosystème fortement convoité et qui se trouve au centre de l'univers du Big Data [3].

Hadoop fait partie des projets de la fondation de logiciel Apache depuis 2009. Il est destiné à faciliter le développement d'applications distribuées et scalables, permettant la gestion de milliers de nœuds ainsi que des pétaoctets de données. En 2011, Yahoo! crée Horton works, sa filiale dédiée à Hadoop. L'entreprise se concentre sur le développement et le support d'Apache Hadoop. De la même manière, Cloudera, créé au début de l'année 2009, se place comme l'un des plus gros contributeurs au projet Hadoop, au côté de MapR (2009) et Hortonworks.

II.3. Typologie d'un cluster Hadoop :

Hadoop repose sur un schéma dit « maître-esclave » et peut être décomposé en cinq éléments. (Figure II.1) :

- **Le nom du noeud (Name Node) :** Le « Name Node » est la pièce centrale dans le HDFS, il maintient une arborescence de tous les fichiers du système et gère l'espace de nommage. Il centralise la localisation des blocs de données répartis sur le système. Sans le « Name Node », les données peuvent être considérées comme perdues car il s'occupe de reconstituer un fichier à partir des différents blocs répartis dans les différents « Data Node ». Il n'y a qu'un « Name Node » par cluster HDFS.
- **Le gestionnaire de tâches (Job Tracker) :** Il s'occupe de la coordination des tâches sur les différents clusters. Il attribue les fonctions de MapReduce aux différents « TaskTrackers ». Le « Job Tracker » est un « Daemon » cohabitant avec le « Name Node » et ne possède donc qu'une instance par cluster.
- **Le moniteur de tâches (Tasktracker) :** Il permet l'exécution des ordres de mapReduce, ainsi que la lecture des blocs de données en accédant aux différents « Data Nodes ». Par ailleurs, le « TaskTracker » notifie de façon périodique au « Job Tracker » le niveau de progression des tâches qu'il exécute, ou alors d'éventuelles erreurs pour que celui-ci puisse reprogrammer et assigner une nouvelle tâche. Un « TaskTracker » est un « Deamon » cohabitant avec un « Data Node », il y a donc un « TaskTracker » par « Data Node ».
- **Le noeud secondaire (Secondarynode) :** N'étant initialement pas présent dans l'architecture Hadoop, celui-ci a été ajouté par la suite afin de répondre au problème du point individuel de défaillance (SPOF- Single point of failure). Le « Secondary Node » va donc périodiquement faire une copie des données du « Name Node » afin de pouvoir prendre la relève en cas de panne de ce dernier.
- **Le nœud de données (Data Node) :** Il permet le stockage des blocs de données. Il communique périodiquement au « Name Node » une liste des blocs qu'il gère. Un HDFS contient plusieurs nœuds de données ainsi que des répliquions d'entre eux. Ce sont les nœuds esclaves [5].



Figure II.1: Architecture d'un Cluster Hadoop

Un Cluster Hadoop peut être constitué de machines hétérogènes, que ce soit au niveau du hardware comme au niveau software (système d'exploitation). Cependant il est bien plus simple d'administrer un cluster de type homogène [4]. Hadoop est aujourd'hui l'un des outils les plus pertinents pour répondre problèmes du Big Data [4]. C'est la première technologie qui vient à l'esprit lorsque l'on évoque

Aujourd'hui ce sujet [5].

Les points forts d'Hadoop se résument dans ses caractéristiques suivantes :

- **évolutif**, car pensé pour utiliser plus de ressources physiques, selon les besoins, et de manière transparente ;
- **rentable**, car il optimise les coûts via une meilleure utilisation des ressources présentes
- **souple**, car il répond à la caractéristique de variété des données en étant capable de traiter différents types de données ;

- et enfin, **résilient**, car pensé pour ne pas perdre d'information et être capable de poursuivre le traitement si un nœud du système tombe en panne [3].

Hadoop n'a pas été conçu pour traiter de grandes quantités de données structurées à grande vitesse. Cette mission reste largement l'apanage des grands systèmes de Datawarehouse et de Datamart reposant sur des SGBD traditionnelles et faisant usage de SQL comme langage de requête. La spécialité d'Hadoop, ce serait plutôt le traitement à très grande échelle de grands volumes de données non structurées tels que des documents textuels, des images, des fichiers audio,... même s'il est aussi possible de traiter des données semi-structurées ou structurées avec Hadoop [5].

Exemple d'usage :

Ce qui fait la spécificité de Hadoop est qu'il est conçu pour traiter un énorme volume de données en un temps record. A titre d'exemple, les Laboratoires de Yahoo! ont trié l'équivalent de 500 GB de données en 59 secondes sur un Cluster de 1400 nœuds (Avril 2009). Sa vocation première est donc d'implémenter des traitements batchs performants, particulièrement lorsqu'ils impliquent un volume de données très important. En dehors de Yahoo!, citons les cas d'utilisation de deux sociétés prestigieuses :

- la plateforme musicale last.fm, met en œuvre Hadoop pour générer les statistiques hebdomadaires (Tops artistes et Top titres) ou mesurer les tendances musicales.
- Facebook l'utilise pour la production de rapport à usage interne, comme la plateforme des campagnes publicitaires opérées par la plateforme sociale, ou des statistiques diverses (croissances du nombre des utilisateurs, consultation des pages, temps moyen de consultation du site, etc.) [W1]

II.4. HDFS:

HDFS (Hadoop Distributed File System) est un système de fichiers distribué, inspiré du système GFS développé par Google. Il se démarque des autres systèmes de fichier distribués par sa grande tolérance aux fautes [6] et le fait qu'il soit conçu pour être déployé sur des machines à faible coût. HDFS fournit un haut débit d'accès aux données et est adapté pour les applications qui nécessitent de grands groupes de données. Il a été à l'origine conçu pour le projet de moteur de recherche web Apache Nutch [7].

HDFS a une architecture de type maître/esclave. Un Cluster HDFS est constitué d'un unique NameNode, un serveur maître qui gère le système de fichier et notamment les droits d'accès aux fichiers. A cela s'ajoute des DataNodes, en général un par nœud dans le Cluster, qui gère le stockage des données affectés au nœud sur lequel elle se trouve (Voir Figure II.2).

HDFS est conçu pour tourner sur des machines simples sous GNU/Linux, et est programmé en Java. Toute machine qui dispose de Java peut donc faire tourner un NameNode ou un DataNode. Les nœuds communiquent entre eux via SSH. Il faut donc entrer la clé publique de chaque DataNode dans le fichier `authorized_keys` du NameNode, afin qu'il puisse se connecter aux autres nœuds via SSH sans avoir besoin de taper un mot de passe à chaque fois [7].

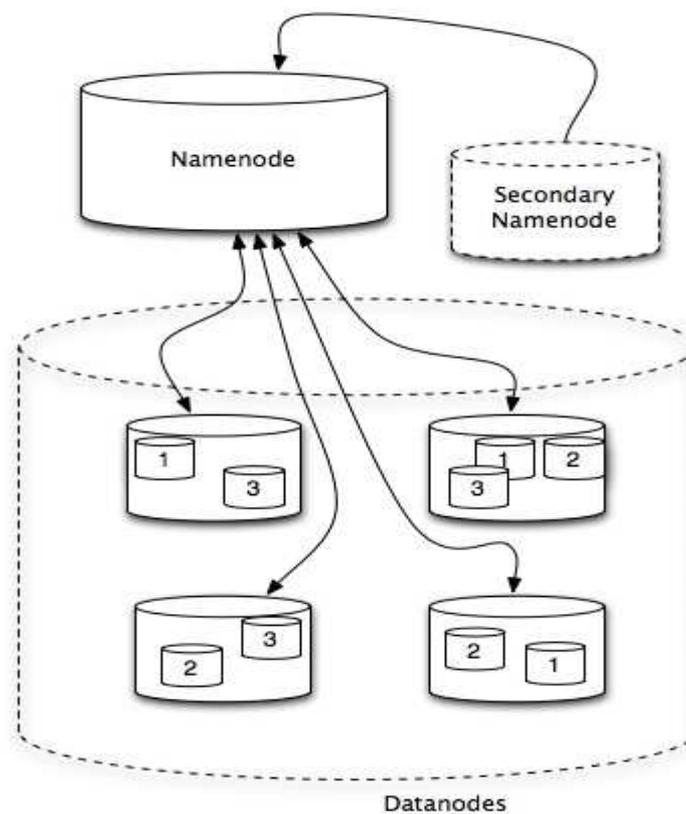


Figure II.2 : Schéma HDFS

Le NameNode dispose d'un fichier de configuration dans lequel il entre l'adresse de chaque machine sur laquelle tourne un DataNode, et se connecte ensuite à chacune de ces machines via SSH au démarrage de HDFS. Par ailleurs, le NameNode est l'arbitre et le dépositaire de toutes les métadonnées du HDFS. HDFS est conçu de telle sorte qu'aucune donnée de l'utilisateur ne transite par le NameNode (voir Figure II.3). [7]

Dans Hadoop, les différents types de données, qu'elles soient structurées ou non, sont stockées à l'aide du HDFS. Le HDFS va prendre les données en entrée et va ensuite les partitionner en plusieurs blocs de données. Afin de garantir une disponibilité des données en cas de panne d'un nœud, le système fera un réplica des données. Par défaut les données sont répliquées sur trois nœuds différents, deux sur le même support et

Sur un support différent. Les différents nœuds de données peuvent communiquer entre eux pour rééquilibrer les données (Figure II.3) [4].

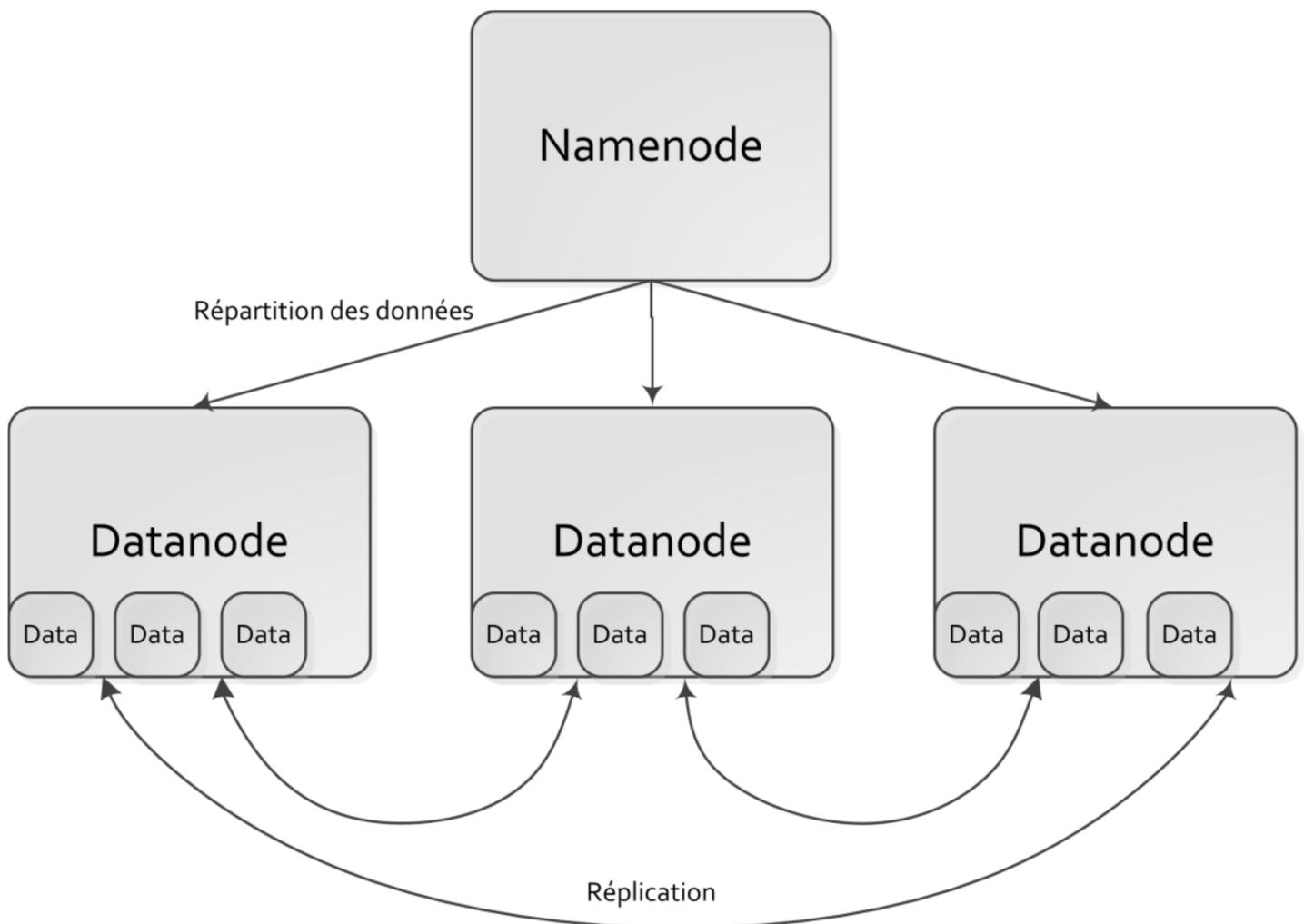


Figure II.3: Réplication des données à l'aide du HDFS

HDFS est fait pour stocker de gros fichiers (chaque fichier est divisé en bloc de 64 MB par défaut). D'après la documentation officielle, un fichier typique dans HDFS fait une taille de l'ordre du GB voire du TB (1 TB = 1012 Octet = 1000 gigabytes) [W2]. HDFS est fait pour une très grande scalabilité et tourne déjà

en production sur des très gros Clusters (plus de 1000 machines) dans des sociétés comme Facebook ou Yahoo [7].

De plus HDFS est avant tout conçu pour être utilisé avec Hadoop (Figure II.4). La philosophie de l'implémentation d'Hadoop repose sur le principe qu'il est plus efficace de déplacer de la capacité de calcul que des données. HDFS fournit donc une interface pour pouvoir déplacer le traitement des données au plus près de celles-ci.

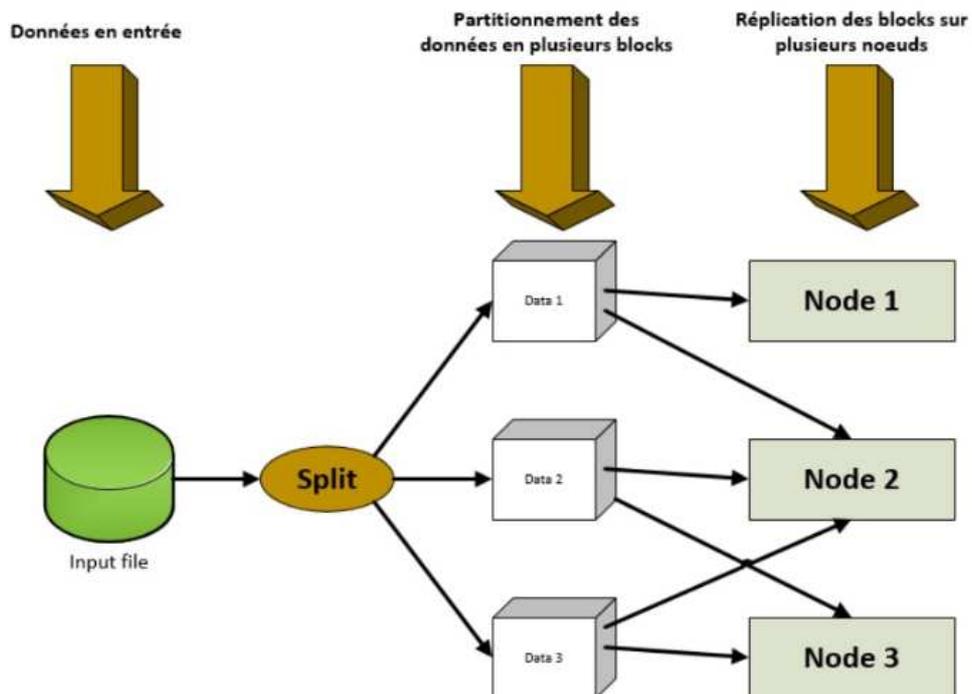


Figure II.4. Hadoop Distributed File System (HDFS)

II.5. Le composant MapReduce :

MapReduce est une technique qui permet d'effectuer des tâches sur de grandes quantités de données en utilisant un cluster de machines. Elle est implémentée au cœur du projet Hadoop

MapReduce est le second composant majeur d'Hadoop qui gère la répartition et l'exécution des requêtes sur les données stockées par le Cluster. Le Framework MapReduce est conçu pour traiter des problèmes parallélisables à très grande échelle en s'appuyant sur un très grand nombre de nœuds.

L'objectif de MapReduce et de son mécanisme avancé de distribution de tâches est de tirer parti de la localité entre données et traitements sur le même nœud de façon à minimiser l'impact des transferts de données entre les nœuds du Cluster sur la performance [5].

Le moteur MapReduce de Hadoop réceptionne les jobs de la part des applications puis les divise en une série de tâches, qu'il assigne ensuite aux différents nœuds de calcul [6]. Les jobs MapReduce se structurent toujours autour de trois parties : deux classes statiques «Mapper» et «Reduce», et une méthode main qui dirige l'application.

L'exécution des jobs se fait à l'aide d'un JobTracker et de Tasktrackers : lors de son exécution le job est soumis au JobTracker qui s'occupe de le distribuer au Tasktracker qui tourne sur chaque noeud. Le JobTracker choisit toujours les TaskTracker qui sont les plus proches de l'emplacement de la donnée à traiter [7].

Une fois que la partie « Map » est terminée, c'est la partie Reduce qui commence à faire remonter les différents résultats et les consolider en un seul résultat final (Figure II.5) [4].

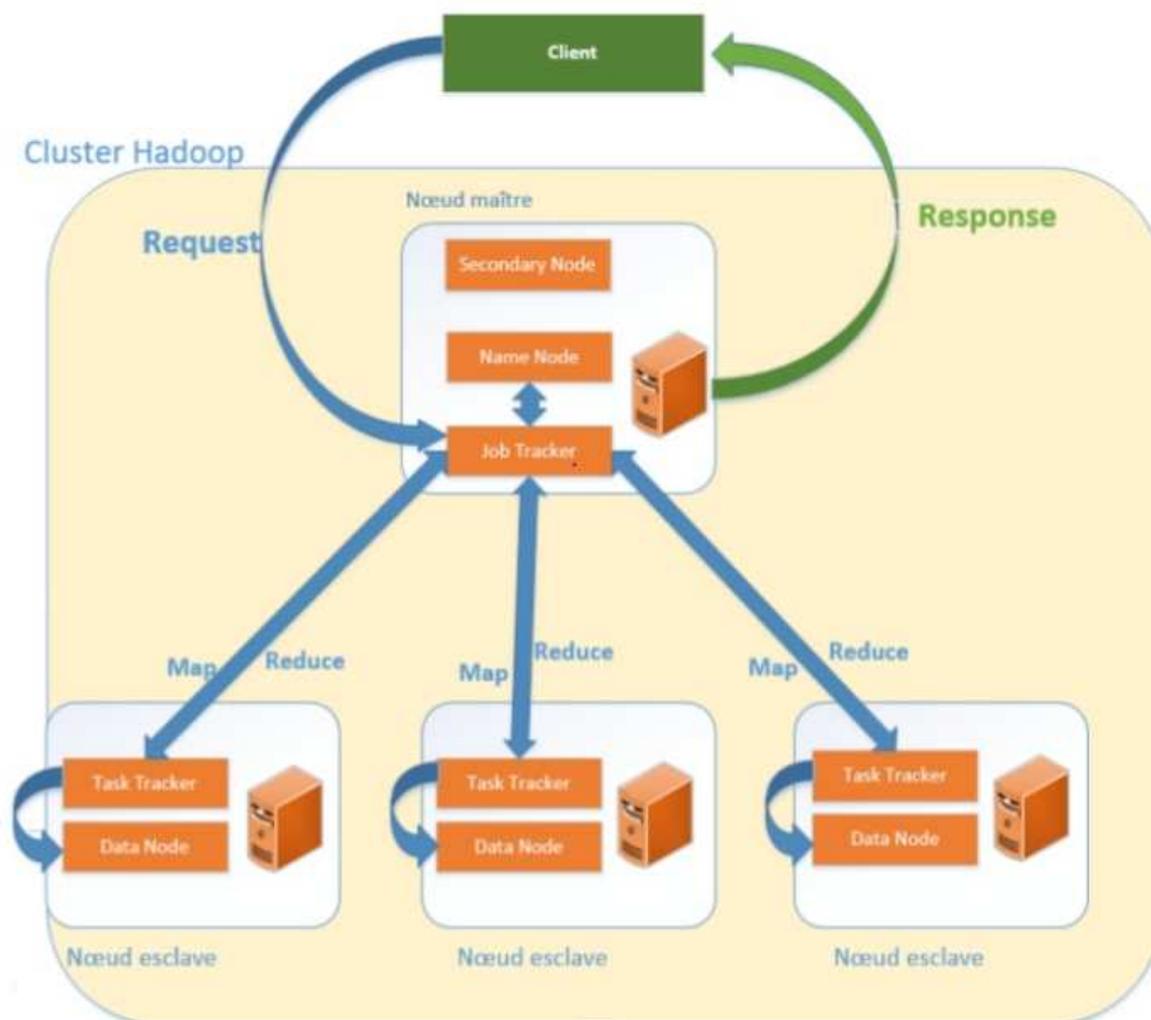


Figure II.5. Fonctionnement de MapReduce dans Hadoop

III. MapReduce :

III.1 : Définition :

MapReduce est un patron d'architecture de développement informatique, inventé par Google¹, dans lequel sont effectués des calculs parallèles, et souvent distribués, de données potentiellement très volumineuses, typiquement supérieures en taille à 1 téraoctet. Les termes « map » et « reduce », et les concepts sous-jacents, sont empruntés aux langages de programmation fonctionnelle utilisés pour leur construction (map et réduction de la programmation fonctionnelle et des langages de programmation tableau).

MapReduce permet de manipuler de grandes quantités de données en les distribuant dans un cluster de machines pour être traitées. Ce modèle connaît un vif succès auprès de sociétés possédant d'importants centres de traitement de données telles Amazon.com ou Facebook. Il commence aussi à être utilisé au sein du Cloud computing. De nombreux frameworks ont vu le jour afin d'implémenter le MapReduce. Le plus connu est

Hadoop qui a été programmé par Apache Software Foundation. Mais ce framework possède des inconvénients qui réduisent considérablement ses performances notamment en milieu hétérogène. Des frameworks permettant d'améliorer les performances de Hadoop ou les performances globales du MapReduce, tant en termes de vitesse de traitement qu'en consommation électrique, commencent à voir le jour.

III.2. Principe de MapReduce :

Le principe de MapReduce est simple (voir Figure II-5): il s'agit de découper une tâche manipulant un gros volume de données en plusieurs tâches traitant chacune un sous-ensemble de ces données. Dans la première étape (Map) les tâches sont donc dispatchées sur l'ensemble des nœuds. Chaque nœud traite un ensemble des données. Dans la deuxième étape, les résultats sont consolidés pour former le résultat final du traitement (Reduce).

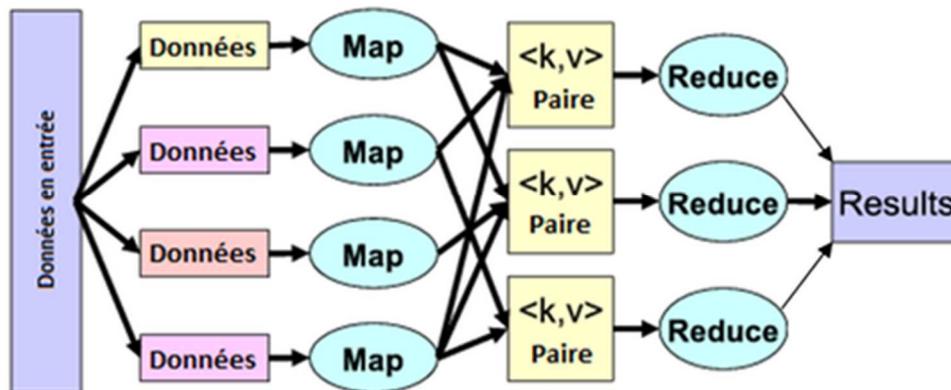


Figure III.1: les étapes Map et Reduce

III.3 : Le Modèle de Programmation MapReduce :

Programmation MapReduce consiste en deux fonctions `map ()` et `reduce ()`.

Dans l'étape Map, le nœud analyse un problème, la découpe en sous-problèmes, et les délègue à d'autres nœuds (qui peuvent en faire de même récursivement). Les sous-problèmes sont ensuite traités par les différents nœuds à l'aide de la fonction Reduce qui à un couple (clé, valeur) associe un ensemble de nouveaux couples (clé, valeur) :

$\text{Map}(\text{clé1}, \text{valeur1}) \rightarrow \text{List}(\text{clé2}, \text{valeur2})$

Fonction Map :

```
//En pseudo code cela donnerait Map (void * document)

{int cles = 1; for each mot m in document calculIntermediaire(m,cles); }
```

Vient ensuite l'étape Reduce, où les nœuds les plus bas font remonter leurs résultats au nœud parent qui les avait sollicités. Celui-ci calcule un résultat partiel à l'aide de la fonction Reduce (réduction) qui associe toutes les valeurs correspondantes à la même clé à une unique paire (clé, valeur). Puis il remonte l'information à son tour.

À la fin du processus, le nœud d'origine peut recomposer une réponse au problème qui lui avait été soumis :

Reduce (key2, list (valeur2)) → list (valeur2)

Fonction Reduce :

```
//En pseudo code cela donnerait Reduce (entier cles, Iterator values)

{int result = 0; for each v in values result += v; }
```

MapReduce peut être utilisé pour un grand nombre d'applications, dont grep distribué, tri distribué, inversion du graphe des liens Web, vecteur de terme par hôte, statistiques d'accès au Web, construction d'index inversé, classification automatique de documents, apprentissage automatique. [W1]

Le MapReduce a émergé en 2004 comme un important modèle de programmation pour les applications utilisant d'énormes quantités de données grâce à sa répartition efficace du travail sur différents nœuds de calcul. Il commence notamment à être utilisé dans le Cloud Computing, car son nombre de données stockées et manipulées ne cesse de croître. Il est donc nécessaire d'avoir un moyen d'améliorer le traitement des données au sein du Cloud. [W1]

Conclusion :

Nous avons présenté dans ce chapitre les différents concepts liés au Cloud Computing à savoir les types de services Cloud connus jusqu'à présent et ses différents modes de déploiement. Les différents composants d'Hadoop, principalement, HDFS et le modèle de programmation MapReduce. Il est principalement utilisé pour la manipulation et le traitement d'un nombre important de données et le plus adapté pour l'analyse de grande quantité de vidéo.

Introduction :

De nos jours, le recours à la surveillance vidéo par les organisations est omniprésente et on la retrouve dans de nombreux secteurs d'activité (banque, transports, industrie grande distribution, etc..) ou lieux de vie (villes, immeubles de bureau, équipements collectifs, etc.). L'évolution de la technologie et la baisse considérable des prix ont rendu la surveillance vidéo de plus en plus accessible à une foule d'organisations. La sécurité et le contrôle de la criminalité sont les motifs les plus souvent évoqués pour justifier l'usage de caméras de surveillance. Les détaillants utilisent des caméras dans le but de décourager les vols et d'identifier les suspects. La plupart des responsables souhaitent accroître la sécurité en protégeant les biens et les personnes par de la vidéosurveillance. Aujourd'hui, le développement technologique permet d'obtenir de très bons résultats lorsqu'on sait exprimer un besoin et lui faire correspondre un matériel efficace et évolutif afin de pérenniser l'installation.

Dans ce chapitre, on présentera les différents types de caméra de surveillance et les différents systèmes ainsi que ses technologies et ses domaines d'application.

I. Définition :

La vidéosurveillance ou vidéo-protection est un Système de caméras et de transmission d'images disposé dans un espace public ou privé pour le surveiller à distance .il s'agit donc d'un type de télésurveillance.les images obtenues avec ce système peuvent être traitées automatiquement et/ou visionnées puis archivées ou détruites. La surveillance a pour but de contrôler les conditions de respect de la sécurité, le principe de filmer et de contrôler un phénomène et pour visualiser et /ou enregistrer un endroit centralisé tous les flux de personnes au sein d'un lieu ouvert ou public pour surveiller les allées et venues, prévenir les vols, agressions, fraudes et gérer les incidents et mouvement de foule.



Figure I.1 : Les vidéos de surveillance

II. Les différents types de caméra de surveillance :

Il existe plusieurs types de caméras qui correspondent à plusieurs contextes d'utilisation. Chaque caméra possède des caractéristiques variées, il est donc important de déterminer l'utilisation que vous souhaitez en faire afin de choisir une caméra adaptée à l'agencement de votre établissement, aux conditions environnementales (intérieur, extérieur, de jour ou de nuit) pour garantir une utilisation optimale du produit.



Caméras Anti vandalisme ou Mini dôme

vidéosurveillance de zones sensibles, en intérieur comme en extérieur. En effet, la caméra mini-dôme est sobre, discrète, esthétique, avec un montage facile au mur ou au plafond ce qui lui permet de se fondre dans le décor sans donner d'indication sur la zone filmée. Elle possède également un boîtier renforcé et une demi-sphère anti-vandale pour résister aux chocs et d'agressions de tous genres. Le mini-dôme fournit des images claires des événements tout en bénéficiant d'une installation facile et rapide.



Caméras boîtiers, classiques, intérieures ou extérieures

Les caméras boîtiers sont des caméras que l'on appelle « classiques », car elles sont le plus répandues et possèdent un design « classique ». Sur ces caméras il faut rajouter des objectifs de

différents zooms, angle de vue. Elles s'adaptent à tous les environnements intérieurs mais également extérieurs grâce à un large panel de caisson de protection (étanches, thermostatés,...)

Caméras infrarouges (IR) :



La camera de surveillance infrarouge est capable de filmer aussi bien le Jour que la nuit. Le jour les images sont en couleurs comme n'importe quelle autre caméra ; la nuit les images sont en noir et blanc et d'aussi bonne qualité que le jour. En effet, l'objectif de la caméra est entouré d'ampoules LED qui s'allument automatiquement dès que la luminosité extérieure est insuffisante. Ce système permet de filmer à 0 lux, c'est-à-dire dans une obscurité totale. Selon les modèles de caméra, ces ampoules LED ont une portée plus ou moins importante et vous permettent de bénéficier d'une transmission d'images parfaite jusqu'à 80 mètres dans les zones les plus obscures. Lors de la sélection de votre caméra, il vous faut déterminer avec la plus grande attention le réglage de l'objectif et la portée de la vision nocturne (de 5m à 50m) qui sont les deux critères de performances.

Dômes motorisés



Les dômes motorisés sont préconisés lorsque la zone à surveiller est très étendue (parking, dépôts...) car ce type de caméra peut effectuer une rotation à 360° en vous assurant une transmission d'images de haute qualité. Vous pouvez les installer sur un mur ou un plafond. Vous pouvez également programmer des rondes pour rendre votre dôme complètement autonome ou tout simplement piloter votre caméra à distance grâce à un joystick ou une souris. Pour manipuler la rotation, l'inclinaison et le zoom de ce type de caméra il vous suffit simplement de la raccorder à votre réseau informatique doté d'une connexion Internet. Vous pouvez également bénéficier d'un contrôle total de votre caméra depuis votre mobile !

Caméras espions



Détecteur de fumée, de mouvement, stylos, ce type de caméras revête l'apparence que vous souhaitez pour vous assurer une vidéosurveillance en toute discrétion. Les caméras espionnes disposent des mêmes caractéristiques techniques que les caméras d'intérieures avec une qualité d'image respectable et la possibilité d'être dissimulées dans n'importe quel objet afin d'assurer une vidéosurveillance sans se faire remarquer.

Caméras IP



Les caméras de surveillance IP vous permettent de réaliser des installations de vidéosurveillance haut de gamme. Selon les modèles, ces caméras vous apportent une résolution d'image jusqu'à 5 millions de pixels.

Caméras sans fils



Les caméras sans fils vous apportent un confort d'utilisation et une surveillance discrète. Vous pouvez déplacer la caméra d'une zone à surveiller à l'autre, sans aucune installation.

Caméras factices



La caméra factice est la solution pour dissuader à moindre frais. Dotées d'une esthétique réaliste, elles vous permettent de mieux sécuriser le périmètre. Selon les modèles, vous pouvez les utiliser en intérieur comme en extérieur et même insérer des piles pour illuminer les ampoules LED.

III .les différents types de systèmes:

III.1 Système sur réseaux IP :

Ce système relie un réseau de caméras IP, qui peut compter de nombreuses unités, à un système d'enregistrement numérique. D'une part, cela permet de pouvoir stocker une quantité importante d'images, sans perte de qualité, tout en pouvant les consulter rapidement grâce à des logiciels de traitement. D'autre part, le fait d'informatiser un système de surveillance permet de profiter des technologies de communication comme Internet. Ainsi, les caméras sont « visibles » et gérables depuis

n'importe où dans le monde. L'évolution des téléphones mobiles a créé la "vidéosurveillance mobile" avec l'accès aux vidéos via Internet mobile sur PDA ou via GSM GPRS sur téléphone GSM doté de Java. Cette technologie permet également d'économiser et de mutualiser les câbles réseaux qui sont généralement disponibles dans les bâtiments récents.

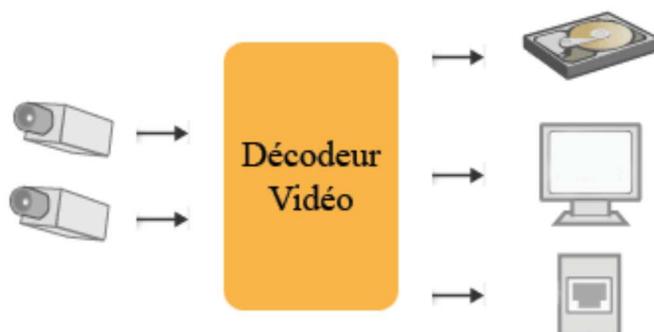


Figure III.2 : Vidéosurveillance sur réseau IP

III.2 Kit de vidéosurveillance :

On entend par « kit » le genre de caméras utilisées dans les petits magasins, par exemple. Il regroupe en général une ou deux caméras et un moniteur. Ces systèmes sont plutôt utilisés à titre de prévention et n'enregistrent pas ce qu'ils voient. C'est en quelque sorte de la vidéosurveillance bon marché qui est proposée comme une solution de sécurité peu coûteuse.

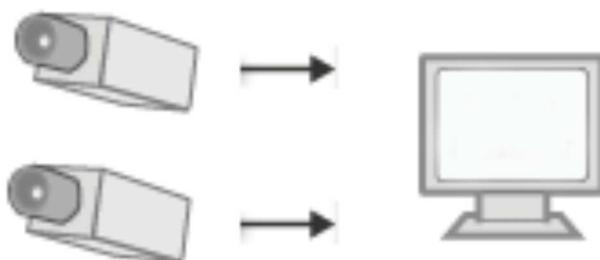


Figure III.3 : Kit de vidéosurveillance

III.3 Réseau « classique » de vidéosurveillance :

Le réseau est basé sur un système analogique, avec dans la plupart des cas un enregistrement limité dans la durée. Il s'agit là d'une des méthodes les plus anciennes donc également des plus répandues dans un grand nombre d'établissement. Cependant, ces systèmes ne répondent plus, à de très rares exceptions près, aux nouvelles exigences techniques de l'arrêté du 3 août 2007.

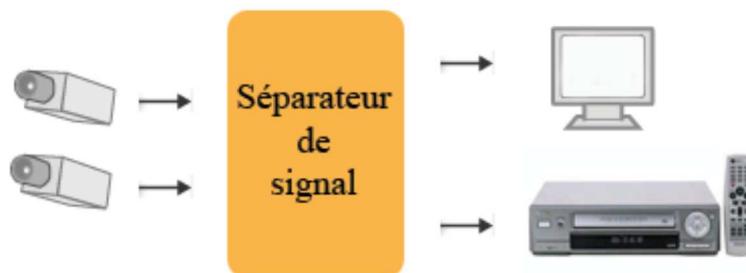


Figure III.4 : Système de vidéosurveillance analogique

III.4 Système « hybride » de vidéosurveillance :

Les systèmes hybrides intègrent les systèmes classiques de vidéosurveillance basés sur les caméras analogiques et les caméras en réseau. Il permet d'intégrer aisément les deux types de systèmes en place sur un seul serveur ou de faciliter l'évolution d'un système de vidéosurveillance analogique vers le numérique, sans remettre en cause l'existant, et introduire de nouvelles fonctions comme la détection de disparition / apparition d'objet et le comptage d'objets ou de personnes.

Exemples de quelques systèmes existants :

➤ Le système ASCAM-2E2I :

La solution ASCAM intègre un logiciel de vidéo surveillance sur réseau IP, un logiciel d'enregistrement sur mouvement avec les fonctions magnétoscope et finalement des caméras réseau IP jour / nuit pour extérieur en caisson étanche et pour intérieur. L'accès aux caméras et aux enregistrements sur alarme par Internet permet d'assurer une surveillance à distance. Les logiciels permettent un monitoring temps réel sur site et à distance via Internet ainsi que l'enregistrement sur détection de mouvement, planning et alarme La détection de mouvement peut être configurée selon les critères suivants :

- Champs de vision des caméras de surveillance
- Zones de détections, masques de détection
- Objet manquant (détection de vol)
- Nouvel objet statique (colis douteux)
- Focus dérégulé, caméra obstruée (vandalisme)

Enfin la notification d'alarmes comprend l'envoi d'email d'alerte, appel téléphonique et également une alarme sonore

➤ **Le système ASCAM-NUUO :**

La solution ASCAM-NUUO intègre un serveur de vidéo surveillance permettant une visualisation à distance via internet sur PC et sur téléphone mobile 3GPP. Elle fournit un ensemble de fonctionnalités que l'on peut résumer dans la liste suivantes :

- Visualisation des flux vidéo en local sur moniteur
- Alarme et visualisation en levée de doute, fenêtre pop up
- Console d'administration des paramètres de gestion
- Gestion des droits d'accès: profils, horaires, à distance
- Alarme sur détection de mouvement
- Enregistrement possible sur Planning (Schedule)

IV. Technologie :

IV.1 Architecture des systèmes de vidéosurveillance :

IV.1.1 Composantes d'un système de vidéosurveillance :

Le schéma de principe :

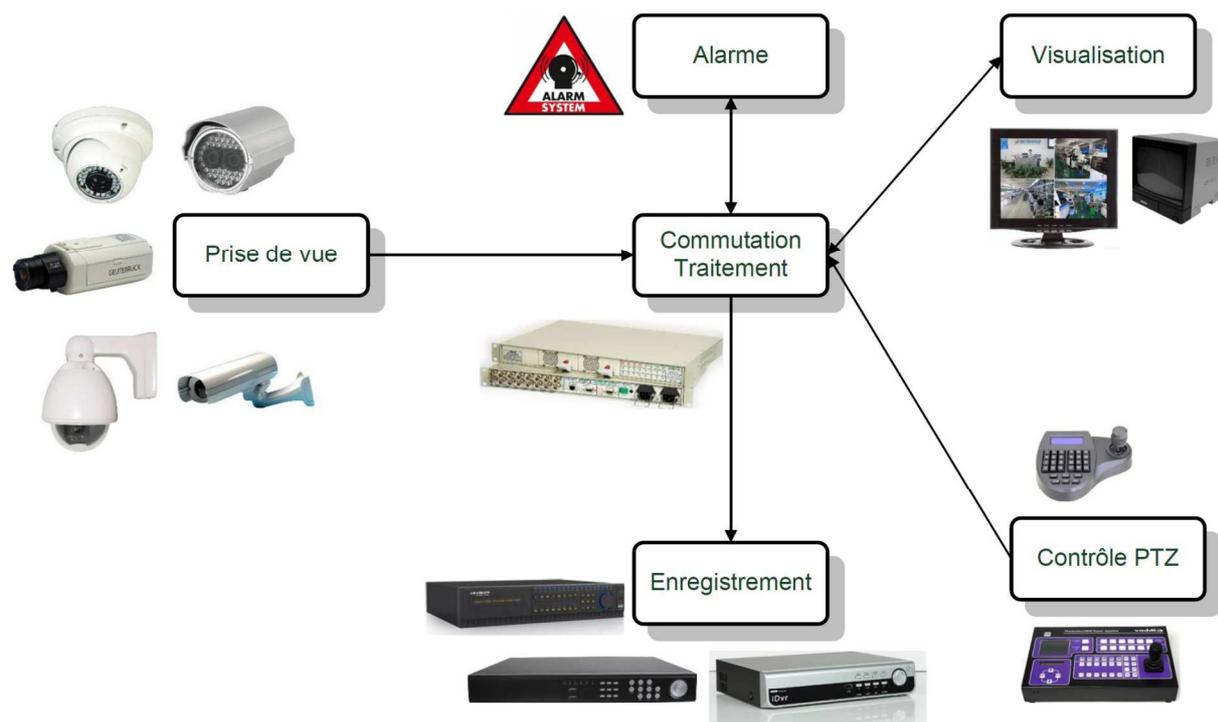


Figure IV .5 : Le schéma de principe

Cette section présente, de façon sommaire, les différentes composantes matérielles et logicielles des systèmes de vidéosurveillance. Une description plus détaillée de l'infrastructure des systèmes de vidéosurveillance, ainsi que des principes guidant les choix matériels, peut être retrouvée dans de nombreux ouvrages et guides de références, incluant [8] et [9].

IV.1.1.1 Acquisition :

Il existe une panoplie de modèles de caméras répondant à différents besoins de surveillance. Elles sont analogiques ou numériques et peuvent être motorisées ou non. Plus spécifiquement, on retrouve les types de caméras suivants.

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

- **PTZ (Pan-Tilt-Zoom)** : Motorisée, elle peut être actionnée, manuellement ou automatiquement, dans des mouvements panoramique/inclinaison/zoom. Elle sert à suivre des objets ou des individus se déplaçant dans la scène ou à zoomer sur des régions d'intérêt (par exemple, sur une plaque d'immatriculation).
- **Dôme** : Recouverte d'un caisson hémisphérique, ce qui la rend discrète et, dans certains modèles, résistante au vandalisme et aux intempéries. Elle peut être fixe ou mobile. Les versions motorisées couvrent une zone très large, grâce à leur balayage horizontal de 360° et de 180° à la verticale. Bien qu'en « tour de garde », elle puisse remplacer dix caméras fixes en balayant l'aire à surveiller, elle n'observe qu'une seule direction à la fois.
- **Mégapixel** : Offre une résolution plus élevée que les caméras standards, allant de 1 à 16 mégapixels¹⁷. Elle permet soit de capter une image plus détaillée, soit de couvrir un plus large champ visuel, réduisant le nombre de caméras nécessaires pour couvrir une aire à surveiller. Lorsqu'utilisée avec un grand angle, elle possède un espace de visualisation allant généralement de 140° à 360°. Offrant la possibilité de zoomer de façon logicielle dans l'image, elle peut ainsi devenir une alternative à la caméra PTZ mécanique qui entraîne l'usure des pièces. Sa résolution élevée contribue à l'amélioration de la performance des algorithmes de détection et de reconnaissance exigeant un haut niveau de détails, telles que la lecture de plaques d'immatriculation et la reconnaissance de visage.
- **Infrarouge et thermique** : Sensible au rayonnement infrarouge (IR), elle est capable de produire une image de bonne qualité dans le noir pour une surveillance nocturne. De nuit, elle filme en noir et blanc, mais elle peut produire une image couleur le jour. Certaines caméras infrarouges sont équipées de leur propre source de lumière IR, allumée lorsque le niveau d'éclairage chute sous un certain seuil. Des projecteurs IR séparés (lampe ou LED¹⁸) peuvent aussi être utilisés. Les caméras thermiques enregistrent le rayonnement de chaleur des objets. Elles ne requièrent aucune source d'illumination.
- **Panoramique** : Grâce à une optique spéciale, elle offre 360° de visibilité avec une seule caméra¹⁹. Elle permet un PTZ virtuel dans l'image. Les principales technologies panoramiques pour la surveillance sont le fisheye, la lentille à miroirs et la lentille panomorphe. Toutefois, la résolution de ces caméras est souvent insuffisante pour des analyses nécessitant un niveau de détail élevé.

VI.1.1.2 Transmission :

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

La **vidéo filaire** prédomine largement dans les systèmes de vidéosurveillance. Elle offre une plus grande bande passante et une meilleure fiabilité que les connections sans fil, à un coût inférieur. Cependant, la **transmission vidéo sans fil** s'impose parfois comme solution, par exemple dans le cas de surveillance de grands périmètres où l'installation de câblage s'avérerait trop coûteuse, ou lorsque les zones à surveiller sont impossibles à rejoindre par câble. Qu'il transite par fil ou sans fil, le signal vidéo peut être analogique ou numérique. Encore aujourd'hui, la majeure partie des transmissions vidéo pour la surveillance sont **analogiques**. Néanmoins, les réseaux informatiques (LAN, WAN ou Internet) sont de plus en plus utilisés pour transporter la vidéo grâce au **protocole IP**. Les caméras IP peuvent se connecter directement sur ces réseaux, tandis que les flux vidéo émergeant de caméras analogiques doivent, au préalable, être numérisés par un **encodeur**, aussi appelé **serveur vidéo**, pour passer par les réseaux IP.

IV.1.1.3 Compression :

La vidéo numérisée représente une grande quantité de données à transmettre et à archiver. L'envoi d'une séquence vidéo peut nécessiter jusqu'à 165 mégabits de bande passante et la vidéo d'une seule caméra pour une journée peut occuper sept gigaoctets d'espace disque. C'est pourquoi la vidéo de surveillance doit être compressée grâce à des **codecs**²⁰, algorithmes permettant de réduire la quantité de données en supprimant les redondances, par image ou entre les trames d'une séquence, ainsi que les détails imperceptibles à l'œil humain. Selon le type de compression, l'usage du processeur requis pour l'exécution du codec est plus ou moins intensif. Un compromis s'impose donc entre le taux de compression et les ressources du processeur qui sont accaparées.

Il existe deux grands groupes de standards internationaux de compression : JPEG, créés par le Joint Photographic Experts Group, et MPEG, élaborés par le Moving Photographic Experts Group. Dans le premier groupe, on retrouve les formats JPEG pour les images fixes, et MJPEG pour les séquences vidéo. Le second groupe comprend les formats MPEG-1, MPEG-2, MPEG-4 et H.264/21. À l'heure actuelle, MJPEG et MPEG-4 sont les standards les plus répandus en vidéosurveillance. Toutefois, avec les améliorations en qualité et efficacité (taux de compression, latence, résistance à l'erreur) qu'il apporte, H.264 devrait bientôt remplacer MPEG-4. En effet, sans affecter la qualité de

l'image, l'encodeur H.264 permet de réduire la taille de celle-ci de plus de 80 % par rapport à la compression MJPEG, et de 50 % par rapport à la compression MPEG-4 [10].

IV.1.1.4 Traitement :

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

- **Enregistreur vidéo numérique (DVR22)** : Appareil qui dispose d'un disque dur interne pour l'enregistrement numérique de la vidéo et d'un logiciel intégré de traitement de la vidéo. Il n'accepte que les flux provenant de caméras analogiques, qu'il numérise. Les modèles récents permettent de visionner la vidéo à distance sur ordinateur. Encore très répandus, ils laissent toutefois peu à peu leur place aux systèmes supportant la vidéo IP de bout en bout.
- **Enregistreur vidéo hybride (HDVR23)** : Similaire à l'enregistreur numérique, mais accepte à la fois le branchement de caméras analogiques et IP. Il est possible de rendre hybrides plusieurs modèles d'enregistreurs vidéo numériques par l'ajout d'un logiciel.
- **Enregistreur numérique réseau (NVR24)** : Conçu pour les architectures réseaux IP de vidéosurveillance, il ne peut traiter que les signaux vidéo provenant de caméras IP ou d'encodeurs.
- **Logiciel de vidéosurveillance IP** : Solution purement logicielle de gestion de la vidéo sur un réseau IP. Dans le cas de systèmes de surveillance comportant peu de caméras, un navigateur Web peut suffire à gérer la vidéo. Pour de plus gros réseaux de vidéosurveillance, un logiciel dédié de gestion vidéo doit être utilisé. Celui-ci s'installe sur un ordinateur personnel ou un serveur. Plus complexe à installer, en raison des configurations nécessaires du serveur, il offre une plus grande flexibilité pour le choix et l'ajout de composantes au réseau de vidéosurveillance. Les logiciels de vidéosurveillance IP représentent une tendance forte en gestion vidéo, surtout dans les infrastructures comportant un grand nombre de caméras. Les plateformes ouvertes permettent d'intégrer facilement des caméras et composantes matérielles de différents manufacturiers.

IV.1.1.5 Archivage :

La période d'archivage des séquences vidéo varie selon les besoins de surveillance, allant de quelques jours à quelques années. En moyenne, les organisations conservent les preuves vidéo entre 30

et 90 jours. Le déploiement de vastes réseaux de caméras et l'utilisation de vidéosurveillance à haute résolution fait exploser les demandes pour les systèmes de stockage. Bien que le coût des supports d'enregistrement ait considérablement baissé dans les dernières années, l'archivage représente souvent une part importante des dépenses d'infrastructure en vidéosurveillance, en raison de la quantité toujours croissante de données vidéo à stocker. Les solutions de stockage sont de deux types :

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

- **Rattaché** : L'archivage se fait sur des appareils externes aux enregistreurs ou serveurs vidéo. De type NAS (Network Attached Storage) ou SAN (*Storage Area Network*), ces systèmes offrent un espace de stockage partagé entre les différents clients du réseau. Sur un système de stockage en réseau NAS, un fichier est archivé sur un même disque dur, alors qu'avec le réseau de stockage SAN, un fichier peut être sauvegardé en fragments répartis sur plusieurs supports de stockage. Ces solutions d'archivage rattachées sont particulièrement avantageuses pour les grands réseaux de vidéosurveillance comportant un grand nombre de caméras²⁵. Bien que plus onéreuses que les systèmes internes d'archivage, ces solutions rattachées sont supérieures en termes d'extensibilité, de flexibilité et de redondance.

IV.1.1.6 Affichage :

Une grande partie de la vidéo captée par les caméras de surveillance n'est jamais regardée. Elle est simplement archivée, au cas où un visionnement soit nécessaire suite à un incident. Traditionnellement, la vidéosurveillance a principalement servi comme outil d'enquête. Toutefois, dans plusieurs cas de surveillance, des agents de sécurité visionnent, en temps réel, les images provenant des caméras de surveillance. Sans nécessairement regarder toute la vidéo captée, les agents peuvent faire une revue périodique des différentes sources vidéo.

La vidéo de surveillance peut être visionnée sur différents appareils. Dans de petites installations, il est possible de regarder la vidéo directement de l'enregistreur, simultanément à son enregistrement. Plus généralement, les images seront regardées à distance, sur un ordinateur ou, de façon mobile, sur un téléphone ou dispositif portable.

Les grands centres d'opérations de sécurité, supervisant des centaines de caméras, utilisent souvent un mur d'écrans vidéo. Ceux-ci offrent une grande surface de visionnement et permettent d'afficher différents flux vidéo.

V .Evolution des systèmes de vidéosurveillance:

Les systèmes de vidéosurveillance existent depuis environ 25 ans. Intégralement analogiques à leurs débuts, ils ont évolué progressivement vers la technologie numérique. Les systèmes actuels ne ressemblent plus guère aux anciennes caméras analogiques branchées sur des magnétoscopes traditionnels. Aujourd'hui, ils utilisent les caméras réseau et les serveurs informatiques pour l'enregistrement vidéo dans un système entièrement numérique. Entre les systèmes entièrement analogiques et les systèmes entièrement numériques, il existe encore néanmoins toute une série de solutions partiellement numériques incluant une quantité variable de composants numériques.

V.1. Systèmes de vidéosurveillance classique CCTV :

Un système de vidéosurveillance classique CCTV utilisant des caméras analogiques avec sorties coaxiales sont reliées au moniteur. Un opérateur doit être toujours présent devant la télé pour le contrôle et l'intervention aux cas anormaux

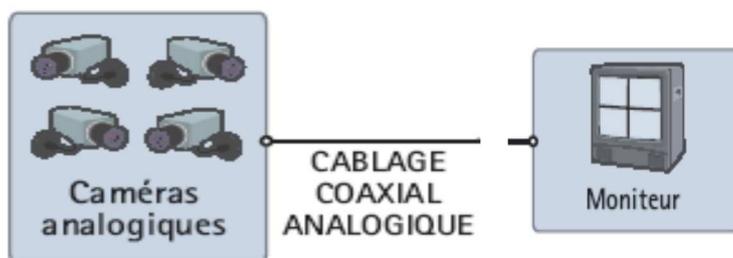


Figure V.6 : vidéosurveillance classique CCTV

Les avantages :

- Installation très simple et non professionnel
- Manipulation des données assez simple et disponible à n'importe quelle personne

Les inconvénients :

- Le fonctionnement de cette surveillance est très limité
- Il faut réserver un opérateur pour le contrôle
- Dans ce type de surveillance on ne parle pas d'enregistrement
- Pas de déclenchement de surveillance

V.2 Systèmes de vidéosurveillance analogique avec magnétoscopes traditionnels :

Un système de vidéosurveillance analogique utilisant un magnéscope traditionnel (VCR) est un système entièrement analogique dans lequel les caméras analogiques avec sorties coaxiales sont reliées au magnéscope pour l'enregistrement. Les bandes utilisées sont identiques à celles utilisées par les particuliers. Les séquences vidéo ne sont pas compressées. Dans le cas d'un enregistrement à vitesse maximale, une cassette a une durée maximale de 8 heures. Dans les systèmes de plus grande envergure, un quad/multiplexeur peut être connecté entre la caméra et le magnéscope. Le quad/multiplexeur permet alors d'enregistrer le contenu de plusieurs caméras sur un même magnéscope, mais selon une fréquence d'image cependant inférieure. La surveillance vidéo nécessite un moniteur analogique.

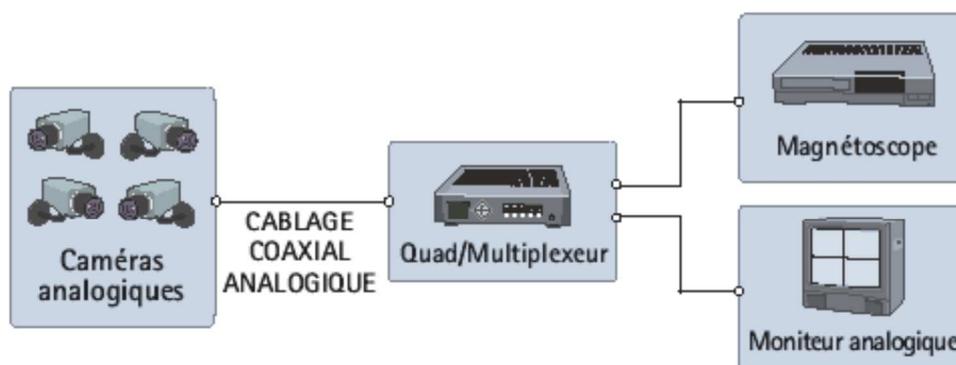


Figure V.7 : vidéosurveillance analogique avec magnétoscopes traditionnels

V.2.1 Multiplexeurs :

Définition : Les multiplexeurs remplissent deux fonctions :

- fournir au moniteur des images multiples (4, 9 ou 16)
- envoyer aux magnétoscopes des images encodées.

La fonction duplex permet ces deux utilisations simultanément. Les images enregistrées sont relues exclusivement à travers le multiplexeur et affichées dans tous les formats (1, 4, 9 ou 16).

V.3. Systèmes de vidéosurveillance analogique avec enregistreurs numériques :

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.

Les enregistreurs numériques présentent les avantages suivants :

- Pas besoin de changer de cassette
- Qualité constante de l'image

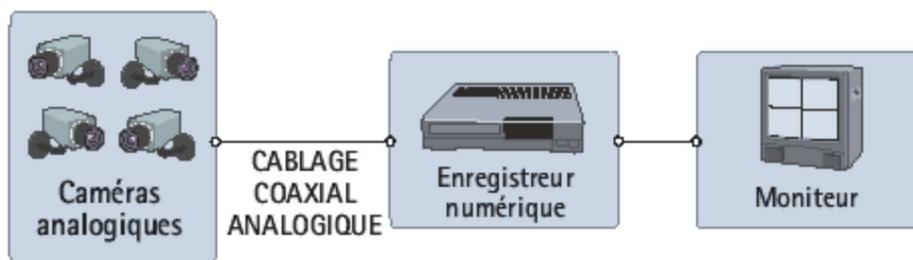


Figure V.8 : vidéosurveillance analogique avec enregistreurs numériques

V.4 .Systèmes de vidéosurveillance analogique avec enregistreurs numériques réseau :

Un système de vidéosurveillance analogique passant par un enregistreur numérique (DVR) réseau est un système en partie numérique comprenant un enregistreur numérique réseau connecté via un port Ethernet. La vidéo étant numérisée et compressée sur l'enregistreur numérique, les images peuvent être transportées sur un réseau informatique à des fins de surveillance sur PC distant. Certains systèmes permettent à la fois la visualisation des séquences en direct et des séquences enregistrées ; d'autres se limitent aux images enregistrées. Sur certains systèmes, la surveillance vidéo requiert en outre un client Windows spécifique, tandis que d'autres nécessitent un simple navigateur web standard, plus flexible pour une visualisation à distance.

Les enregistreurs numériques réseaux présentent les avantages suivants :

- Visualisation vidéo à distance sur PC
- Contrôle du système à distance

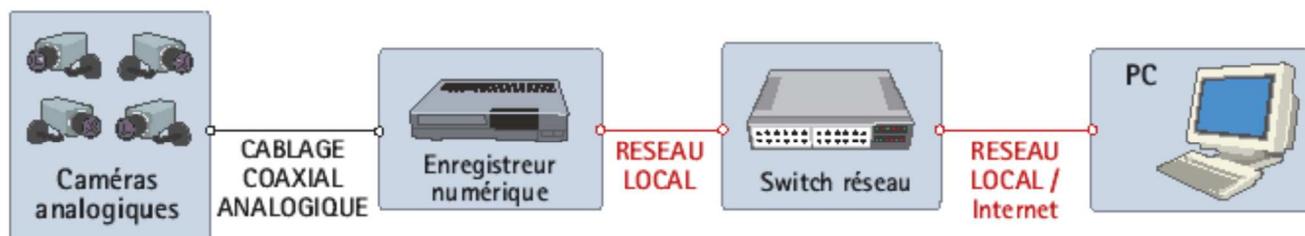


Figure V.9 : vidéosurveillance analogique avec enregistreurs numériques réseau :

V.5. 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.

Ce système utilisant les serveurs vidéo présentent les avantages suivants :

- ✓ Recours à un réseau standard et à un serveur informatique standard pour 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

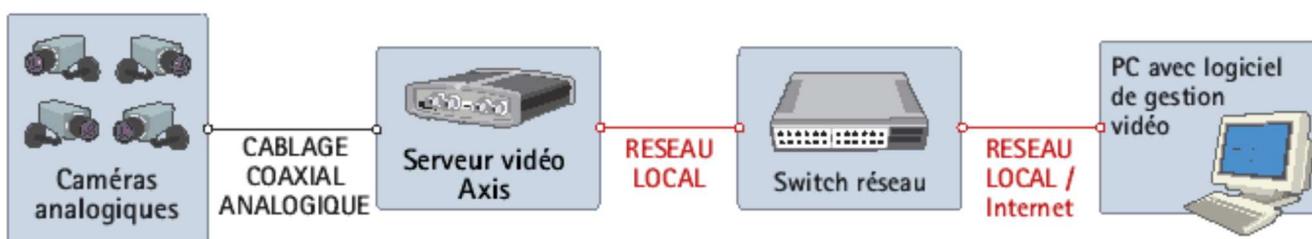


Figure V.10 : vidéo sur IP avec serveurs vidéo

Ce schéma présente un véritable système vidéo sur IP dans lequel les données vidéo sont transmises en permanence sur un réseau IP.

Le serveur vidéo sert de pierre angulaire pour assurer la migration du système de sécurité analogique vers une solution de vidéo sur IP.

V.5.1. Serveurs vidéo :

Les serveurs vidéo s'intègrent facilement dans un système existant de vidéosurveillance analogique CCTV (télévision en circuit fermé). Un serveur vidéo numérise les signaux vidéo analogiques et distribue des images numériques directement sur un réseau IP (par exemple, un réseau local LAN/intranet/Internet), en transformant les caméras analogiques en caméras réseau et en permettant aux utilisateurs de visualiser des images en direct à partir d'un navigateur web depuis tout ordinateur du réseau, en tout lieu et à tout moment.

Les avantages :

- ✓ Faible coût total de possession grâce à l'utilisation de l'infrastructure réseau et des équipements existants
- ✓ Fonctions complémentaires à celles d'un système de vidéo analogique CCTV Intégration facile dans les installations existantes et évolution future garantie
- ✓ Accès distant aux images en direct, à tout moment, en tout lieu, à partir de tout ordinateur agréé et autorisé muni d'un navigateur web
- ✓ Facilité de stockage d'images vidéo numériques sur des supports informatiques
- ✓ Large gamme de logiciels d'applications

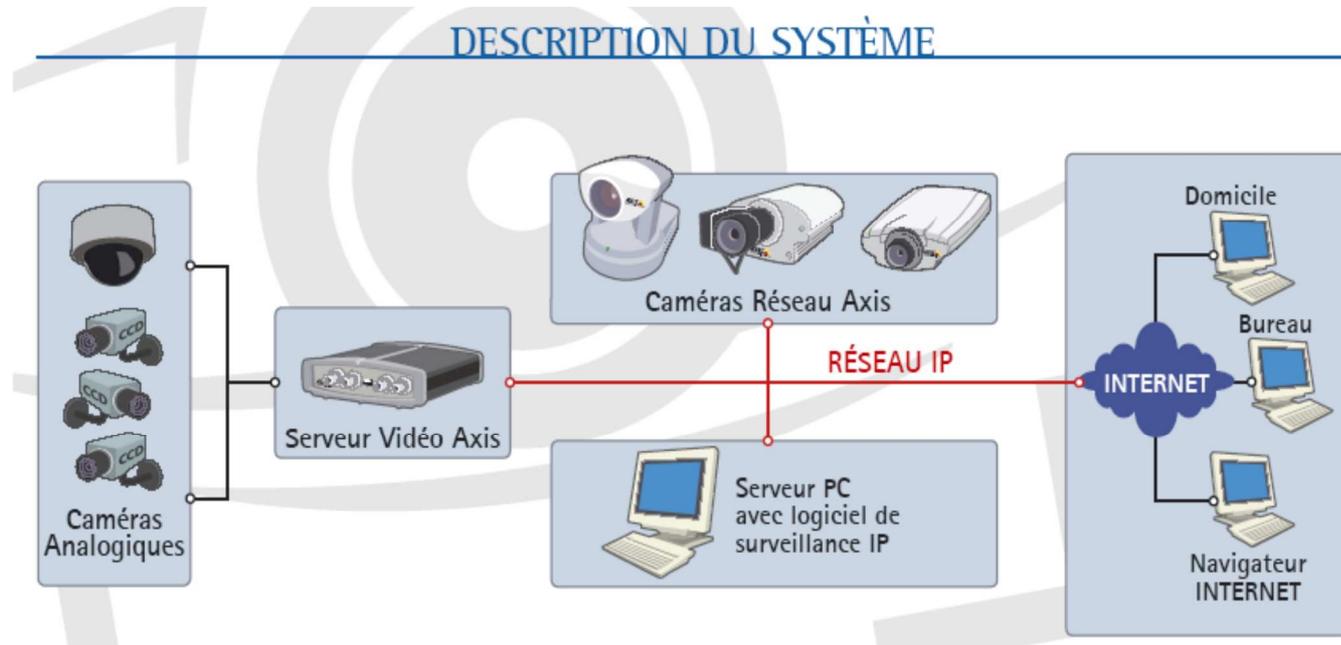


Figure V. 11 : Description du système

Un serveur vidéo comprend au moins une entrée vidéo analogique, un numériseur d'images, un compresseur d'images et un serveur web avec interface réseau. En associant un système de CCTV analogique à un système de vidéo réseau, vous disposez de toutes les fonctions et de tous les avantages de la technologie numérique : accès à distance, économie, souplesse, évolutivité, fonctions d'intégration et d'évolutivité et exceptionnelle qualité d'image.

V.5.2 Décodeur vidéo :

Dans certains cas, il est nécessaire de pouvoir surveiller les flux vidéo et audio IP sur un équipement analogique existant. Un décodeur vidéo IP permet dans ce cas de transformer les flux vidéo et audio du réseau en signaux analogiques qui seront interprétés par les écrans de télévision classiques, les moniteurs analogiques et les commutateurs vidéo. Un encodeur/décodeur est un moyen très économique de transmettre de la vidéo analogique sur de grandes distances (analogique – numérique- analogique)

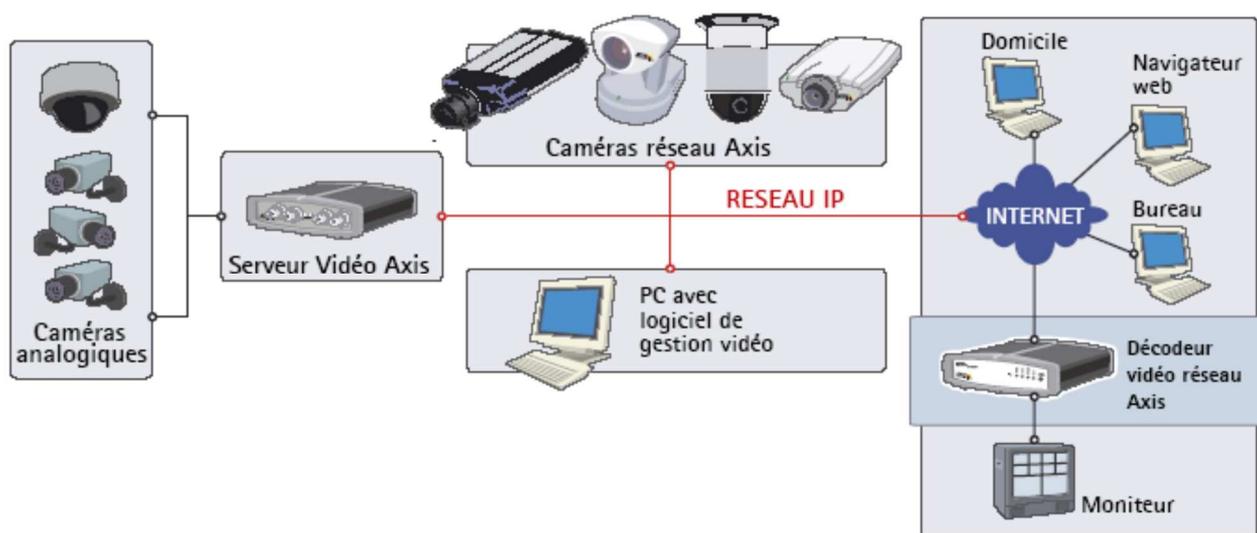


Figure V.12 : Décodeur vidéo

Grâce au décodeur vidéo, les moniteurs analogiques peuvent recevoir des informations vidéo et audio en provenance de caméras ou de systèmes analogiques comme s'ils étaient installés en local auprès de l'opérateur, alors qu'en réalité ils se trouvent par exemple dans une autre ville.

V.6 Systèmes de vidéo sur IP avec cameras réseau :

Une caméra réseau associe une caméra et un ordinateur. Permettant la numérisation et la compression vidéo, elle est en outre équipée d'un connecteur réseau. La vidéo est acheminée par réseau IP via les commutateurs réseau, pour être enregistrée sur un PC/serveur standard à l'aide d'outils de gestion vidéo. Il s'agit d'un système de vidéo sur IP à part entière, doublé d'un système entièrement numérique n'utilisant aucun composant analogique.

Les systèmes de vidéo sur IP reposant sur l'utilisation de caméras réseau présentent les avantages suivants :

- ✓ Caméras haute résolution (méga pixels)
- ✓ Qualité constante de l'image

- ✓ Fonction d'alimentation par câble Ethernet (Power over Ethernet) et réseau sans fil
- ✓ Fonctions panoramique/inclinaison/zoom, audio, entrées et sorties numériques sur IP
- ✓ Grandes flexibilité et évolutivité

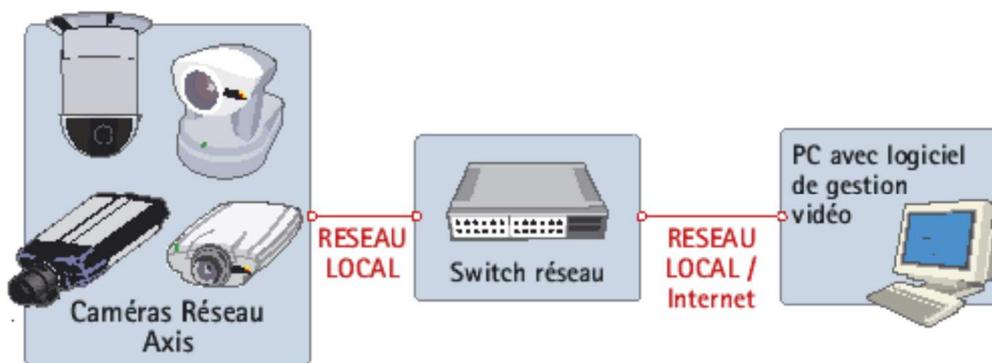


Figure V.13 : Systèmes de vidéo sur IP avec caméras réseau

Ce schéma présente un système de vidéo sur IP dans lequel la vidéo est transmise en permanence sur réseau IP à l'aide de caméras réseau. Ce système permet de profiter pleinement de tous les avantages offerts par la technologie numérique tout en offrant une qualité d'image constante entre la caméra et l'écran, quel que soit l'endroit où ils se trouvent.

V.7. Réseau sans fil :

Même si les réseaux filaires prévalent actuellement dans la plupart des bâtiments, une solution sans fil peut s'avérer intéressante pour l'utilisateur, tant financièrement que sur le plan fonctionnel.

Songons par exemple à certains bâtiments classés, où l'installation d'un câblage endommagerait inévitablement l'intérieur, ou à certains sites (commerces par exemple) pour lesquels la caméra doit être régulièrement déplacée et où l'on ne souhaite pas devoir tirer chaque fois de nouveaux câbles.

Une autre utilisation courante de la technologie sans fil concerne les bâtiments ou les sites que l'on souhaite relier sans pour autant devoir entreprendre de lourds et coûteux travaux au sol.

La technologie sans fil s'applique à la fois aux systèmes de vidéo sur IP et aux systèmes analogiques.

Elle dépasse donc le périmètre strict des réseaux.

La transmission sans fil se divise en deux catégories principales :

LAN sans fil (Wireless LAN, ou WLAN) :

Le LAN définit un réseau local, c'est-à-dire sur de courtes distances et en principe à l'intérieur.

Les normes LAN sont aujourd'hui bien définies et les périphériques de marques différentes sont généralement compatibles entre eux.

Ponts sans fil

Lorsque certains bâtiments ou sites doivent être reliés par une liaison rapide, une liaison point à point longue distance et à grande vitesse est nécessaire. Les technologies micro-ondes et laser sont couramment utilisées.

Réseau typique avec connexion filières et sans fils.

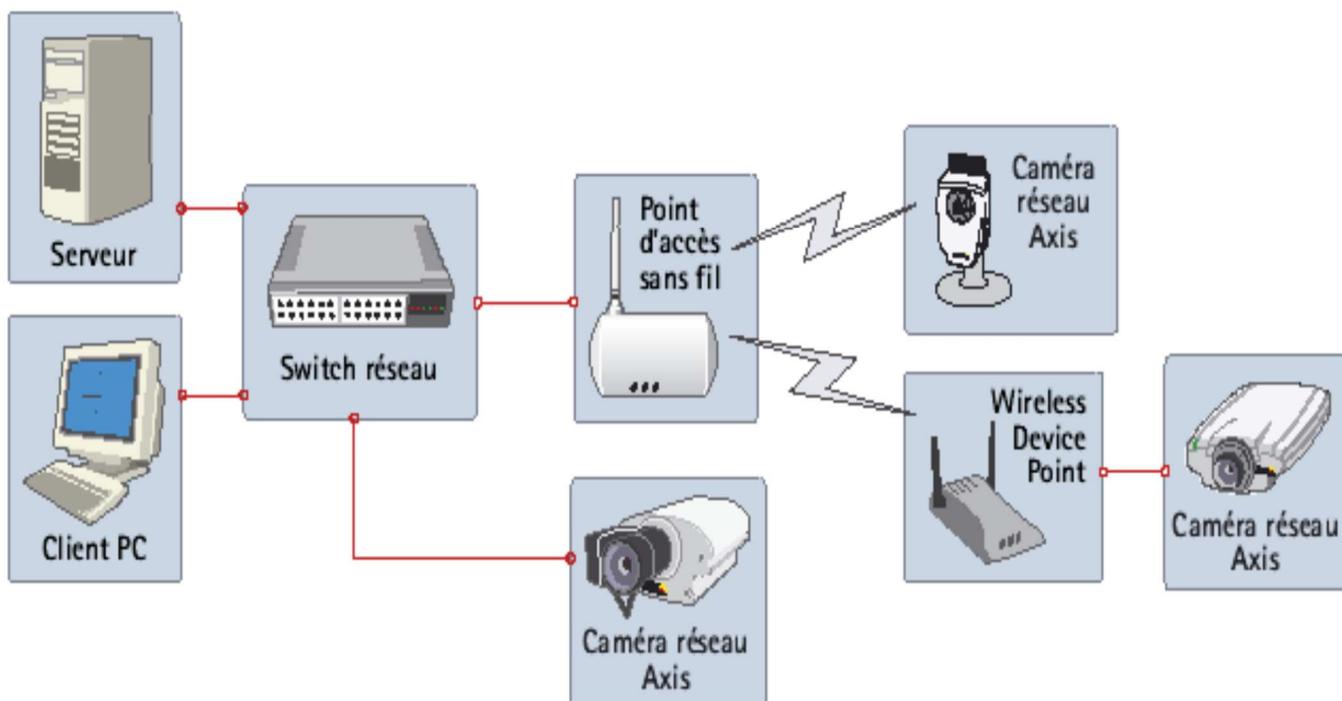


Figure V.14 : Réseau sans fil

Ce schéma présente un usage combiné des deux technologies :

- ✓ Le point central est un Switch/commutateur réseau. A gauche, un serveur (en haut) et un client PC (en bas) sont connectés via un câblage Ethernet. A côté du Switch se trouve un point d'accès sans fil. Celui-ci gère tous les appareils sans fil à proximité.
- ✓ Deux appareils sans fil sont représentés dans ce schéma :

La Caméra Réseau AXIS 206W. Cette caméra intègre le support pour les communications sans fil.

Un Wireless Device Point. Cet appareil fournit la communication sans fil et se connecte directement au point d'accès sans fil.

Il est également possible de connecter le serveur et le PC sans fil. La bande passante en mode sans fil étant limitée par rapport aux communications filaires, il est préférable d'utiliser les réseaux câblés aussi souvent que possible.

VI .Domaines d'applications :

La tendance de l'économie mondiale actuelle exige aux entreprises d'être réactive devant les demandes de plus en plus gourmandes de moyen de connectivité et infrastructure de communication et de marketing, la technologie de la vidéo sur réseau IP redynamise les applications de vidéosurveillance par de nombreuses fonctionnalités comme le contrôle à distance, la vidéo en temps réel. Ce qui rend plusieurs secteurs d'activités interactives. On dénombre trois grandes catégories publiques dans lesquelles l'on retrouve ces systèmes de surveillance :

- Les aéroports, les transports publics et les gares.
- Les lieux publics et les parkings. Qui se verront principalement doter de systèmes classiques pour la surveillance globale bien que les aéroports commencent à adopter les mesures biométriques.
- Le trafic autoroutier. Qui pour sa part privilégiera les caméras qui ont la possibilité de reconnaître les véhicules.

Les installations privées importantes concernent les casinos et autres salles de jeux qui font régulièrement appel à des systèmes d'identification faciale pour reconnaître les fraudeurs.

VI- Le traitement d'image :

Le traitement d'images est un domaine très vaste qui a connu, et qui connaît encore, un développement important depuis quelques dizaines d'années.

On désigne par traitement d'images numériques l'ensemble des techniques permettant de modifier une image numérique afin d'améliorer ou d'en extraire des informations.

De ce fait, le traitement d'images est l'ensemble des méthodes et techniques opérant sur celles-ci, dans le but de rendre cette opération possible, plus simple, plus efficace et plus agréable, d'améliorer l'aspect visuel de l'image et d'en extraire des informations jugées pertinentes.

VI.1 Définition d'une vidéo :

Une vidéo est une succession d'images à une certaine cadence. L'œil humain a comme caractéristique d'être capable de distinguer environ 20 images par seconde. Ainsi, en affichant plus de 20 images par seconde, il est possible de tromper l'œil et de lui faire croire à une image animée. On caractérise la fluidité d'une vidéo par le nombre d'images par secondes (en anglais frame rate), exprimé en FPS (Frames per second, en français frames par seconde). D'autre part la vidéo au sens multimédia du terme est généralement accompagnée de son, c'est-à-dire de données audio.

VI.2. Définition d'une image :

Une image est plutôt difficile à décrire d'une façon générale. Une image est une représentation du monde. En traitement d'image, la majorité du temps, on considère qu'il s'agit d'une fonction mathématique de $R \times R$ dans R où le couplet d'entrée est considéré comme une position spatiale, le singleton de sortie comme l'intensité (couleur ou niveaux de gris) du phénomène physique. Il arrive cependant que l'image soit dite "3D" donc la fonction est de $R \times R \times R$ dans R . Les images couleurs peuvent être représentées soit par trois images représentant les trois couleurs fondamentales, soit par une image de $R \times R$ dans $R \times R \times R$ [w3].

L'image numérique est l'image dont la surface est divisée en éléments de tailles fixes appelés cellules ou pixels, ayant chacun comme caractéristique un niveau de gris ou de couleurs prélevé à l'emplacement correspondant dans l'image réelle, ou calculé à partir d'une description interne de la scène à représenter [W4].

VI.3. Caractéristiques d'une image:

L'image est un ensemble structuré d'informations caractérisé par les paramètres suivants:

A. Dimension

C'est la taille de l'image. Cette dernière se présente sous forme de matrice dont les éléments sont des valeurs numériques représentatives des intensités lumineuses (pixels). Le nombre de lignes de cette matrice multiplié par le nombre de colonnes nous donne le nombre total de pixels dans une image [W4].

B. Résolution

C'est la clarté ou la finesse de détails atteinte par un moniteur ou une imprimante dans la production d'images. Sur les moniteurs d'ordinateurs, la résolution est exprimée en nombre de pixels par unité de mesure (pouce ou centimètre). On utilise aussi le mot résolution pour désigner le nombre total de pixels affichables horizontalement ou verticalement sur un moniteur; plus grand est ce nombre, meilleure est la résolution [W4].

C. 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 [W4].

D. Contraste:

C'est l'opposition marquée entre deux régions d'une image, plus précisément entre les

$$C = \frac{L1 - L2}{L1 + L2}$$

régions sombres et les régions claires de cette image. Le contraste est défini en fonction des luminances de deux zones d'images. Si L_1 et L_2 sont les degrés de luminosité respectivement de deux zones voisines A_1 et A_2 d'une image, le contraste C est défini par le rapport [W4] :

E. Images à niveaux de gris:

Le niveau de gris est la valeur de l'intensité lumineuse en un point. La couleur du pixel peut prendre des valeurs allant du noir au blanc en passant par un nombre fini de niveaux intermédiaires. Donc pour représenter les images à niveaux de gris, on peut attribuer à chaque pixel de l'image une valeur correspondant à la quantité de lumière renvoyée. Cette valeur peut être comprise par exemple entre 0 et 255. Chaque pixel n'est donc plus représenté par un bit, mais par un octet. Pour cela, il faut que le matériel utilisé pour afficher l'image soit capable de produire les différents niveaux de gris correspondant.

Le nombre de niveaux de gris dépend du nombre de bits utilisés pour décrire la "couleur" de chaque pixel de l'image. Plus ce nombre est important, plus les niveaux possibles sont nombreux [W4].

F. Images en couleurs:

Même s'il est parfois utile de pouvoir représenter des images en noir et blanc, les applications multimédias utilisent le plus souvent des images en couleurs. La représentation des couleurs s'effectue de la même manière que les images monochromes avec cependant quelques particularités. En effet, il faut tout d'abord choisir un modèle de représentation. On peut représenter les couleurs à l'aide de leurs composantes primaires. Les systèmes émettant de la lumière (écrans d'ordinateurs,...) sont basés sur le principe de la synthèse additive : les couleurs sont composées d'un mélange de rouge, vert et bleu (modèle R.V.B.) [W4].

VI.4. Les caractéristiques locales et globales :

1) Les caractéristiques globales :

Cette première étape dans notre traitement de l'image bas niveau consiste à représenter sous forme d'un vecteur caractéristique les données issues d'une chaîne de segmentation en régions d'images puis à calculer la distance entre ces images. Dans la recherche d'images similaires par le contenu, diverses statistiques (basés sur la couleur, la texture et la forme) forment la description de l'image. Ces statistiques, que nous appelons ici descripteurs peuvent être mises sous la forme d'un vecteur. La distance entre deux images est alors la distance dans un espace à n dimensions, où n est le nombre de descripteurs choisis. Ce mode de représentation implique un choix définitif dans la nomenclature utilisée pour représenter les images : chaque image doit avoir le même nombre de

descripteurs et chaque descripteur possède une place précise au sein du vecteur. La couleur, la forme et la texture sont les trois familles de descripteurs les plus couramment utilisés. Un vecteur caractéristique global caractérise l'image de façon générale.

Dans le cas de l'outil SIMILAR, du fait de la normalisation et du filtrage passe-bas sur les images en amont de la chaîne de segmentation, la texture n'est plus pertinente, les petits motifs ayant disparu. L'outil caractérise donc l'arrière-plan et les deux plus grandes régions en termes de couleur, de forme et de localisation dans l'image. La procédure retourne la liste des images considérées comme proches de l'image requête. Cette liste est ordonnée de façon croissante en fonction de l'indice de proximité de chaque image par rapport à l'image requête.

Le traitement est simplifié et nous permet d'obtenir le sens global d'une image. Par calcul de similarité des paquets d'images, nous pouvons maintenant associer manuellement les images à nos concepts. Cela dit cette approche nous confronte actuellement à deux problèmes distincts :

- Si les résultats sont concluants sur des images couleurs vu le choix des descripteurs de l'outil SIMILAR développé initialement pour traiter des affiches publicitaires couleur, nous ne pouvons pour le moment valider cette approche sur le fonds iconographique mis à disposition par la MIDR car près de 80% de ces images sont en niveau de gris. La chaîne de traitement de l'outil SIMILAR est en cours de modification pour prendre en compte un segmenteur en niveau de gris.
- L'identification de plusieurs éléments (concepts) distincts pour composer une scène donnée est également problématique actuellement. Prenons l'exemple de l'image figure 14, où l'on ne pourrait choisir les deux concepts Kiosque et Statue pour obtenir l'image donnée.

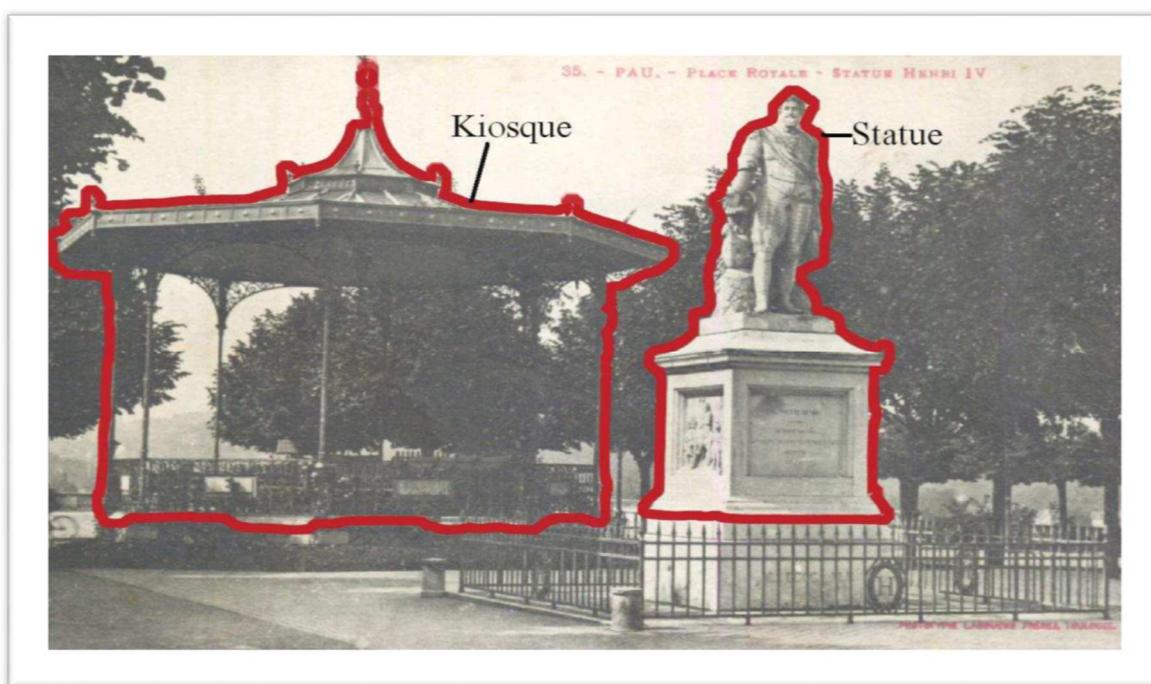


Figure VI.15 : Composition de concepts constituant une image

2) Les caractéristiques locales :

La solution envisagée consiste à identifier ces éléments principaux constituant l'image et d'extraire de l'image les vecteurs caractéristiques de chacun plutôt que de recalculer une signature générale de l'image qui nous fera perdre de l'information pertinente. Un vecteur de caractéristiques local est alors un ensemble de caractéristiques d'une partie de l'image et leur poids dans l'image.

VI.5. Technique de détection de caractéristiques locales et caractéristiques globales :

A la base, les caractéristiques présentées précédemment sont calculées de façon globale sur l'image. Cependant, un système basé uniquement sur des caractéristiques globales ne peut pas donner les résultats désirés. Soit une image composée de plusieurs objets ayant des caractéristiques, couleurs et textures, très différentes, le vecteur de caractéristiques global extrait à partir de l'image entière perd les informations locales (les objets) et ne produit qu'une moyenne grossière du contenu de cette image. Par contre, l'analyse uniquement basée sur des caractéristiques locales risque de perdre le sens global de l'image, en submergeant celui-ci dans un flot de petits détails inutiles. Par conséquent, un compromis doit être trouvé, différent selon les applications et selon les requêtes individuelles, entre caractéristiques globales et caractéristiques locales.

Deux approches différentes peuvent être employées pour calculer les caractéristiques locales. La première approche consiste à diviser une image en utilisant une grille et les caractéristiques de chaque bloc sont calculées pour chaque case de cette grille. Dans la figure 3(a), l'image est divisée en 12 zones de même taille et les caractéristiques sont calculées pour chacune de ces zones.

La deuxième approche consiste à segmenter l'image pour la diviser en zones locales plus proches des objets constituant l'image et ensuite calculer les caractéristiques pour chacune des régions extraites. La segmentation est une étape qu'on souhaiterait éviter, en raison de tous les problèmes pour choisir une "bonne" méthode de segmentation valide pour toutes les images de la base d'images. Cependant, la division d'une image en régions ou objets d'intérêt est souvent nécessaire pour pouvoir s'attaquer ensuite à l'extraction d'informations sémantiques à partir de l'image. Plutôt que de viser une segmentation exacte (en termes d'interprétation de l'image), on préfère souvent parler de « groupement de pixels » (Forsyth et al. 2003) ou encore de « segmentation faible » (Medioni et al. 2005) selon les auteurs. Dans les deux cas, il s'agit simplement d'oublier le découpage précis en termes d'objets, mais plutôt de découper l'image en régions similaires du point de vue caractéristiques mesurées sur l'image (voir figure 3b). Cette dernière idée apparaît comme plus logique, car l'idée même d'objet est une vue sémantique de l'image, or il est facile de conclure à la lumière des

innombrables travaux en segmentation des dernières décennies que segmentation bas niveau (basée uniquement sur les caractéristiques issues de l'image) et découpage sémantique de l'image sont deux idées complètement différentes. Un objet tel que nous le percevons dans l'image n'est pas forcément homogène en termes de caractéristiques, tandis qu'une zone homogène de l'image peut très bien contenir plusieurs objets différents.

Dans notre système de recherche d'images, nous combinons approche globale et approche locale dans la mesure de la similarité entre deux images. Du point de vue symbolique, les deux approches sont complémentaires dans la mesure de la similarité. En effet, tant que la requête n'est pas mieux définie (voir la section 4.3 sur l'interaction pour la recherche sémantique), aucune information ne permet de conclure que les images recherchées doivent être semblables dans leur globalité, ou dans leurs composants.

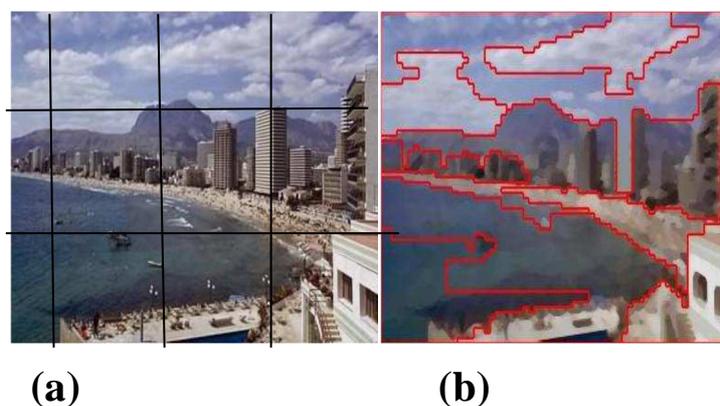


Figure.VI.16. Deux découpages différents de l'image. (a) Découpage en grille de l'image. (b) Découpage par la segmentation division-fusion de l'image.

L'approche locale utilisée repose soit sur la division d'une image, soit par découpage en grille ou par segmentation, en quelques zones générales d'intérêt, sans rechercher le calcul exact de toutes les régions. Pour ce faire, nous utilisons une technique très connue qu'est la segmentation par division-fusion.

Nous recherchons des régions homogènes, mais utilisons comme critère d'arrêt de l'algorithme la division de l'image en 3 à 10 régions maximum (voir figure 3b). Encore une fois, sans précision sur la requête, il est inutile de chercher le découpage « parfait » pour deux raisons : (1) la définition même de perfection dans ce cas est subjective à chaque personne voire à chaque requête et (2) les zones similaires recherchées dans la base d'images ne sont que similaires et pas identiques.

VI.6.Descripteurs de caractéristiques d'image :

Une fois les points d'intérêt détectés, il faut extraire les factures de ces points d'intérêt. Une grande variété de descripteurs ont été développés à cette fin. Les recherches récentes se sont concentrées sur la possibilité de rendre la combinaison détecteur-descripteur invariante aux transformations de l'image.

La technique la plus simple pour décrire une région d'une image est le vecteur de pixels. La mesure de la cross-corrélation permet de calculer un score de similarité entre deux descripteurs. Néanmoins ce descripteur de grande dimension rend les calculs de similarité assez lourds. Cette technique est donc applicable pour trouver les correspondances entre deux images, mais pas à l'analyse d'une base de données d'images. L'application d'une analyse en composantes principales permet éventuellement de réduire la dimension des vecteurs²¹.

a) Descripteurs basés sur les distributions :

Le principe de ces descripteurs est d'utiliser des histogrammes pour représenter différentes caractéristiques d'apparence ou de forme. Une méthode simple est de représenter la distribution des intensités des pixels sous la forme d'histogrammes. Une approche plus évoluée²² est la création d'histogrammes en deux dimensions, représentant la distance par rapport au centre de la région d'intérêt et la valeur du pixel.

Lowe²³ propose un modèle (SIFT, Scale Invariant Feature Transform) qui combine un détecteur de points d'intérêt invariant aux changements d'échelle et un descripteur basé sur la distribution des gradients calculés sur la région d'intérêt. Comme illustré par la figure 11, SIFT calcule l'amplitude et l'orientation du gradient pour une sélection de points des régions d'intérêt (échantillonnage dépendant de la taille de la région traitée). Ces amplitudes sont ensuite pondérées par une gaussienne (le cercle bleu sur l'image correspond à un contour d'équiprobabilité) qui permet de prendre en compte la distance par rapport au centre de la région. Ces gradients sont ensuite accumulés dans des histogrammes d'orientation pour chaque sous-région (associée à un bin de l'histogramme spatial). On se retrouve donc avec des histogrammes tridimensionnels (deux dimensions pour la position et une dimension pour l'orientation). La figure 11 montre une région de 8x8, découpée en 2x2 sous-régions mais, en pratique, SIFT travaille sur des régions de 16x16 (après échantillonnage dépendant de l'échelle) découpées en 4x4 sous-régions. Le résultat est stocké dans un vecteur de 128 dimensions, correspondant aux 8 orientations pour chacune des 16 bins de position. La quantification de la position et de l'orientation des gradients rend le descripteur robuste aux petites déformations géométriques et à l'imprécision de la détection des points d'intérêt.

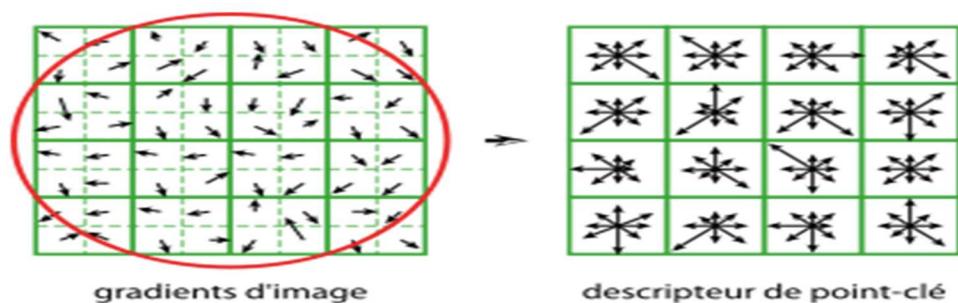


Figure 17 : SIFT

b) Descripteurs basés sur l'analyse spatio-fréquentielle :

Plusieurs techniques permettent de décrire le contenu fréquentiel d'une image. La transformée de Fourier décompose l'image en une somme de fonctions de base, mais cette approche est difficile à mettre en œuvre car les relations spatiales entre les points ne sont pas explicites et la base de fonction est infinie. Il est donc difficile de généraliser cette analyse à des descripteurs locaux. Les filtres de Gabor²⁸ et la théorie des ondelettes proposent des solutions à ces problèmes et sont très largement utilisés pour l'analyse de textures.

c) Descripteurs différentiels :

Un ensemble de dérivées de l'image calculées jusqu'à un certain ordre fournissent une approximation du voisinage d'un point. Les propriétés des dérivées locales (*local jet*) ont été étudiées par Koenderick et Van Doorn²⁹. Freeman et Adelson³⁰ développent l'idée des "steerable filters" qui orientent les dérivées dans la direction du gradient et rendent la méthode invariante aux rotations. Une estimation des dérivées est donnée par la convolution avec les dérivées d'une gaussienne.

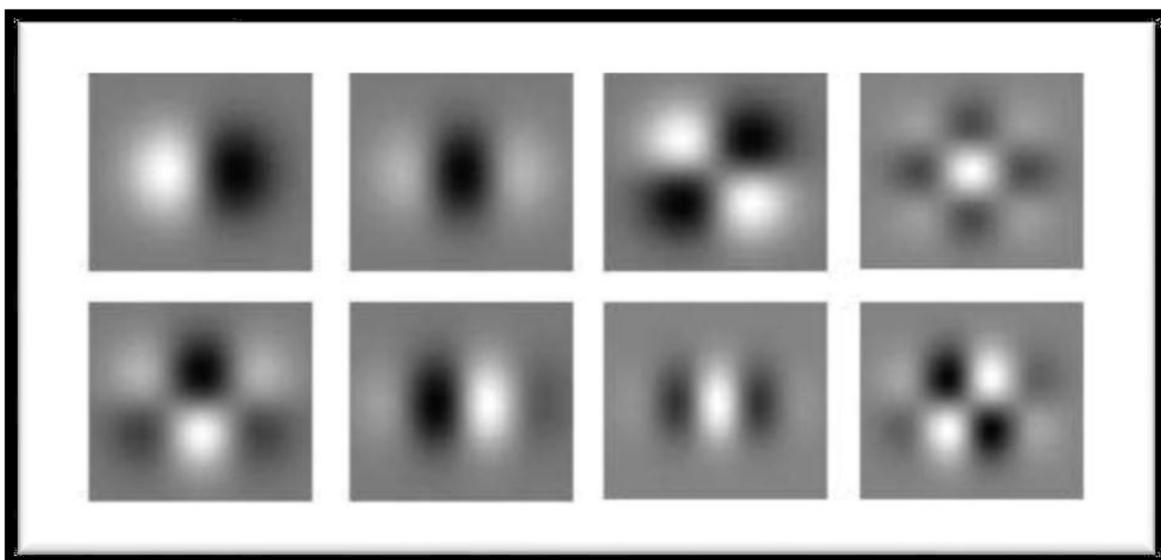


Figure VI.8: Dérivées de gaussienne, jusqu'à l'ordre 4

Conclusion :

la technologie de la vidéosurveillance à été étudié dans ce chapitre voir son historique depuis son invention jusqu'à nos jours , connaître les différents types de système de vidéosurveillance et leur composition ,les différents type de caméra de surveillance qui nous a permis d'avoir leur domaine d'application par la suite nous avons présenté quelques notion de base sur le traitement d'image dans le chapitre III nous allons plus détaillée sur ces nouvelles technologie qui nous a permis d'avoir plus clair et de comprendre profondément comment ce fonctionne la vidéosurveillance

I. Introduction :

Le développement rapide de la technologie de capture et l'édition vidéo numérique a conduit à une croissance rapide des données vidéo qui créent le besoin de techniques efficaces et avancées pour récupération de la vidéo. À cette époque, par contrainte de temps, les utilisateurs ne disposent pas de suffisamment de temps pour regarder la vidéo entière. Dans de tels cas, l'utilisateur peut consulter le résumé de la vidéo qui donne un aperçu de la vidéo originale. Beaucoup des gens passent leur temps libre à regarder ou jouer différents sports comme le football, le cricket, le basket-ball, etc., mais il est impossible de regarder chaque match en raison d'une longue période de jeu. Donc, ces gens préfèrent regarder les grands moments du jeu. Ainsi, selon l'exigence de l'utilisateur, il existe un besoin croissant pour résumer les vidéos. Annotation de la vidéo est un processus qui facilite la navigation rapide parmi les grandes collections de vidéos. Il permet également l'indexation de contenu plus efficace. Annotation de la vidéo se réfère à la création d'un résumé d'une vidéo qui porte sur trois points principaux.

- ✓ Le résumé vidéo doit contenir des scènes et des événements non seulement le plus court possible de la vidéo, mais aussi le plus important. Par exemple, dans un match de football, le résumé doit contenir des objectifs, des fautes, des limites de tir, tentative de but, et quelques autres scènes importantes.
- ✓ Le résumé vidéo devrait maintenir une connexion continue entre les scènes. Cela signifie que le résumé vidéo ne doit pas contenir des segments vidéo connectés d'une manière aveugle.
- ✓ La vidéo résumée ne doit contenir aucune redondance.

Dans ce chapitre, on va donner un aperçu général sur l'annotation des vidéos, ainsi qu'un ensemble de méthodes et de techniques utilisé pour la réaliser. À la fin, on va présenter quelques travaux d'annotation de vidéo à grande échelle.

II. Présentation de l'annotation vidéo :

L'annotation de vidéo est un mécanisme pour produire un court résumé d'une vidéo pour donner à l'utilisateur un résumé visuel et utile de la séquence vidéo, il peut être une des images (images clés) ou des images (écumes vidéo) en mouvement. Quand à regarder et navigation, un bon résumé de la vidéo permettra à l'utilisateur d'obtenir un maximum d'informations à propos de la séquence vidéo cible dans une contrainte de temps spécifiée ou des informations suffisantes dans le minimum de temps [11]. Ces résumés générés peuvent aider les utilisateurs à naviguer sur de grandes

archives vidéo et de prendre des décisions de manière plus efficace en ce qui concerne la sélection, la consommation, le partage ou la suppression de contenu [12]. Les vidéos résumées peuvent également être utilisés comme un produit final à partager, digéré, et apprécié par l'utilisateur. Retenir uniquement les informations essentielles d'une séquence vidéo améliore le stockage, la bande passante et le temps d'écoute [11]. Les techniques développées dans la vidéo annotation touchent divers domaines, tels que les films, sports, vidéos, e-Learning, etc., Le processus d'abstraction de vidéo à généralement trois phases:

Analyse de l'information de la vidéo, la sélection du clip significative et la synthèse production.

Pour analyser les informations vidéo, il est nécessaire de détecter les contours, les structures ou les motifs dans la composante visuelle, la composante audio et le composant textuel comme des sous-titres.

II. l'annotation vidéo basée sur les images clé:

L'annotation vidéo basée sur les images clés commence par extraire les frames de la séquence vidéo. Dans la deuxième étape les images extraites sont regroupées en se basant généralement sur la ressemblance entre les frames successives, si la différence dépasse un certain seuil on crée un nouveau groupe de frame sinon on ajout le nouveau frame au groupe en court de traitement. La sélection des images clés est effectuée par l'étape trois en sélectionnant en général les bords des groupes c'est-à-dire soit le premier ou le dernier frame. La procédure entière présentée sur la Figure qui suit.

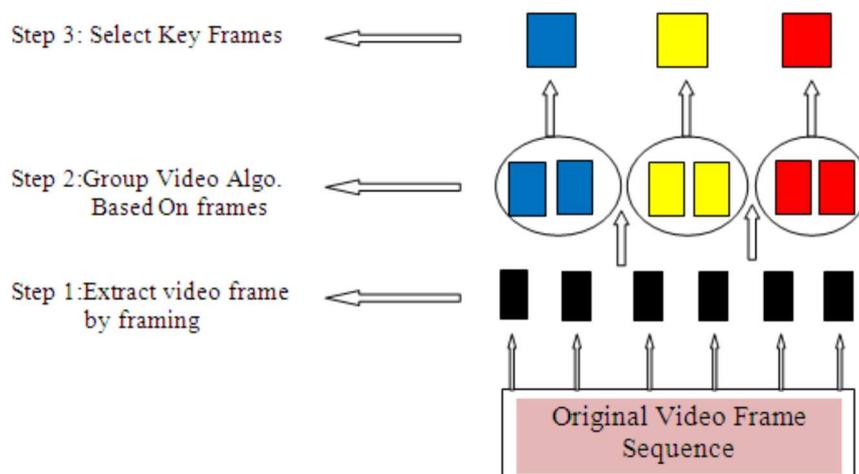


Figure I.1: l'annotation vidéo basée sur les images clés [19]

En général, l'annotation vidéo basée sur les images clés peut être effectuée en trois façons différentes. Ceux-ci sont les suivantes.

1) Une annotation basée sur l'échantillonnage:

Cette méthode choisit des images clés uniformément ou aléatoirement sous-échantillonnage, sans tenir compte du contenu vidéo [13]. Le résumé produit par ces méthodes ne représente pas toutes les parties vidéo et peut provoquer une certaine redondance des images clés avec des contenus similaires.

2) Annotation basée sur une segmentation de scène:

Cette méthode extrait les images clés en utilisant la détection des scènes, la scène comprend toutes les pièces avec un lien sémantique dans la vidéo ou dans le même espace ou dans le même temps. L'inconvénient de ces techniques est la production d'un résumé, qui ne tient pas compte de la position temporelle des images.

3) Annotation basée sur la segmentation du shot:

Cette méthode extrait les images clés en utilisant la détection des shots, et les images clés choisis sont la première et la dernière image de chaque shot.

Ces méthodes sont efficaces pour un shot avec une petite variation du contenu, mais ils ne fournissent pas une représentation adéquate du shot avec des mouvements forts.

III. Etat de l'art sur les techniques et les méthodes d'annotation vidéo :

Devant le volume grandissant des données audiovisuelles, L'annotation de vidéo a pour objectif de fournir des informations pertinentes et concises afin d'aider l'utilisateur à naviguer ou à organiser ses fichiers vidéos plus efficacement. Deux méthodes d'annotation vidéo peuvent être retrouvées dans la littérature [14, 15] que nous appellerons en français : annotation statique (video summary) et annotation dynamique (video skimming). L'annotation statique consiste à sélectionner les images les plus représentatives de la vidéo. Ces images appelées images clés se présentent en général sous la forme d'un scénarimage (storyboard). L'annotation dynamique se compose d'extraits de la vidéo et correspond à une version courte de la vidéo originale.

III.1. L'annotation vidéo statique :

L'annotation statique de vidéo se compose d'un ensemble d'images qui représente le contenu de la vidéo. Beaucoup de travaux ont été réalisés ces dernières années et quatre familles de d'annotation statique ont été dégagées dans [14] : méthodes reposant sur l'échantillonnage, sur les plans, sur les scènes et autres. Nous avons repris ces quatre familles d'annotation que nous avons complétées avec des travaux plus récents.

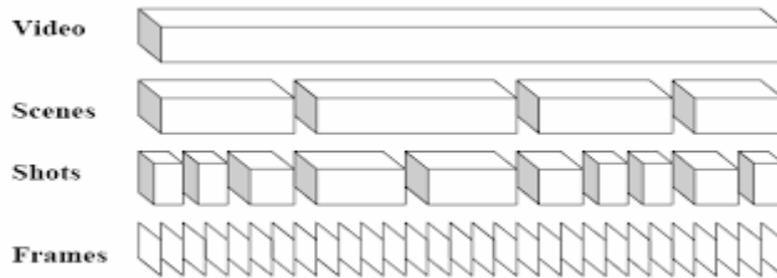


Figure III.1 : La structure hiérarchique d'une séquence vidéo

III.1.1 Méthodes basées sur l'échantillonnage :

Les premiers travaux [16, 17] concernant les annotations des vidéos consistent à choisir les images clés en sous-échantillonnant uniformément ou aléatoirement la séquence originale. L'inconvénient des méthodes qui n'étudient pas le contenu, est le non représentation de certaines parties de la vidéo et la possible redondance de certaines images clés avec des contenus similaires.

III.1.2 Méthodes basées sur les plans :

Des travaux plus élaborés tentent d'extraire des images clés en s'adaptant au contenu de la vidéo. La détection des plans est réalisée pour mieux ajuster la sélection des images clés au contenu de la vidéo. De nombreuses méthodes existent pour détecter les plans et une description détaillée. Une façon simple pour représenter les différents plans de la vidéo est d'extraire la première image du plan comme image clé [18, 19] ou les première et dernière images du plan [20].

D'autres travaux choisissent alors de représenter le contenu des vidéos en employant des caractéristiques visuelles de bas niveau comme la couleur, le mouvement ou la texture.

Dans [21, 22], le nombre d'images clés dépend du contenu des plans. La première image du plan est sélectionnée comme image clé. Puis, si la distance entre l'histogramme couleur de la dernière image clé sélectionnée et l'image courante est supérieure à un seuil alors une nouvelle image clé est choisie. La sélection de la première image n'est pas forcément judicieuse puisqu'elle peut être soumise aux effets de transition (fondu enchaîné) entre les plans. L'approche décrite dans [23] compare les images d'un plan suivant leurs histogrammes couleur puis réalise le rassemblement des images en plusieurs groupes. Seuls les groupes de taille assez importante sont conservés et les images les plus proches du

centre de gravité de chaque groupe sont alors choisies comme images clés. Dans [24], l'histogramme couleur est utilisé comme vecteur caractéristique pour représenter les images et la courbe dans cet espace caractéristique est considérée.

Des travaux ont également étudié différentes caractéristiques de bas niveau. Dans [25], plusieurs descripteurs (histogrammes couleur et mouvement) sont extraits pour chaque image du plan. Les vecteurs caractéristiques des différents descripteurs sont alors réunis pour créer un unique vecteur pour chaque image et former une courbe temporelle dans l'espace des caractéristiques. Les images clés sont obtenues en sélectionnant des points appropriés sur la trajectoire pour la caractériser. Une méthode similaire est discutée dans [26, 27], où une courbe est définie et caractérise des différences d'images suivant trois descripteurs (histogrammes couleur, histogrammes des orientations et analyse par ondelettes pour décrire la texture).

L'inconvénient de travailler au niveau des plans est que le nombre d'images clés peut être trop important pour représenter la vidéo.

III.1.3 Méthodes basées sur les scènes ou macro-segments :

Des travaux ont été réalisés en ne considérant pas comme unité de la vidéo le plan mais en définissant des unités de plus haut niveau. Suivant les auteurs, différents niveaux de hiérarchie sont étudiés pour créer l'annotation vidéo. Par exemple, le regroupement de plans par similarité peut être considéré comme un niveau plus élevé que la segmentation en plans et aura pour conséquence de sélectionner moins d'images clés.

Des méthodes d'annotation ont été conçues pour décrire seulement les plans les plus représentatifs de la vidéo. L'approche décrite dans [28] consiste d'abord à assigner un groupe à chaque image, puis à réunir les deux groupes les plus similaires de manière itérative. L'algorithme s'arrête quand la similarité entre les groupes est inférieure à un seuil. Pour chaque groupe créé, les images adjacentes sont déterminées et constituent un segment. Une mesure d'importance est alors calculée pour chaque segment suivant sa longueur et sa rareté. Tous les segments de petite taille sont alors supprimés et seuls les segments ayant une mesure d'importance suffisante sont conservés. Puis, l'image la plus proche du centre de gravité de chaque segment est extraite comme image clé. Le résumé est ensuite présenté en ajustant la taille des images clés suivant leur importance. Un algorithme pour organiser les images clés selon leur taille a été proposé dans [29] afin d'optimiser l'affichage du résumé. Gong et al. [30] utilisent une décomposition en valeur singulière (SVD) pour créer le résumé de vidéo. Ils créent une matrice où chaque colonne contient un vecteur caractéristique associé à chaque image de la vidéo. Une SVD est ensuite réalisée pour réduire l'espace des caractéristiques et elle est à l'origine

d'une mesure sur le changement visuel d'un groupe d'images. Ils l'emploient dans un algorithme de regroupement des vecteurs caractéristiques dans l'espace réduit afin de former des groupes d'images avec des contenus similaires.

Finalement, les plans les plus longs de chaque groupe sont conservés et l'image la plus proche du centre du groupe est choisie comme image clé. Dans [31], une mesure de pertinence est associée à chaque plan suivant différentes caractéristiques d'identification d'objets et les plans sont sélectionnés dans l'ordre décroissant de cette mesure jusqu'à atteindre la taille du résumé souhaitée.

L'inconvénient des approches qui travaillent sur la totalité de la vidéo est qu'elles peuvent s'avérer très contraignantes par le temps de calcul et la mémoire sollicitée.

Pour s'adapter à la demande des utilisateurs, des méthodes ont été élaborées pour créer un résumé de vidéo hiérarchique. Par exemple, Ferman et al. [32] proposent un résumé hiérarchique avec deux niveaux de résolution. Le premier est réalisé en appliquant un algorithme de regroupement (fuzzy c-means) pour chaque plan de la vidéo afin d'extraire un ensemble d'images clés non redondantes qui représente les différents plans de la vidéo. Le second consiste à calculer une matrice de similarité entre les images clés et à regrouper les similaires suivant le nombre d'images demandé par l'utilisateur. De manière générale, différents algorithmes de regroupement par similarité ont été utilisés et parfois adaptés pour créer un résumé hiérarchique (algorithme des k-moyens [33, 34], algorithme « fuzzy c-means »

III.2 L'annotation vidéo dynamique :

Comparé au l'annotation statique, l'annotation dynamique conserve les propriétés dynamiques de la vidéo et par conséquent il est plus agréable à regarder que le résumé statique. De plus l'information audiovisuelle est préservée et donc fournit une représentation plus proche de la vidéo originale. Dans [15], deux types d'annotation dynamique sont distingués : résumé suivant des extraits clés (highlight) et résumé donnant une vue d'ensemble de la vidéo (summary sequence). Ce dernier est utilisé pour donner à l'utilisateur une impression globale du contenu de la vidéo alors que le résumé par extraits clés contient seulement les passages les plus intéressants de la vidéo, comme une bande annonce d'un film qui montre les scènes les plus attrayantes sans révéler la fin de l'histoire.

La création de résumé suivant des extraits clés (highlight) est une tâche difficile sans connaissance a priori sur la nature de la vidéo. Dans [35], une méthode a été conçue pour retenir certaines scènes importantes de la vidéo (scènes contenant des objets ou des personnes, scènes d'action, les scènes où un dialogue est reconnu, suppression des scènes à la fin du film. . .). Toutes les scènes qui ont été sélectionnées forment la bande annonce du film.

Cependant, la création d'annotation suivant des extraits clés (highlight) repose en général sur des hypothèses trop élémentaires pour être véritablement efficaces. Une amélioration possible pour détecter les événements importants est la recherche des explosions, des fusillades ou des gros plans qui peuvent être considérés comme des extraits clé.

Beaucoup d'approches essaient de résoudre le problème du résumé suivant des extraits clés en exploitant des connaissances a priori sur le type de la vidéo, comme les vidéos de sport ou des vidéos avec des caractéristiques spécifiques. Pour les vidéos de sport, les travaux sont basés sur la détection d'événements comme dans des vidéos de football [36], de baseball [37] ou de basketball [38]. Par exemple, dans [39], un modèle de chaînes de Markov, entraîné par un algorithme EM (Espérance-Maximisation), est utilisé pour détecter les phases de jeu et de pause dans une vidéo de football. Ils développent également une technique pour détecter les lancers du baseball en construisant une hiérarchie suivant les phases du lancer. L'annotation dynamique est alors conçu en sélectionnant les phases importantes. Pour les journaux télévisés, un système est également proposé dans [40]. La plupart des approches qui élaborent un résumé dynamique (ou statique) reposent sur des caractéristiques de bas niveau. Elles ne garantissent pas que les résumés créés contiennent tout le contenu de la vidéo, et donc que ces résumés aient un sens pour l'utilisateur. Des informations sémantiques ont alors besoin d'être ajoutées pour améliorer la performance du résumé. Mais peu de travaux ont été réalisés pour annoter le contenu des vidéos dans un cadre général. Pour collecter ce genre d'informations, le contenu de la vidéo est parfois annoté manuellement. , les résumés suivant des extraits clés restent un procédé subjectif dont une extraction automatique et efficace est encore loin d'être réalisée.

La plupart des recherches sur l'annotation dynamique se focalisent sur le résumé donnant une vue d'ensemble de la vidéo (summary sequence). Dans [41], la vidéo est découpée en sous-plans où un index sur l'information de mouvement est calculé. La sélection d'images clés s'effectue sur ces sous-plans et le résumé dynamique est ensuite créé par interpolation des images clés. Le système « CueVideo » décrit dans [42] consiste à changer la vitesse de lecture de la vidéo. Il délivre des passages plus rapides en présence de scènes statiques et moins rapides pour les scènes dynamiques. Néanmoins, ce type d'approches ne permet pas d'avoir des compressions élevées de la vidéo.

Comme la compréhension automatique du contenu est encore loin d'être atteinte, des approches essaient de diminuer la redondance visuelle de la vidéo. Une métrique basée sur l'entropie est définie dans [43] pour mesurer la redondance du contenu de la vidéo et un résumé est obtenu en optimisant cette mesure. L'approche développée dans [44, 45] consiste à créer un résumé basé sur un graphe orienté. Ils construisent un graphe orienté où les noeuds sont les plans et les arrêtes symbolisent la

distance entre les plans. Le chemin le plus long forme alors l'annotation dynamique. Dans [46], les auteurs essaient de modéliser l'attention portée par un observateur. Plusieurs caractéristiques sont extraites comme le contraste, le mouvement, le son et le texte, elles sont utilisées pour définir des mesures d'attention. Elles sont ensuite combinées pour obtenir une courbe temporelle d'attention. L'annotation dynamique est alors créée en sélectionnant les portions de la vidéo où la courbe d'attention est maximale.

Un résumé statique peut également être obtenu et une méthode [47] a été développée

Pour optimiser la visualisation du résumé. Ngo et al. [48] construisent une méthode de résumé qui repose sur des graphes pour remonter aux scènes de la vidéo et sur des mesures d'attention pour sélectionner dans les scènes, les sous-plans appropriés pour le résumé.

V. Traitement vidéo de la base de données sur hadoop :

V.1.MapReduce pour traitement la base de données de la vidéo :

Map et Reduce les fonctions de MapReduce sont toutes deux définies par rapport à des données structurées en paires clé-valeur. En bref, nous pouvons exécuter le traitement distribué en créant des paires clé-valeur sous forme MapReduce. Cependant, pour des données non structurées telles que des données vidéo, on peut supposer qu'il est plus difficile de créer des paires de clés valeurs et exécuter le traitement de données structurées.

Dans cette expérience, afin d'utiliser le langage de programmation Ruby, nous utilisons un package d'extension de Hadoop, à savoir Hadoop streaming. En vue d'effectuer un traitement parallèle distribué sur les formes MapReduce, nous avons besoin de créer des programmes à utiliser comme Map et Reduce les fonctions dans le langage de programmation Ruby. Le traitement de données vidéo est exécuté en divisant les données dans une base de données vidéo et de création de paires de clés valeurs. Par exemple, le numéro d'image peut être utilisé comme une clé pour une image vidéo. Dans le cas d'un traitement parallèle d'une image vidéo, l'image vidéo est divisée en plusieurs parties, et les références peuvent être les clés (identifiants) pour ces différentes parties.

Le sortie est exécuté à l'aide du numéro de la clé, et joignant cadres séparés ou séparés des pièces est réalisée par la fonction Reduce. La figure V.3 montre un exemple de flux de traitement à l'aide MapReduce. Dans cette figure, chaque image vidéo est divisée en quatre parties, et chaque partie à un numéro de clé unique.

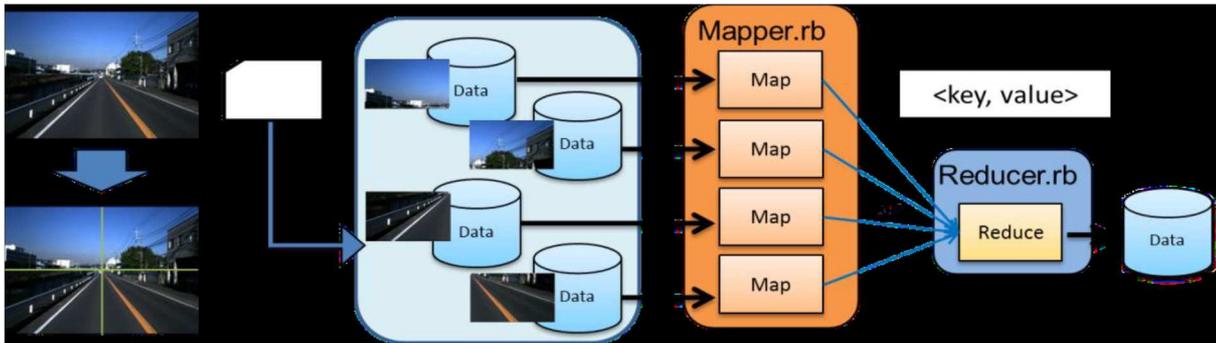


Figure V.3 : Flux de Traitement d'Image en utilisant MapReduce

V.2.Traitement MapReduce de base de données vidéo:

Nous avons implémenté les étapes de traitement de base de données vidéo suivante en utilisant Hadoop en streaming. Plusieurs images vidéo séquentielles sont entrées, et le traitement d'image est effectué pour chaque image vidéo. Et implémenté une fonction de Map pour traitement de l'image. L'entrée de la fonction de Map est une seule image vidéo, et la fonction de Map produit une image vidéo comme production .Et traitons des images vidéo en parallèle avec les serveurs esclaves, qui utilisent HDFS. La base de données vidéo est stockée dans HDFS. Chaque processus de Map renvoie également à HDFS. La figure :v.4 illustre un exemple de flux de traitement de données vidéo en utilisant la fonction de Map et HDFS.

Le processus de sélection image-clé est le seuil rapide, gratuit et les images clés extraites sont représentatives. Dans l'algorithme descripteurs image visuel [49] trois caractéristiques visuelles: couleur histogramme, statistiques ondelettes et direction de bord histogramme sont utilisés pour la sélection d'images clés. Des mesures de ressemblance sont calculées pour chaque descripteur et combinés pour former une mesure de différence d'image. Fidélité, Degré Reconstruction du shot, Compression Ratio qualités sont utilisées pour évaluer le résumé vidéo [49].

Dans le modèle de l'Attention de mouvement [50] les shots sont détectés en utilisant la répartition des couleurs et le bord taux de recouvrement qui augmentent la précision de la détection du shot. Les images clés sont extraites de chaque prise de vue en utilisant le modèle de l'attention de mouvement. Ici la première et la dernière image de tous les shots sont considérées comme image clé et les autres sont extraits en adoptant le modèle de l'attention de mouvement [51] [50]. Ces images clés sont ensuite regroupées et une valeur de priorité est calculée en estimant l'énergie de mouvement et de la variation de la couleur des shots de feu. Dans le descripteur visuel Multiple Caractériser l'algorithme [52], les images clés sont sélectionnées en construisant le graphique cumulatif pour les valeurs de différence des images. Les images à la forte pente indiquent le changement visuel significatif; par conséquent, ils sont sélectionnés et inclus dans le résumé final.

Mouvement méthode mise au point [53] se concentre sur un seul mouvement à vitesse constante et aligne les images vidéo en fixant le mouvement ciblé dans une situation statique. Un résumé est généré contenant tous les objets en mouvement et intégré avec des informations spatiales et de mouvement. Soustraction de l'origine et min shote sont principalement utilisés dans la convergence du mouvement. Dans mouvement Camera et l'objet de [54], la vidéo est segmenté en utilisant des classes à base de mouvement de la caméra: panoramique, zoom avant, zoom arrière et fixé. Finales des sélections d'images clés de chacun de ces segments sont extraites basés sur la confiance valueformulated pour le zoom, le panoramique et segments stables.

Conclusion :

L'annotation vidéo joue un rôle important dans de nombreuses applications vidéo. Une enquête sur les différentes méthodes pour l'image clé à base d'annotation vidéo à été réalisée. Mais il n'y a pas une méthode universellement disponible pour l'acceptation d'annotation de la vidéo qui donne un meilleur rendement dans toutes sortes de vidéos. A travers ce chapitre, Nous avons présentant annotation de vidéo, cité différents méthode et techniques, le traitement des donnée vidéo sur hadoop est proposé quelque travaux. Notre prochain chapitre sera consacré à mettre en œuvre une plate forme qui nous permet la réalisation de notre application.

I. Introduction :

Les clusters sont souvent reliés par des réseaux longue distance, formant des grilles, qui permettent d'offrir une quantité de ressources très importante aux utilisateurs. MPI est la bibliothèque de communication standard utilisée pour écrire des applications parallèles pour ces grilles. Les utilisateurs souhaitent exécuter sur des grilles les applications qu'ils avaient écrites pour des clusters afin de disposer du plus grand nombre de ressources. Par exemple, des applications comme ray2mesh, une application sismique de projection de rayons dans un maillage 3D de la Terra, ou des applications médicales comme Simri, un simulateur d'imagerie à résonance magnétique en 3D ont été portées sur la grille avec succès. Cependant, MPI a été initialement écrit pour les clusters et ne prend pas en considération les spécificités des grilles. L'exécution des applications MPI pose essentiellement trois problèmes quand on veut les porter sur des grilles de calcul. Premièrement, les implémentations de MPI doivent gérer de manière efficace les liens longue distance entre les différents sites. La forte latence entre les sites est très coûteuse, particulièrement pour les petits messages. Les communications inter-sites prennent plus de temps que les communications intra-sites. Les liens inter-sites offrent des débits plus élevés que les communications intra-sites.

Deuxièmement, les implémentations MPI doivent prendre en compte l'hétérogénéité des différents réseaux rapides présents dans les grilles. Il s'agit d'une part de permettre des communications intra-site (par exemple communiquer entre un réseau Myrinet et un réseau Infiniband) et d'autre part de gérer les communications inter-sites c'est-à-dire entre deux clusters Myrinet séparés par un WAN (Wide Area Network) sur lequel TCP est utilisé.

L'annotation de la vidéo fait partie de la vision par ordinateur et il y a des solutions qui ont été déjà proposées dans ce domaine pour faciliter la recherche d'informations dans une base de données vidéo. Les traitements peuvent être effectués séparément ou conjointement sur les différents canaux. Nous avons choisi de nous consacrer aux informations visuelles des vidéos. L'extraction de caractéristiques vise à indexer de manière efficace les vidéos et à créer des résumés pertinents. L'intérêt du résumé de vidéo réside aussi dans la grande variété des applications qui en découlent comme la recherche, la classification et la navigation dans des bases de vidéos. et donc elle est utilisée pour l'annotation vidéo ainsi que notre programme est basé sur MapReduce sur Hadoop. Hadoop est une infrastructure de traitement par lots distribuée à grande échelle idéalement adaptée pour le traitement de données à grande échelle. Hadoop inclut un système de fichiers distribué qui rompt des

données en entrée et envoie les fractions des données originales à plusieurs machines pour que les données puissent être traitées en parallèle.

II. Architecture de Système :

Applications nécessitent souvent plus de ressources que sont disponibles sur une machine peu coûteuse [55]. La nécessité pour les modèles efficaces et efficaces de calcul parallèle est évidente en raison de l'existence de très grands ensembles de données qui ne peuvent être traitées sans utiliser plusieurs ordinateurs. Le modèle de programmation MapReduce de Google [56] fournit un cadre efficace pour le traitement de grands ensembles de données d'une manière extrêmement parallèle. Le système de fichiers Google [48] qui sous-tend MapReduce permet le stockage de données distribuées efficace et fiable nécessaire pour des applications impliquant de grands ensembles de données. La fonction de base du modèle de MapReduce est d'itérer sur les paires d'entrée, clés de calcul / valeur de chaque partie de l'entrée, le groupe toutes les valeurs intermédiaires par clé, puis itérer sur les groupes résultant et enfin réduire chaque groupe.

Le modèle prend en charge efficacement le parallélisme.

Le programmeur peut faire abstraction de la question de la répartition et la programmation parallèle, MapReduce offre de mise en œuvre avec des questions telles que l'équilibrage de charge, la performance du réseau, la tolérance aux pannes, etc. [56].

Le projet Apache Hadoop [58] est la mise en œuvre la plus populaire et largement utilisé open-source de l'écriture MapReduce de Google en java pour fiable, évolutive, l'informatique distribuée.

Le modèle MapReduce est basée sur deux étapes distinctes dans une application [55];

-Map: Une étape de transformation initiale, dans lequel les enregistrements d'entrée individuels sont traités en parallèle.

- Réduire: Une étape de synthèse, dans lequel tous les enregistrements associés sont traités ensemble par une seule entité.

La mise en œuvre du modèle Hadoop MapReduce est présentée dans la Figure II.1 montre qui divisent l'entrée en morceaux logiques et chaque morceau est traité de façon indépendante par une tâche de Map. Les résultats de ces morceaux de traitement peuvent être divisés en

Proposition d'algorithme Scalable pour annotation vidéo

ensembles physiquement séparés qui sont ensuite triés. Chaque morceau trié est passé à une tâche de Reduce.

Le système de fichiers Hadoop (HDFS) est un système de fichiers conçu pour les emplois MapReduce qui lisent l'entrée en gros morceaux, le processus, et écrire des morceaux de sortie. Pour la fiabilité, les données de fichier est répliqué sur plusieurs nœuds de stockage [55]. HDFS services sont fournis par deux processus:

- **NameNode:** Gère les métadonnées du système de fichiers, et fournit des services de gestion et de contrôle.
- **DataNode:** Fournit le stockage de bloc et services de récupération.

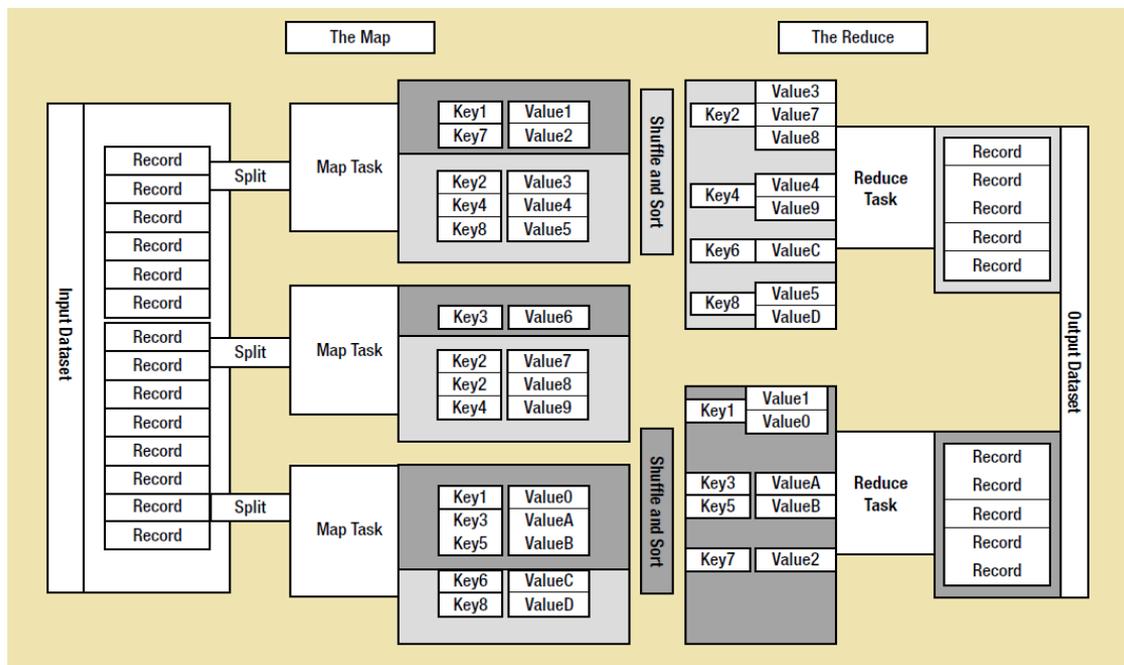


Figure II.1 : Architecture de Hadoop

II. La méthode d'annotation utilisée

Nous allons nous intéresser uniquement au résumé statique pour les raisons suivantes. Nous avons constaté que les méthodes les plus sophistiquées sur le résumé dynamique reposent en général sur la transcription de la bande son. Par ailleurs, les méthodes qui utilisent des caractéristiques de bas niveau et qui n'ont pas de connaissance a priori sur la vidéo réalisent souvent un résumé dynamique à partir d'un résumé statique. Comme nous travaillons uniquement avec le contenu visuel des vidéos et que nous n'avons aucune connaissance a

Proposition d'algorithme Scalable pour annotation vidéo

priori sur les vidéos, nous allons étudier le résumé statique. De plus, les applications comme l'indexation, la recherche et la navigation dans les vidéos sont principalement obtenues à partir des résumés statiques. Dans le manuscrit, le terme « résumé » fera donc référence à un Annotation statique qui utilisera l'algorithme K-means pour créer un résumé de vidéo hiérarchique.

K-means :

L'algorithme k-means mis au point par McQueen en 1967, un des plus simples algorithmes d'apprentissage non supervisé, appelée algorithme des centres mobiles [Celeux et al, 1989], il attribue chaque point dans un cluster dont le centre est le plus proche. Le centre est la moyenne de tous les points dans le cluster, ses coordonnées sont la moyenne arithmétique pour chaque dimension séparément de tous les points dans le cluster chaque cluster est représentée par son centre de gravité.

Principe de K-means :

Comme pour la régression, voici une simplification du programme k-means. Il s'agit cette fois d'itérer un nombre fixé a priori d'étapes MapReduce ou bien jusqu'à convergence de l'algorithme.

- _ La première étape map initialise l'algorithme en répartissant aléatoirement les observations en k classes puis gère l'affectation des observations aux centres les plus proches,
- _ l'étape reduce calcule les nouveaux centres.

Algorithme de K-means :

Dans cette version simplifiée, le nombre d'itérations est fixé a priori et les numéros des classes finalement obtenus ne sont pas listés. Ceux-ci sont stockés comme clefs dans la base pour d'autres utilisations.

IV. Algorithme d'annotation avec MapReduce :

MapReduce est une approche de programmation parallèle utilisée pour l'extraction et l'analyse des informations non structurées de stockage des grandes données [60].

Dans MapReduce : une fonction de Map qui traite une paire clé / valeur pour générer un ensemble de paires clé / valeur intermédiaire, et Reduce est une fonction qui IMMIX tous les valeurs intermédiaires associées à la même clé intermédiaire [59].

Mapping (M) se fait sur des données d'entrée pour obtenir des paires clé / valeur intermédiaire comme le montre la figure IV.2, alors ces données intermédiaires sont regroupées par clé toutes les valeurs v avec k1 clé dans un groupe, toutes les valeurs v avec clé K2 etc.

Proposition d'algorithme Scalable pour annotation vidéo

Ces données regroupées est réduite pour donner sortie suivante à savoir

{K1, 4} {k2, 1} {k3, 2}

{K4, 3} et {k5, 1}

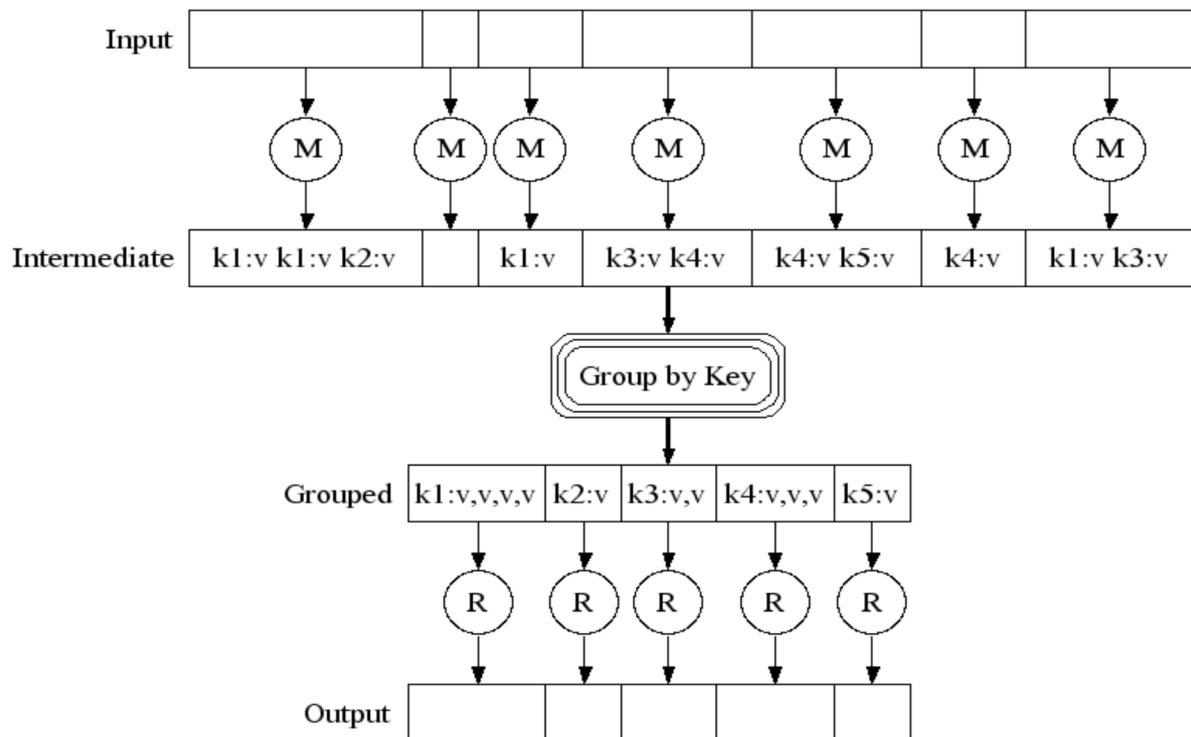


Figure IV.1: structure MapReduce [50]

Ci-dessus exemple de données de couleur sur lequel MapReduce est effectuée.

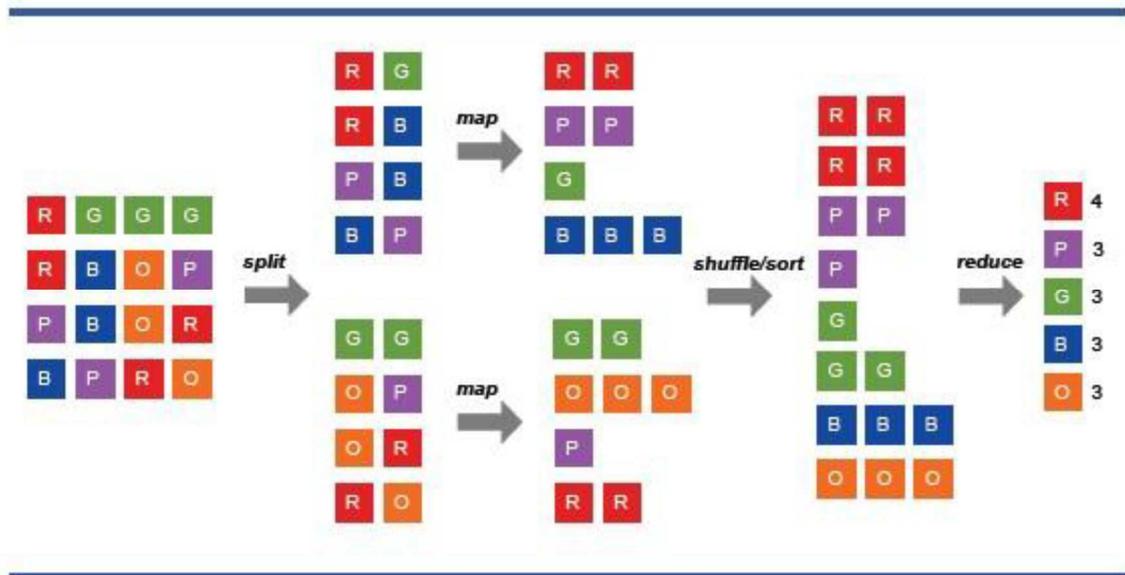
Les 16 blocs sont divisés en deux ensembles de 8 blocs dans chaque et il est mappé pour organiser des blocs par rapport aux couleurs où R est rouge, P est rose, G est vert, B est bleu et O est orange.

Le premier ensemble contient 2 rouges, 2 rose, 1 vert, 3 ensemble bleu et second contient 2 vertes, 3 orange, 1 rose, 2 rouge.

Ces 2 ensembles sont mélangés ou triés pour obtenir un ensemble unique et est réduite pour donner sortie suivante {R, 4} {P, 3} {G, 3} {B, 3} {O, 3}.

Voici un code pour la carte-réduire: lecture s enregistrements d'entrée Le mappeur et produit <mot, 1> sous forme de paires intermédiaires.

Après brassage, les comptes intermédiaires associés au même mot sont transmis à un réducteur, qui ajoute les chiffres ensemble pour produire la somme .



Exemple de MapReduce

V. La plateforme utilisé:

Afin de mener à bien ce projet, nous avons utilisé un ensemble de matériels dont les principales caractéristiques sont les suivantes :

- _ Processeur: Intel(R) Core™ i3-M370 @2.4 GHZ
- _ RAM: 4.00 GO
- _ OS : Linux Ubuntu 12.04 LTS
- _ Compilateur : Gnu gcc, Gnu Make

1. Machine Virtuelle :

Une machine virtuelle est un logiciel qui tourne sur l'OS hôte, ce logiciel permettant de lancer un ou plusieurs OS invités, c'est l'archétype de la solution de virtualisation par empilement de systèmes. La machine virtualise le matériel (ce qui passe généralement par une émulation partielle) pour les systèmes d'exploitation invités : les systèmes d'exploitation invités croient dialoguer directement avec le matériel. En pratique on a recours à une émulation logicielle des périphériques, et parfois aussi de tout ou partie de la machine. [60]

2- VirtualBox :

2.1 Présentation :

VirtualBox est un produit qui se présente comme "seule solution professionnelle de virtualisation distribuée sous licence GPL". Il se présente accompagné d'utilitaires Destinés à

faciliter la création de machines virtuelles, disques et montages CD/DVD (Fichiers .iso testés). Cependant, le site d'Innotek est très clair sur les limites de la Version Open Source. La version Open Source est disponible packagée, il faut le noter, Pour un nombre important de distributions (dont OpenSuse, Mandriva, Fedora, RHEL et Ubuntu).

2.2 Fonctionnement global de VirtualBox :

Sous VirtualBox, la manipulation des machines virtuelles nécessite plusieurs étapes :

Création d'un disque dur virtuel (VDI) : Nous pouvons soit créer un disque virtuel de taille fixe, soit utiliser un live-cd. Cette deuxième solution permet d'obtenir un disque virtuel n'occupant que peu de place.

Créer une nouvelle machine virtuelle : Un fichier de description de la Machine, comportant des différents paramètres de configuration, est créé.

Rattacher le disque VDI et l'image ISO : du système d'exploitation invité à la Machine virtuelle.

Configurer le réseau : de la machine virtuelle.

Lancer la machine virtuelle.

Pour exécuter ces actions, il existe plusieurs possibilités :

- ❖ Grâce à **l'interface graphique**, on peut réaliser ses actions sans difficultés. On peut également manipuler les disques VDI et les images ISO rattachées aux machines virtuelles via le gestionnaire de disque virtuel.
- ❖ On peut aussi utiliser la **commande** VboxManage.

2.3 Créations d'une machine virtuelle ubuntu :

Avant de créer la machine virtuelle ubuntu, Il faut télécharger les paquets pour utiliser dans l'installation de virtualbox sur la machine linux quand l'installation est terminée il suffit de redémarrer la machine pour obtenir une application qui se trouve dans les outils système (oracle VM virtualBox) donc nous pouvons créer une machine virtuelle.

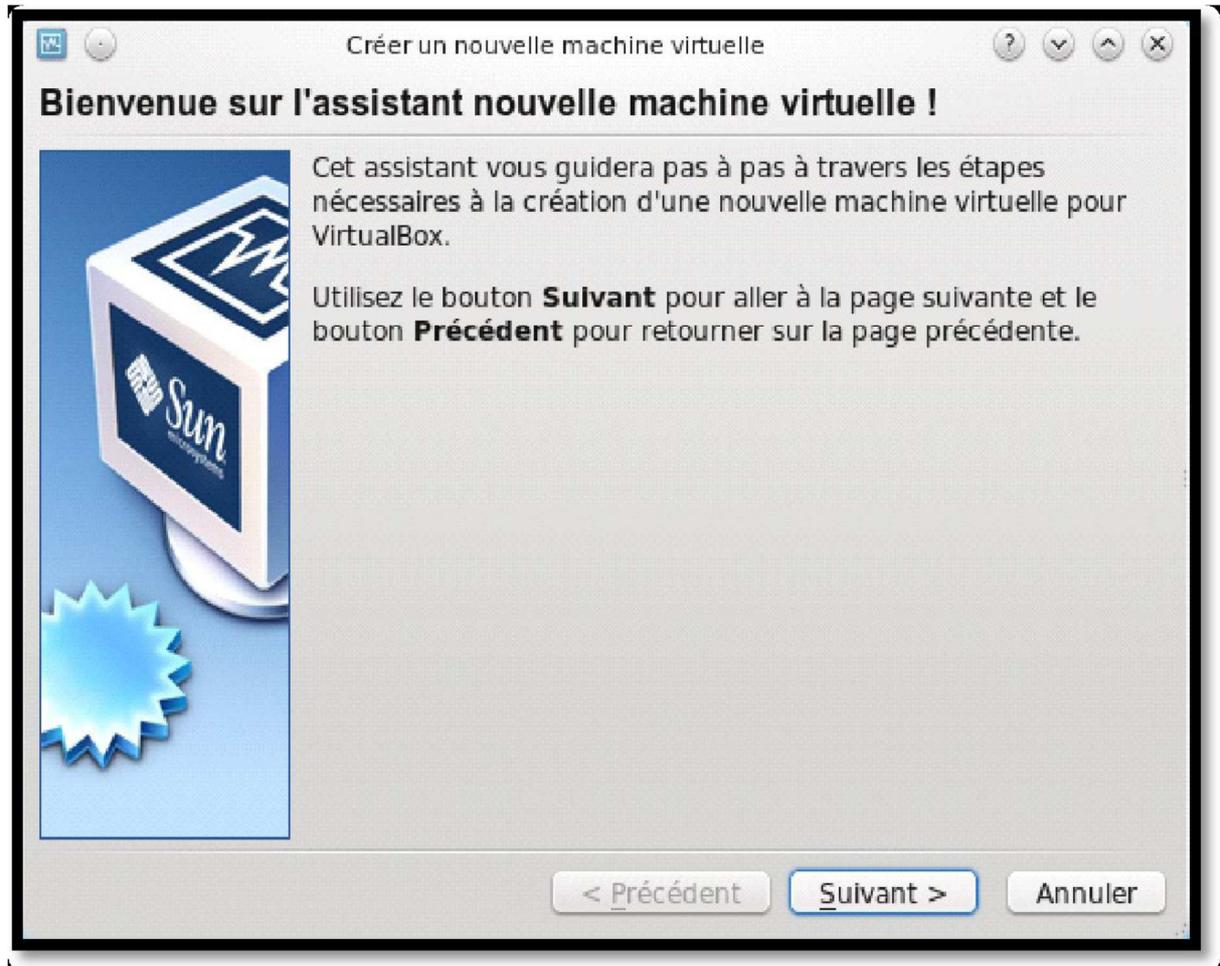


Figure 1 : VirtualBox

4. OpenImaj

OpenImaj est un primé ensemble de bibliothèques et d'outils pour l'analyse de contenu multimédia et la génération de contenu. OpenImaj est très large et contient tout de la vision state-of-the-art (par exemple EIPD descripteurs, la détection de la région saillante, détection de visage, etc.) et le regroupement des données avancées, par le biais d'un logiciel qui effectue une analyse sur le contenu, la présentation et la structure des pages web.

OpenImaj fournit une suite d'outils de ligne de commande qui permettent d'accéder facilement aux fonctionnalités de base dans la bibliothèque OpenImaj. Alors que nouvelles fonctionnalités sont ajoutées à la bibliothèque, ils vont probablement être exposés comme des outils afin de permettre l'intégration et l'essai des outils et des techniques mises en œuvre facile. Les outils offrent la possibilité de construire un pipeline de récupération d'image entière, qui est :

Proposition d'algorithme Scalable pour annotation vidéo

- ✓ Image téléchargé
- ✓ Extraction de caractéristique
- ✓ La génération de dictionnaire
- ✓ Quantification et indexation

Il existe également des versions Hadoop de ces outils qui permettent des versions évolutives de chaque tâche. Nous fournissons également des outils simples d'aide pour fonction/visualisation des images.

VI. Réalisation avec des captures d'écran :

Une longue séquence vidéo contient trop d'informations (plusieurs thèmes, différentes personnes, plusieurs lieux, plusieurs objets, etc.). Conséquemment, on ne peut pas la décrire telle qu'elle. Afin d'obtenir des informations utiles de cette vidéo, on doit d'abord la diviser en ensemble des frames pour extraction des keyframes selon les étapes suivantes :

Étape 1: c'est de tester l'installation de logiciel ffmpeg avec la commande ffmpeg qui fait la transformation d'une vidéo à une séquence d'images

Étape 2 :

Création d'un dossier qui va contenir l'ensemble des images avec la commande :mkdir /home/m1/Images/image4 après c'est divisé la vidéo de la figure1 en ensemble d'frames et les mettre dans ce dossier de la figure 2 avec la commande ffmpeg -i /home/m1/Images/clip1.mpg -r 20/1 /home/m1/Images/image4/\$frame%04d.jpg

Figure1 :clip1.mp

Proposition d'algorithme Scalable pour annotation vidéo

Chargement de fichier.seq sur HDFS prés l'lancement da Hadoop avec la commande : start-all.sh

```
m1@m1-VirtualBox: ~
m1@m1-VirtualBox:~$ start-all.sh
This script is Deprecated. Instead use start-dfs.sh and start-yarn.sh
16/07/09 15:27:59 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
Starting namenodes on [localhost]
localhost: starting namenode, logging to /usr/local/hadoop/logs/hadoop-m1-namenode-m1-VirtualBox.out
localhost: starting datanode, logging to /usr/local/hadoop/logs/hadoop-m1-datanode-m1-VirtualBox.out
Starting secondary namenodes [0.0.0.0]
0.0.0.0: starting secondarynamenode, logging to /usr/local/hadoop/logs/hadoop-m1-secondarynamenode-m1-VirtualBox.out
16/07/09 15:29:10 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
starting yarn daemons
starting resourcemanager, logging to /usr/local/hadoop/logs/yarn-m1-resourcemanager-m1-VirtualBox.out
localhost: starting nodemanager, logging to /usr/local/hadoop/logs/yarn-m1-nodemanager-m1-VirtualBox.out
```

Voir la racine de Hadoop avec la commande :

Hadoop fs -ls /

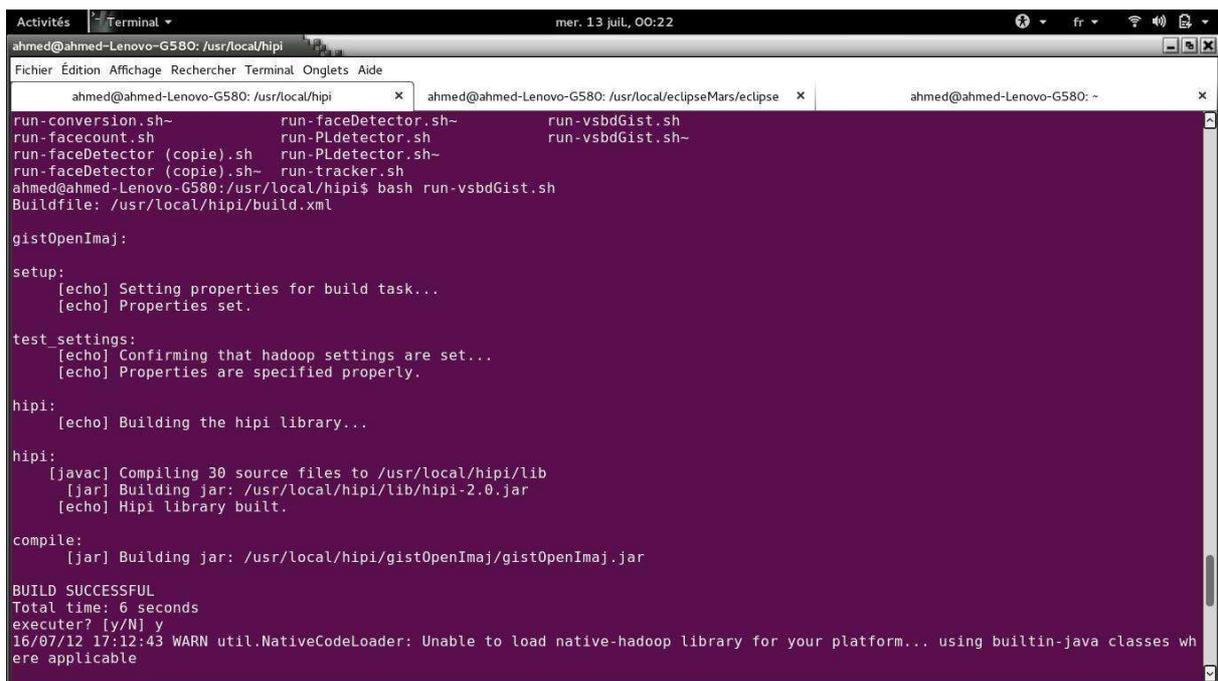
Voir le contenu de répertoire tracker avec la commande **hadoop fs -ls /tracker**

```
m1@m1-VirtualBox:~$ hadoop fs -ls /tracker
16/07/09 15:36:35 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
Found 3 items
-rw-r--r--  1 m1 supergroup  189325509 2016-06-27 14:28 /tracker/image2.seq
-rw-r--r--  1 m1 supergroup  189325509 2016-06-27 14:37 /tracker/image5.seq
drwxr-xr-x  - m1 supergroup          0 2016-06-07 12:44 /tracker/output
```

Proposition d'algorithme Scalable pour annotation vidéo

```
Found 1 items
drwxr-xr-x  - m1 supergroup          0 2016-06-07 12:44 /tracker/
er/output
m1@m1-VirtualBox:~$ hadoop fs -put image5.seq /tracker/
16/07/09 16:07:58 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
put: `image5.seq': No such file or directory
m1@m1-VirtualBox:~$ hadoop fs -put /home/m1/Musique/image5.seq /tracker
16/07/09 16:13:01 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
m1@m1-VirtualBox:~$ hadoop fs -ls /tracker
16/07/09 16:13:59 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
Found 2 items
-rw-r--r--  1 m1 supergroup 189325509 2016-07-09 16:13 /tracker/image5.seq
drwxr-xr-x  - m1 supergroup          0 2016-06-07 12:44 /tracker/output
m1@m1-VirtualBox:~$
```

Pour exécuter avec Hadoop



```
ahmed@ahmed-Lenovo-G580: /usr/local/hipi
Fichier Édition Affichage Rechercher Terminal Onglets Aide
ahmed@ahmed-Lenovo-G580: /usr/local/hipi x ahmed@ahmed-Lenovo-G580: /usr/local/eclipseMars/eclipse x ahmed@ahmed-Lenovo-G580: ~ x
run-conversion.sh~ run-faceDetector.sh~ run-vsbdGist.sh
run-facecount.sh run-PLdetector.sh run-vsbdGist.sh
run-faceDetector (copie).sh run-PLdetector.sh~
run-faceDetector (copie).sh~ run-tracker.sh
ahmed@ahmed-Lenovo-G580:/usr/local/hipi$ bash run-vsbdGist.sh
Buildfile: /usr/local/hipi/build.xml

gistOpenImaj:

setup:
[echo] Setting properties for build task...
[echo] Properties set.

test_settings:
[echo] Confirming that hadoop settings are set...
[echo] Properties are specified properly.

hipi:
[echo] Building the hipi library...

hipi:
[javac] Compiling 30 source files to /usr/local/hipi/lib
[jar] Building jar: /usr/local/hipi/lib/hipi-2.0.jar
[echo] Hipi library built.

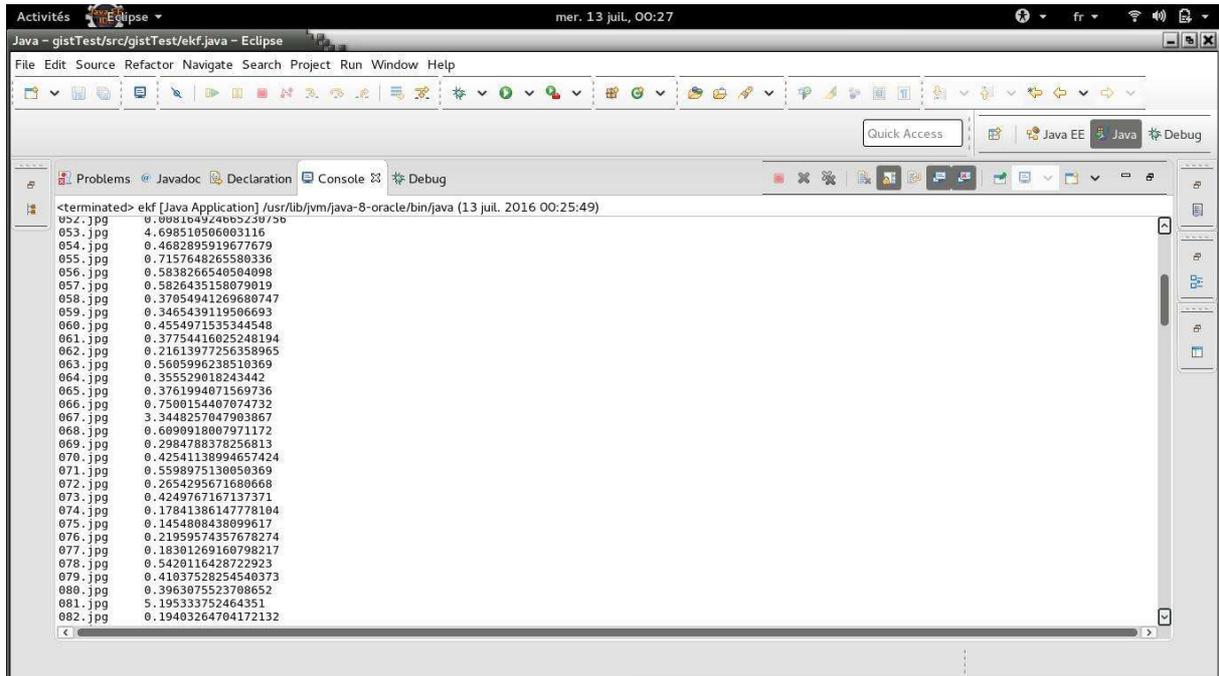
compile:
[jar] Building jar: /usr/local/hipi/gistOpenImaj/gistOpenImaj.jar

BUILD SUCCESSFUL
Total time: 6 seconds
executer? [y/N] y
16/07/12 17:12:43 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes wh
ere applicable
```

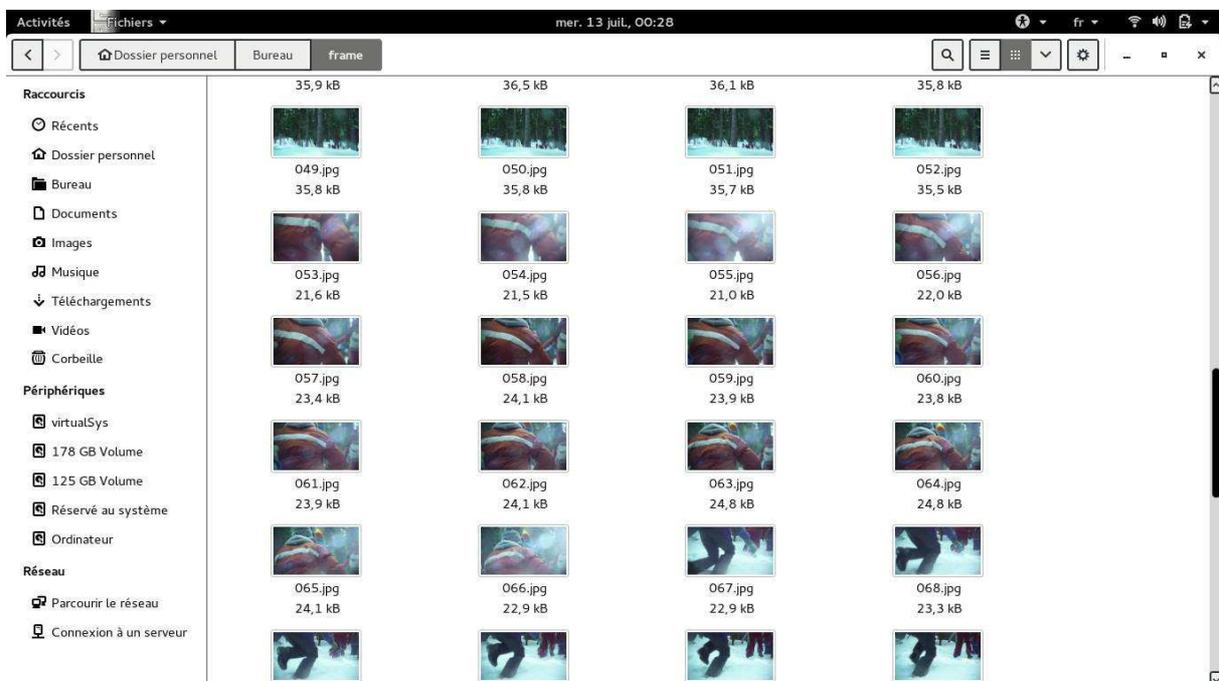
Excute sans hadoop :

Proposition d'algorithme Scalable pour annotation vidéo

Pour l'exécution sans hadoop :



Les images à traiter :



Conclusion :

Dans ce chapitre nous avons présenté brièvement Hadoop avec son architecture de système, et nous à intéresser à l'un des méthodes annotation vidéo qui est la méthode statique basé sur l'algorithme K-means ,et la partie applicative qui nous a permis de concevoir notre application, ça nous a aussi permis de bien comprendre dans l'ensemble les étapes d'annotation vidéo avec le déploiement d'hadoop et ses composantes MapReduce et HDFS.

Conclusion générale

L'objectif de cette étude est de concevoir et réaliser l'annotation vidéo basée sur MapReduce dans le domaine vidéosurveillance. Les résultats obtenus à partir de notre test montrent que parmi les méthodes utilisées on a basé sur la méthode statique qui est efficace pour réduire la dimensionnalité de vecteur caractéristique et l'extraction des images clés.

Dans ce travail, nous avons voulu trouver des solutions à des problèmes de grande quantité d'image vidéo qui exigent énormément de temps pour les traiter.

Au cours de ce mémoire, contribué à la réalisation d'annotation vidéo, Pour la réalisation de cette dernière nous avons jugé utile de décomposer notre travail en quatre chapitres

- Le premier chapitre concerne la présentation de cloud computing Hadoop et MapReduce.
- Le deuxième chapitre concerne vidéosurveillance.
- Le troisième chapitre concerne l'annotation de vidéo de notre système en utilisant la méthode statique.

Et enfin nous avons terminé par la présentation des différents outils de développement utilisés lors de la mise en œuvre effective de l'application, ainsi que la description de son fonctionnement illustré par ses différentes interfaces.

Le travail présenté nous a permis d'affronter des problèmes concrets, en mettant en pratique nos connaissances acquises lors de notre cursus de Master en informatique, et nous a permis en outre, de bénéficier d'une première expérience dans le monde professionnel.

Arrivés à terme de ce travail, nous souhaitons que les utilisateurs trouvent satisfaction à l'utilisation de cette étude.

Références bibliographiques

- [1] Olivier BENDAVID, Bi in the Cloud, 18/06/10.
- [3] Mickael CORINUS, Rapport d'étude sur le Big Data, 2012.
- [4] Adriano Girolamo PIAZZA, NoSQL Etat de l'art et benchmark Travail, 2013.
- [5] TechTarget, Tout savoir sur Hadoop : Vulgarisation de la technologie et les stratégies de certains acteurs.2014.
- [6] Intel Corporation Maitriser les technologies Big Data pour obtenir des résultats en quasi-temps réel ,2013.
- [7]Julien Gerlier et Siman Chen, Prototypage et évaluation de performances d'un service de traçabilité avec une architecture. Distribuée basée sur Hadoop, 2011
- [8] Nilsson, F. "Intelligent Network Video: Understanding Modern Video Surveillance System". Boca Raton: CRC Press, 2009. 389 p.
- [9] Honovich, J., SecurityManager'sGuide to Video Surveillance. Version 2.0, IPVideoMarket.info, November 2008, 131 p.
- [10] Axis Communications, *H.264vidéocompressionstandard. New possibilities within videosurveillance*.Document technique. 2008. 9 p.
- [11]B. T. Truong and S. Venkatesh, "Video Abstraction: A Systematic Review and Classification," ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM), vol. 3, no. 3, February 2007.
- [12] E. Asadi and N. M. Charkari, "Video Summarization Using Fuzzy CMeans Clustering," in 20th Iranian Conference on Electrical Engineering, (ICEE2012). IEEE, 15-17 May 2012, pp. 690 – 694.
- [13]Sabbar, W.; Chergui, A.; Bekkhoucha, A.,"Video summarization using shot segmentation and local motion estimation," *InnovativeComputing Technology (INTECH), 2012 Second International Conference on*, vol., no., pp.190, 193, 18-20 Sept. 2012
- [14] Y. Li, T. Zhang, and D. Tretter. An overview of video abstraction techniques. Technical report, HP Laboratory Technical Report, HPL-2001-191, July 2001.

- [15] S. Li. Content Analysis and Summarization for Video Documents. PhD thesis, Research Associate, VIEW Lab, The Chinese University of Hong Kong, Department of Computer Science and Engineering, December 2004.
- [16] M. Mills, J. Cohen, and Y. Y. Wong. A magnifier tool for video data. In Proceedings of the SIGCHI conference on Human factors in computing systems, pages 93–98, Monterey, California, USA, May 1992.
- [17] Y. Taniguchi, A. Akutsu, Y. Tonomura, and H. Hamada. An intuitive and efficient access interface to real-time incoming video based on automatic indexing. In Proceedings of the 3rd ACM international conference on Multimedia, pages 25–33, San Francisco, California, USA, November 1995.
- [18] A. Nagasaka and Y. Tanaka. Automatic video indexing and full-video search for object appearances. In Visual Database Systems, pages 113–127, Budapest, Hungary, September/October 1991.
- [19] Y. Tonomura, A. Akutsu, K. Otsuji, and T. Sadakata. Videomap and video spaceicon: Tools for anatomizing video content. In Proceedings of the SIGCHI conference on Human factors in computing systems, pages 131–136, April 1993.
- [20] H. Ueda, T. Miyatake, and S. Yoshizawa. Impact : An interactive natural-motionpicture dedicated multimedia authoring system. In Proceedings of the SIGCHI conference on Human factors in computing systems, pages 343–350, April/May 1991.
- [21] S. W. Smoliar D. Zhong H. J. Zhang, C. Y. Low. Video parsing, retrieval and browsing : An integrated and content-based solution. In Proceedings of the 3rd ACM International Conference on Multimedia, pages 15–24, San Francisco, California, USA, November 1995.
- [22] B. Günsel, Y. Fu, and A. M. Tekalp. Hierarchical temporal video segmentation and content characterization. In Proc SPIE Multimedia Storage and Archiving Systems II, volume 3229, pages 46–56, 1997.
- [23] Y. Zhuang, Y. Rui, T. S. Huang, and S. Mehrotra. Adaptive key frame extraction using unsupervised clustering. In IEEE International Conference on Image Processing (ICIP'98), pages 886–870, Chicago, USA, October 1998.
- [24] L. Zhao, W. Qi, S. Z. Li, S.-Q. Yang, and H. J. Zhang. Key-frame extraction and shot retrieval using nearest feature line (nfl). In Proceedings of the ACM Workshops

on Multimedia, pages 217–220, Los Angeles, California, USA, October/November 2000.

[25] A. D. Doulamis, N. Doulamis, and S. Kollas. Non-sequential video object representation

using temporal variation of feature vectors. *IEEE Transactions on Consumer Electronics*, 46(3) :758–768, August 2000.

[26] G. Ciocca and R. Schettini. Dynamic key-frame extraction for video summarization.

In *Proceedings of SPIE Internet Imaging VI*, pages 137–142, January 2005.

[27] G. Ciocca and R. Schettini. An innovative algorithm for key frame extraction in video summarization. In *Journal of Real-Time Image Processing (in Print)*, 2006.

[28] S. Uchihashi, J. Foote, A. Girgensohn, , and J. Boreczky. Video manga : Generating

semantically meaningful video summaries. In *Proceedings of the 7th ACM international conference on Multimedia*, pages 383–392, Orlando, Florida, USA, October/

November 1999.

[29] A. Girgensohn. A fast layout algorithm for visual video summaries. In *IEEE International Conference on Multimedia and Expo (ICME'03)*, volume 2, pages 77–80, Baltimore, Maryland, July 2003.

[30] Y. Gong and X. Liu. Generating optimal video summaries. In *IEEE International Conference on Multimedia and Expo (ICME'00)*, volume 3, pages 1559–1562, New York, USA, July/August 2000.

[31] S. Kopf, T. Haenselmann, D. Farin, and W. Effelsberg. Automatic generation of video summaries for historical films. In *IEEE International Conference on Multimedia and Expo 2004 (ICME'04)*, volume 3, pages 2067–2070, Taipei, Taiwan, June 2004.

[32] A. M. Ferman and A. M. Tekalp. Two-stage hierarchical video summary extraction

to match low-level user browsing preferences. *IEEE Transactions on Multimedia*, 5(2) :244–256, June 2003.

[33] D. Zhong, H. J. Zhang, and S.-F. Chang. Clustering methods for video browsing and annotation. In *SPIE Storage and Retrieval for Image and Video Databases IV*, pages 239–246, 1996.

- [34] D. Farin, W. Effelsberg, and P. H. N. de With. Robust clustering-based videosummarization
With integration of domain-knowledge. In Proceedings of IEEE International Conference on Multimedia and Expo (ICME'02), volume 1, pages 89–92, Lausanne, Switzerland, 2002.
- [35] S. Pfeiffer, R. Lienhart, S. Fischer, and W. Effelsberg. Abstracting digital movies Automatically. *Journal of Visual Communication and Image Representation*, 7(4) :345–353, December 1996.
- [36] A. Ekin, A. M. Tekalp, and R. Mehrotra. Automatic soccer video analysis and summarization. *IEEE Transactions on Image Processing*, 12(7) :796–807, August 2003.
- [37] Y. Rui, A. Gupta, and A. Acero. Automatically extracting highlights for tv baseball programs. In Proceedings of the 8th ACM international conference on Multimedia, pages 105–115, Marina del Rey, California, US, October/November 2000.
- [38] D. Zhang and D. Ellis. Detecting sound events in basketball video archive. Technical report, Technical Report, Dept. of Electrical Eng., Columbia University, 2001.
- [39] H. Sundaram and S.-F. Chang. Video Analysis and Summarization at Structural and Semantic Levels. Book chapter, *Multimedia information retrieval and management : Technological Fundamentals and Applications*, March 2003.
- [40] H. Yang, L. Chaisorn, Y. Zhao, S. Y. Neo, and T. S. Chua. VideoQA : Question answering on news video. In Proceedings of the 2003 ACM Conference on Multimedia, pages 623–641, Berkeley, CA, USA, November 2003.
- [41] J. Nam and A. H. Tewfik. Video abstract of video. In Proceedings of IEEE 3rd Workshop on Multimedia Signal Processing, pages 117–122, Copenhagen, Denmark, September 1999.
- [42] D. Ponceleon, A. Amir, , S. Srinivasan, T. Syeda-Mahmood, and D. Petkovic. Cuevideo : Automated multimedia indexing and retrieval. In Proceedings of the 7th

ACM international conference on Multimedia, volume 2, Orlando, Florida, US, October/

November 1999.

[43] Y. Gong and X. Liu. Video summarization with minimal visual content redundancies.

In IEEE International Conference on Image Processing (ICIP'01), volume 3, pages 362–365, Thessalonique, Grèce, October 2001.

[44] S. Lu, M. R. Lyu, and I. King. Video summarization by spatial-temporal graph optimization.

In Proceedings of the International Symposium on Circuits and Systems (ISCAS'04), volume 2, pages 197–200, Vancouver, Canada, May 2004.

[45] S. Lu, I. King, and M. R. Lyu. A novel video summarization framework for document preparation and archival applications. In IEEE Aerospace Conference, Big Sky, Montana, USA, March 2005.

[46] Y.-F. Ma, L. Lu, H.-J. Zhang, and M. Li. A user attention model for video summarization.

In Proceedings of the 10th ACM International Conference on Multimedia (ACM'02), pages 533–542, Juan Les Pins, France, December 2002.

[47] Y.-F. Ma and H.-J. Zhang. Video snapshot : A bird view of video sequence. In Proceedings of the 11th International Multimedia Modelling Conference (MMM'05), pages 94–101, Melbourne, Australia, January 2005.

[48] C.-W. Ngo, Y.-F. Ma, and H.-J. Zhang. Automatic video summarization by graph modeling. In Proceedings of 9th International Conference on Computer Vision (ICCV'03), volume 1, pages 104–109, Nice, France, October 2003.

[49] S. Chris, L. Liu, A. Sean, and L. Jason, HIPI: A hadoop image processing interface for image-based map reduces tasks, B.S. Thesis. University of Virginia, *Department of Computer Science*, 2011.

[50] J. Venner, Pro Hadoop, Springer, 2009, pp. 1–407.

[51] R. Lämmel, Google's mapreduce programming model — revisited, *Sci. Comput. Program.* 70 (Jan.2008) 1–30.

[52] Oracle In-Database Hadoop: When MapReduce Meets RDBMS Xueyuan Su Computer Science Yale University New Haven, CT 06520 xueyuan.su@yale.edu Garret Swart Oracle

Corporation Redwood Shores, CA 94065 garret.swart@oracle.com SIGMOD'12, May 20–24, 2012,

[53] Map-Reduce: Simplified Data Processing on Large Clusters Jeffrey Dean and Sanjay Ghemawat jeff@google.com, sanjay@google.com Google, Inc.

[54] <http://webpages.iust.ac.ir/hsalimi/Courses/89-90-1/CC/CloudComputing.html>:- Cloud Programming Models.pptx By Khalili, Sh.

Les Sites Internet:

[W1] <http://fr.wikipedia.org/wiki/MapReduce>,28/11/2014.

[W2] <http://en.wikipedia.org/wiki/Terabyte>

[W3] http://fr.wikipedia.org/wiki/Traitement_d%27images

[W4] http://fr.wikipedia.org/wiki/Images_numérique

