

# **Chapitre 1**

## **TRAITEMENT DE SIGNAL MULTIMEDIA**

### **Introduction**

## Introduction

- **Traitement d'images et vidéos : un nouveau domaine technologique**
  - Mathématiques de l'information
  - Traitement du signal
  - Systèmes électroniques et optroniques
  - Architectures machines et micro-processeurs (VLSI, DSP)
- **De nombreuses méthodes utilisées**
  - Choix de celles qui sont adéquates et bien fondées
  - Besoin d'une méthodologie
- **Vaste champ d'applications**

On désigne par *traitement d'images numériques* l'ensemble des techniques permettant de modifier une image numérique dans le but de l'améliorer (qualitativement ou quantitativement avec le codage à compression) ou d'en extraire des informations. Le traitement d'images numériques est une discipline nouvelle qui s'est développée rapidement grâce à l'émergence des nouvelles technologies de l'information. Il s'appuie notamment sur les mathématiques liées à l'information, le traitement du signal, les systèmes électroniques, et sur l'avancée des capacités de calcul des microprocesseurs, notamment ceux qui sont développés exclusivement pour le traitement de signal et qui offrent de grandes capacités et vitesse de calculs (DSP, ...).

La jeunesse du traitement d'images numériques et le vaste champ d'application qu'il s'est vu attribué ont fortement contribué à la nécessité de créer une méthodologie et de séparer les domaines d'applications. Le traitement d'image s'est ainsi distingué en fonction de quatre domaines d'applications majeurs : l'analyse, la synthèse, le codage et l'amélioration.

Dans un premier temps, la description de ces quatre fonctions principales permettra de mieux appréhender ce qu'est le traitement d'images. Puis, dans la suite, nous verrons comment acquérir les images numériques (numérisation d'images analogiques) que nous souhaitons traiter.

## Introduction

*« Une image vaut mille mots,  
et une vidéo vaut mille phrases »*

**Beaucoup d'informations dans une donnée visuelle**

**Exemples d'images autour de nous**

- les photographies de scènes naturelles
- les dessins artistiques et industriels
- les images scientifiques (satellite, médical, ...)

**« Images animées » => Vidéo**

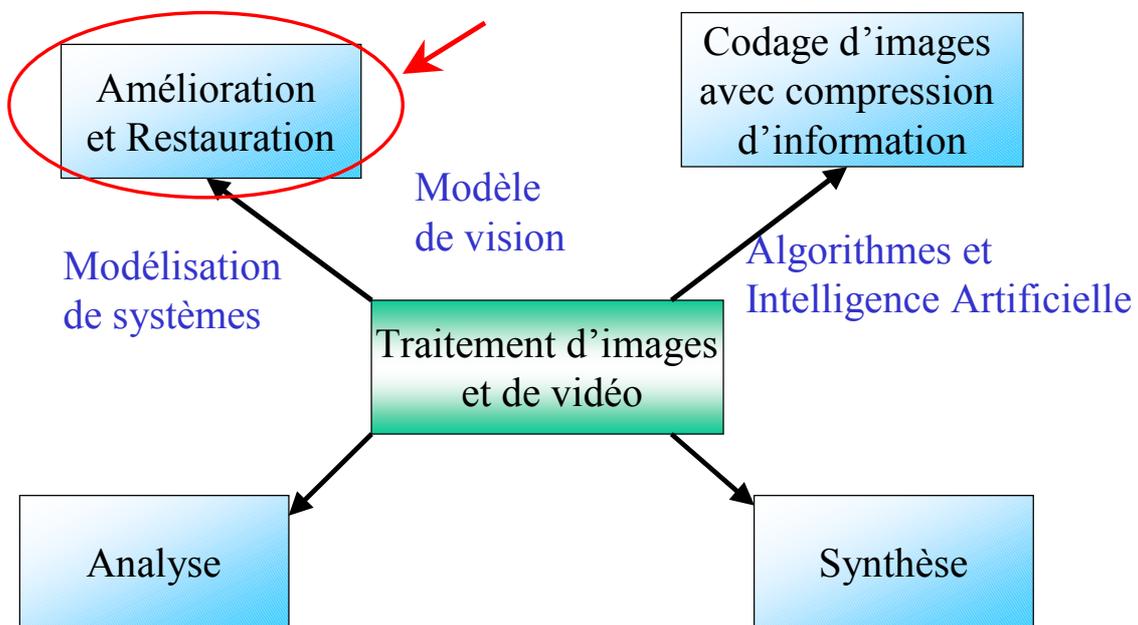
- les films, programmes TV
- les vidéos familiales
- la vidéo surveillance (autoroutes, trains, lieux publics,...)



Quotidiennement nous rencontrons des images de toutes sortes dans notre environnement : des photographies de paysages, de personnes, des peintures, des dessins par ordinateur, des images de radiologie médicale, des images prises par des satellites, ... Certaines de ces images (satellite, médical, ...) ne peuvent être observées directement, d'autres images présentent des caractéristiques à extraire automatiquement, à stocker et envoyer... les traitements envisageables sur les images sont très variés, car les images que nous rencontrons dans notre environnement sont diverses tant par leur nature et leurs caractéristiques que par la scène qu'elles décrivent. Elles sont toutes différentes et il n'est bien évidemment pas envisageable de créer un traitement spécifique pour chacune d'entre elles. Il a donc fallu créer une classification non pas en fonction des caractéristiques des images, mais en fonction de l'objectif du traitement. Quatre types de domaines d'application du traitement d'images numériques se distinguent :

- la restauration et l'amélioration d'images,
- l'analyse d'images,
- le codage avec compression d'information,
- la synthèse d'images.

## 4 Types de traitements d'images et vidéos



Définissons maintenant plus en détails les quatre domaines d'application liés au traitement numérique des images.

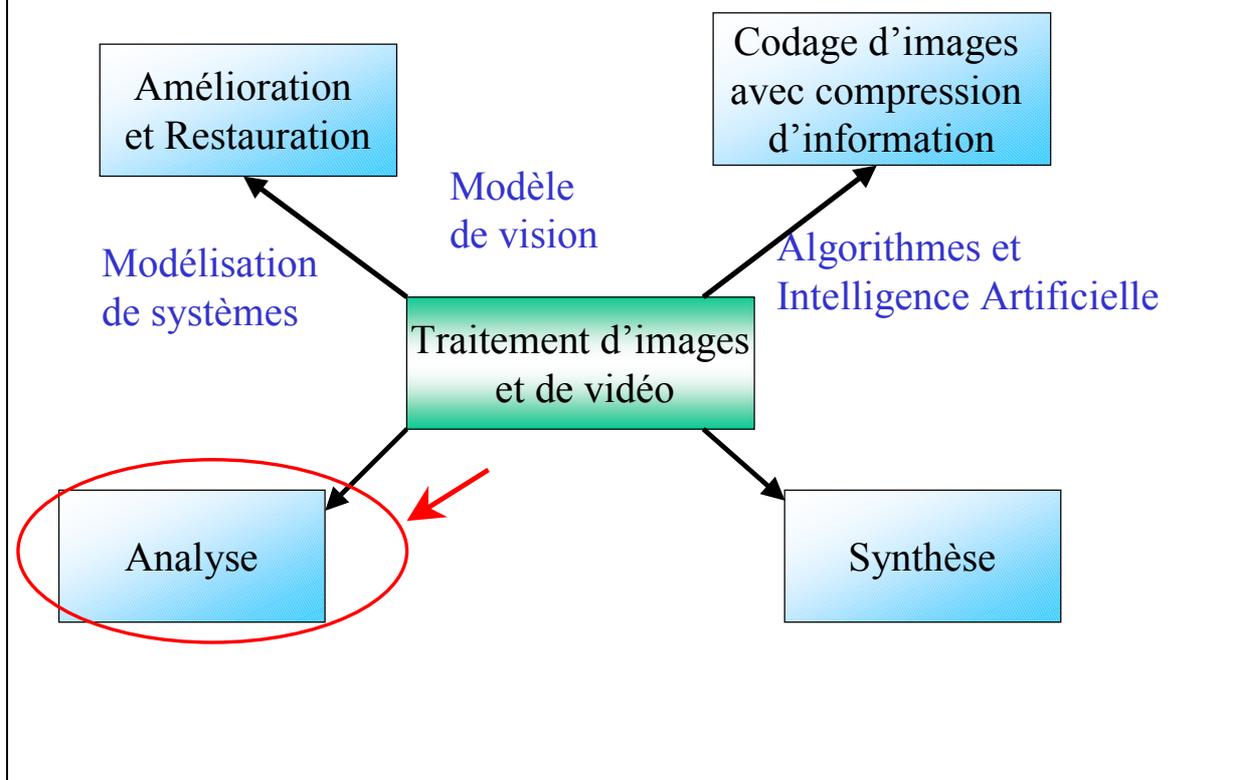
### L'amélioration et la restauration :

Si on considère une image observée  $I_0$  à laquelle on associe le signal  $s_0$  que l'on modélise par :  $s_0 = f(s_u, d, b)$

- où :
- $s_u$  est le signal utile d'image provenant d'une image idéale (sans dégradation) ;
  - $d$  est la fonction de distorsion qui opère sur l'image idéale (flou, ...) ;
  - $b$  est le bruit ;
  - $f$  est une fonction d'observation dépendante des deux signaux  $s_u$  et  $b$  et de la fonction de distorsion.

Le traitement effectué sur  $I_0$ , qui donnera en sortie l'image transformée  $I_T$ , doit permettre d'exploiter l'information contenue dans  $I_T$  de manière plus efficace que l'information dans l'image directement observée  $I_0$ . Dans le cas d'une modification des caractéristiques de présentation de l'image, on parlera de « traitement d'amélioration », alors que dans le cas d'une inversion partielle du phénomène de dégradation on parlera de « traitement de restauration » (exemples : filtrage linéaire 2-D, filtrage 2-D adaptatif, ...).

## 4 Types de traitements d'images et vidéos



### L'analyse d'images :

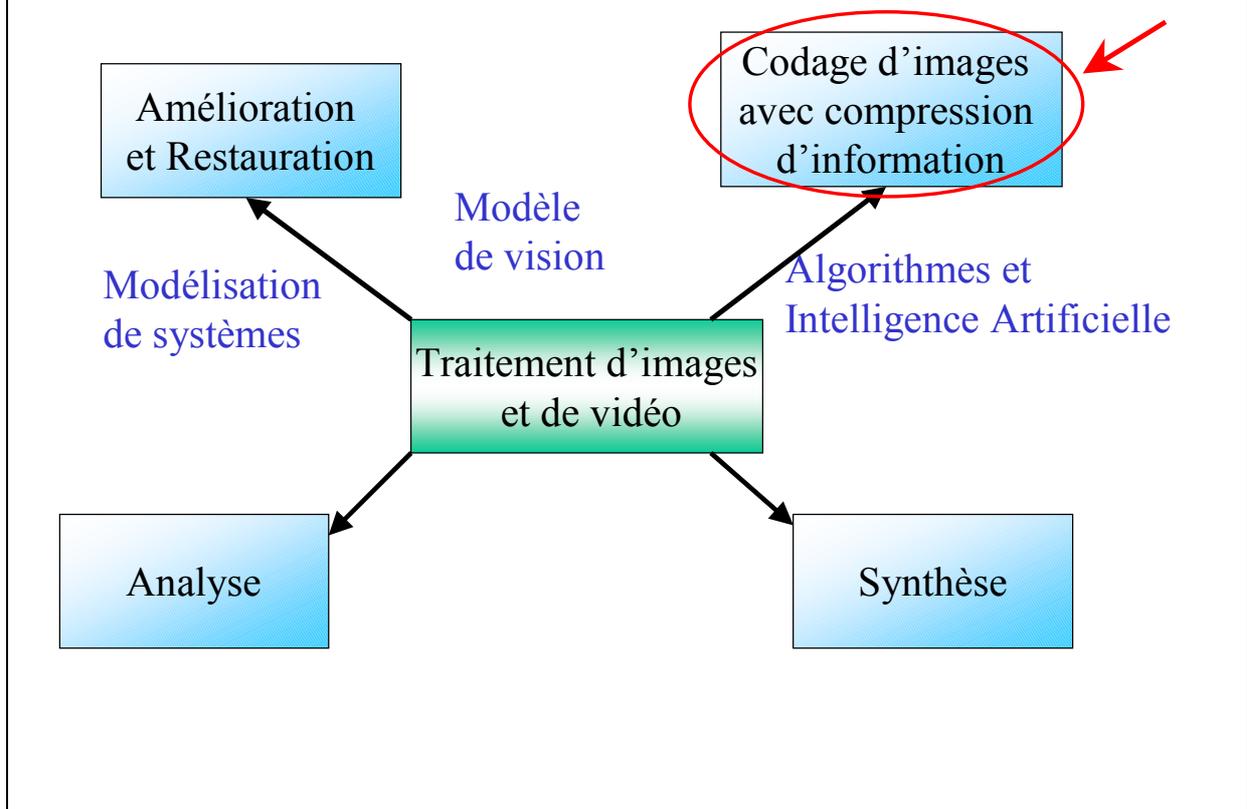
Il s'agit de décrire partiellement ou complètement la scène à partir de l'image observée (objets présents, dimensions, position dans l'espace, ...).

Classiquement ce processus d'analyse se décompose en 3 phases successives :

- pré-traitement et extraction des traits caractéristiques,
- classification et reconnaissance de formes,
- description de l'image et, éventuellement, interprétation.

Il faut cependant noter que l'analyse d'images varie également en fonction du support : les problématiques soulevées par l'analyse d'images 2-D, l'analyse d'images 3-D, ou l'analyse d'images animées varient, si bien qu'on différencie ces dernières en raison des techniques mises en œuvres et de la nature des objectifs recherchés.

## 4 Types de traitements d'images et vidéos



### Le codage d'images avec compression d'information :

La représentation de base d'une image numérique correspond à un tableau 2-D rectangulaire d'éléments appelés pixels. Il faut donc mémoriser un grand nombre de pixels. Pour une simple image monochrome (en niveaux de gris), il faut mémoriser typiquement  $512 \times 512$  pixels (soit 256 000 pixels) et 8 bits par pixel pour obtenir une bonne résolution (8 bits pour coder un pixel, ce qui donne par pixel un nombre de  $2^8 = 256$  valeurs possibles). Pour des images couleurs de haute résolution, et pour des images animées le nombre de bits nécessaires à la représentation de base devient vite très important pour mémoriser ou transmettre.

Le codage d'une image a donc pour but d'obtenir de cette dernière, une représentation qui ne nécessite qu'un nombre très réduit de bits en comparaison de l'image de base.

Pour mesurer cette réduction, on utilise le taux de compression  $\tau$  défini par :

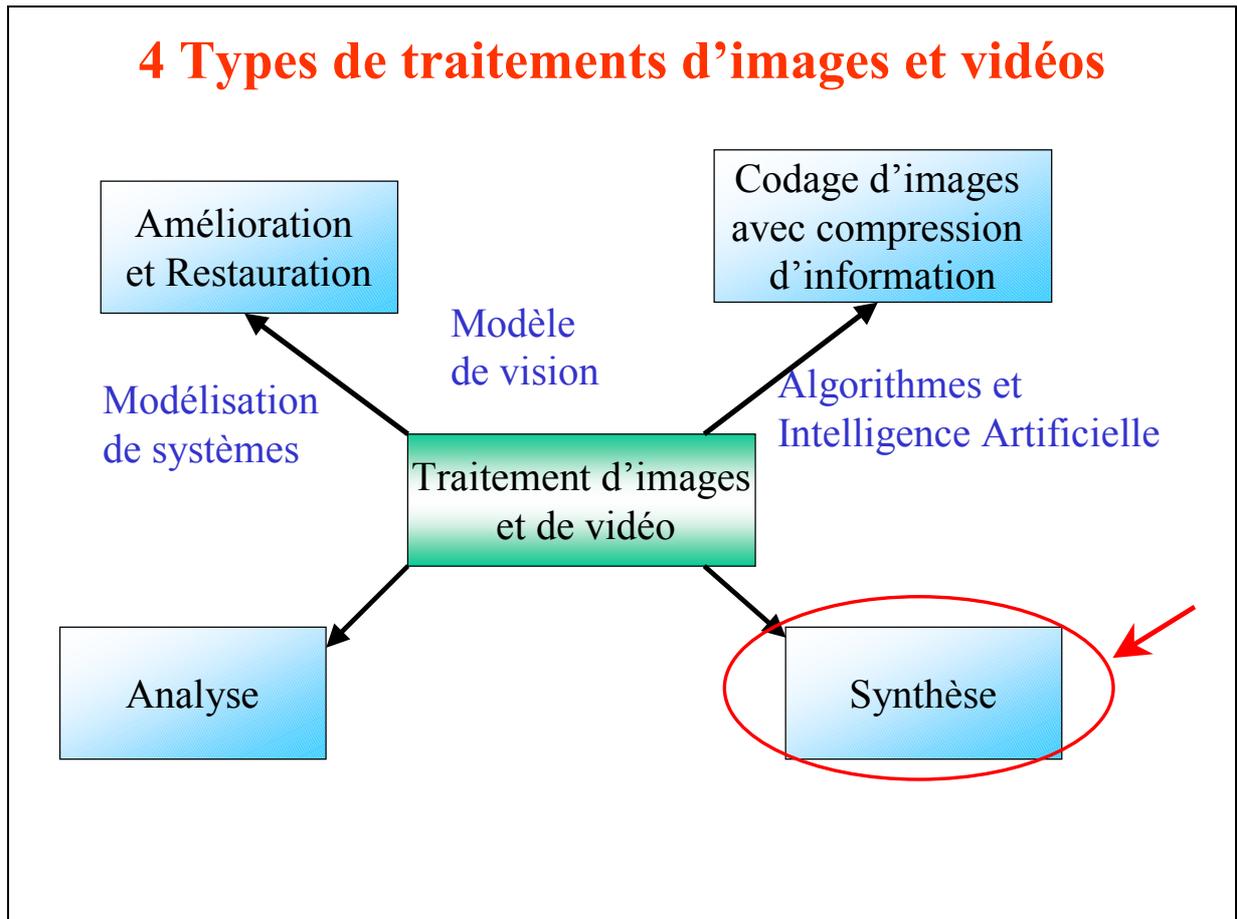
$$\tau = \frac{\text{nombre de bits pour la représentation de l'image de base}}{\text{nombre de bits pour la représentation de l'image après codage}}$$

Ce taux doit être supérieur à 1 pour effectuer vraiment une compression.

Le nombre de bits pour la représentation de l'image de base étant fixé, le taux de compression est en fait inversement proportionnel au nombre de bits pour la représentation de l'image après codage. Si on souhaite une reconstruction exacte de l'image après décodage, il faut alors utiliser un codage réversible. Ceci génère une contrainte telle que le taux de compression est

très souvent peu élevé. Pour augmenter significativement le taux de compression, il suffit de se contenter d'une représentation visuellement exacte. Dans ce cas, l'œil ne distingue aucune différence entre l'image d'origine et l'image reconstituée après décodage. De plus, la complexité du codage et du décodage doit être limitée. Le codage d'une image numérique implique donc de chercher un juste équilibre pour obtenir un taux de compression assez important afin de faciliter la mémorisation et la transmission, mais suffisamment bas pour ne pas engendrer de dégradations gênantes, tout en gardant à l'esprit que la complexité du décodage doit être maîtrisée.

## 4 Types de traitements d'images et vidéos



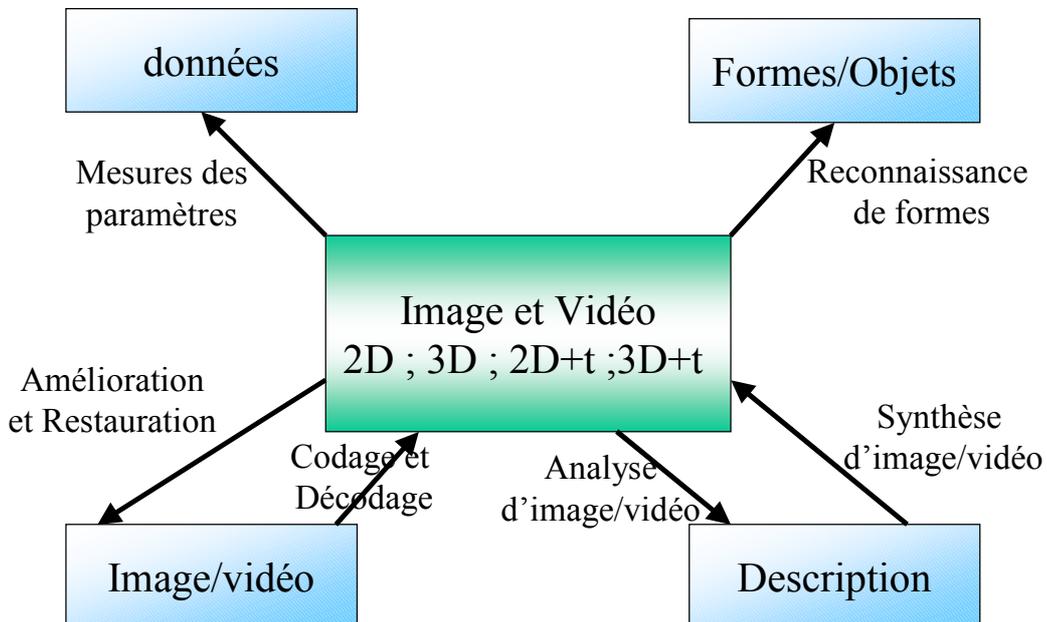
### La synthèse d'images :

Le but est de reconstruire une image qui ressemble à l'image de la scène simulée, à partir d'une description de la scène, des objets qui la composent, des caractéristiques de l'éclairage (incidence, intensité, ...) et, enfin, du dispositif de prise de vue (caméra CCD, CMOS, ...). Cette scène reconstruite peut être ressemblante à la réalité ou purement imaginaire. Les premières applications liées aux images de synthèse furent d'abord orientées vers les simulateurs d'entraînement (vol, engins terrestres, ...) puis se sont amplement diversifiées (audiovisuel, cinéma, art, ...).

Enfin, il convient de signaler que le traitement d'images s'appuie également sur les études liées à la **structure des machines de traitement**. De fait, une image comporte un nombre de données très important. Il y a en effet, pour une image animée,  $N \times M \times P$  échantillons par seconde à traiter ( $N$  points par lignes,  $M$  lignes et  $P$  images par secondes). Le traitement d'image nécessite donc une importante puissance de calcul. Il demande des architectures très performantes à haut degré de parallélisme et à haute vitesse de traitement.

Le traitement numérique des images vient d'être présenté en fonction de ses quatre grands domaines d'application. Il est également possible d'aborder différemment cette présentation en s'intéressant maintenant à la nature des résultats du traitement.

## Vue globale du traitement d'images et de la vidéo



On peut caractériser le traitement d'image, non plus directement en fonction de ses domaines d'application, mais en fonction de la nature des résultats obtenus en sortie. Deux types d'entrées peuvent être considérés : une image ou une description.

- En partant d'une image en entrée

La nature de la sortie peut être :

- de type image  
C'est le cas du codage avec compression d'information, de l'amélioration d'image et de la restauration d'images dégradées ; tous trois ayant été présentés précédemment.
- de type donnée  
C'est le cas lorsqu'on effectue une analyse d'image de façon élémentaire. On s'intéresse par exemple aux dimensions spatiales d'un objet de la scène, à sa position, ou à sa couleur.
- De type forme  
C'est également le cas de l'analyse d'images dans une version plus élaborée. Il s'agit d'extraire et de reconnaître les formes observées dans la scène.

- De type description de scène

Il s'agit également d'une sortie possible lors d'une analyse d'images, mais dans sa version la plus avancée. L'image est entièrement décomposée, afin que chaque objet présent dans la scène puisse être reconnu. La scène est totalement décrite, et peut être interprétée.

- En partant d'une description en entrée

Pour la sortie, la seule nature à prendre en compte est alors de type image. Le domaine d'application mis en jeu est alors « la synthèse d'image ». On souhaite reconstruire l'image en fonction de la description donnée : quels objets sont présents ? quelles sont leurs dimensions ? leur emplacement dans la scène ? comment sont-ils éclairés ? Quels sont les paramètres (focale, angle de vue, ...) du capteur qui effectue la prise de vue ? ...

Le traitement d'image étant maintenant présenté sous deux aspects différents, mais néanmoins fortement liés entre eux, il convient de s'attarder sur les caractéristiques du signal que l'on souhaite traiter : l'image.

## Propriétés de base des Images et Vidéos

- L'image (resp. la vidéo) est un signal 2-D (resp. 2-D + t) scalaire ou vectoriel qui est complexe car:
  - Non stationnaire
  - Non gaussien
  - Non isotrope
- Deux (*images fixes*) ou trois (*vidéo*) caractéristiques principales
  - les contours: changement abrupt dans une caractéristique importante
  - la texture: variation spatiale (souvent la luminance) du signal 2-D pour laquelle la valeur moyenne ne change pas
    - Les bords sont localement des signaux 1-D, alors que la texture est toujours un signal 2-D*
  - les mouvements: pour la vidéo (variation temporelle du signal à une même coordonnée spatiale)

L'image est un signal 2-D scalaire (image monochrome) ou vectoriel (image couleur par exemple). La vidéo, qui est une suite d'images ordonnées temporellement, est un signal 2-D+t (t : temps) scalaire (vidéo monochrome) ou vectoriel (vidéo couleur).

Le signal « Image » est complexe car :

- il est non stationnaire : son contenu en fréquences spatiales change avec les coordonnées spatiales ;
- il est non gaussien : ses propriétés statistiques ne suivent pas une loi de probabilité gaussienne ;
- il n'est pas isotrope : les propriétés du signal d'image ne sont pas les mêmes avec l'orientation (par exemple, dans les images prises au sol, les directions horizontales et verticales sont plus fréquentes pour les contours que les directions obliques).

Les méthodes et les outils classiques, utilisés en traitement du signal, sont très fréquemment conçus pour des signaux stationnaires, gaussiens ou isotropes (Transformée de Fourier Discrète, ...). Ils ne peuvent donc pas être appliqués directement aux images tels quels.

Par ailleurs, les images sont principalement caractérisées par deux types d'éléments : les *contours* et les *textures* :

- les contours correspondent à des changements rapides de caractéristiques importantes quand on passe d'une zone A à une zone B de la scène (valeur moyenne, description de texture). Les bords peuvent être considérés localement comme des signaux 1-D ;
- la texture correspond aux variations spatiales du signal 2-D d'image autour de sa valeur moyenne locale ;
- Le mouvement des objets dans une scène entraîne des modifications temporelles dans les images successives d'une vidéo.

Les images et le traitement d'images numériques sont maintenant présentés dans leur globalité. Dans la suite de ce chapitre, on s'intéressera d'une part aux méthodes de représentation des images (numérisation puis codage) qu'il faut employer afin de pouvoir effectuer un traitement numérique, d'autre part à des exemples concrets de résultats liés aux différents domaines d'application vus précédemment.