

1. Introduction

La vision industrielle est une discipline qui vise à remplacer l'observateur humain par une machine ayant la capacité de **voir** afin d'**automatiser** les tâches de contrôle. Elle est basée principalement sur l'image acquise.

Une image consiste en un jeu de valeurs représentables, sur un écran par exemple. Pour comprendre la nature de ces valeurs et par la suite de l'image elle-même afin de pouvoir effectuer différents types de traitements pour différents objectifs, il faut tout d'abords distinguer les éléments suivants :

- **la perception de l'image** : elle s'articule autour des caractéristiques du système visuel humain. L'œil est sensible à certaines fréquences du spectre électromagnétique ; ces fréquences déterminent la couleur de la lumière en général. Ce que nous appelons couleur est en réalité le résultat de l'action conjuguée de plusieurs paramètres comme l'illustre la figure 1. 1:

- la source lumineuse utilisée;
- la géométrie d'observation (angles d'éclairement et d'observation) ;
- la scène et ses caractéristiques physiques;
- l'œil de l'observateur, avec les qualités et les défauts propres à chaque individu;
- le cerveau de l'observateur, dont la capacité de discernement des couleurs évolue en fonction de l'âge et de l'expérience acquise.

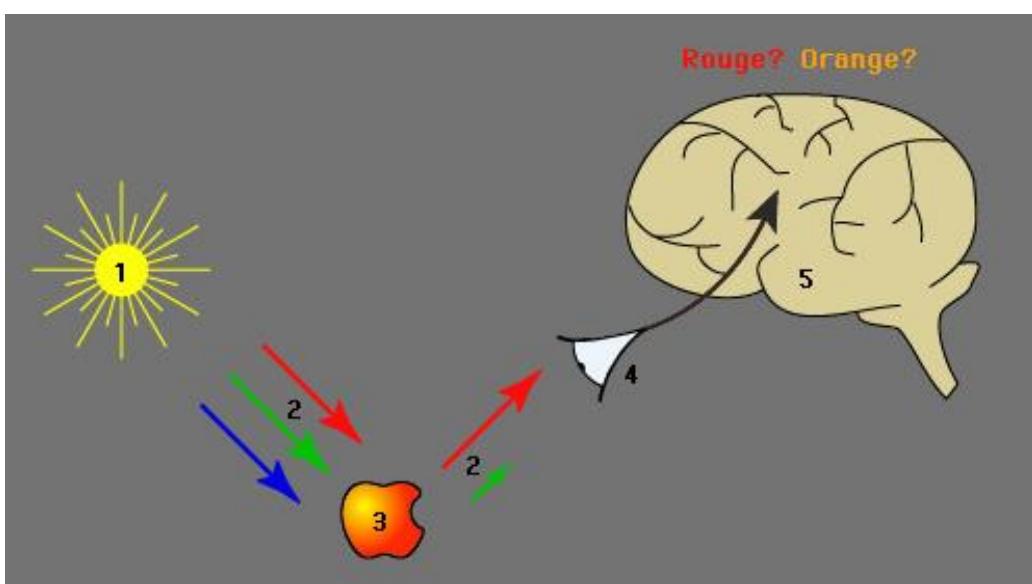


Figure 1. 1 : perception de la couleur

- **la représentation de l'image** : c'est la transformation d'une entité physique (onde électromagnétique issue du corps) sous une forme électrique ou une forme informatique (binaire). La représentation joue un rôle essentiel dans une chaîne de traitement car elle conditionne la capacité de stockage nécessaire ainsi que la mise en œuvre.
- **les traitements d'image** : les traitements sont des opérations qui interprètent ou affectent la représentation de l'image.

2. Historique

Les grandes dates qui ont marqué l'évolution de ce nouveau domaine peuvent être divisées en quatre étapes :

- **1^{ère} étape** : Des premières images à leur traitement (1950-1970)

Origine du traitement d'images : 1950 (avec les premières analyses d'images dans les chambres à bulles). Application : Rayons X, OCR, ... Les images de cette époque étaient de mauvaise qualité et très volumineuses. Par la suite les travaux de recherche consistent en l'amélioration de la qualité et de la compression du volume mémoire.

- **2^{ème} étape** : Du traitement à l'interprétation d'images (1970 - 1980) :

Une évolution naturelle vers l'extraction automatique d'informations et par la suite de nombreuses applications en contrôle qualité, en milieu industriel. On trouve l'apparition de la notion de description structurelle (ensemble des informations contenues dans une image et que l'on peut extraire sans aucune connaissance a priori de l'application). Des nouveaux thèmes de recherche apparaissent:

- seuillage, segmentation, extraction de contours, morphologie mathématique (pour les images binaires), ...

- Interprétation d'images : un engouement explosif au moment de l'apparition des systèmes experts suivi par une déception proportionnelle à l'engouement (Les raisons de l'échec ; il n'y a pas d'expert, le savoir est trop complexe pour être modélisé et par conséquent le modèle perceptif est délaissé!!).

- **3^{ème} étape** : La vision par ordinateur (1980 - 1999) :

C'est le passage de l'image 2D aux modèles tridimensionnels, l'analyse du mouvement, la vision pour la robotique de manutention puis mobile (mouvement, 3D, détection d'obstacle, trajectoire).

- 4^{ème} étape : (1999 à l'heure actuelle)

Un nouvel élan : les applications sont devenues de plus en plus nombreuses : les bibliothèques numériques, transmission (codage Internet et réseaux), multimédia, militaire, médecine, spatial, environnement, communication, culture, sécurité, sport...

2. Définition de la vision industrielle

Un système de vision est un ensemble de processus fortement liés, opérant des transformations progressives des informations sensorielles, en référence à différents modèles (perceptuel, fonctionnel, etc.) intégrant en particulier les connaissances spécifiques aux différents niveaux de traitement. D'une façon générale nous pouvons dire qu'un système de vision est un ensemble capable de :

1. Saisir une image ou une représentation spatiale d'une scène
2. Extraire les informations intéressantes (relativement aux objectifs)
3. Interpréter ces informations pour décider d'une action à engager.

De ce fait, un système de vision industrielle vise à remplacer l'observateur humain par une machine ayant la capacité de voir afin d'**automatiser** les tâches de contrôle. Il est composé principalement par une caméra remplaçant l'œil humain, un dispositif d'éclairage remplaçant les rayons lumineux et une unité de traitement remplaçant le cerveau humain. La figure I. 2 illustre un exemple sur les différents composants qui se trouvent dans un système de vision industrielle.

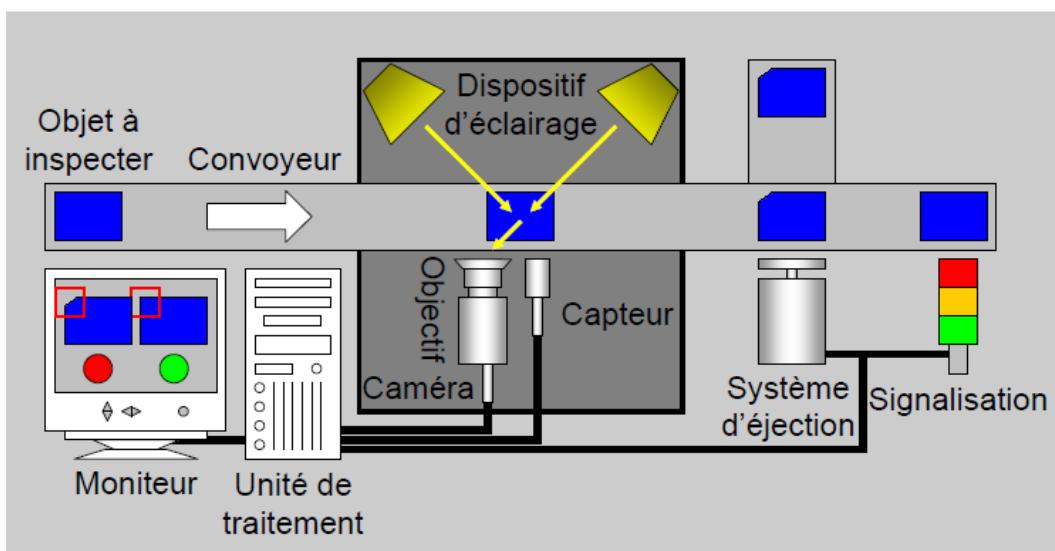


Figure I. 2 : Système de vision industrielle

La chaîne d'analyse d'image en vision industrielle est composé de 4 étapes classiques que nous allons expliciter dans ce cours et qui sont illustré sur la figure 1. 3.

- Acquisition d'une scène
- Traitement de l'image numérique brute (prétraitement)
- Analyse de l'image (traitement)
- Interprétation / Décision

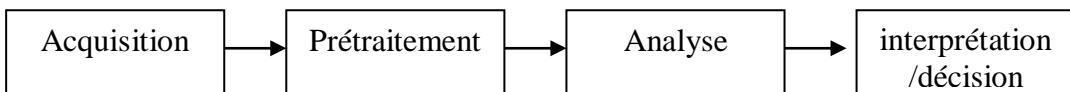


Figure 1.3 : *chaîne d'analyse d'image*

3. Acquisition et représentation des images numériques

La première étape dans la chaîne d'analyse est l'acquisition. Elle consiste à transformer le signal lumineux (onde électromagnétique) réfléchi de l'objet éclairé en un signal numérique 2D composé de points élémentaires appelés pixel.

3. 1. Acquisition des images

Les capteurs (composants de base d'une caméra) permettent de transformer les photons lumineux en signal électrique. Il existe deux grandes technologies : CCD (Charge Coupled Devices) et CMOS (Complementary Metal Oxide Semiconductor). Ils sont formées en une multitude de cellules photosensibles capables de transformer la lumière en un courant d'intensité variable, la différence réside dans la façon, pour chaque cellule, de transférer ses informations au processeur de l'appareil : tandis que CMOS identifie la valeur de chaque cellule séparément, CCD travaille sur l'intensité d'une ligne/colonne de cellules directement. La figure I. 4 représente les modes de fonctionnement des deux types de capteurs. L'obtention de la représentation informatique de l'image se fait dans deux espaces :

- L'espace spatial où l'image est numérisée suivant l'axe des abscisses et des ordonnées : on parle d'échantillonnage. Les échantillons dans cet espace sont des pixels (une unique valeur) et leurs nombre vont constituer la définition de l'image.
- L'espace des couleurs où les différentes valeurs de luminosité que peut prendre un pixel sont numérisées pour représenter sa couleur et son intensité ; on parle de

quantification (la limitation du nombre de valeurs différentes que peut prendre un échantillon). La précision dans cet espace dépend du nombre de bits sur lesquels on code la luminosité et est appelée profondeur de l'image.

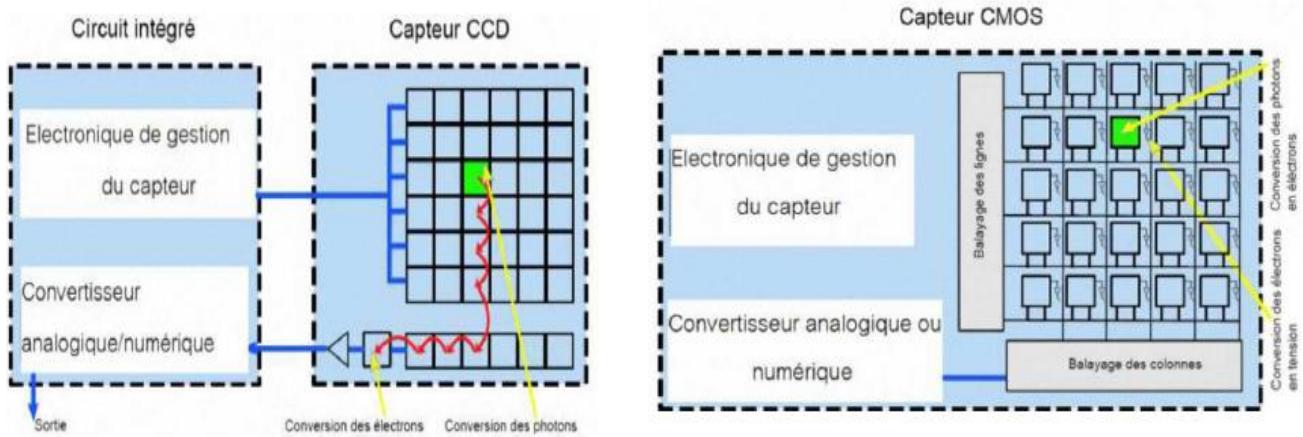


Figure I. 4 : Capteur CCD et Capteur CMOS

3. 1. 1. Représentation de la couleur

La couleur est à la fois un phénomène psychophysique faisant intervenir la physique de la matière, notamment les interactions des ondes électromagnétiques avec les matériaux physiques. C'est également un phénomène psycho-physiologique par l'interprétation des phénomènes psychophysiques par le système visuel constitué notamment de l'œil et du cerveau. Il existe deux types de synthèses différentes pour composer les couleurs, selon que la lumière est remise ou absorbée. La synthèse additive basée sur les propriétés des sources lumineuses émettrices de lumière et résulte de la juxtaposition de lumières colorées. La synthèse soustractive basée sur les propriétés réflexives et d'absorption de la matière.

Il existe plusieurs espaces de couleurs (associés des nombres aux couleurs visibles) qui sont utilisés dans différents domaines. L'espace CIE (Commission Internationale de l'Eclairage), les espaces luminance-chrominance, et les espaces répondant à la perception humaine.

a- espace CIE : on trouve principalement l'espace RGB et l'espace XYZ.

*L'espace RGB (Red, Green, Blue, en français RVB Rouge Vert Bleu) : mis au point en 1931 par la commission CIE. Il consiste à représenter l'espace des couleurs à partir de trois rayonnements monochromatiques de couleurs (rouge, vert et bleu). Le RGB est un espace de couleur additif c.-à-d., ses trois couleurs primaires (rouge, vert, et

bleu) se combinent d'une manière additive pour produire toutes les couleurs désirées. Chacune de ces valeurs varie dans l'intervalle $[0\{255]$.

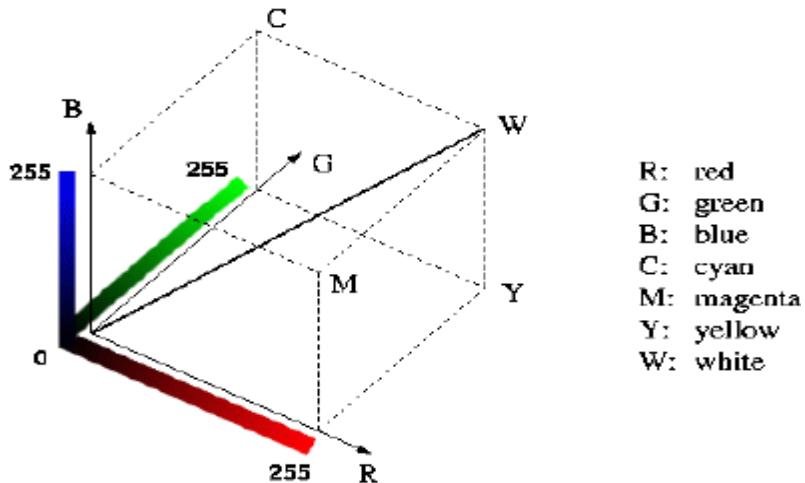


Figure I. 6 : Représentation des couleurs dans l'espace RGB

*L'espace XYZ : pour certaines valeurs de (longueur d'onde) la valeur R du RGB est négative et donc certaines couleurs ne sont pas représentées. L'espace RGB ne représente donc pas tout le spectre visible. Pour remédier à ce problème l'espace CIE XYZ a été développé. Ce dernier caractérise une couleur en fonction de deux composantes X et Z perpendiculaires représentant la chromacité (tonalité et saturation), et d'une composante Y perpendiculaire aux autres axes qui représente la luminance (clarté) ; l'ensemble définissant une couleur. Le passage de l'espace RGB vers l'espace XYZ et vice versa, se fait grâce à ces deux matrices 3x3 inversibles.

$$\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = \begin{bmatrix} 0.618 & 0.177 & 0.205 \\ 0.299 & 0.587 & 0.114 \\ 0 & 0.056 & 0.944 \end{bmatrix} \begin{pmatrix} R \\ G \\ B \end{pmatrix} \quad \begin{pmatrix} R \\ G \\ B \end{pmatrix} = \begin{bmatrix} 1.876 & -0.533 & -0.343 \\ -0.967 & 1.998 & -0.031 \\ 0.054 & -0.118 & 1.061 \end{bmatrix} \begin{pmatrix} X \\ Y \\ Z \end{pmatrix}$$

b- espace luminance-chrominance : Cette famille des espaces couleurs dissocie l'information chromatique des couleurs de leur intensité lumineuse. Les composantes de ces espaces sont calculées à partir des composantes trichromatiques RGB.

* Le modèle YUV définit un espace colorimétrique en trois composantes. Le paramètre Y représente la luminance (c.-à-d l'information en noir et blanc), tandis que U et V permettent de représenter la chrominance, c.-à-d l'information sur la couleur. Ce modèle a été mis au point afin de permettre la transmission des informations colorées aux téléviseurs couleurs, tout en s'assurant que les téléviseurs noir et blanc

existant continuent d'afficher une image en gris. Le passage de l'espace RGB à l'espace YUV se fait par les deux équations suivantes:

$$\begin{pmatrix} Y \\ U \\ V \end{pmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ -0.147 & -0.289 & 0.436 \\ 0.615 & -0.515 & -0.100 \end{bmatrix} \begin{pmatrix} R \\ G \\ B \end{pmatrix} \quad \begin{pmatrix} R \\ G \\ B \end{pmatrix} = \begin{bmatrix} 1 & 0 & 1.140 \\ 1 & -0.395 & -0.581 \\ 1 & 2.032 & 0 \end{bmatrix} \begin{pmatrix} Y \\ U \\ V \end{pmatrix}$$

c-espace répondant à la perception humaine : Cette famille de systèmes couleurs décrit la couleur selon la perception humaine de la couleur. Elle regroupe deux catégories de systèmes : les systèmes uniformes au sens de la perception visuelle et les systèmes de type Luminance, Teinte et Saturation.

*L'espace HSV : c'est un modèle de représentation dit naturel, c.-à-d proche de la perception physiologique de la couleur par l'œil humain. Il consiste à décomposer la couleur selon des critères physiologiques : la teinte (la perception de la couleur) ; la saturation (la pureté de la couleur, vif ou terne) et la luminance (clair ou sombre). Il existe deux modèles de représentation de cet espace : hexagonal et double-hexagonal.

3. 2. Image Numériques

La numérisation est un procédé de codage de l'information qui se fait par des chiffres en mode binaire. En effet, elle est basée sur l'analyse intégrale d'un support d'information, (image, dessin ou texte) qui est fictivement découpé en éléments de surface, échantillonnage spatial, auxquels on associe une valeur numérique, échantillonnage chromatique ou quantification. Cette analyse consiste à mesurer point par point la quantité de lumière transmise ou réfléchie du document pour former en fin un ensemble de points élémentaires appelé pixels (figure I. 5).

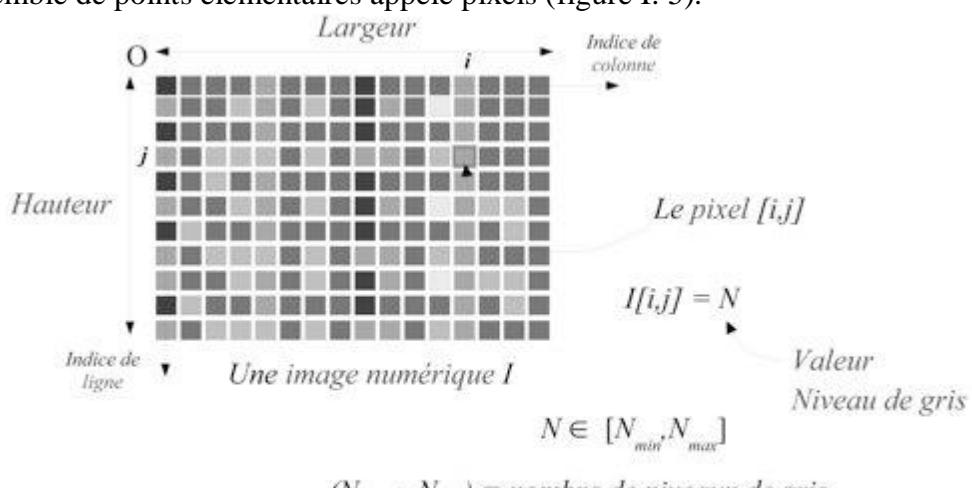


Figure I.5 : image numérique (Echantillonnage+ Quantification)

- La discréttisation des coordonnées spatiales (x, y) est appelée **échantillonnage** spatial. Le résultat de cette opération est la représentation binaire de l'image qui va permettre de travailler sur un nombre finis d'échantillons de ce signal 2D.
- La discréttisation de l'amplitude du signal lumineux est appelée la **quantification**. En fait, Les amplitudes liées au signal analogique s'étendent également sur un intervalle continu qu'il va falloir discréttiser selon différents critères.

3. 2. 1. Caractéristiques d'une image numérique

L'image numérique est repérée par les caractéristiques suivantes:

- **Pixel** : Le pixel (Picture Eléments) est le plus petit élément de l'image, c'est une valeur numérique représentative des intensités lumineuses d'une image.
- **Résolution** : elle correspond au niveau de détail qui va être représenté sur cette image. C'est le nombre de pixels par unité (en pouce : 2.54 cm) de longueur dans l'image à numériser. Elle est en dpi (dots per inch) ou en ppp (points par pouce). Plus le nombre de pixels est élevé par unité de longueur de l'image à numériser, plus la quantité d'information qui décrit l'image est importante et plus la résolution est élevée.
- **Dimension ou définition** : 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.
- **Texture** : c'est une région dans une image numérique qui a des caractéristiques homogènes. La texture est composée de Texel, l'équivalent des pixels.
- **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.

3. 2. 2. Types d'images numériques

Les différents types d'images numériques manipulées sont :

- a. **image à niveau de gris** : les images en niveaux de gris (N. G) renferment 256 teintes de gris. Par convention la valeur zéro représente le noir (intensité lumineuse nulle) et la valeur 255 le blanc (intensité lumineuse maximale). Le nombre 256 est lié à la quantification de l'image. En effet chaque entier représentant un niveau de gris est codé sur 8 bits.

b. image noir et blanc (binaire) : C'est l'exemple des images les plus simples : un pixel peut prendre uniquement les valeurs noir ou blanc. C'est typiquement le type d'image que l'on utilise pour numériser du texte quand celui-ci est composé d'une seule couleur.

c. image en couleurs : l'espace couleur est basé sur la synthèse des couleurs, c.-à-d. que le mélange de trois composantes donne une couleur. Un pixel est codé par trois valeurs numériques. La signification de ces valeurs dépend du type de codage choisi. Le plus utilisé est l'espace RGB (Red Green Bleu).

3. 2. 3. Formats des images couleurs

Un format d'image est sa représentation informatique, où est associée des informations sur la façon dont l'image est codée et fournissant éventuellement des indications sur la manière de la décoder et de la manipuler. On distingue deux types d'images numériques : les images matricielles et les images vectorielles.

- **Une image matricielle** (image en mode point ou image bitmap) est composée comme son nom l'indique d'une matrice (tableau) de points à plusieurs dimensions, chaque dimension représentant une variable telle que (hauteur, largeur, profondeur), durée ou autre (par exemple, un niveau de résolution). Les photos numériques et les images scannées sont de ce type.

Les formats bitmap les plus utilisés sont : TIFF (Tagged Image File Format), BMP (Bitmap - windows) et JPEG (Joint Photographic Experts Group).

- **Une image vectorielle** (ou image en mode trait) est une image numérique composée d'objets géométriques individuels (segments de droite, polygones, arcs de cercle, ...) définis chacun par divers attributs de forme, de position, de couleur, etc. L'intérêt est de pouvoir redimensionner l'image à volonté sans aucun effet d'escalier.

L'inconvénient est que pour atteindre une qualité photoréaliste, il faut pouvoir disposer d'une puissance de calcul importante et de beaucoup de mémoire. Ce type d'images est aussi utilisé pour les animations Flash, utilisées sur Internet pour la création de bannières publicitaires, l'introduction de sites web, voire des sites web complets. On donne quelques formats de ce type : PS/EPS (Postscript /Encapsulated Postscript), SVG (Scalable Vector Graphics) et PICT (Picture).