

La vidéo en multimédia

Introduction

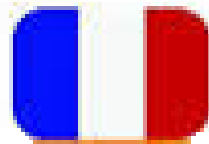
Le mot vidéo vient du latin vidéo qui signifie «je vois». C'est un terme qui regroupe l'ensemble des techniques permettant la visualisation ou l'enregistrement d'images animées accompagnées de son, sur un support électronique.

Une vidéo est une succession d'images à une certaine cadence. L'œil humain est capable de distinguer environ 20 images par seconde.

Introduction

On caractérise la fluidité (vitesse) d'une vidéo par le nombre d'images par secondes (en anglais frame rate), exprimé en FPS (Frames per second).

La fréquence du courant électrique définissait directement le nombre d'images maximum qu'on pouvait afficher pour une vidéo analogique. En France avec le courant EDF à 50 Hertz, les téléviseurs affichaient 25 images par secondes (50 divisé par 2). La norme **PAL** européenne est à 25 images par seconde.



Introduction

Aux Etats-Unis et au Japon où le courant électrique a une fréquence de 60 hertz, la norme est à 30 images par secondes (60 divisé par 2). La norme vidéo américaine NTSC est à 30 images par secondes.



Vidéo numérique et analogique

On distingue généralement deux grandes familles de vidéo :

La vidéo analogique est un signal vidéo pouvant prendre n'importe quelle valeur située entre deux extrémités. Le principe de l'analogique est de reproduire le signal à enregistrer sous forme similaire sur un support magnétique. Le signal est en forme d'onde, il est très sensible aux perturbations extérieures. Ces perturbations peuvent entraîner une modification importante du signal.

Vidéo numérique et analogique

Il existe plusieurs normes pour la vidéo analogique. Les trois principales sont :

- PAL
- NTSC
- SECAM

La vidéo numérique désigne l'ensemble des techniques qui permettent d'enregistrer, de traiter et de reproduire les images vidéo sous forme de données numériques.

Vidéo numérique/ stockage

La vidéo numérique consiste à afficher une succession d'images numériques. Puisqu'il s'agit d'images numériques affichées à une certaine cadence, il est possible de connaître le débit nécessaire pour l'affichage d'une vidéo, c'est-à-dire le nombre d'octets affichés (ou transférés) par unité de temps.

Ainsi le débit nécessaire pour afficher une vidéo (en octets par seconde) est égal à la taille d'une image que multiplie le nombre d'images par seconde.

Vidéo numérique/ stockage

Une séquence vidéo peut être caractérisée par trois principaux paramètres :

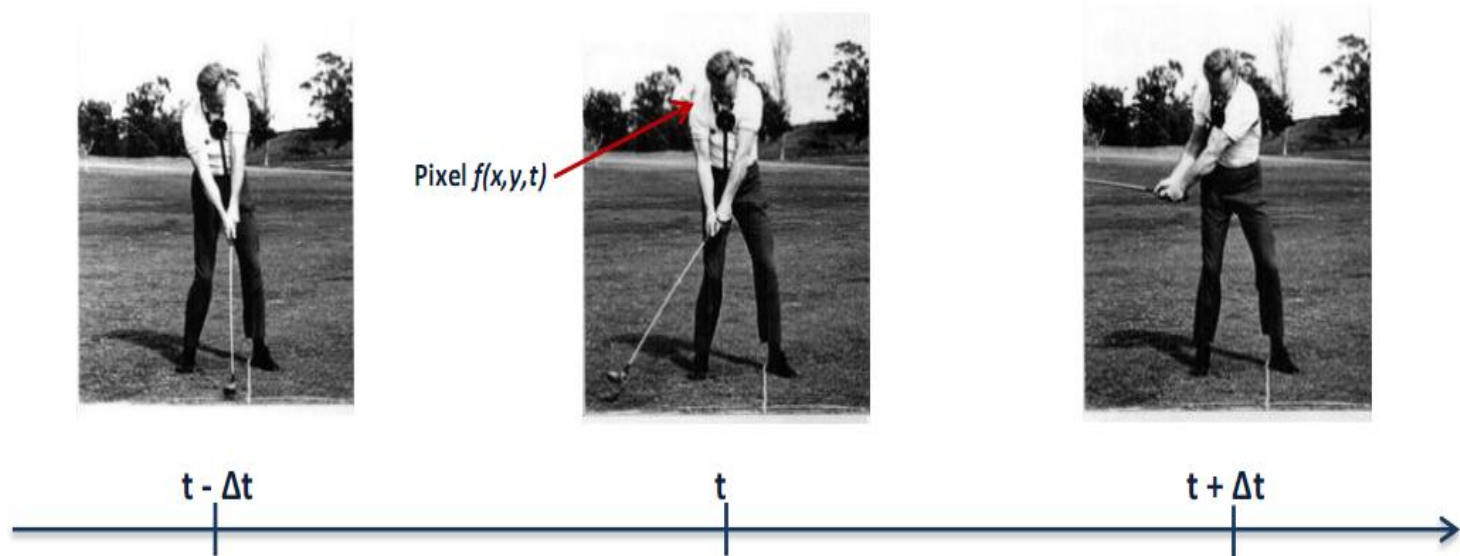
Résolution en luminance	}	+ son audio
Résolution spatiale		
Résolution temporelle		

La résolution en luminance détermine le nombre de nuances ou de couleurs possibles pour un pixel. Celle-ci est généralement de 8 bits pour les niveaux de gris et de 24 bits pour les séquences en couleurs.

Vidéo numérique / stockage

La résolution spatiale définit le nombre de lignes et de colonnes de la matrice de pixels.

La résolution temporelle est le nombre d'images par seconde.



Vidéo numérique / stockage

La valeur de ces trois paramètres détermine l'espace mémoire nécessaire pour stocker séquence vidéo.

Soit une image trois couleurs (24 bits) ayant une définition de 640 pixels par 480. Pour afficher correctement une vidéo possédant cette définition il est nécessaire d'afficher au moins 30 images par seconde, c'est-à-dire un débit égal à :

$$900 \text{ Ko} * 30 = 27 \text{ Mo/s}$$

Vidéo numérique/ Formats

Le terme "format" est utilisé pour désigner un fichier de données aussi bien que son contenu. Une vidéo est une succession d'images (idéalement entre 24 et 30 images par seconde) à laquelle s'ajoute du son, des titres et des sous-titres, qu'on appelle « **métadonnées** ». L'ensemble de ces éléments est ensuite placé et compressé dans un fichier appelé «conteneur» identifié par son nom comme AVI, MOV, FLV, ...etc.

Vidéo numérique/ C-A-N

Le monde physique est par nature analogique. Il est perçu via des signaux analogiques qui peuvent être traités par des systèmes analogiques. Depuis une trentaine d'années, le traitement numérique des données prend le pas sur les approches purement analogiques.

L'interface nécessaire entre le monde analogique et un traitement numérique donné est réalisé par des convertisseurs analogique – numérique (CAN, ou ADC en anglais). Le rôle d'un CAN est de convertir une tension (ou un courant) en un nombre binaire qui lui est proportionnel.

Vidéo numérique/ C-A-N

L'entrée est une tension analogique comprise entre U_{Emin} et U_{Emax} . La sortie est numérique (n bits) : $N = (a_{n-1} \dots a_1 a_0)_2$

Signal analogique : signal continu en temps et en amplitude.

Signal numérique : signal échantillonné et quantifié, discret en temps et en amplitude.

Conceptuellement, la conversion analogique – numérique peut être divisée en trois étapes : l'**échantillonnage** temporel, la **quantification** et le **codage**.

Vidéo numérique/ C-A-N

La figure.1 présente successivement ces trois étapes pour un CAN dont la sortie du signal numérique est sur 3 bits :

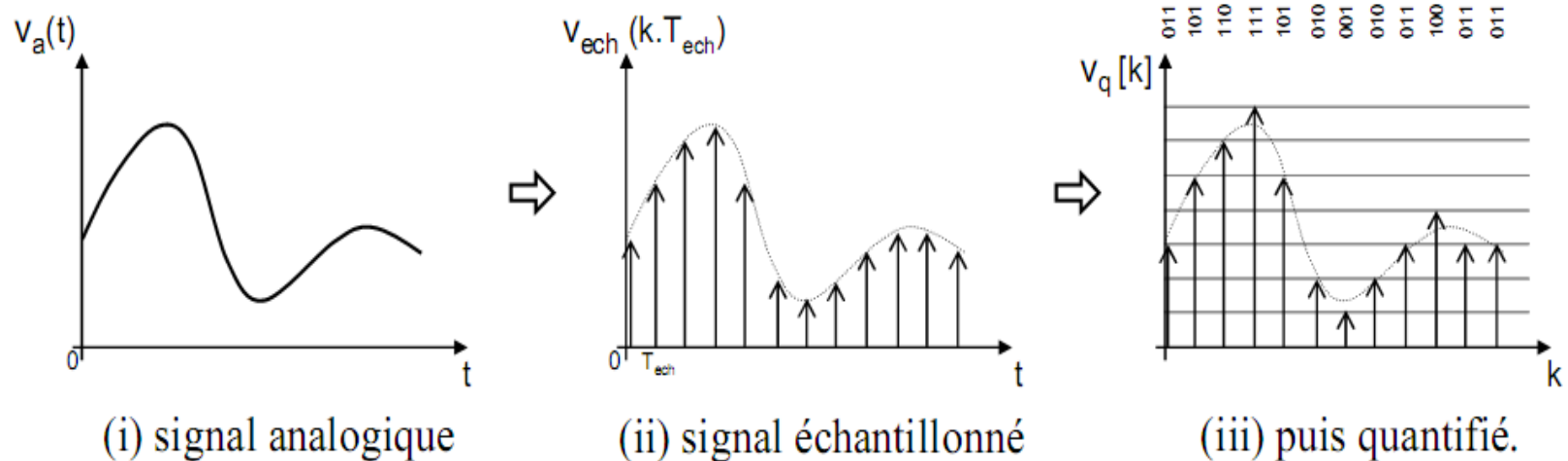


Figure.1

Vidéo numérique/ C-A-N

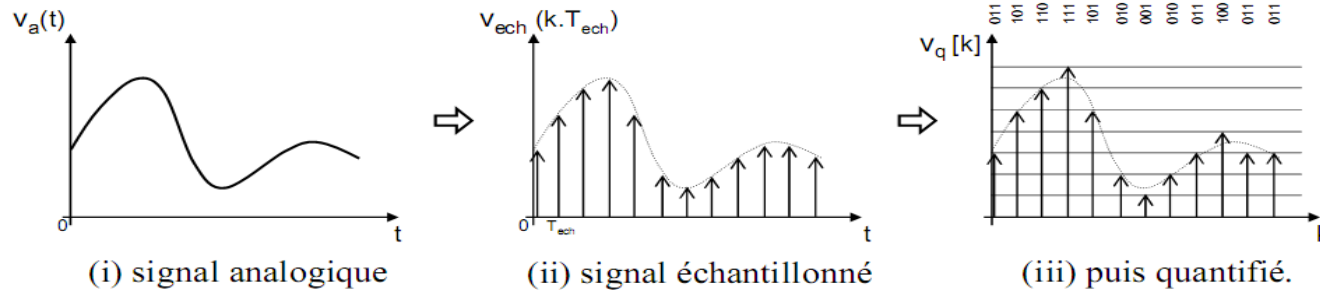


Figure.1

Un signal analogique, $v_a(t)$ continu en temps et en amplitude (i) est échantillonné à une période d'échantillonnage constante T_{ech} . On obtient alors un signal échantillonné $v_{ech}(k.T_{ech})$ discret en temps et continu en amplitude (ii). Ce dernier est ensuite quantifié, on obtient alors un signal numérique $v_q[k]$ discret en temps et en amplitude (iii). La quantification est liée à la résolution du CAN (son nombre de bits) ; dans l'exemple précédent $v_q[k]$ peut prendre huit amplitudes différentes (soit 2^3 , 3 étant le nombre de bits du CAN).

Vidéo numérique/ Compression vidéo

Avec un débit d'environ 20 Mo/sec, la vidéo numérique non-compressée exigerait donc plus de 1.2 Go d'espace disque pour capter 1 seule minute de vidéo.

Le problème auquel on doit actuellement faire face avec la vidéo numérique est le transfert fiable de ce volume important de données, c'est pour cela qu'on utilise un algorithme de compression CODEC pour la compression et décompression.

Le principe est de diminuer au maximum le nombre de bits ou d'octets utilisés pour représenter une image, en minimisant l'impact sur la qualité visuelle de la vidéo. Cela permet de réduire l'espace de stockage nécessaire et la bande passante de transmission.

Vidéo numérique/ Compression vidéo

Certains principes peuvent s'appliquer à l'ensemble de la vidéo sans distinction des images individuelles. Par exemple, il est possible de diminuer la fréquence d'image, déterminée par le nombre d'images par seconde (généralement entre 25 et 30 suivant les formats). Moins d'images, c'est moins d'information à stocker mais on perd généralement en qualité au niveau de la fluidité du mouvement. Il est également envisageable de diminuer la résolution : le nombre de pixels constituant l'image est diminué par 2 ou par 4 grâce à des techniques de ré-échantillonnage.

Vidéo numérique/ Compression vidéo

Les algorithmes de compression vidéo se basent aussi sur une notion simple : il existe généralement au sein d'une vidéo de nombreuses redondances d'information, c'est-à-dire une similitude au sein d'une même image ou entre les images constituant le flux vidéo. Aussi, le principe de la compression vidéo a pour objectif de réduire voire d'éliminer cette redondance.

Vidéo numérique/ Compression vidéo

A l'échelle d'un pixel, on peut noter des corrélations sur deux axes :

- L'axe spatial : il existe une similitude entre les pixels adjacents d'une même image. On parle alors de codage intra-image, intra-frame ou intra-trame.
- L'axe temporel : il existe des informations qui se ressemblent ou se répètent dans le temps, même si leur position dans l'image a changé, c'est à dire une similitude entre les pixels d'une séquence d'images analogues appelée groupe d'images ou GOP (Group Of Pictures). On parle alors de codage inter-image, inter-frame ou inter-trame.

Vidéo numérique/ Compression vidéo

Les algorithmes de compression vidéo **intra-frame** s'appuient sur les principes classiques de compression d'image (par exemple une compression JPEG).

□ Compression inter-frame :

En examinant une vidéo, il est facile de remarquer qu'il existe dans certaines séquences des éléments redondants fixes ou mobiles mais suivant un mouvement logique. A l'échelle du pixel, on parle donc de corrélation temporelle entre des images successives. Dès lors, une image est décrite en fonction de sa différence par rapport à la précédente. Les zones redondantes sont reprises des images précédentes. Cela implique différents mécanismes d'identification des redondances.

Vidéo numérique/ Compression vidéo

Le premier mécanisme est le codage différentiel, où une image est comparée à une image de référence et seuls les pixels qui ont changé par rapport à cette image de référence sont codés. Cela permet de réduire le nombre de valeurs codées et stockées pour chaque pixel. Le codage différentiel compare des éléments statiques d'une image à l'autre.

Image de référence



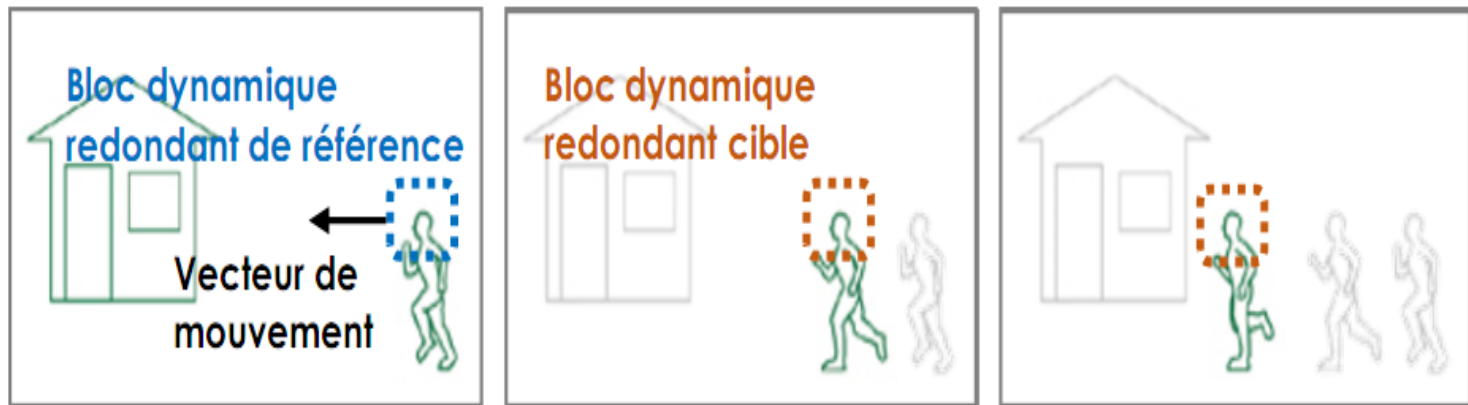
Vidéo numérique/ Compression vidéo

Le deuxième mécanisme, plus complexe, est la compensation de mouvement utilisée afin de réduire encore davantage la quantité de données. Cette technique tient compte du fait qu'une bonne partie de ce qui compose une nouvelle image dans une séquence vidéo se retrouve dans une image adjacente, mais pas forcément au même endroit. Il est donc possible de composer ou de « prédire » une nouvelle image bloc par bloc en recherchant un bloc identique dans une image de référence. Si un bloc référent est trouvé, l'encodeur code l'emplacement dans l'image de référence où se situe le bloc identique ainsi que le vecteur de mouvement qui utilise moins d'espace que le codage du contenu réel d'un bloc.

Vidéo numérique/ Compression vidéo

Le processus de recherche de vecteur de mouvement est connu sous le nom d'estimation de mouvement.

Image de référence



Vidéo numérique/ Compression vidéo

Avec la prédiction **inter-images**, chaque image d'une séquence est classifiée comme un certain type d'image (image **I**, image **P** ou image **B**).

Une image **I** est une image de référence également appelée **Intra-frame**. C'est une image autonome c'est-à-dire qui peut être décodée sans référence à d'autres images. La première image d'une séquence vidéo est toujours une image **I**. Elle fait généralement l'objet d'une compression spatiale **intra-image** à l'aide de l'algorithme JPEG. L'intégration régulière d'image **I** garantit la qualité résultant de la décompression.

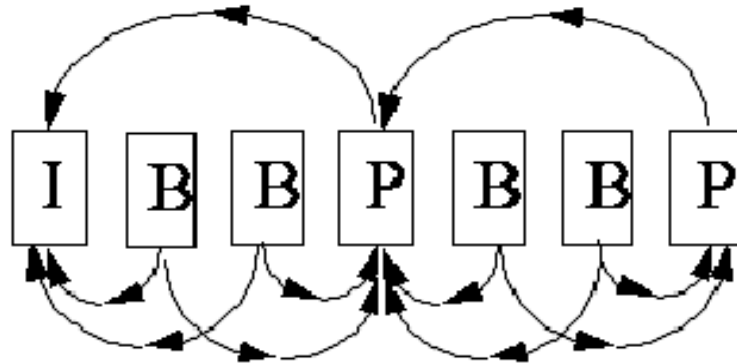
Vidéo numérique/ Compression vidéo

Une **image P** est une image Prédictive qui fait référence à l'image précédente de type P ou I pour le codage de l'image. Une image P nécessite généralement moins de bits qu'une image I, mais elle peut être sensible aux erreurs en raison de la dépendance complexe vis-à-vis des images P et/ou I antérieures.

Une **image B** est une inter-image Bi-prédictive qui fait référence à la fois à une image antérieure et à une image future. Une image devant être compressée avec la méthode B est comparée à toutes les images de type I ou P précédentes ou suivantes, à la recherche de la plus ressemblante et on définit les changements et vecteurs de mouvement.

Vidéo numérique/ Compression vidéo

La figure suivante représente les dépendances entre les images.



Cette succession d'images (GOP) constitue des paquets IBBPBBP. La première image d'un GOP est nécessairement une image I, c'est la référence des autres et c'est la seule à être fiable et d'assez bonne qualité. Les autres images ne sont que des dégradations de celle-ci.

Vidéo numérique/ Compression vidéo

Ensuite on encode les images suivantes de manière B ou P en fonction du niveau de qualité souhaité. De la même façon, lorsqu'un décodeur vidéo rétablit une vidéo en décodant le flux image par image, le décodage doit donc toujours commencer par une image I. Les images P et B, si elles sont utilisées, doivent être décodées avec les images de référence. D'ailleurs, lorsque vous faites avance rapide sur certains lecteurs vidéos et sur certains formats de vidéo utilisant la compression temporelle, vous pouvez apercevoir les images clés défiler.

Vidéo numérique/ MPEG

MPEG (Motion Pictures Expert Group) a été formé par l'ISO pour formuler un ensemble de normes relatives à une gamme d'applications multimédia impliquant l'utilisation de la vidéo avec le son. Au début des années 1990 le groupement MPEG a standardiser le premier système numérique de qualité VHS (Video Home System), le MPEG1.

MPEG1 est une norme de compression vidéo en couches, basée sur DCT qui se traduit par un flux vidéo compressé de qualité VHS qui a un débit d'environ 1,5 Mbps à une résolution d'environ 352x240.

Vidéo numérique/ MPEG

Les limitations du MPEG1 furent vite apparentes. Sa résolution limitée et sa qualité assez basse n'en faisait pas un standard capable de durer dans le temps face à l'évolution rapide des moyens informatiques et numériques.

Le MPEG2 apparût ainsi pour résoudre ces problèmes, et reste à ce jour le format le plus utilisé, depuis les DVD Vidéo, jusqu'à la télévision Haute Définition (HDTV), en passant par les transmissions vidéo par satellite. Ce format laisse un large choix de matériel et de logiciel pour le décodage, le rendant aisément utilisable, ainsi qu'une grande adaptabilité.

Vidéo numérique/ MPEG

L' utilisation de norme MPEG2 reste prévue à long terme puisque la télévision HDTV aux Etats-Unis prévoit sa globalisation sur tout son territoire en 2014. La recherche et le développement de compressions vidéo ne s'arrête cependant pas là, et depuis 1995, le standard MPEG4 commence à émerger du point de vue théorique. Au début de 1998 le standard MPEG4 fait son apparition. Le but du MPEG4 est d'englober les codecs MPEG existant et de leur rajouter une nouvelle dimension permettant un standard beaucoup plus malléable et étendu.

Vidéo numérique/ MPEG

Les usages de MPEG-4 englobent toutes les nouvelles applications multimédias comme le téléchargement et le streaming sur Internet, le multimédia sur téléphone mobile, la radio numérique, les jeux vidéo, la télévision et les supports haute définition.

Le MPEG4 utilise une technologie plus complexe et possède donc de bien meilleurs résultats, mais son utilisation reste à l'heure actuelle essentiellement expérimentale.

Le MPEG1, le MPEG2 et le MPEG4 sont les seuls formats MPEG à être réellement standardisés et utilisés.