

## Codage

# Codage numérique

- Comment intégrer dans un document Word une photo papier ?
- Récupérer le texte d'un article de journal sans devoir le saisir au clavier
- Enregistrer de la parole ou de la musique pour la page d'accueil d'un site Web

## 1. Principe et unités du codage

Tout ordinateur, du plus rudimentaire au plus puissant, n'est qu'une machine électronique qui en tant que telle n'est capable de traiter que 2 informations : le courant électrique passe (symbolisé par le chiffre 1) ou ne passe pas (symbolisé par le chiffre 0). Ces informations sont structurées selon un format (cf. *Formats*)

Les informations qu'un ordinateur peut traiter, aussi complexes soient-elles, se ramènent toujours à un ensemble de 0 ou 1, et tout objet devant faire l'objet d'un traitement informatique doit donc être codé sous cette forme.

De même que les 26 caractères de l'alphabet latin complétés des caractères accentués et de ponctuation permettent de coder tous les textes possibles en français, on peut considérer les chiffres 0 et 1 comme l'alphabet à 2 caractères qui permet de coder n'importe quelle information ou commande pouvant être traitée par un ordinateur.

L'unité élémentaire d'information, qui peut prendre la valeur 0 ou 1, est appelée **bit** (contraction de *binary digit* qui signifie chiffre binaire en anglais). A partir de cette unité élémentaire on compose des unités plus complexes, qui sont récapitulées dans le tableau 1.

Unité	Notation	Valeur
<b>Bit</b>		
<b>Octet</b>	<b>o</b>	8 bits consécutifs
<b>Kilo Octet</b>	<b>Ko</b>	1000 octets ( $10^3$ octets)
<b>Méga Octet</b>	<b>Mo</b>	1 million d'octets ( $10^6$ octets)
<b>Giga Octet</b>	<b>Go</b>	1 milliard d'octets ( $10^9$ octets)
<b>Téra Octet</b>	<b>To</b>	1000 milliards d'octets ( $10^{12}$ octets)

Tableau 1 - Unités de dénombrement de la quantité d'information



**Remarque :** Les unités de dénombrement de la quantité d'information présentées dans le tableau ci-dessus sont les unités normalisées par l'organisme international IEC afin d'être en accord avec le système international d'unités. Ce sont ces unités qui devraient être utilisées. Cependant dans l'ancienne notation, qui continue à être assez largement utilisée en dépit des recommandations officielles, on considère que  $1 \text{ Ko} = 2^{10} \text{ octets} = 1024 \text{ octets}$ ,  $1 \text{ Mo} = 2^{10} \text{ Ko} = 1\,048\,576 \text{ octets}$ ,  $1 \text{ Go} = 2^{10} \text{ Mo}$ , etc., alors que ces unités basées sur les puissances de 2 devraient être notées Kio, Mio et Gio. Si l'ordre de grandeur reste le même entre ces 2 systèmes d'unités, ceci peut expliquer certaines divergences dans le calcul de tailles de fichiers.

## 2. Codage informatique d'objets physiques

La numérisation désigne la conversion d'un objet réel (par exemple un paysage ou encore des ondes radio) en un objet virtuel, c'est-à-dire une suite de 0 et de 1 organisée de façon à pouvoir être décodée. La structuration de cette suite de bits est définie par le format de fichier choisi pour enregistrer les données numérisées (cf. fiche *Formats*). La numérisation d'un type d'objet donné se fait au moyen de périphériques spécifiques (par exemple une carte d'acquisition) et de logiciels spécialisés.

Numériser un objet physique ne revient pas à représenter toutes les facettes de cet objet, mais simplement à modéliser sous forme numérique les propriétés de cet objet jugées utiles en fonction de

Fiches L2 :

Codage>Codage numérique

l'usage prévu de l'objet virtuel et que l'on est capable de mesurer. Par exemple, lorsqu'on photographie un paysage, on se contente de capturer les couleurs réfléchies par les objets placés en face de l'objectif et suffisamment grands pour faire réagir les capteurs de l'appareil photo. Cependant un grand nombre d'informations ne sont pas mesurées : couleurs des objets placés hors-champ, température, son, odeurs, vitesse du vent, etc.

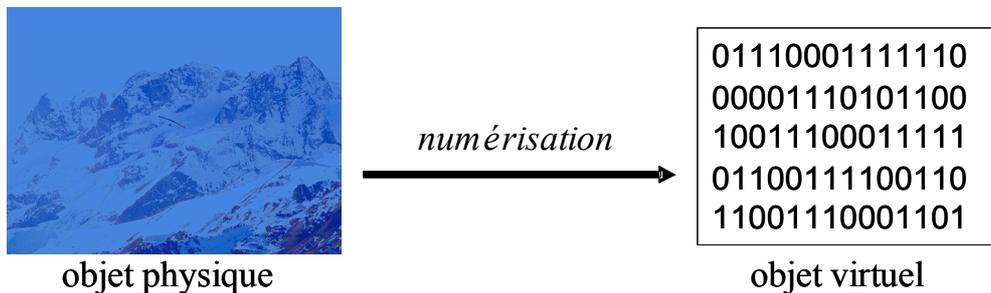


Figure 1 – Principe général de la numérisation



La numérisation d'un objet physique se fait par l'intermédiaire de capteurs qui permettent de mesurer les propriétés jugées utiles de l'objet à numériser (on appelle **acquisition** le transfert vers l'ordinateur des informations mesurées par le capteur) : microphone qui permet de mesurer les variations de pression dans le cas de l'acquisition de sons, capteurs photosensibles qui réagissent à la lumière réfléchiée par les objets dans le cas d'un appareil photo numérique ou d'un scanner, etc. Le principe de la numérisation ne se limite donc pas à la conversion en fichiers informatiques de textes, sons et images, mais s'applique également à la gestion informatique de capteurs divers (par exemple capteurs de pression dans l'industrie, suivi de la température d'un patient à l'hôpital, etc.).

La numérisation d'un texte revient à la numérisation d'une image, l'identification des caractères présents dans l'image numérisée étant ensuite prise en charge par un logiciel spécialisé de reconnaissance de formes.

Objet réel	Intermédiaire (capture / acquisition)	Objet virtuel
scène 3D statique	appareil photo numérique	image numérique (cf. <i>Formats d'image</i> )
photographie papier	scanner	image numérique (cf. <i>Formats d'image</i> )
document imprimé	scanner + application de reconnaissance des caractères (OCR)	document texte brut ou texte mis en forme (cf. <i>Formats texte</i> )
son (variations de la pression de l'air)	microphone + carte son	son numérique (cf. <i>Son et vidéo</i> )
scène 3D en mouvement	caméra numérique + carte d'acquisition	vidéo numérique (cf. <i>Son et vidéo</i> )

Tableau 2 - Applications courantes de la numérisation

## Références

- Codage binaire : <http://www.commentcamarche.net/base/binaire.php3>
- Numérisation : <http://fr.wikipedia.org/wiki/Num%C3%A9risation>

## Terminologie

- **Bit** = chiffre binaire = 0 ou 1
- **Octet** = byte
- **Numérisation** = digitalisation = conversion analogique/numérique