

7. Codage de contenu média



Principes de fonctionnement des ordinateurs

Jonas Lätt

Centre Universitaire d'Informatique



Trouvé une erreur sur un transparent? Envoyez-moi un message

- sur Twitter [@teachjl](#) ou
- par e-mail jonas.latt@unige.ch



Contenu du cours

Partie I: Introduction

1. Introduction

2. Histoire de l'informatique

3. Information digitale et codage de l'information

4. Codage des nombres entiers naturels

5. Codage des nombres entiers relatifs

6. Codage des nombres réels

7. Codage de contenu média

8. Portes logiques

9. Circuits logiques combinatoires et algèbre de Boole

10. Réalisation d'un circuit combinatoire

11. Circuits combinatoires importants

12. Principes de logique séquentielle

13. Réalisation de la bascule DFF

14. Architecture de von Neumann

15. Réalisation des composants

16. Code machine et langage assembleur

17. Réalisation d'un processeur

18. Performance et micro-architecture

19. Du processeur au système

Partie II: Codage de l'information

Partie III: Circuits logiques

Partie IV: Architecture des ordinateurs



Codage de texte

Codage de texte: de quoi parle-t-on?



Représentation de texte

Codage pour chacun des caractères.



Représentation d'un document Word

Codage pour les caractères, les indications de format et de police, la mise en page, les images et autres objets insérés, etc.

Que peut-on coder avec une taille de mot fixe?



1 Octet	Alphabète anglais
1 Octet	Alphabète anglais + quelques langues latines
1 Octet	Alphabètes cyrilliques
2 Octets	Alphabète chinois (quasi-totalité)
2 Octets	Langues européennes/américaines courantes, quelques langues asiatiques courantes, alphabet chinois (basique)
3 Octets	«Toutes» les langues

Historique: le codage de texte ASCII



- Le **codage ASCII** (American Standard Code for Information Interchange) est l'un des plus anciens standards pour l'échange d'information.
- Il est codé sur **7 bits** (8ème bit: vérification du transfert).
- **ASCII étendu**: utilisation du 8ème bit pour représenter 255 caractères au lieu de 127 (exemple: le codage `latin1`).
- **Désavantages** du ASCII ou ASCII étendu: on ne peut pas représenter différentes regions linguistiques en même temps, ou une langue avec plus de 255 caractères.

Codage de texte: ASCII



	Y	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F
X																	
0		□	□	□	□	□	□	□	□	□	□	□	□	□	□	□	□
1		□	□	□	□	□	□	□	□	□	□	□	□	□	□	□	□
2		□	!	"	#	\$	%	&	'	()	*	+	,	-	.	/
3		0	1	2	3	4	5	6	7	8	9	:	;	<	=	>	?
4		@	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
5		P	Q	R	S	T	U	V	W	X	Y	Z	[\]	^	_
6		'	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o
7		p	q	r	s	t	u	v	w	x	y	z	{		}	~	□



- **Unicode** est un standard moderne dans l'industrie du calcul qui permet d'**uniformiser et standardiser** le codage de caractères dans "toutes" les langues.
- Actuellement, Unicode représente à peu près 100'000 caractères.
- Chaque caractère est associé à un **point de code**, une valeur entre 0 et $10FFFF_{(16)}$: Place pour plus d'un million de caractères.
- Les points de code s'écrivent par des nombres hexadécimaux avec préfixe "**U+**". Exemple: le point de code de la lettre É est le U+C9.



Unicode n'est pas un codage

Unicode n'est pas un codage

... c'est un standard qui permet de définir différents codages de manière uniformisée

Un **codage Unicode** choisit soit l'intégralité des points de code, soit un sous-ensemble, et leur attribue une valeur entière (la valeur de codage)

Points de code: quelques exemples



Point de Code	Caractère	Origine
U+C9	É	Latin-1 étendu
U+42F	Я	Cyrillique (Russe)
U+E09	฿	Thailandais
U+13EA	Ꮝ	Cherokee
U+211E	℞	Symboles similaires aux lettres
U+21CC	⇔	Flèches
U+282F	⠆	Braille
U+244FA	𐀀	Idéogrammes unifiés Chinois / Japonais / Coréens



Codage sur un octet: ASCII

- ASCII est maintenant un codage Unicode pour l'alphabet anglais.

Codage sur deux octets: UCS-2

- Le codage UCS-2 représente les points de code de U+0000 à U+FFFF en utilisant la valeur même du point de code comme valeur de codage.
- Exemple: Sous Windows, les fonctionnalités du système sont en UCS-2.

Codage sur trois octets: UCS-3

- Tous les points de code, représentés par leur valeur même.
- **Avantage** d'un codage à taille fixe: simplicité et efficacité.
- **Désavantage** d'un codage à taille fixe: gaspillage de mémoire.

Unicode: Codages à taille variable



Exemple: UTF-8 utilise 1, 2, ou 3 octets pour représenter les points de code.

- Les textes écrits en ASCII peuvent alors aussi être interprétés comme des textes UTF-8.
- Le contenu Web moderne est le plus souvent codé en UTF-8.
- Les fonctionnalités système Linux / MAC OS X sont en UTF-8.
- **Avantage** d'un codage à taille variable: gain de mémoire.
- **Avantage** d'un codage à taille variable: généralité.
- **Désavantage** d'un codage à taille variable: lenteur.

Mauvaises interprétations de codage



Une erreur fréquente est qu'un texte Unicode est interprété comme un texte ASCII.

Exemple: jusqu'en 2014, tous les emails officiels de l'Université de Genève envoyés par la liste de mailing "Unilist" se terminaient par un message de ce genre:

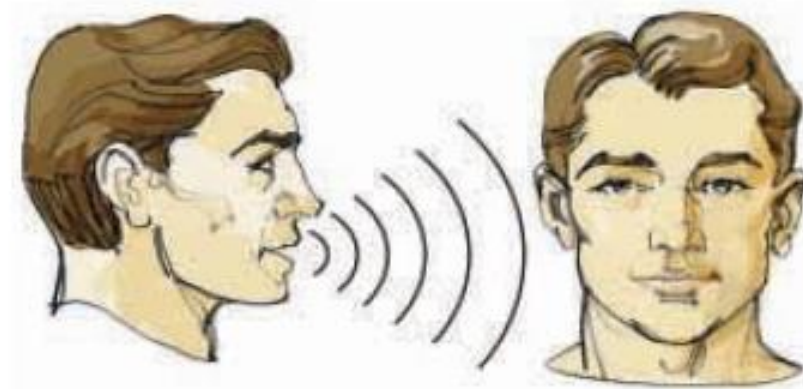
Ce message est envoyÃ© Ã :
- (Structure collaborateurs=Toutes ET Fonction collaborateurs=Toutes)



Codage du son

Perception du son

- Nous percevons le son comme une vibration d'une membrane dans nos oreilles à la suite d'ondes de compression dans l'air.



Représentation analogique du son



Microphone: traduit le son en vibrations d'une membrane, puis en signal électrique analogique.

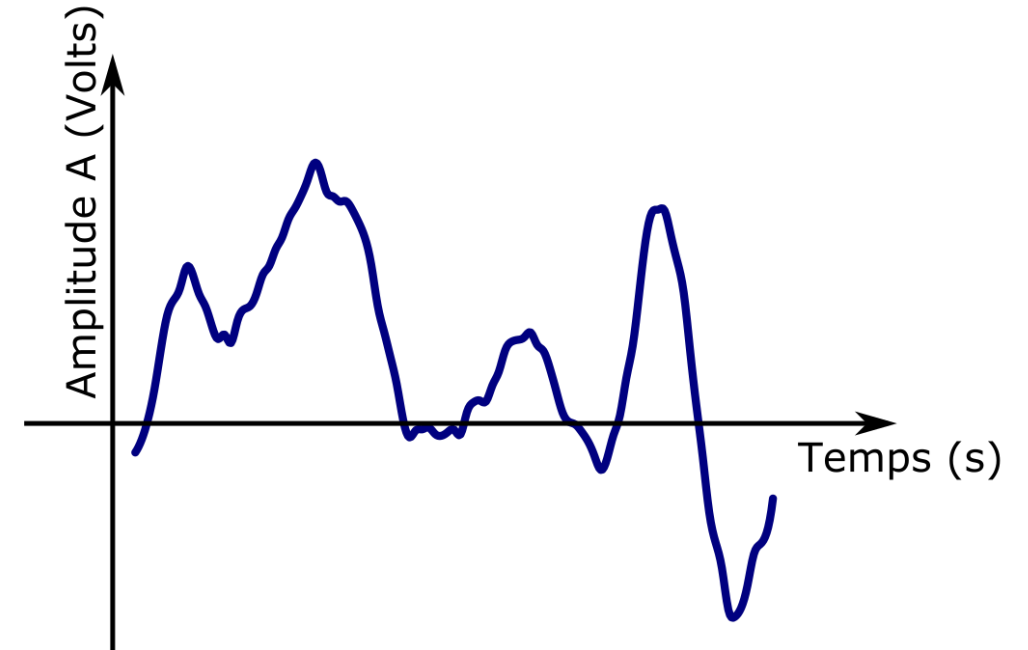
Onde sonore



Mouvement de membrane



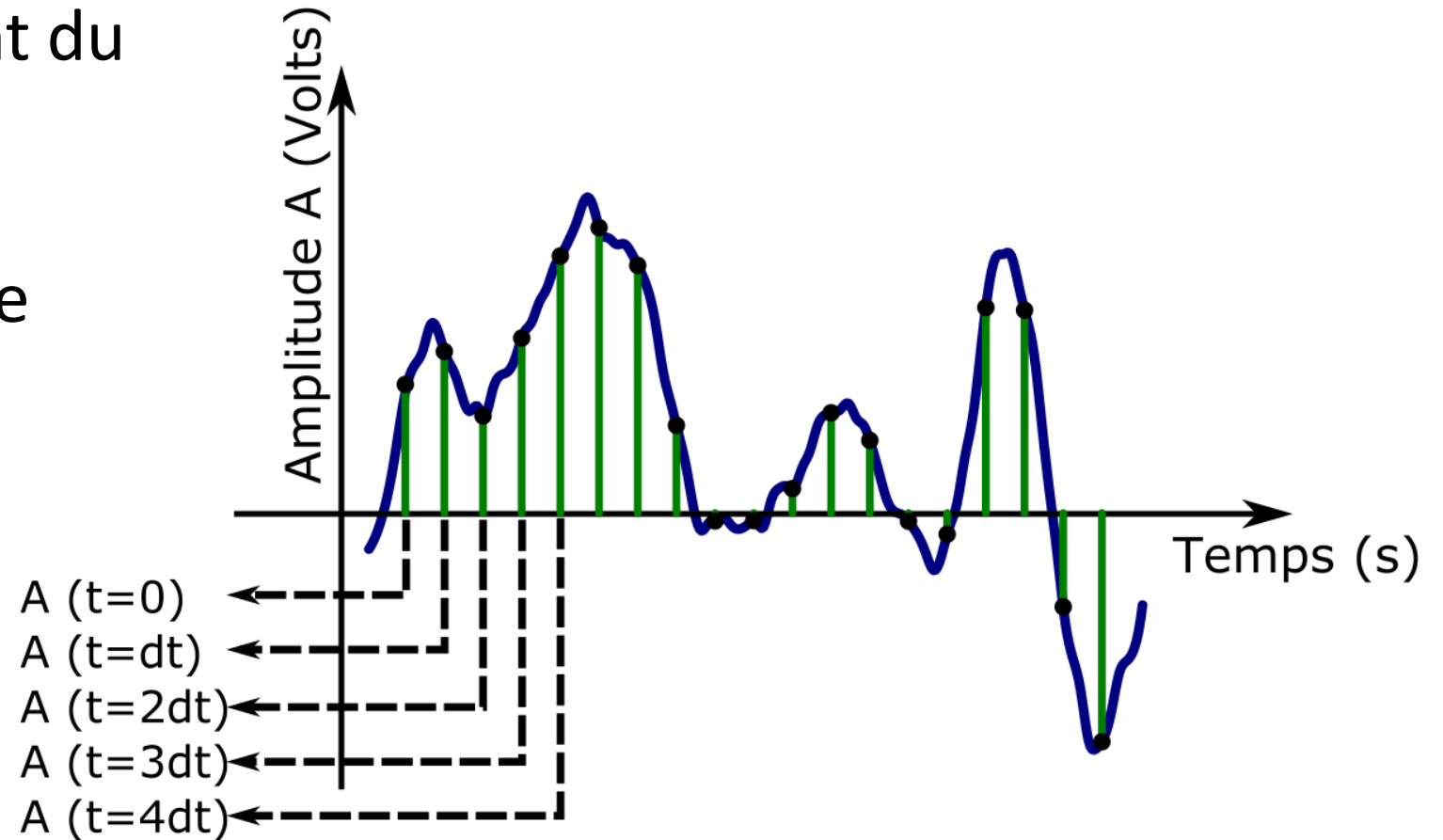
Signal électrique



Codage digital du son



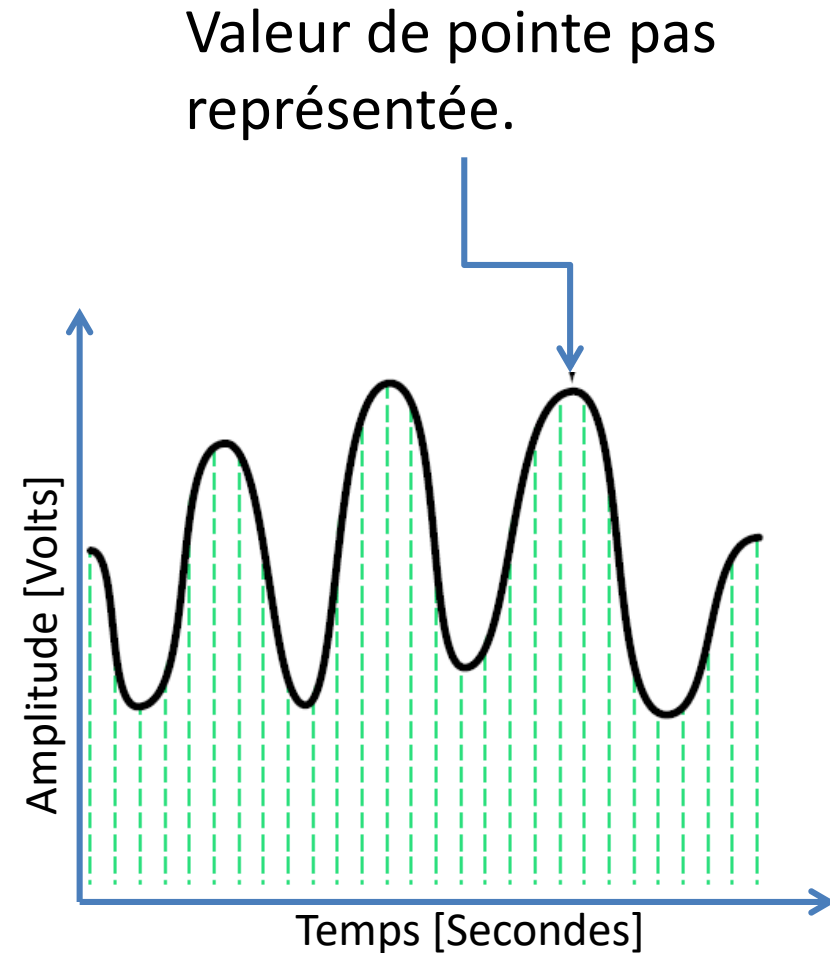
- **Sampling:** prélèvement du signal à des moments précis dans le temps.
- Permet de **digitaliser** le signal.



Codage digital du son



- Lors du sampling on perd de l'information.
- Bonne qualité à un **taux de sampling** de 40'000 valeurs par seconde (**40'000 Hz**).





Formats Audio

- La plupart des formats audio procèdent à une **compression des données**.
- Exemples de formats: WAV, AU, AIFF, VQF, et MP3.
- Format populaire: MP3. Bonne faculté de compression, stocke l'information audio sur peu d'espace.
- Compression MP3 basée sur l'étude du signal sonore, et l'élimination d'information inaudible.



Codage d'images

Perception des couleurs par la rétine des yeux

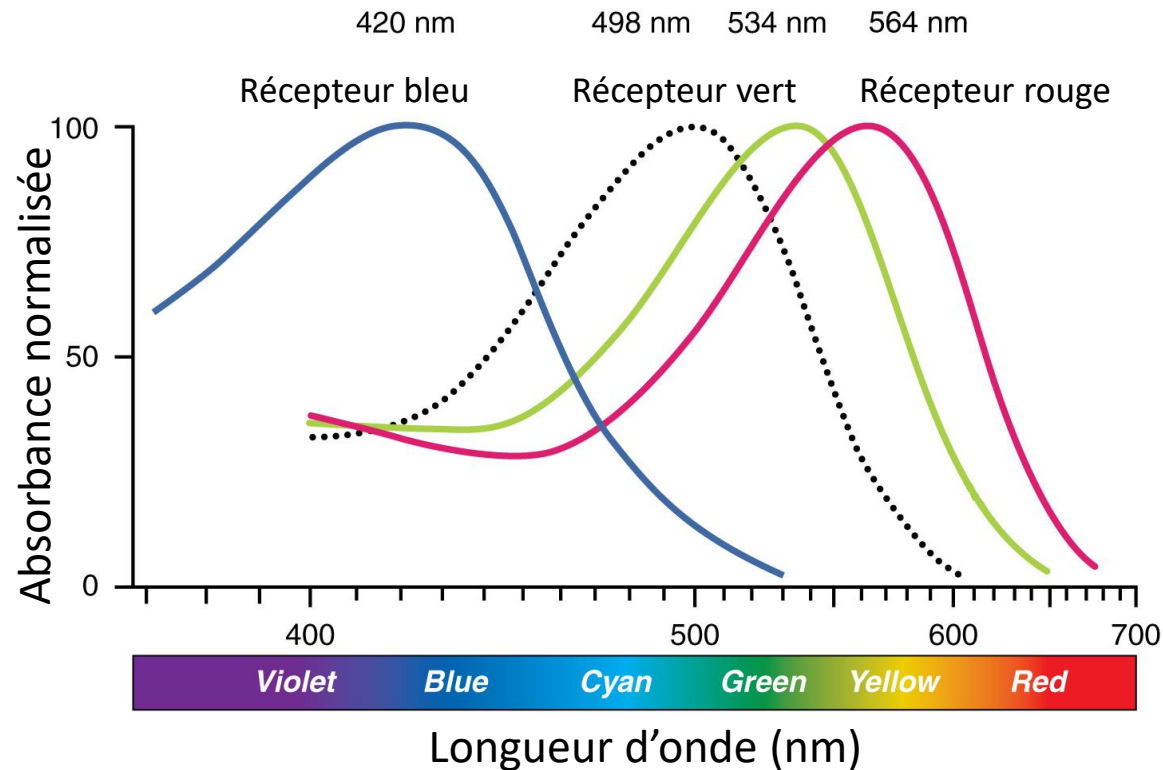


Image: OpenStax College - <http://cnx.org/content/col11496/1.6/>

Trois types de cellules photoréceptrices:

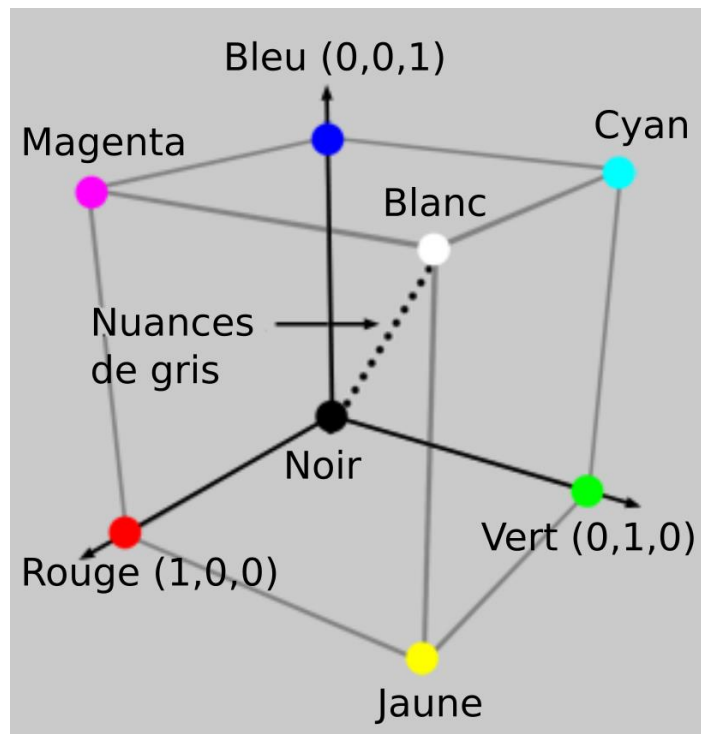
(1) Sensibles aux couleurs proches du rouge, (2) proche du vert, (3) proche du bleu.

Perception et représentation des couleurs



Représentation par un **triplet rouge-vert-bleu (RGB)**. Correspond à

- Les cellules photoréceptrices de trois types.
- Une lumière de trois couleurs émise par des écrans.



Triplets RGB: représentation tri-dimensionnelle de l'espace des couleurs.

Codage des couleurs



- Mémoire attribuée à une couleur: **profondeur de couleur (color depth)**.
- Profondeur de couleur de 24 bits (1 octet par couleur): représentation **TrueColor**. Plus de 16.7 millions de couleurs différentes (l'oeil humain distingue 10 millions de couleurs).

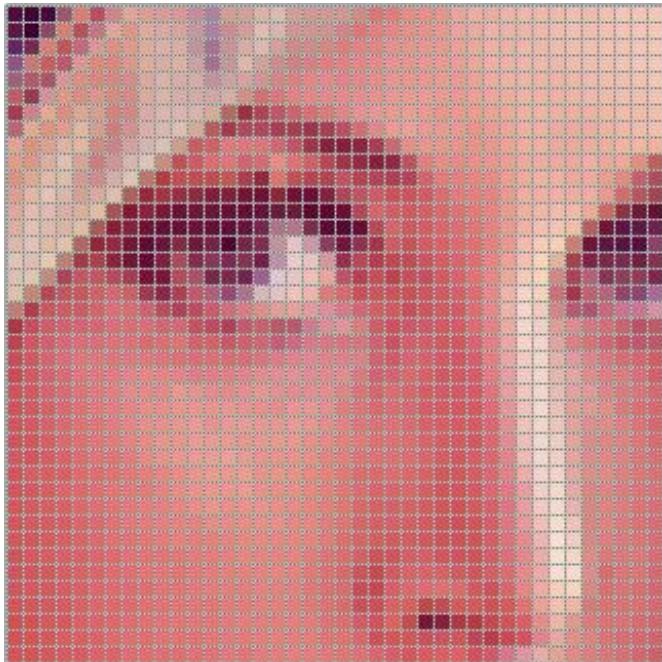
Rouge	Vert	Bleu	Couleur Résultante
0	0	0	
255	255	255	
255	255	0	
255	130	255	
146	81	0	
158	14	64	

Codage digital d'une image, format matriciel ("raster format")



Décomposition de l'image en «points» de couleur unique

- **Pixel**: un des points représentant l'image.
- **Résolution/Définition**: le nombre de pixels utilisés pour représenter l'image.

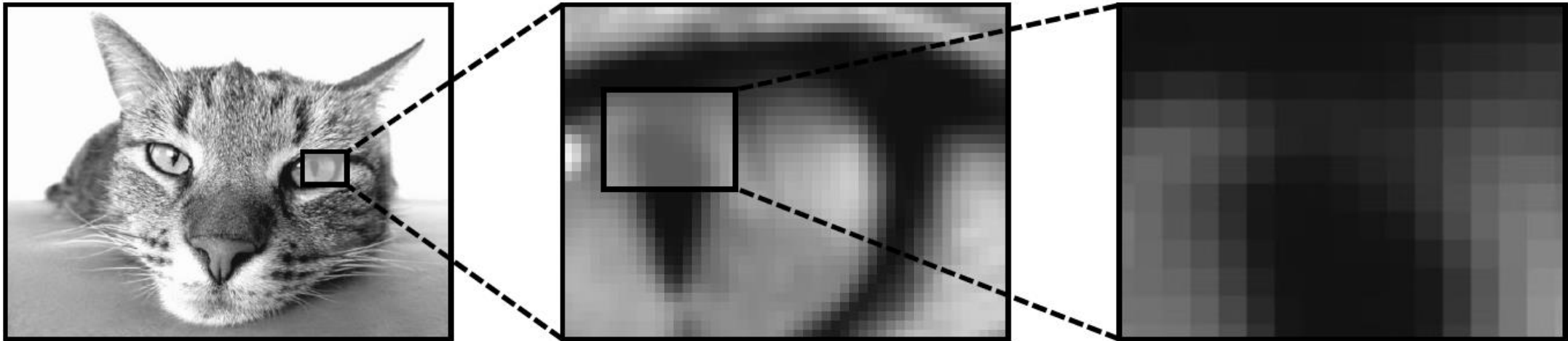


Chaque pixel est unicolore.
On lui attribue un triplet RGB.

Codage digital d'une image, format matriciel



- Une image matricielle est créée à une résolution spécifique.
- Redimensionnement: perte de qualité potentielle.



- Formats courants: GIF, le PNG, et le JPEG.
- Logiciels courants: Adobe Photoshop, GIMP.

Codage digital d'une image, format vectoriel

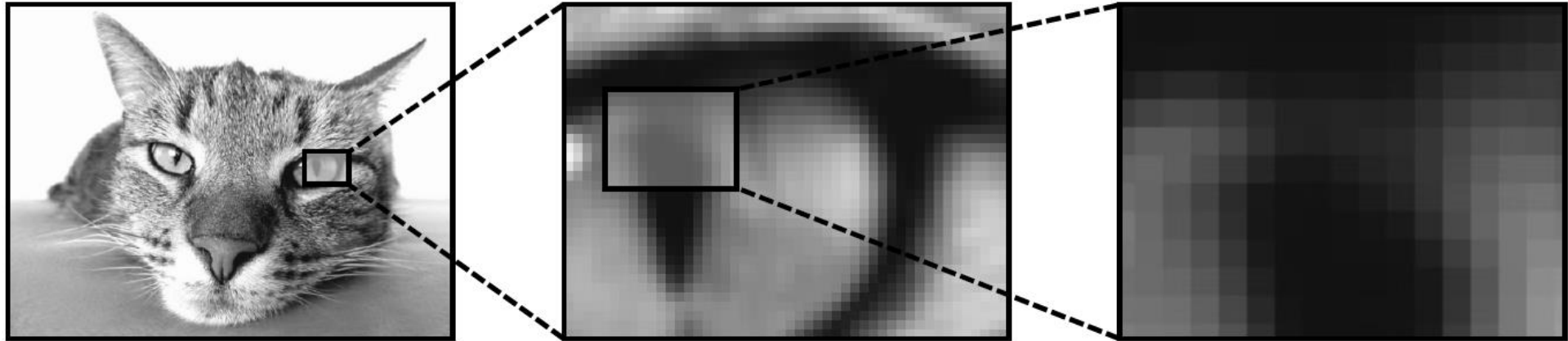


- Composée d'objets géométriques, les **primitives géométriques**: des droites, des arcs de cercle, courbes de Bézier, etc.
- Utilisation: dessins schématiques (des cartes, des graphiques, des illustrations, ...)
- Peut être redimensionnée à n'importe quelle taille, sans perte de qualité.
- Formats: SVG (exemple: pages Web), PDF (exemple: textes).
- Logiciels: Adobe Illustrator, Inkscape.

Redimensionnement: matriciel vs vectoriel



Matriciel



Vectoriel

