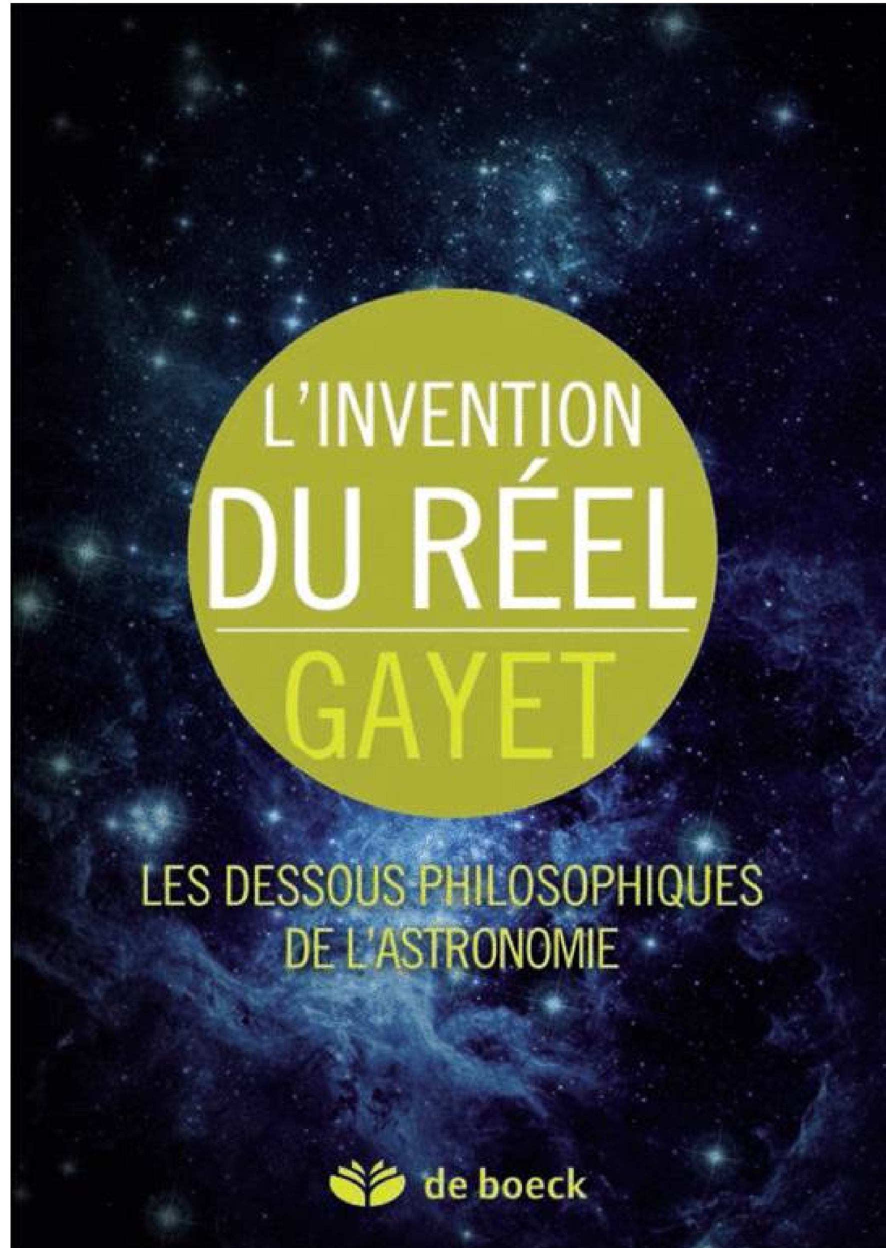
IMAGES NUMERIQUES

-

COMPRESSION JPEG

Stéphane Verjux – IREM de Franche Comté - 2017

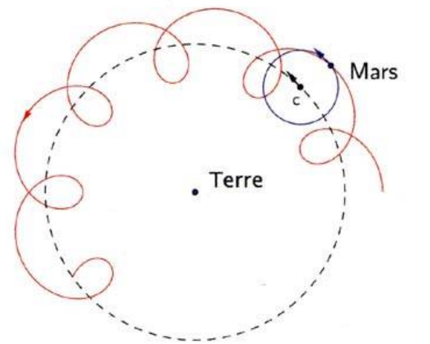
BILBLIOGRAPHIE



La planète, la chaleur et le MP3. L'idée d'Hipparque, qui consiste à décomposer une trajectoire compliquée en une association de cercles bien choisis, réapparaîtra de façon totalement inattendue deux millénaires plus tard. En effet, en 1843, le mathématicien et physicien Joseph Fourier publie son mythique ouvrage *Théorie de la Chaleur,* dans lequel le savant utilise un procédé très similaire de décomposition mathématique pour comprendre la diffusion de la chaleur dans un matériau.

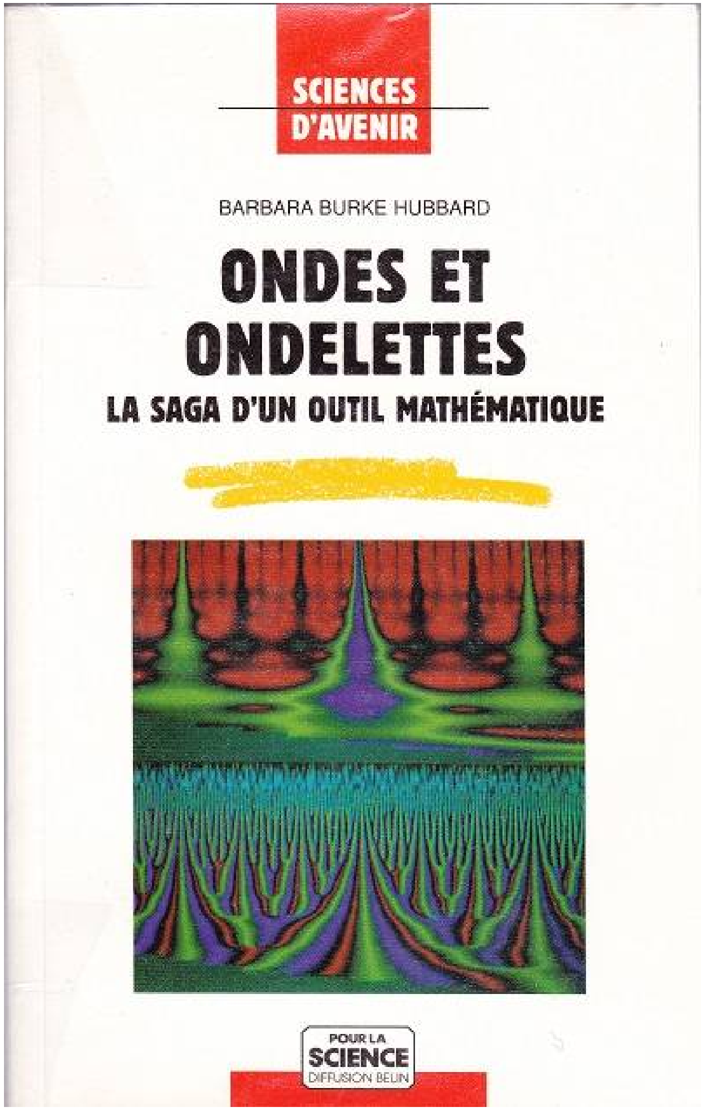
Rapidement, les mathématiciens comme les physiciens se rendront compte de l'universalité de ce mode de pensée, et de l'électricité à la théorie du signal en passant par l'analyse, la descendance d'Hipparque et Fourier sera prodigieuse. Alors qu'on pensait avoir épuisé toutes les applications possibles de cette théorie, l'informatique a offert récemment à Hipparque une nouvelle et singulière descendance : la compression de données. Elle donnera naissance à des enfants que les mystiques grecs et le génie de Fourier auraient eu bien du mal à imaginer : le format JPEG et les fichiers MP3.

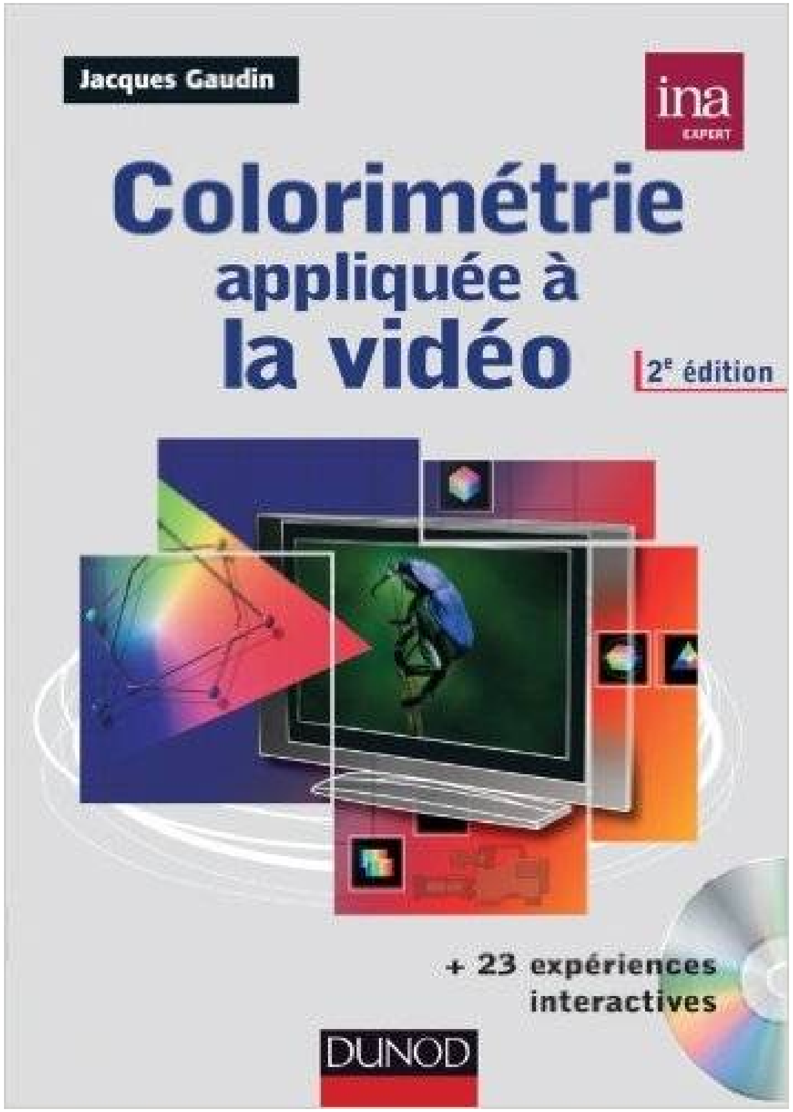
Jean Baptiste Joseph Fourier (1768-1830) croqué (à droite) par Julien-Leopold Boilly, 1820 (à gauche, le mathématicien Adrien-Marie Legendre). (Bibliothèque de l'Institut de France)



La trajectoire de Mars selon la théorie des épicycles. La planète parcourt un petit cercle (ici en bleu) dont le centre C parcourt un grand cercle (ici en pointillé) centré sur la Terre. La trajectoire totale (en rouge) peut être particulièrement complexe.

Expliquons en deux mots pourquoi l'idée d'Hipparque peut s'appliquer à la compression des données. Si l'on veut décrire une trajectoire très complexe comme celle de Mars, à chaque instant il faut associer la position dans le ciel de la planète. Avec trois cercles épicycles Hipparque n'a besoin de connaître que six informations : trois rayons et trois vitesses de rotation. Même si la courbe obtenue par la somme des trois cercles n'est pas tout à fait la trajectoire réelle, elle en est très proche, et le gain est énorme. Le principe du MP3 est semblable : au lieu d'enregistrer l'intégralité d'une onde sonore, on ne retient que les amplitudes des signaux élémentaires qui la composent. Ainsi, le nombre d'informations à conserver est beaucoup plus faible.


SCIENCES
D'AVENIR
BARBARA BURKE HUBBARD
ONDES ET
ONDELETTES
LA SAGA D'UN OUTIL MATHÉMATIQUE
POUR LA
SCIENCE
DIFFUSION BELIN


Jacques Gaudin
ina
EXPERT
Colorimétrie
appliquée à
la vidéo
2e édition
+ 23 expériences
interactives
DUNOD

Accromath Vol. 7 • été – automne 2012

https://fr.wikipedia.org/wiki/JPEG

Les sites Web que nous visitons sur la Toile sont maintenant inondés d'images. Cela constitue naturellement un problème de taille, car une quantité énorme d'informations doit être transférée du serveur jusqu'à notre ordinateur. Pour accélérer le traitement de ces images, il faut compresser celles-ci. Ce procédé diminue le poids de l'image en ne sacrifiant pas ou presque pas la qualité.

Les images sur la Toile, un défi de taille

Marc Bergeron
Cégep de Ste-Foy

Il existe plusieurs formats de compression d'images : JPEG, PNG, GIF, RAW, ... Chacun possède ses avantages. Par exemple, avec la compression au format PNG, il n'y a pas de perte de qualité et celui-ci gère la transparence de l'image (très utile pour de beaux effets visuels sur votre site préféré), mais le fichier sera habituellement plus volumineux que si nous avions utilisé la compression au format JPEG. Ce dernier est justement l'un des plus utilisés, car il est bien adapté aux images ayant plusieurs couleurs différentes comme les photographies. Regardons de plus près l'algorithme de compression JPEG dans lequel les mathématiques interviennent à plusieurs étapes.

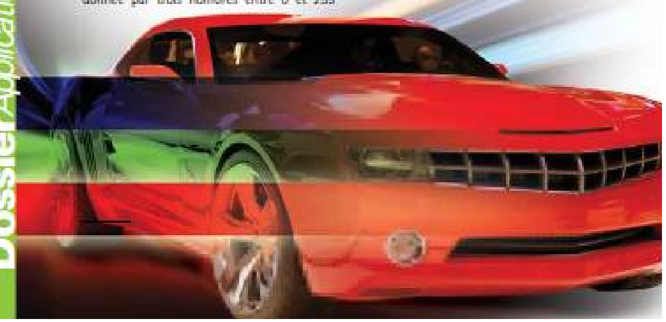
Première étape :

Changement de repère

Sur l'écran d'un ordinateur, une image est divisée en points appelés pixels. La couleur du pixel affichée sera habituellement donnée par trois nombres entre 0 et 255 (norme *RVB*) : le premier représente l'intensité de rouge; le deuxième, l'intensité de vert et le dernier, l'intensité de bleu. Le rouge, le vert et le bleu sont les longueurs d'onde détectées par les 3 types de cônes situés sur la rétine de nos yeux. Le mélange de ces trois couleurs permet de recréer toutes les couleurs détectables par l'humain. Une image peut donc être décomposée en trois images élémentaires (voir illustrations). Remarquons que dans l'image ii), la carrosserie de l'automobile nous semble presque noire. Cela est dû au fait qu'il n'y a pratiquement pas de teinte de vert sur celle-ci.

Avant de procéder à la compression au format JPEG, il faudra effectuer un changement de repère[1] du système à trois composantes *RVB* (rouge, vert, bleu) vers le système *LMN* (luminance, chrominance1, chrominance2).

1. *Les changements de repère ont déjà été partiellement expliqués dans Accromath, été-automne 2011, p. 22.*



JPEG

JPEG (acronyme de **Joint Photographic Experts Group**) est une norme qui définit le format d'enregistrement et l'algorithme de décodage pour une représentation numérique compressée d'une image fixe.

JPEG	
Extension	`.jpg, .jpeg, .jpe, .jfif`
Type MIME	`image/jpeg`
Type de format	image matricielle

Une photo de fleur compressée en JPEG, avec des compressions de plus en plus fortes, de gauche à droite.

Sommaire

Introduction au JPEG

JPEG est l'acronyme de *Joint Photographic Experts Group*. Il s'agit d'un comité d'experts qui édite des normes de compression pour l'image fixe. La norme communément appelée JPEG, de son vrai nom ISO/CEI 10918-1 UIT-T Recommendation T.81, est le résultat de l'évolution de travaux qui ont débuté dans les années 1978 à 1980 avec les premiers essais en laboratoire de compression d'images.

Le groupe JPEG qui a réuni une trentaine d'experts internationaux, a spécifié la norme en 1991. La norme officielle et définitive a été adoptée en 1992. Dans la pratique, seule la partie concernant le codage arithmétique est brevetée, et par conséquent protégée par IBM, son concepteur.

JPEG normalise uniquement l'algorithme et le format de décodage. Le processus d'encodage quant à lui est laissé libre à la compétition des industriels et des universitaires. La seule contrainte est que l'image produite doit pouvoir être décodée par un décodeur respectant le standard. La norme propose un jeu de fichiers de tests appelés fichiers de conformance qui permettent de vérifier qu'un décodeur respecte bien la norme. Un décodeur est dit conforme s'il est capable de décoder tous les fichiers de conformance.

Un brevet concernant la norme JPEG a été déposé par l'entreprise Forgent[1], mais a été remis en cause par le bureau américain des brevets (USPTO), qui l'a invalidé le 24 mai 2006 pour antériorité existante à la suite d'une plainte de la *Public Patent Foundation*[2]. Mais depuis le 27 septembre 2007, la société Global Patent Holdings, filiale d'Acacia Research Corporation, a à son tour revendiqué la paternité de ce format.

JPEG définit deux classes de processus de compression :

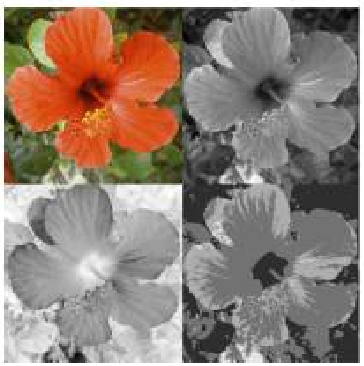
- avec pertes ou compression irréversible. C'est le JPEG « classique ». Il permet des taux de compression de 3 à 100[réf. nécessaire] http://images.math.cnrs.fr/Le-traitement-numerique-des-images

IMAGES DES MATHÉMATIQUES

Echos de la recherche (-Articles-mathematiques-.html)

LE TRAITEMENT NUMÉRIQUE DES IMAGES

Piste bleue (spip.php?page=mot&id_mot=21) le 28 novembre 2011 - Rédigé par Gabriel Peyré (_gpeyre_.html)



Les appareils numériques photographient de manière très précise le monde qui nous entoure. L'utilisateur souhaite pouvoir stocker avec un encombrement minimal ses photos sur son disque dur. Il souhaite également pouvoir les retoucher afin d'améliorer leur qualité. Cet article présente les outils mathématiques et informatiques qui permettent d'effectuer ces différentes tâches.

Cet article présente quelques concepts du traitement (http://fr.wikipedia.org

Pour la Science, Septembre 1987

L'analyse par ondelettes

Cet outil mathématique s'ajoute aux méthodes classiques d'analyse du signal. Il met l'accent sur les caractéristiques importantes du signal et semble en outre correspondre à des réalités physiologiques du traitement des signaux acoustiques et lumineux chez l'homme.

par Yves Meyer, Stéphane Jaffard et Olivier Rioul

L'analyse du signal porte sur un vaste ensemble de phénomènes et sur des réalités physiques diverses : la variation de la pression de l'air en un lieu donné en fonction du temps est un signal sonore ; l'évolution de l'intensité du courant en un point d'un réseau est un signal électrique ; la vibration du sol est un signal sismique et les fluctuations de l'indice de la bourse constituent un signal économique. Ces signaux dépendent d'une seule variable (ici le temps), mais ce n'est pas toujours le cas. Une photographie en noir et blanc peut être interprétée comme une quantité numérique (le niveau de gris) fonction de deux variables (les coordonnées du point considéré). L'analyse du signal consiste à extraire dans chacun de ces cas l'information pertinente, la nature de celle-ci différant selon la nature physique du signal.

Le géophysicien Jean Morlet s'est penché sur ce problème pour étudier certains signaux en sismique-réflexion, une méthode de recherche pétrolière qui consiste à émettre un signal vibro-sismique à la surface du sol à l'aide de camions vibrateurs ; on produit ainsi des ondes modulées en fréquence et de très faible intensité, qui se propagent dans le sous-sol et sont réfléchies différemment selon les couches géologiques. L'écho de ces ondes est capté par un système d'écoute disposé à la surface du sol et enregistré par un camion laboratoire.

L'analyse de ce signal doit renseigner sur la composition du sous-sol et l'éventuelle présence de couches de pétrole. Pour ce problème, l'analyse de Fourier classique s'était depuis longtemps montrée inadaptée et J. Morlet s'était rendu compte que la méthode de Gabor, que nous examinerons plus loin, était inadéquate, car elle ne permettait pas d'obtenir une résolution suffisante. Indépendamment d'un travail effectué au début des années 1960 par le mathématicien argentin Alberto Calderón, à l'Université de Chicago, pour résoudre des problèmes entièrement différents, J. Morlet a proposé en 1983, un procédé révolutionnaire, l'analyse et la synthèse par les ondelettes, qui permet d'analyser efficacement des signaux où se combinent des phénomènes d'échelles très différentes.

La transformation en ondelettes, créée pour résoudre des problèmes posés par la sismique-réflexion, a été ensuite appliquée à l'analyse des sons, des images, et de toute forme de signal.

Les débuts de l'analyse du signal

Le traitement du signal consiste à dégager des « informations » contenues dans un signal qui se déroule au cours du temps. Ce signal s est décrit par une fonction $s(t)$ du temps t et peut avoir des origines très diverses (sons musicaux, voix humaine, ondes sismiques, cardiogrammes, etc.). Quelles informations peut-on tirer d'un signal ? Lorsqu'on observe son évolution au cours du temps, on repère bien son commencement, sa fin et la durée de ses éléments caractéristiques, ainsi que des discontinuités, des changements de rythme, etc. En revanche, cette représentation temporelle du signal renseigne peu sur ses périodicités, donc sur ses fréquences. Depuis le XVIII^e siècle, de nombreux mathématiciens étudient la représentation en fréquence des signaux. La technique des séries de Fourier constitue sans doute le point de départ de cette approche qui a abouti à l'analyse par ondelettes.

Les séries de Fourier sont utilisées pour l'analyse des signaux périodiques, c'est-à-dire ceux qui, après un laps de temps, se répètent identiques à eux-mêmes, et ceci indéfiniment. Un tel signal est la superposition d'une onde sinusoïdale fondamentale, dont la fréquence est appelée fréquence fondamentale, et de divers harmoniques dont les fréquences sont des multiples entiers de la fréquence fondamentale. On calcule les amplitudes de ces différentes fréquences par des formules connues depuis le début du XIX^e siècle. Ces amplitudes s'appellent les coefficients de Fourier.

Les séries de Fourier ne permettent d'analyser que des phénomènes périodiques ; pour les phénomènes non périodiques on a recours à une intégrale de Fourier (ou somme continue) : cette méthode consiste à représenter le signal étudié par une superposition d'ondes sinusoïdales de toutes les fréquences possibles ; les amplitudes associées à chaque fréquence représentent les importances respectives des diverses ondes sinusoïdales dans le signal global. Ces amplitudes forment alors une fonction de la fréquence f que les physiciens appellent « spectre continu des fréquences du signal » : c'est la transformée de Fourier du signal $s(t)$, notée $S(f)$. Cette transformée est égale à l'intégrale pour toutes les valeurs du temps du produit du signal $s(t)$ par la fonction $e^{2i\pi ft}$. On la calcule à l'aide de l'intégrale de Fourier :

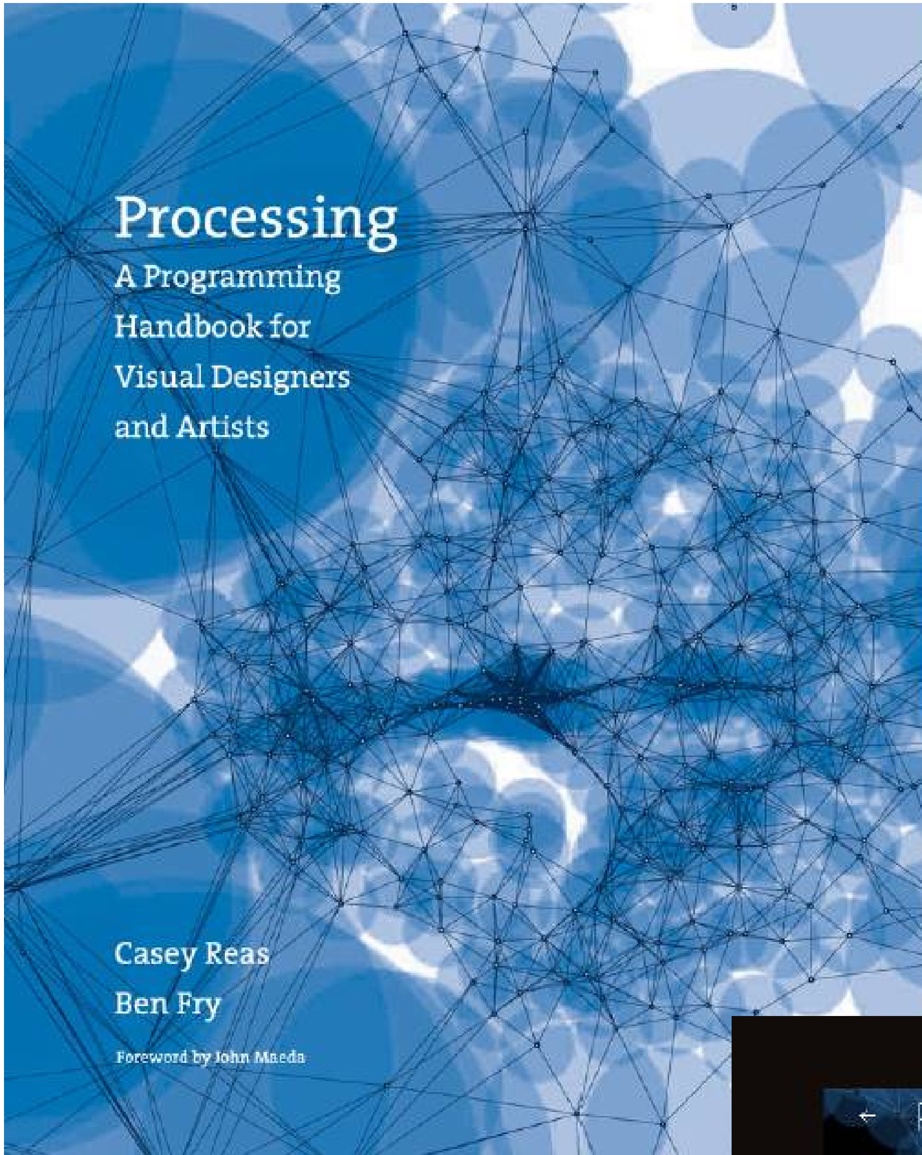
$$S(f) = \int_{-\infty}^{+\infty} s(t)e^{2i\pi ft}\, dt$$

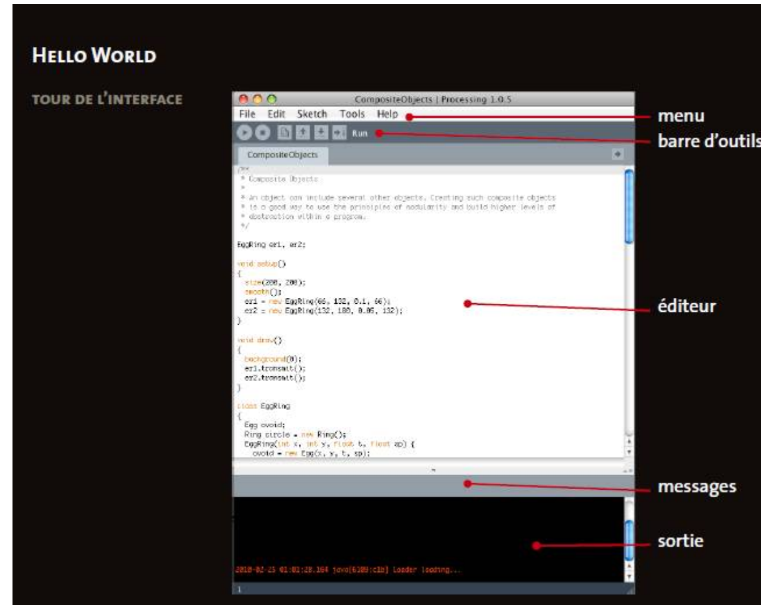
La quantité $e^{2i\pi ft}$ est égale à $\cos 2\pi ft + i \sin 2\pi ft$ où i est un nombre « imaginaire » dont le carré est égal à -1. Le nombre complexe $S(f)$ s'identifie, pour une fréquence f donnée, à un point M du plan qu'on repère par ses coordonnées cartésiennes (partie réelle ou partie imaginaire) ou encore par sa distance à l'origine O (module) et l'écart angulaire (phase) de la droite OM par rapport à l'horizontale.

En étudiant la partie réelle de la transformée de Fourier du signal décrit mathématiquement par $s(t) = \sin 2\pi t + 1/2 \sin 4\pi t + 1/2 \sin 6\pi t + ...$, on reconnaît sur le spectre les fréquences des divers harmoniques d'un signal périodique.

L'intégrale de Fourier est très générale, car on peut l'appliquer sans faire d'hypothèse sur l'origine physique du signal : elle s'est révélée très fructueuse aussi bien d'un point de vue théorique que numérique et tout particulièrement depuis que l'on a mis au point, pour la calculer, un algorithme extrêmement économique, la transformée de Fourier rapide.

Cette méthode ne permet pas d'analyser correctement tous les types de signaux, comme le montre l'exemple suivant, tiré de la musique. Si nous


Processing
A Programming Handbook for Visual Designers and Artists
Casey Reas
Ben Fry
Foreword by John Maeda


Hello World
Tour de l'interface
File Edit Sketch Tools Help
menu
barre d'outils
éditeur
messages
sortie


Processing
Workshop __ Beaux arts de Besançon _ 2 / 3 / 4 mars 2010

La compression JPEG

Le processus de compression et de décompression JPEG irréversibles comporte six étapes principales représentées ci-dessous :

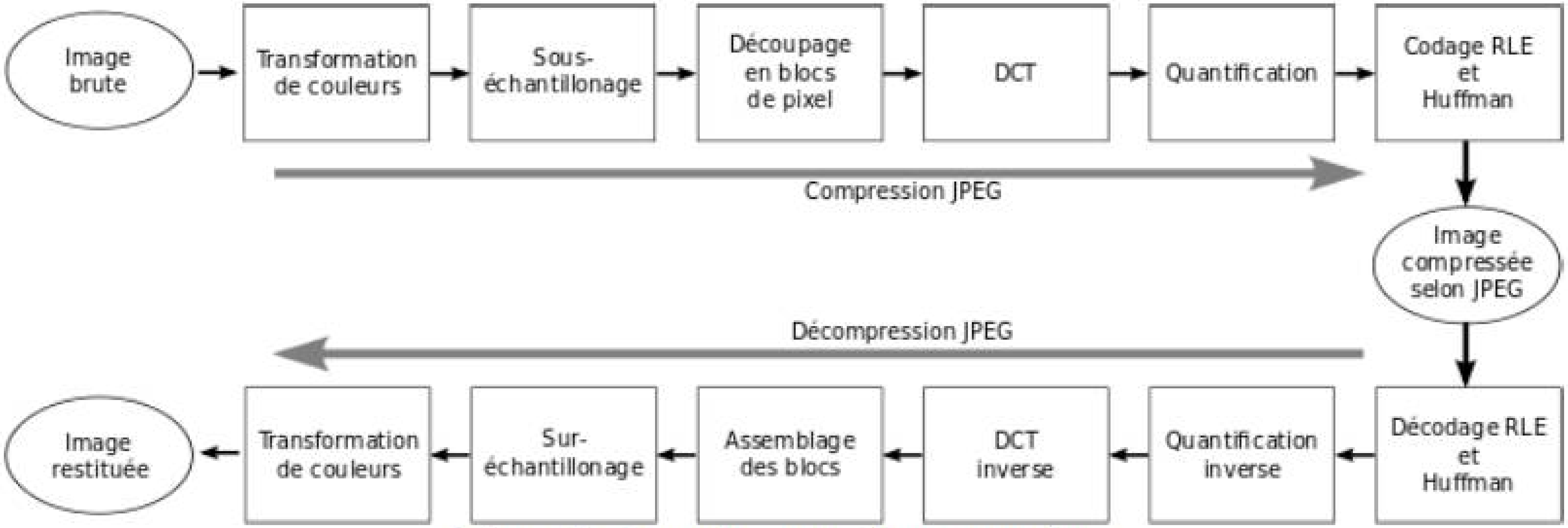


Figure 1 : Organigramme de compression.

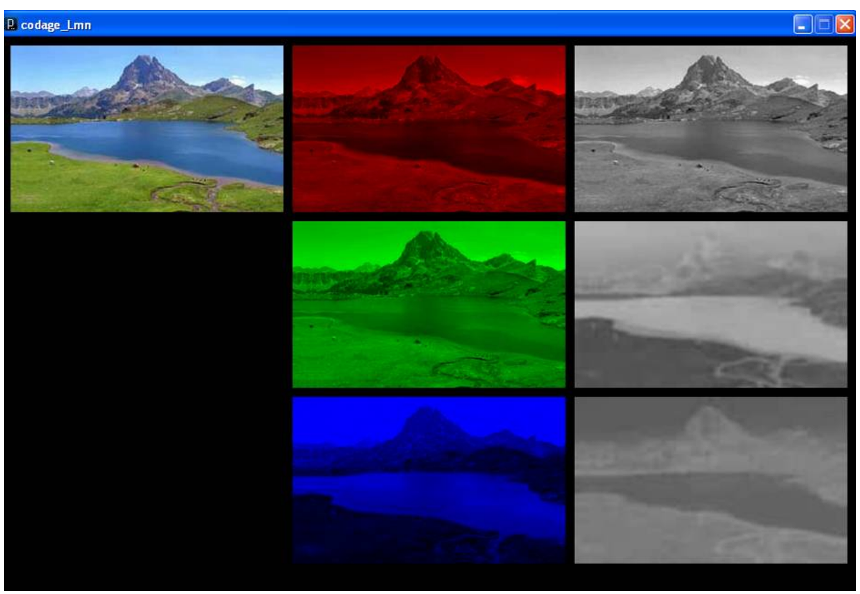
Etape 1. Transformation des couleurs

Les chrominances

La chrominance1 (notée M) varie linéairement avec l'écart entre le niveau de bleu et la luminance du pixel. Donc,

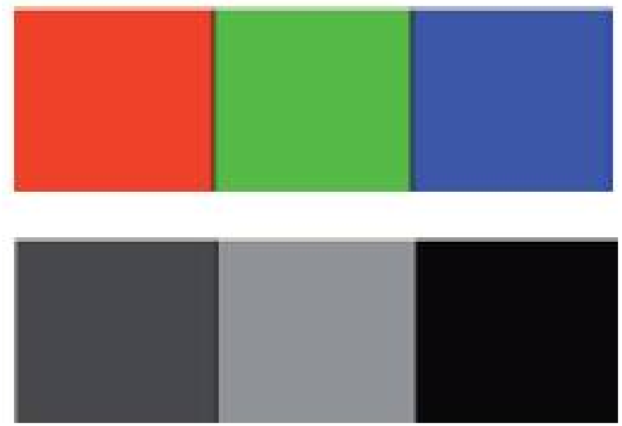
$$M = a(B - L) + b. \qquad (^{**})$$

La chrominance2, pour sa part, contient l'information de l'écart entre le niveau de rouge et la luminance ($a(R - L) + b$, où $a = 0{,}71327$ et $b = 128$).



La luminance

La luminance[2] contient l'information des niveaux de gris de l'image. Pour la définir, il faut remarquer que la sensibilité de nos yeux n'est pas la même pour les couleurs de base (rouge, vert et bleu). La figure suivante représente les trois couleurs en noir et blanc (plus précisément en niveaux de gris) :



Luminance apparente du rouge, du vert et du bleu

Étant donné que le vert est presque au milieu du spectre visible, cette couleur aura la plus grande luminance apparente, suivie par le rouge et le bleu. Après des tests expérimentaux effectués par la Commission internationale de l'éclairage (CIE), la luminance (L) a été définie par la formule suivante :

$$L = 0{,}299\,R + 0{,}587\,V + 0{,}114\,B, \qquad (^{*})$$

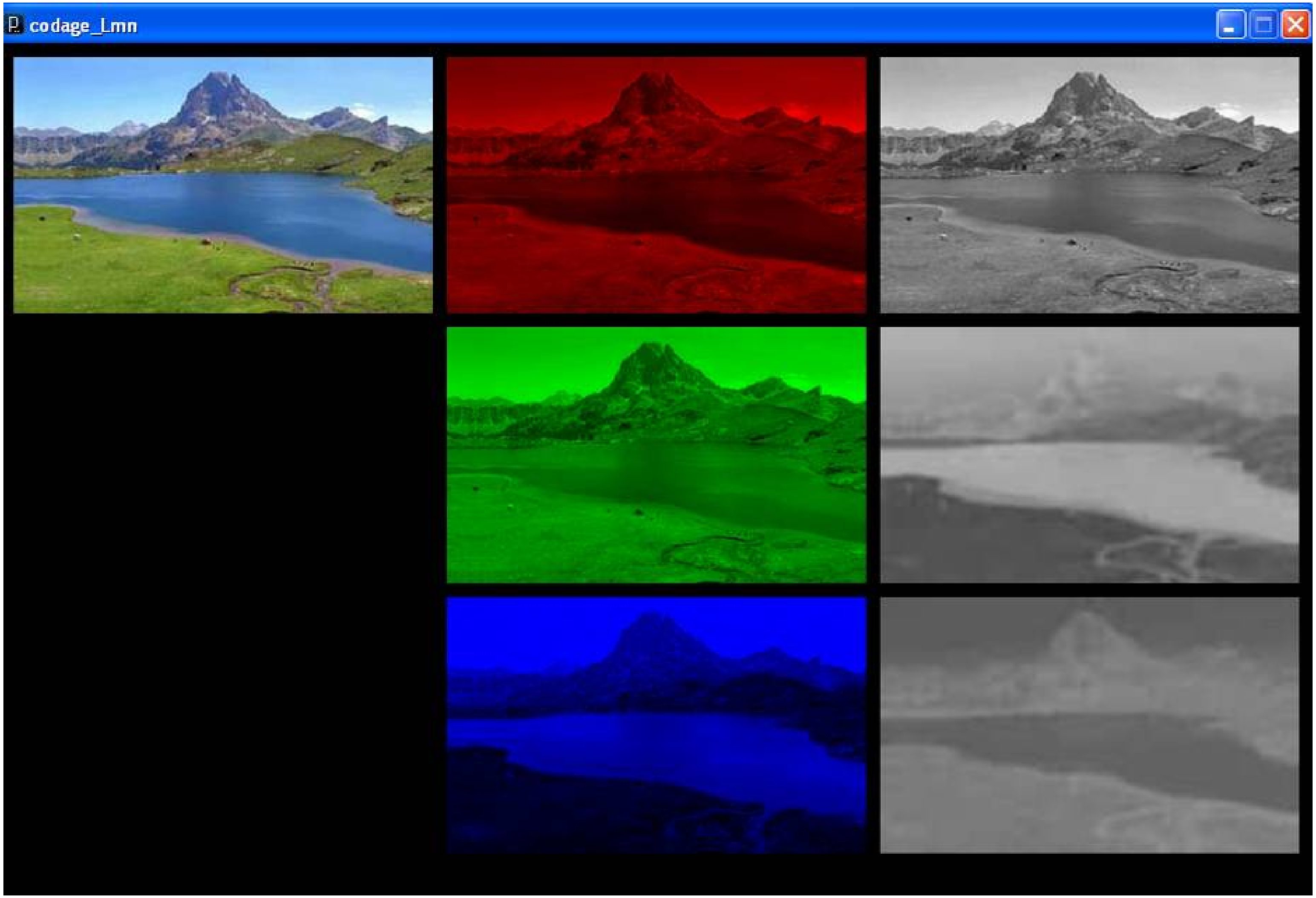
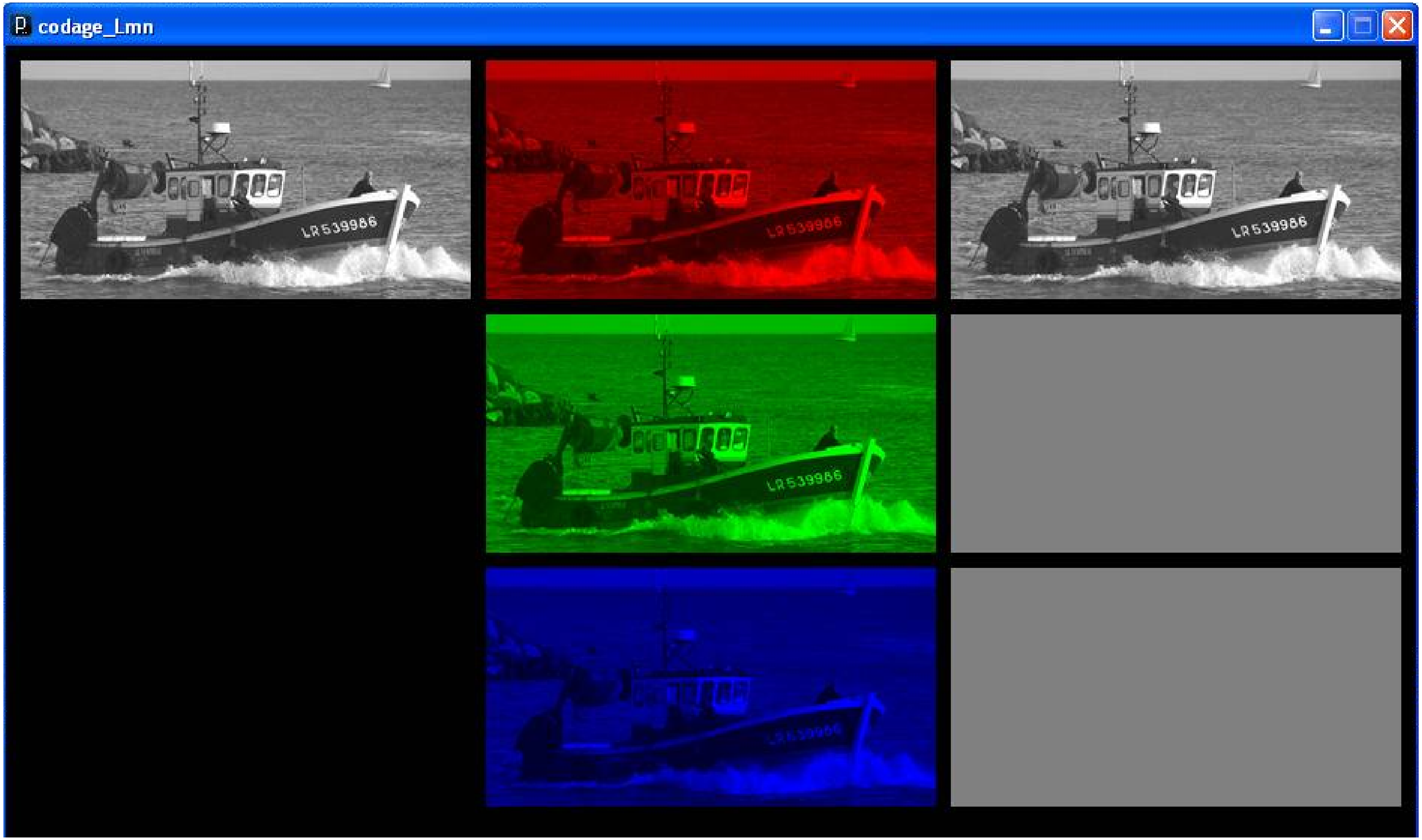


Image brute => Couches R,V et B => Couches Luminances, Chrominances M et N.

Les couches de chrominances comportant peu d'informations, on va donc les **sous-échantillonner (étape 2)**.


codage_Lmn

Changement de repère

Le changement de repère du système *RVB* vers le système *LMN* se fait facilement en utilisant les matrices et les opérations sur celles-ci :

$$\begin{pmatrix} L \\ M \\ N \end{pmatrix} = \underbrace{\begin{pmatrix} 0{,}299 & 0{,}587 & 0{,}114 \\ -0{,}168\,74 & -0{,}331\,26 & 0{,}5 \\ 0{,}5 & -0{,}418\,69 & -0{,}081\,31 \end{pmatrix}}_{T} \begin{pmatrix} R \\ V \\ B \end{pmatrix} + \begin{pmatrix} 0 \\ 128 \\ 128 \end{pmatrix}$$

Pour passer du système *LMN* au système *RVB*, lors de la décompression, il suffit de prendre l'inverse de la matrice T :

$$\begin{pmatrix} R \\ V \\ B \end{pmatrix} = \underbrace{\begin{pmatrix} 1 & 0 & 1{,}402 \\ 1 & -0{,}344\,13 & -0{,}714\,14 \\ 1 & 1{,}772 & 0 \end{pmatrix}}_{T^{-1}} \begin{pmatrix} L \\ M-128 \\ N-128 \end{pmatrix}$$

Etape 3. Découpage en bloc et transformation DCT

La transformée DCT s'exprime mathématiquement par :

$$\mathrm{DCT}(i,j) = \frac{2}{N} C(i) C(j) \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} \mathrm{pixel}(x,y) \cos\left[\frac{(2x+1)i\pi}{2N}\right] \cos\left[\frac{(2y+1)j\pi}{2N}\right]$$

Équation 1 : Transformée DCT directe.

Et la transformée DCT inverse s'exprime par :

$$\mathrm{pixel}(x,y) = \frac{2}{N} \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} C(i) C(j)\, \mathrm{DCT}(i,j) \cos\left[\frac{(2x+1)i\pi}{2N}\right] \cos\left[\frac{(2y+1)j\pi}{2N}\right]$$

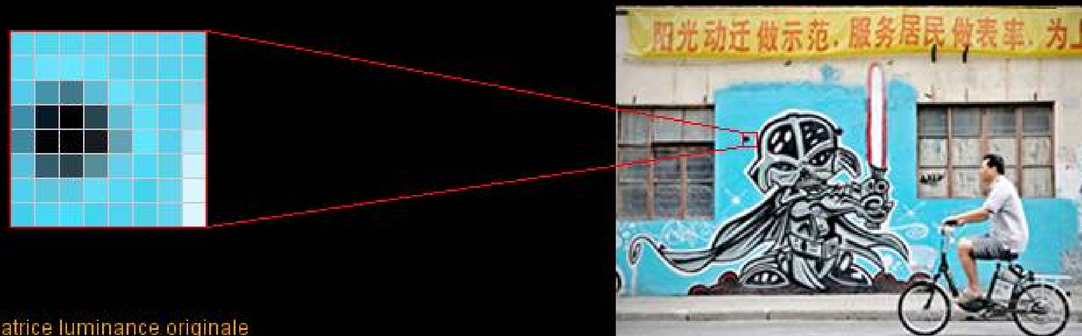
Équation 2 : Transformée DCT inverse.

Dans les deux cas, la constante C vaut :

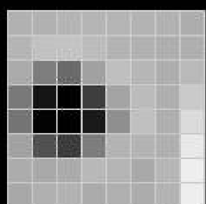
$$C(x) = \begin{cases} \frac{1}{\sqrt{2}} & \text{pour } x = 0 \\ 1 & \text{pour } x > 0 \end{cases}$$

Équation 3 : Définition de la constante C.

compression_jpeg_01



Matrice luminance originale

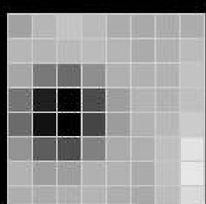


177	177	180	180	180	179	178	173
181	193	195	188	179	177	175	175
170	125	106	161	191	182	174	185
121	20	0	62	164	187	175	202
116	3	0	25	143	192	175	219
162	81	60	125	179	179	176	232
172	169	171	184	180	169	182	238
168	176	178	171	175	172	179	238

252	-22	5	4	4	0	0	0
0	4	-2	0	-1	0	0	0
15	10	-2	-5	-2	0	0	0
-1	-1	0	0	0	0	0	0
-4	-2	0	1	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0

Matrice TCD

Matrice luminance JPEG



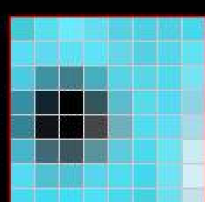
161	182	195	185	172	169	171	172
182	181	185	187	180	171	176	187
165	121	107	143	175	176	178	194
116	31	1	76	158	179	178	191
107	18	-12	69	156	178	184	203
148	93	78	129	170	175	191	226
178	158	158	176	178	171	192	230
174	171	175	180	173	168	188	217

taux de compression : 4

Matrice erreur

16	-4	-14	-5	7	10	7	0
0	11	9	0	-1	5	0	-12
5	4	0	18	16	6	-3	-8
4	-10	-1	-14	5	8	-3	11
8	-15	11	-43	-12	13	-9	16
14	-11	-18	-3	8	3	-15	6
-5	10	12	8	1	-1	-10	8
-5	4	3	-9	1	4	-8	21

Matrice couleur JPEG



Puis, on applique la transformée en cosinus discrète (*TCD*) à chacun des éléments de ces petits tableaux.

On obtient :

$$\begin{pmatrix} 1\,041 & -7 & 13 & -6 & 26 & -17 & -9 & -19 \\ 10 & -25 & 21 & -4 & -7 & 10 & 15 & 6 \\ -21 & -18 & -3 & 8 & -22 & -1 & -18 & 6 \\ -16 & 1 & 10 & -15 & -9 & -9 & -1 & 15 \\ -6 & 15 & 7 & 0 & -7 & 15 & 19 & 19 \\ 6 & -2 & -15 & 11 & 12 & -5 & 1 & -7 \\ 21 & -13 & -7 & 10 & -1 & -11 & 1 & -2 \\ 8 & -20 & -5 & 6 & -5 & -6 & -1 & 3 \end{pmatrix}$$

Tableau 2 : chaque élément est calculé avec la formule $TCD(x,y)$.

C'est à cette étape que la luminance est réellement compressée. On choisit d'abord un facteur de compression (une valeur facteur comprise entre 1 à 25). Plus ce facteur est grand, plus on élimine des valeurs du tableau 2, donc, plus la luminance est compressée. Pour déterminer les valeurs qui seront éliminées, il faut considérer deux critères. Premièrement, plus une valeur obtenue dans le tableau 2 est petite, plus cette donnée pourra être éliminée. Deuxièmement, plus la valeur est près du coin inférieur droit, plus cette donnée pourra aussi être éliminée. Ce dernier critère provient du fait que les valeurs de la *TCD* situées près du coin inférieur gauche correspondent à des irrégularités de l'image initiale moins perceptibles par notre œil. Pour tenir compte de ces deux critères, il suffit de diviser chaque valeur $TCD(x, y)$ par

$$1 + \text{facteur} * (x + y - 1)$$

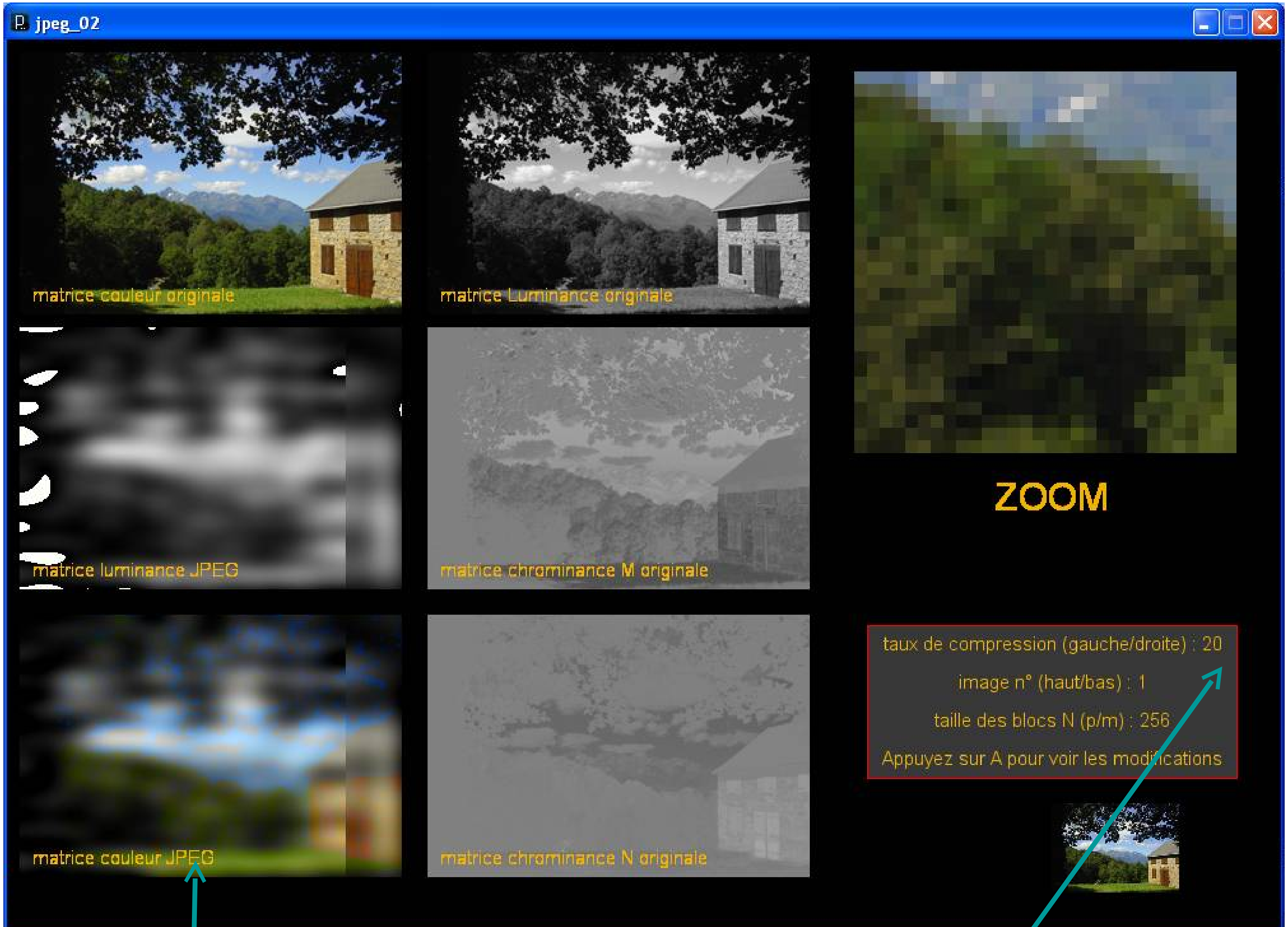
et d'arrondir à l'entier le plus près.

La compression consiste à éliminer les valeurs nulles du tableau 3 situées à la fin du zigzag comme illustré dans la figure suivante :

$$\begin{pmatrix} 208 & -1 & 1 & 0 & 1 & -1 & 0 & -1 \\ 1 & -2 & 1 & 0 & 0 & 0 & 0 & 0 \\ -2 & -1 & 0 & 0 & -1 & 0 & 0 & 0 \\ -1 & 0 & 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Tableau 4 : chemin de lecture des valeurs en zigzag


compression_jpeg_02
matrice couleur originale
matrice Luminance originale
matrice luminance JPEG
matrice chrominance M originale
matrice couleur JPEG
matrice chrominance N originale
ZOOM
taux de compression (gauche/droite) : 21
image n° (haut/bas) : 0
taille des blocs N (p/m) : 8
Appuyez sur A pour voir les modifications



Filtrage passe-bas d'une image par compression Jpeg

Merci !

FIN…