

PROCES-VERBAL

FORMATION EN REALISATION LINEAIRE ET NON LINEAIRE

« DU PIXEL A L'IMAGE »

Le 04 octobre 2023 par vidéoconférence

Destinée aux organismes du CIRTEF :

**ORTB Bénin, RTB Burkina Faso, ORTC Comores, MBC Ile Maurice, RTN Niger, Radio
Lomé Togo – CRTV Cameroun (pas de désignation)**

Présentateur : Ali Oumarou, Directeur Afrique

Formateurs : Roger Roberts, CIRTEF/TITAN

Damien Duriau, RTBF

Liste des participants :

ORTB Bénin	Mr Lionel Gbegonnoude (point focal)
	Mme Christelle Sagbohan
	Mr Arnaud Togbenon
	Mr Armel Gbeyihin
	Mr Janvier Tonon
	Mr Thimothée Houton
RTB Burkina Faso	Mme Delphine Yerbenga (point focal)
	Mme Clémence Tuina (point focal)
	Mr Gilbert Tougma
ORTC Comores	Mr Chahalane Ahamed Saïd Ali Mohamed (point focal – absent)
	Mr Inzoudine Ahmed
	Mr Irham Mohamed
	Mr Mohamed Hila
	Mr Ahamadi Houmadi
	Mr Noufroi Bent Nidhoime
	Mr Saindou Abdillah Libasse

	Mr Ali Maoulida Mr Hadidja Abdillah Mr El Arif Mohamed Soilih Mr Abdallah Mohamed Mr Sainda Ali
MBC Ile Maurice	Mr Rakesh Gooljar (point focal) Mr Mahadeven Mootoosawmy
RTN Niger	Mme Hamey Ramatou Doulla (point focal)
TVT Togo	Mr Fidel Mawuena (point focal) Mme Rachel Kpizing Mr Egome Amah Mr Mathias Gbegbeni N'Lipibe Mr Dissirama Kossi Sodhar
VOV Vietnam	Mme Dieu Van Mme Thi Thu Ha Mme Thuy Van Nguyen (excusée)
CSTIC/ISTIC (Institut des Sciences et Techniques de l'Information et de la Communication du Burkina Faso)	Mme Mafarma Sanogo
CRPF Cotonou	Mr Anselme Awannou
CRTF Yaoundé	Mr Hubert Atangana
CIRTEF	Mme Lucia Verdone Mr Daniel Brouyère (excusé)

1. CONTENU

Au nom de l'Equipe Formation « CIRTEF », j'ai le plaisir de vous transmettre les documents relatifs à la formation « image » ... bien plus proches d'un quotidien dans l'opérationnel.

En 1666, lors de ses expériences d'optique, Isaac Newton fait passer de la lumière solaire blanche à travers un prisme de façon à la décomposer en rayons lumineux de différentes couleurs. Newton est le premier à comprendre que la lumière blanche est faite d'un mélange de rayons de lumière de couleurs différentes. C'est le début de notre aventure !

La couleur est en effet le résultat de l'interaction de la lumière et de la matière. La lumière blanche est composée de plusieurs ondes électromagnétiques qui se décomposent en spectre coloré lorsqu'elles traversent un prisme ou un fluide. La matière absorbe, réfléchit, émet ou transmet certaines longueurs d'onde de la lumière, ce qui crée la sensation de couleur perçue par l'œil et le cerveau. La couleur d'un objet dépend donc de la lumière qui l'éclaire et de l'observateur qui le regarde.

Cette seconde formation traite donc des normes et standards des formats de codage employés pour spécifier les images fixes et animées.

En photographie : la dimension de la photo en représentation, le type d'affichage (en transparence – en réflexion), la définition (le nombre de points ou pixels significatifs de la photo ligne horizontale / nombre de lignes verticales), la résolution (pixels par centimètre carré), la profondeur de couleur (sur 8 ou 10 bits) ;

Pour la vidéo, toutes les spécifications de la photographie, avec en plus le rapport hauteur/largeur de l'image (4:3, 16:9 par exemple), la fréquence images par seconde (25, 30, 50, 60, ...), le balayage entrelacé ou progressif ; ...

La transformation de la photo ou de la vidéo comporte en général trois phases :

- une phase d'échantillonnage où des caractéristiques d'un document sont explorées à intervalles temps régulières ;
- une phase de quantification qui attribue une valeur représentant l'objet aux points d'échantillonnage (pixel pour l'image) ;
- une phase de numérisation qui confère à chacune de ces valeurs une valeur numérique (par exemple sous forme binaire) !

L'ensemble de ces traitements effectués constitue un fichier de données qui va stocker l'ensemble des données numérisées relative à un document, cela implique de comprendre le taux de compression, le débit numérique et les algorithmes exploités.

La suite dans les pages qui suivent ...

Si vous souhaitez visionner le PPT présenté en séance, veuillez bien à le télécharger et l'enregistrer car le lien expirera **le 04 janvier 2024** : <https://we.tl/t-n4kRCoBW1p>

Le document textuel reprenant une transcription de la présentation PTT figure à la suite de ce PV ainsi que le document présenté par M. Oumarou concernant la présentation sur les formats.

Le lien pour télécharger le PPT de M. Oumarou sur les formats est le suivant :

<https://we.tl/t-0du3K0Tztd>

Attention ce lien expirera lui aussi **le 04 janvier 2024**.

2. ENREGISTREMENT DE LA FORMATION

1^{ère} partie :

https://us02web.zoom.us/rec/share/sF_L9D8_H4DTDDkZszlA9ulm8UKuYyVhm1QuNRh6Y7lgLCN36YTYQ1yakHVATviR.vMcrMkJWqQsy_dli?startTime=1696410221000

2^{ème} partie :

https://us02web.zoom.us/rec/share/sF_L9D8_H4DTDDkZszlA9ulm8UKuYyVhm1QuNRh6Y7lgLCN36YTYQ1yakHVATviR.vMcrMkJWqQsy_dli?startTime=1696413613000

3. PROCHAINE FORMATION

DU SON A L'EMOTION : mercredi 11 octobre 2023

HORAIRE EN GMT :

De 8H45 à 9H00 : connexion

De 09H00 à 11H00 : partie théorique

De 11H00 à 13H00 : pause

De 13H00 à 16H30 : partie pratique

Voici déjà le lien qui vous permettra de vous y connecter :

<https://us02web.zoom.us/j/3923235558?pwd=am1aSVBLNDEvTVVHRk90UGFVaXdyUT09>

CIRTEF aisbl - TITAN asbl

Formation en réalisation non linéaire et linéaire : ‘du pixel à l’image’

Ali Oumarou & Roger Roberts
04/10/2023

Colophon :

Base : Formation image V2023_05_31

PPT Formation réalisation pour le linéaire et le non-linéaire V2023_01_11

Texte : Formation en réalisation non linéaire et linéaire Part 2 Ive rro V2023_01_31

Transcription textuelle du PPT : Ive

Caractéristiques :

- Texte : Calibri corps 11
- Titre : 4 niveaux
- Sous-titres : Calibri corps 10
- Lien :

Ajout des formats film et vidéo le 31 mai

Base ddu du 6 juin

Travaux été 2023

Version finale présentée le 4 octobre

Compléments : fichier/dossier – JPEG et HDR (high dynamic Range) dans cette version textuelle.

Table des matières

Préface.....	5
1 La Lumière et la couleur :	6
1.1 Le spectre visible : les standards	6
1.1.1 Le spectre visible vs la luminosité d'un écran	7
1.1.2 La plage dynamique en photographie :	8
1.1.3 La luminosité d'un écran : OLED vs QLED :	9
1.2 L'espace couleur	9
1.2.1 Mode Couleur RVB additif (en transparence) :	11
1.2.2 Mode Couleur soustractif (en réflexion):	12
1.2.3 L' espace couleur : le 'gamut'	12
2 Les fondamentaux de l'image animée :	14
2.1 L'aspect ratio : le cinéma vs la télévision :	14
2.2 Le balayage de l'image : mode entrelacé – progressif :	14
2.3 La fréquence image (25 ou 30 images / seconde ... et plus)	15
2.4 La définition vs la résolution.....	17
2.4.1 De la définition :	17
2.4.2 De la résolution de l'image :	18
2.4.3 De l'UHD à Ultra HD !	19
2.4.4 De la profondeur de champs :	19
3 La transformation des essences : Codage – Transcodage – Décodage :	21
3.1 La numérisation :	21
3.2 Codage – transcodage – décodage :	22
3.3 L'encapsulation des données :	22
3.3.1 La définition du fichier :	23
3.3.2 La définition du dossier :	23
4 La transformation des essences (image – son -vidéo – texte):.....	25
4.1 La transformation des images le pixel (picture x element)	26
4.2 Le codage "informatique" d'un pixel dans une couleur :	26
5 Les systèmes de codage des couleurs (RVB – YUV)	28
6 Les différents modes d'échantillonnage des images (4.2.2 – 4.1.1 – 4.2.0)	29
7 La compression numérique (DCT – DWT – fractales) :	32
7.1 Le codage des images : la compression numérique :	33
7.2 Compression des images : la DCT	33
7.3 Compression des images : la DWT (Discrete Wavelet Transform).....	34

7.4	Compression des images : les fractales	36
8	Le codage des images fixes : JPEG (fruit du Joint Photographic Expert Group)	37
9	La transformation des images animées.....	39
9.1	Codage vidéo : l'échantillonnage des couleurs (rappel)	39
9.2	Codage vidéo : Compression intra – inter image.....	39
9.3	Codage vidéo : le traitement de l'AV : le MPEG 1 - 2 (Moving Picture Expert Group).....	40
9.4	Le codage vidéo : la prédiction du mouvement	41
10	Les normes de format et de compression numériques :	42
10.1	du MPEG-2 au MPEG-4.....	42
10.2	Autre format : le JPEG2K :	42
11	En développement : l'Ultra HD – HDR (High Dynamic Range)	43
11.1	UHD vs HDR :	43
11.1.1	Le HDR10 - 10+ - HR10+ Adaptative	45
11.1.2	Le format HDR Dolby Vision	46
12	Conclusion :	47

Illustrations :

Figure 1-1 : Flux lumineux lumen(Lm) – candela (Cd) – lux (Lx)	6
Figure 1-2 : le spectre de l'œil vs le capteur image	7
Figure 1-3 : de la vision humaine à l'écran	8
Figure 1-4 : La perception humaine vs la reproduction photographique.....	8
Figure 1-5 : La luminosité d'un écran : OLED vs QLED.....	9
Figure 1-6 : la décomposition de la lumière blanche à travers un prisme	10
Figure 1-7 : sensibilité de l'œil aux couleurs.....	11
Figure 1-8 : le système couleur additif	11
Figure 1-9 : le système couleur soustractif	12
Figure 1-10 : le gamut de la couleur.....	13
Figure 2-1 : le transcodage des formats film/vidéo en 16/9	14
Figure 2-2 : le balayage de l'image : mode entrelacé – progressif	15
Figure 2-3 : La fréquence image : 25 ou 30 images / seconde	16
Figure 2-4 : Evolution de la définition d'une image de télévision	17
Figure 2-5 : La taille de l'écran vs le recul du spectateur.....	18
Figure 2-6 : La profondeur de champ	20
Figure 3-1 : codage - décodage.....	21
Figure 3-2 : opto/électrique (camera) vs électrique/opto (vers écran).....	22
Figure 3-3 : le MXF (Material eXchange Format) OP ATOM	24
Figure 5-1 : le découpage image en "pixel"	26
Figure 5-1 : le codage R V B sur 8 bits	28
Figure 5-2 : le système RVB vs le système YUV.....	28
Figure 6-1 : le découpage du signal en macroblocs.....	29
Figure 6-2 : échantillonnage en 4.4.4	30
Figure 6-3 : échantillonnage en 4.2.2	30
Figure 6-4 : échantillonnage en 4.1.1	31
Figure 6-5 : échantillonnage en 4.2.0 – une image sur deux.....	31
Figure 7-1 : les principes de la compression numérique	32
Figure 7-2 : Le concept de macro-bloc : découpe en blocs de 8 pixels sur 8 pixels.	33
Figure 7-3 : DCT : Réduction de niveaux dans les hautes fréquences	34
Figure 7-4 : DWT : une transformée en ondelettes discrètes.....	35
Figure 7-5 : DWT : les zones d'intérêts	35
Figure 7-6 : une transformée en fractales	36
Figure 9-1 : un échantillonnage en 4.2.2	39
Figure 9-2 : La compression vidéo inter-image (prédiction spatiale)	40
Figure 9-3 : le codec découpe le flux en séquences de l'ordre de 12 à 16 images.....	41
Figure 9-4 : MPEG 4 : décomposition en blocs pour l'estimation de mouvement.....	41
Figure 12-1 : Segmentation des macro-blocs (plus petits ou plus grands).....	42
Figure 10-2 : un exemple de codage JPEG2K : Transformée en Ondelettes Discrète	42
Figure 11-1 : HDR : comparaison de la plage de luminosité de différents standards	44
Figure 11-2 : espace couleur : le gamut en HDR	45

Préface

Une belle image requiert de la technique, un apprentissage et une sensibilité propre à chaque regard. Derrière tout art, il y a une maîtrise technique ! L'un et l'autre se complète sans s'exclure !

Tout comme le texte, ou le son, la construction d'un espace visuel résulte d'un formalisme ! L'objet de cette formation « du pixel à l'image » est de décrire les différents processus qui structurent ce formalisme !

Dans ce cadre, ce n'est pas l'art qui constitue l'objectif, mais bien l'ensemble des « connaissances » liées aux processus technologiques et numériques en particulier !

C'est l'éclairage de l'objet photographié qui apporte de la profondeur à la scène, grâce à un subtil jeu de contrastes et de couleurs. La lumière blanche est un mélange de plusieurs lumières colorées qui peuvent se séparer ou se superposer.

L'image est par essence un véhicule privilégié de l'intensité. Notre cerveau est plus apte à traiter des données visuelles et ou sonores que textuelles. Couleurs et formes ont un impact significatif sur le message perçu.

La couleur est une perception visuelle de la lumière qui dépend de la répartition spectrale de la lumière, de la matière qui la diffuse et de l'œil qui la perçoit. L'intensité lumineuse influe sur le rendu de la couleur.

La distance entre un objet et la personne exerce une influence sur la perception de la couleur : plus cet objet est éloigné, plus le contraste diminue et moins les couleurs de l'objet sont saturées.

Module 1 : « du pixel à l'image » :

Le pixel est l'unité de base de la définition d'une image numérique matricielle adressable par un contrôleur vidéo. Ce mot provient de la locution anglaise 'picture element', qui signifie « élément d'image ». C'est l'unité utilisée pour spécifier les formats d'affichage (nombre de points x nombre de lignes).

À chaque pixel est associée une couleur, usuellement décomposée en trois composantes primaires par synthèse additive : rouge vert et bleu ...

La suite dans les pages qui suivent !

Bonne lecture.

Ali Oumarou & Roger Roberts
Septembre 2023

1 La Lumière et la couleur :

1.1 Le spectre visible : les standards

La lumière est une onde électromagnétique, un facteur physique capable de rendre visible ou de rendre plus éclairé un objet pour l'œil : sans la lumière, on ne peut pas voir. Comme le son, la lumière est une onde; mais à la différence du son, la lumière peut se propager dans le vide. La lumière parcourt environ 300 000 kilomètres par seconde.

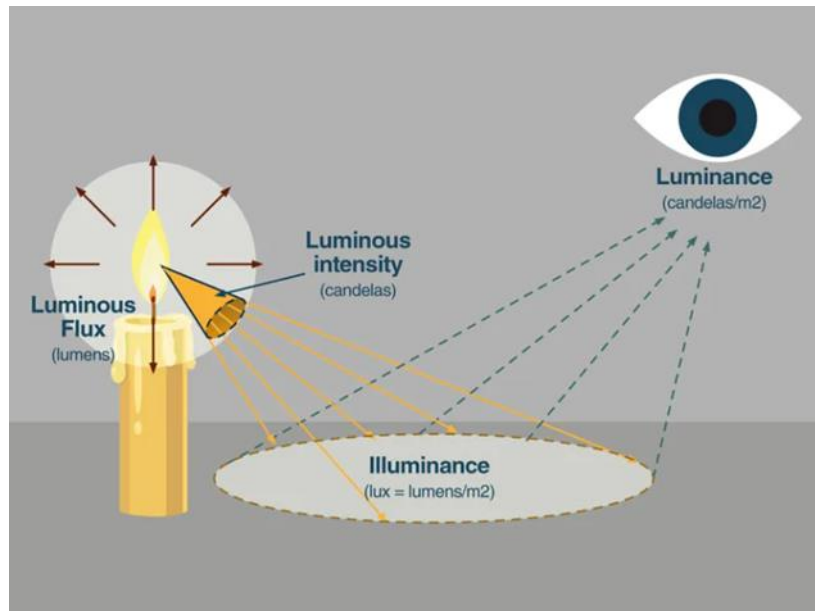


Figure 1-1 : Flux lumineux lumen(Lm) – candela (Cd) – lux (Lx)

Le lumen (symbole Lm) est une unité de mesure du flux lumineux qui mesure de la quantité totale de rayonnement visible émis par seconde. Le lumen correspond à la quantité totale de lumière visible dans un faisceau ou un angle défini, ou émise par une source : c'est la luminosité¹.

Lux (Lx) : est l'équivalent de l'énergie produite par un lumen incident sur une surface de 1 m².

Candela (Cd) : c'est l'unité de base qui mesure l'intensité lumineuse. Une candela est équivalente à 1 lumen par stéradian² (lm/sr). Elle est définie comme l'intensité lumineuse allant dans une direction spécifique, et est donc liée à l'angle d'ouverture vers la lumière.

La **luminance** s'exprime en candela/m² et est directement appréciable par l'être humain. La composante luminosité exprime l'impression de clarté, de brillance de la couleur. En colorimétrie de base, on considère que la luminosité est une perception indépendante (orthogonale) de la chromaticité.

La **chrominance**, ou chroma, est la partie du signal vidéo correspondant à l'information de couleur. Le signal vidéo nécessite une information de luminance et deux informations de chrominance afin de

¹ C'est le flux lumineux d'un rayonnement monochromatique caractérisé par une fréquence de 540x10¹² Hz et un flux d'énergie radiante de 1/683 watts.

² Stéradian : Unité d'angle solide (portion d'espace située dans un cône) correspondant à un angle solide ayant son sommet au centre d'une sphère et découpant sur la surface de cette sphère une aire égale à celle d'un carré qui aurait pour côté le rayon de la sphère (symbole sr).

reconstituer les trois couleurs primaires rouge, vert et bleu (RGB), chacune représentant la quantité de couleur primaire qu'il faut pour faire apparaître la bonne couleur à l'écran.

Nit (vient du mot latin nitere, qui signifie "briller«) est une mesure de la quantité de lumière que l'écran de télévision envoie à vos yeux (luminance) dans une zone donnée. Techniquement parlant, une NIT est la quantité de lumière émise égale à une candela par mètre carré (1 cd/m^2).

1.1.1 Le spectre visible vs la luminosité d'un écran

Le spectre visible est la partie du spectre électromagnétique qui est perceptible par l'humain.

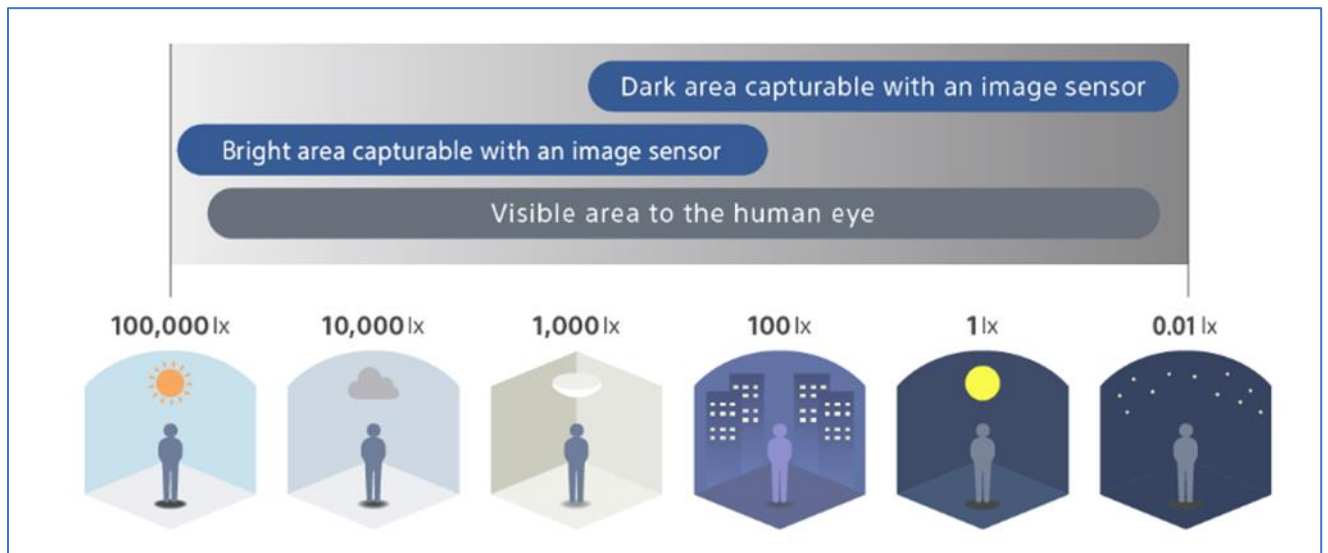


Figure 1-2 : le spectre de l'œil vs le capteur image

Comme le capteur d'image d'une caméra a une plage dynamique de photographie étroite, il ne peut pas capter la plage visible par l'œil humain. Pour un sujet lumineux, les zones sombres seront noircies, et pour un sujet sombre, les zones lumineuses deviendront 'éteintes'.

C'est la raison pour laquelle le résultat d'une photographie peut différer de celui de l'œil nu lors de capture d'une scène avec une grande différence de luminosité comme un contre-jour avec une caméra ou un smartphone.



Figure 1-3 : de la vision humaine à l'écran

L'œil humain est sensible à une très large gamme d'intensités lumineuses, mais à de faibles niveaux d'éclairement, il peut perdre la capacité de différencier certains détails. Pour cette raison, il est très important de pouvoir mesurer correctement les niveaux de luminosité en fonction du type d'écran / espace de vision.

Dans le passé, les téléviseurs avaient une luminosité maximale de 100 nits et le cinéma affichait jusqu'à 48 nits. Le degré de luminosité qu'un écran UHD-HDR peut afficher est de 108 nits.

1.1.2 La plage dynamique en photographie :

Ce que nous percevons est très différent de ce qu'un appareil photo peut reproduire ! La plage dynamique en photographie correspond à la plage de luminosité qu'un appareil est capable d'enregistrer en une seule exposition. L'œil humain peut afficher les détails à la lumière avec une plage dynamique généralement comprise entre 20 et 24 IL (Indice de Lumination) tandis que les appareils photo tournent autour de 10 IL soit 5 à 8 stops.

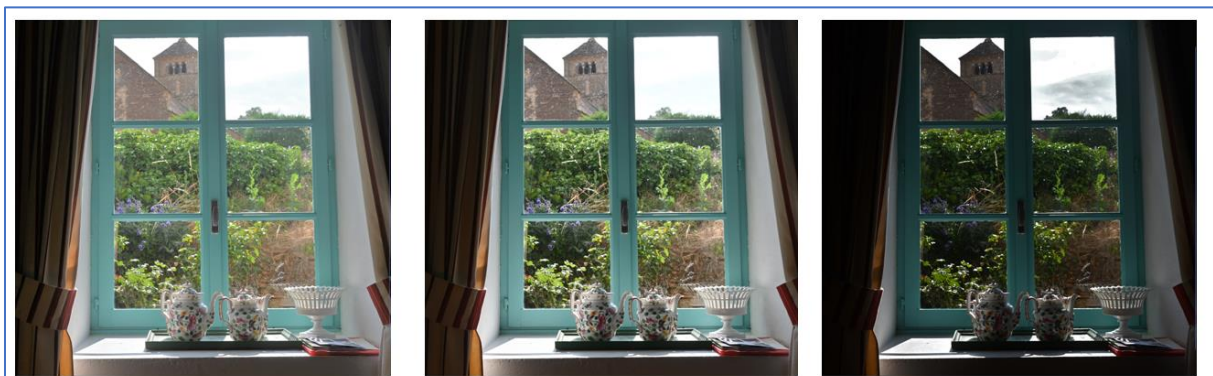


Figure 1-4 : La perception humaine vs la reproduction photographique

La plage dynamique fait référence à la capacité à inscrire des informations simultanément dans les zones d'ombres et de lumières d'une scène. Il y a clairement une zone de la scène qui reçoit plus de

lumière qu'une autre. En l'occurrence, ce sera le ciel là où toute la lumière sera présente, tandis que l'intérieur restera plus sombre.

1.1.3 La luminosité d'un écran : OLED vs QLED :



OLED (Organic Light-Emitting Diode / Diode Electro-Luminescente Organique) est une technologie utilisée dans les téléviseurs afin de produire des images à l'écran.

Lorsqu'il y a un contact avec de l'électricité, chaque diode peut s'allumer individuellement et produire une vaste gamme de couleurs.

Les téléviseurs grand public HDR récents sont basés sur la technologie de rétroéclairage actif ou OLED et ont une luminosité maximale de 500 à 1000 nits.

Figure 1-5 : La luminosité d'un écran : OLED vs QLED



QLED (Quantum dots (points quantiques) Organic Light-Emitting Diode) qui permet aux téléviseurs QLED de se distinguer des téléviseurs classiques. Les points quantiques sont des matériaux semi-conducteurs ultra-fins de l'ordre du nanomètre. Ces points produisent différentes couleurs lumineuses en fonction de la taille de la particule : plus la taille est grande, plus la couleur est rouge, et plus la taille est petite, plus la couleur est bleue.

1.2 L'espace couleur

En 1666, lors de ses expériences d'optique, Isaac Newton fait passer de la lumière solaire blanche à travers un prisme de façon à la décomposer en rayons lumineux de différentes couleurs. Newton est le premier à comprendre que la lumière blanche est faite d'un mélange de rayons de lumière de couleurs différentes.

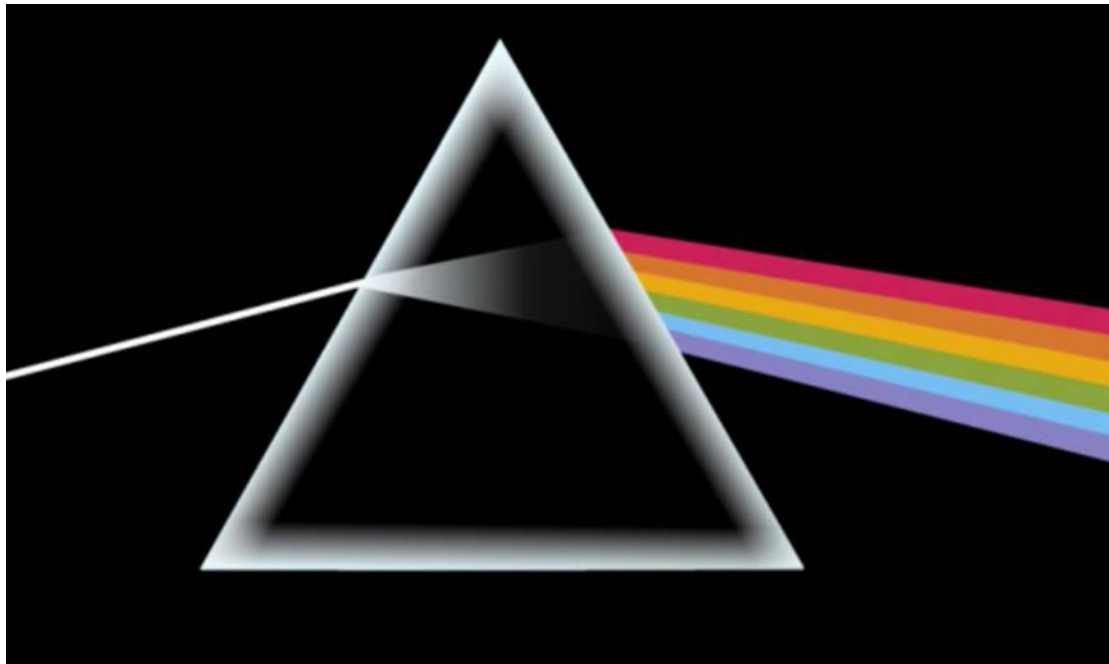


Figure 1-6 : la décomposition de la lumière blanche à travers un prisme

Le spectre visible est la partie du spectre électromagnétique qui est perceptible par l'œil humain. Il se décompose en rayonnements monochromatiques par le passage de la lumière à travers un dispositif disperser tel qu'un prisme ou un réseau diffractant.

L'œil distingue bien toutes les couleurs du spectre visible, du violet au rouge de manière continue. C'est par le mélange de ces trois "couleurs" que le cerveau est capable de "fabriquer", de reconstituer, donc de percevoir toutes les "autres" couleurs.

Voilà pourquoi on a inventé un modèle couleur R(ouge) V(ert) B(leu) : pour nous rapprocher du fonctionnement de l'œil humain. Ainsi, les écrans ou les capteurs des appareils numériques possèdent des pixels Rouge-Vert-Bleu ou RVB et ne traitent que des signaux codés RVB !

Ils se rapprochent du fonctionnement l'œil humain grâce à trois "types" de pixels. Pour l'écran il suffit d'éclairer plus ou moins chaque pixel pour reproduire TOUTES les couleurs visibles !

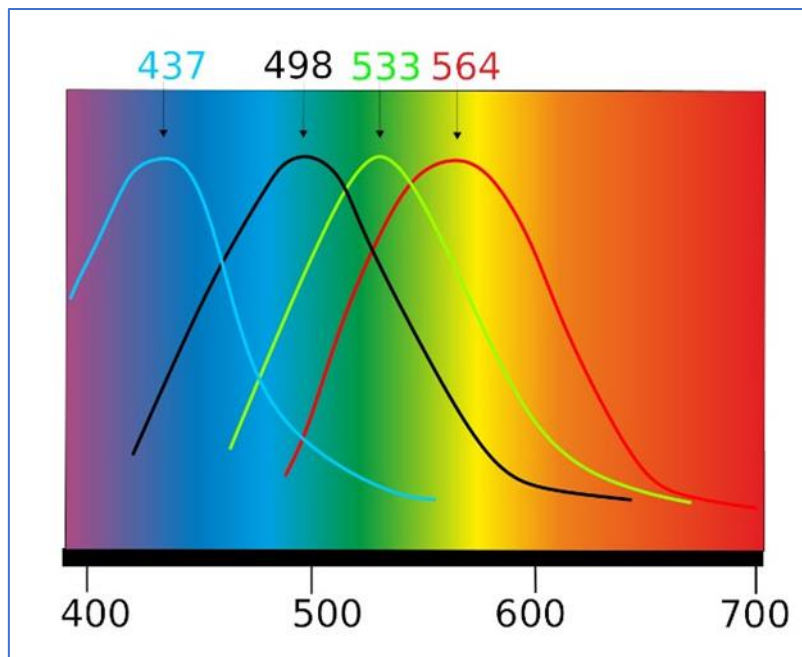


Figure 1-7 : sensibilité de l'œil aux couleurs

Les pixels RVB sont séparés mais ils sont fabriqués suffisamment petits pour que, vus par un être humain à distance normale, ils soient vus comme amalgamés en un seul pixel aux multiples couleurs.

Un moniteur ou une imprimante ne travaillent pas avec les couleurs de la même manière :

- l'écran travaille avec des couleurs dites additives RVB pour donner du blanc par addition,
- alors qu'une imprimante travaille en mode soustractif (Cyan, Magenta, Jaune, Noir) pour donner du noir par soustraction.

1.2.1 Mode Couleur RVB additif (en transparence) :

Mode Couleur RVB additif : sur un écran en mode additif, avec un signal aux pixels ($R : 0, V : 0, B : 0$), l'écran reste noir. Pour obtenir du blanc (255, 255, 255) il faut éclairer chaque pixel Rouge, Vert et Bleu, au maximum. Les autres couleurs sont obtenues en modifiant la valeur de chaque couleur primaire. Pour afficher du jaune par exemple ($R : 255, V : 255, B : 0$), l'écran est bien jaune à l'intersection entre le vert et le rouge car les pixels bleus sont éteints.

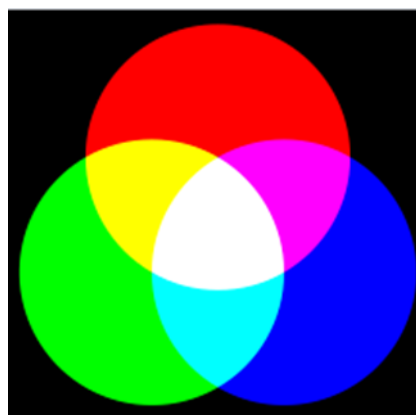


Figure 1-8 : le système couleur additif

Seules trois couleurs ne nécessitent qu'un seul pixel. Les trois couleurs primaires et leurs déclinaisons en luminosité (de 0 à 255). Toutes les autres, c'est-à-dire les millions d'autres, seront un mélange d'au moins deux pixels de couleurs primaires plus ou moins lumineux.

L'œil humain fonctionne comme un écran, c'est-à-dire sur un mode additif.

1.2.2 Mode Couleur soustractif (en réflexion):

Mode Couleur soustractif : avec une imprimante, on part d'une feuille de papier déjà blanche par défaut car elle réfléchit à parts égales toutes les longueurs d'onde de la lumière visible. Donc là, c'est le contraire, si l'on veut du noir il va falloir projeter sur cette feuille de l'encre aux couleurs complémentaires, c'est-à-dire Cyan, Magenta et Jaune qui ont la particularité d'absorber à chaque fois une partie de la lumière visible.



Figure 1-9 : le système couleur soustractif

Ces encres projetées au même endroit vont avoir la propriété d'absorber toutes les radiations lumineuses. Cet endroit apparaîtra donc noir maintenant. Dans la pratique il faudra aussi rajouter de l'encre noire car 100% des trois autres couleurs ne permettent pas d'obtenir un noir suffisamment profond. Si l'on veut que l'impression sur le papier paraisse bleue, il faut projeter au même endroit une combinaison de magenta et de cyan. Toutes les radiations lumineuses, sauf le bleu, seront absorbées. Seules les longueurs d'onde bleues de la lumière seront réfléchies à cet endroit.

1.2.3 L' espace couleur : le 'gamut'

Pour restituer la perception humaine colorée d'une image, différentes méthodes existent, en peinture, en imprimerie, en projection photographique, vidéo ou en affichage électronique et notamment, grâce au principe de la trichromie.

Le 'gamut', c'est l'ensemble des couleurs qu'un écran ou une imprimante, peut reproduire parmi le spectre visible par l'œil humain. Si un œil exercé est effectivement capable de distinguer jusqu'à 200 nuances d'une même couleur primaire, ce même œil n'atteint qu'exceptionnellement cette prouesse dans les deux autres couleurs primaires.

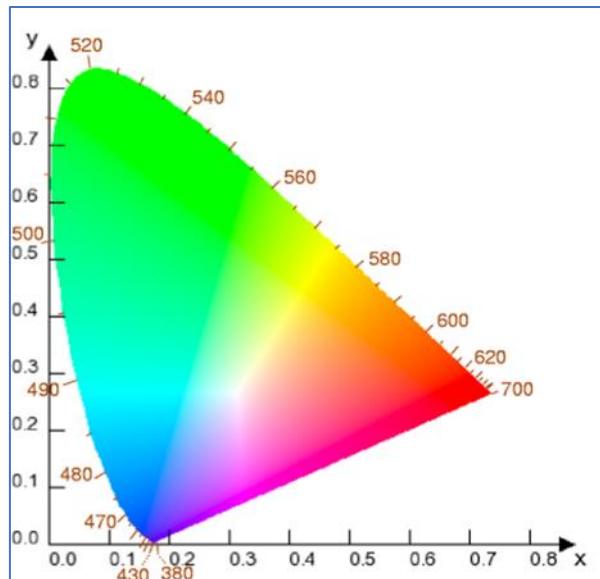


Figure 1-10 : le gamut de la couleur

Ce principe est notamment exploité par un téléviseur, un écran vidéo ou d'ordinateur, lequel reproduit la couleur par synthèse additive, à partir des trois couleurs primaires : rouge, vert et bleu.

À l'écran, chaque élément constitutif de l'image vidéo, cellule ou élément luminophore est associé à une valeur de ces couleurs primaires. Chaque élément utilise notamment une intensité lumineuse, dévolue à chacune des trois couleurs ; le codage RVB permet de définir la valeur de la luminance de chacun des « points » en affichage analogique ou « pixels », en affichage numérique, composant l'image vidéo.

2 Les fondamentaux de l'image animée :

2.1 L'aspect ratio : le cinéma vs la télévision :

Début du 20ème siècle, tant au cinéma qu'en télévision, la forme de l'écran était un rectangle de quatre unités de large sur trois unités de haut.

- dans le monde du cinéma, cette forme était exprimée sous la forme d'un rapport, 1,33:1 (4 divisé par 3) ;
- dans le monde de la vidéo, la même chose était exprimée sous la forme d'un rapport de 4:3.

Début des années 50, avec l'apparition de la télévision, le cinéma a voulu offrir une expérience plus riche en projetant des films au format scope ou anamorphique, des images deux fois plus larges que celles qui étaient produites par le passé. Certains films « CinemaScope » sont tournés à 2,39:1. D'où le recours à des bandes noires latérales ou verticales pour respecter le format natif !

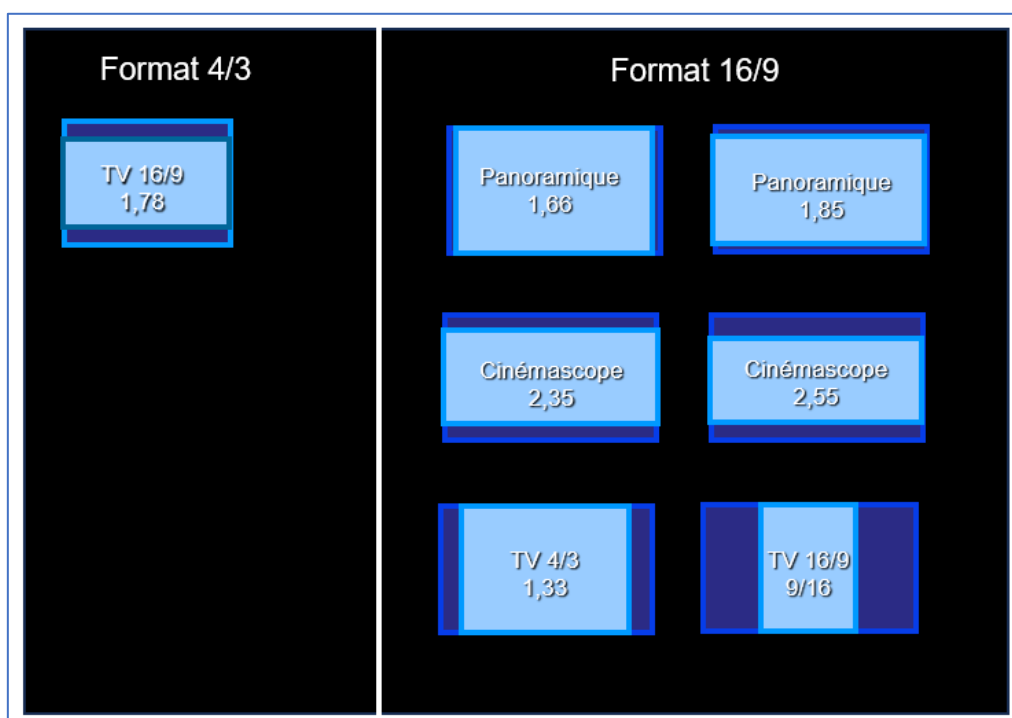


Figure 2-1 : le transcoding des formats film/vidéo en 16/9

En vidéo HD, écran large signifie 16:9, ce qui équivaut à 1,78:1, soit un peu moins que la norme cinématographique de 1,85 pour la même hauteur d'écran. Le choix d'un format d'image détermine des caractéristiques figées ! Un transcoding image ne fait que gérer l'inclusion d'un format dans un autre (avec barres noires) : un rectangle dans le carré (format 4/3) ou le carré dans le rectangle (format 16/9).

Aujourd'hui d'autres formats sont largement utilisés comme le format vertical 9:16 fort répandu dans le monde des réseaux sociaux car la façon naturelle de prendre un smartphone est généralement verticale. Ce n'est pas le format le plus approprié pour une diffusion sur un écran 16/9 ! Comme pour la conversion film / vidéo, il faut recourir à des bandes noires ou des effets miroirs.

2.2 Le balayage de l'image : mode entrelacé – progressif :

Le balayage entrelacé est l'unique mode d'affichage en télévision, alors que le mode progressif commence à apparaître à la fin des années 1970, pour l'univers de l'affichage vidéo informatique.



Figure 2-2 : le balayage de l'image : mode entrelacé – progressif

Le balayage d'une image consiste à afficher les lignes successives qui la composent sur un écran. Il existe deux types de balayage : progressif (format P) et entrelacé (format i pour 'interlaced') :

- le balayage progressif affiche toutes les lignes successivement, ce qui génère une image plus nette et plus fluide, surtout pour les scènes en mouvement ;
- le balayage entrelacé (un standard issu des systèmes de télévision analogique) affiche deux demi-images ou trames, l'une avec les lignes impaires et l'autre avec les lignes paires, en alternance. Cela permet de réduire le scintillement des images à faible fréquence de balayage. Cependant, le balayage entrelacé peut créer des effets de peigne ou de décalage lorsque l'image est en mouvement. L'image apparaît donc légèrement lignée dans le sens horizontal. Il s'agit du mode d'affichage le plus courant en télévision standard.
- Sources entrelacées : postes de télévision classiques : émissions télévisées, événements sportifs, concerts.
- Sources progressives : films issus du cinéma sur pellicule (TV, cassette, DVD), productions réalisées avec une caméra progressive. Le lecteur DVD sépare donc les images progressives du disque en deux trames qui seront affichées successivement. La perception visuelle est donc similaire à celle d'une source entrelacée, avec hélas le même lignage horizontal.

Lorsqu'un téléviseur progressif reçoit un signal vidéo entrelacé, il assemble les deux trames, lignes paires et impaires, pour un affichage progressif. Cet assemblage est hasardeux. En effet, d'une trame à l'autre, le sujet filmé a généralement bougé. Des déchirures ou des stries apparaissent une fois les deux trames jointes. Certains lecteurs de DVD disposent de prises DVI pour se connecter aux téléviseurs et projecteurs dotés d'une telle interface.

Le transcodage d'une vidéo codée en mode entrelacé (une trame paire, une impaire) en progressif (pour un affichage sur un écran informatique) est susceptible de générer un effet de peigne !

2.3 La fréquence image (25 ou 30 images / seconde ... et plus)

La fréquence d'images en télévision est le nombre d'images affichées par seconde. Elle peut varier selon les normes et les pays.

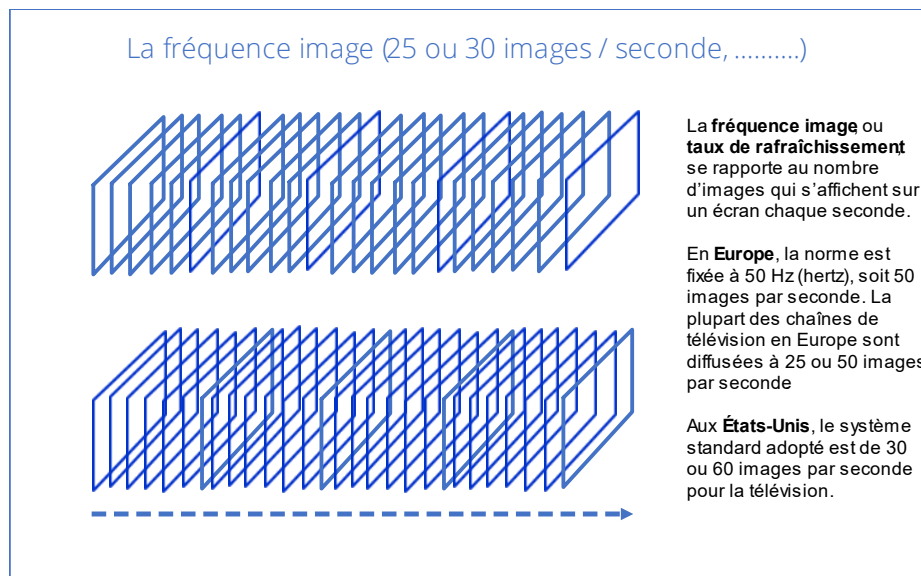


Figure 2-3 : La fréquence image : 25 ou 30 images / seconde

- En Europe, la fréquence d'images est généralement de 25 images par seconde pour la télévision analogique et numérique. **En Afrique ????**
- En Amérique du Nord, la fréquence d'images est plutôt de 30 images par seconde pour la télévision numérique.

La fréquence d'images peut aussi dépendre du type de contenu. Par exemple, les émissions sportives ou en direct peuvent utiliser une fréquence plus élevée pour rendre les mouvements.

Depuis le numérique et surtout depuis les écrans plats, qui sont prévus pour être multi-cadences, les écrans sont aujourd'hui prévus pour diffuser plusieurs cadences en s'adaptant automatiquement :

- **24p** : la fréquence standardisée historique du cinéma sur pellicule. Un écran plat qui affiche cette cadence va alors diffuser un taux de rafraîchissement multiple de cette cadence : 48 Hz, 72 Hz ou 120 Hz2.
- **25i et 50i** : fréquence d'affichage utilisées en Europe en mode entrelacé à 1080i, standardisé pour des motifs d'économie de bande passante vidéo.
- **25p et 50p** : fréquence d'enregistrement vidéo utilisées en Europe par les chaînes TV et caméscopes. Un écran plat qui affiche ces cadences va alors diffuser un taux de rafraîchissement multiple de ces cadences : 50 Hz ou 100 Hz.
- **30p et 60p** : fréquence d'enregistrement vidéo utilisées notamment aux États-Unis et au Japon par les chaînes TV et caméscopes. Un écran plat qui affiche ces cadences va alors diffuser un taux de rafraîchissement multiple de ces cadences : 60 Hz ou 120 Hz. Ces cadences 30p et 60p sont en fait aussi très présentes sur les appareils multimédia (héritiers du 525 lignes). Dans tous les pays, la plupart des écrans de PC portables, smartphones et tablettes fonctionnent par défaut en mode 60 Hz, que les smartphones enregistrent les vidéos dans une cadence de 30p ou 60p, et que les consoles de jeux vidéo proposent une cadence de 30p ou 60p.

La technologie de l'entrelacement va probablement disparaître au profit de fréquences images plus élevées progressives (High Frame Rate) : notamment le 50p (50 images entières par seconde) et le 60p (60 images entières par seconde).

Cette modification permettra de faire disparaître de nombreux défauts comme l’inversion des sens des roues d’une voiture de course³ ! Plus la fréquence des images est faible, plus il y a de chances que la roue ait fait un peu moins qu’un tour complet entre deux images. À 50 images par seconde, la roue doit tourner deux fois plus vite pour produire le même effet. En général, plus la fréquence des images est élevée, plus le mouvement est fluide et réaliste.

2.4 La définition vs la résolution

2.4.1 De la définition :

Le format d’une image est calculé par le nombre de pixels X le nombre de ligne qui la compose.

La télévision à définition standard possède le plus petit nombre de pixels des formats de diffusion. Héritage de l’ère analogique, il existe deux types de définitions standards utilisés dans différentes parties du monde :

- Format “NTSC”, l’image définition standard comporte 480 lignes horizontales de 720 pixels
- Format “PAL”, l’image définition standard comporte 576 lignes horizontales de 720 pixels.

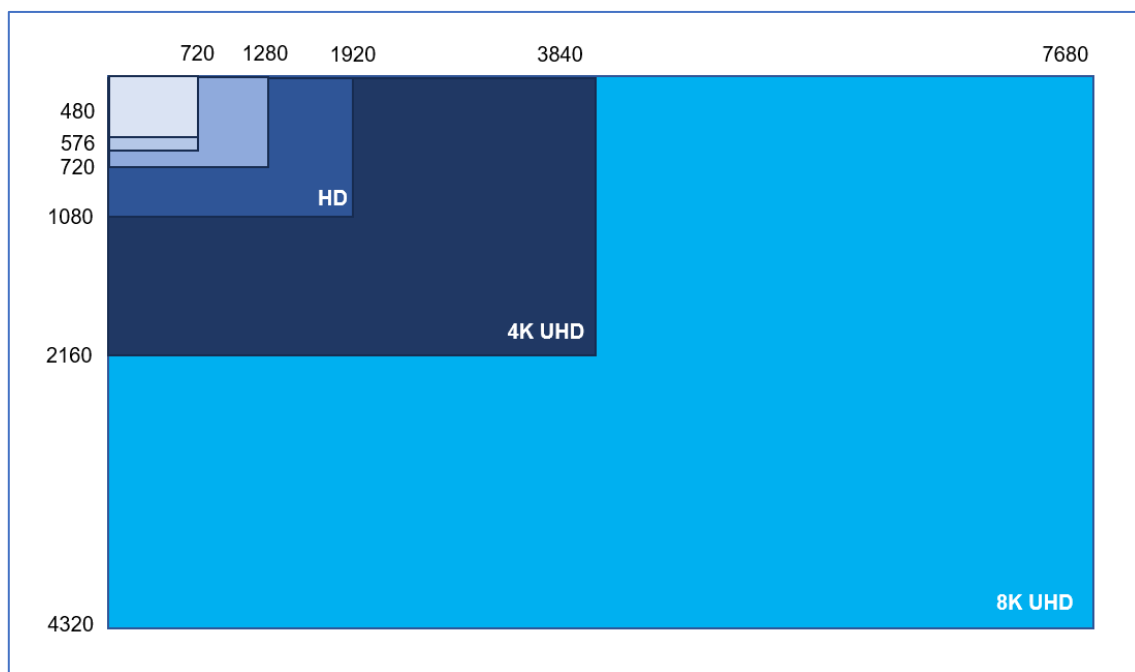


Figure 2-4 : Evolution de la définition d’une image de télévision

Il existe également deux formats de HD :

- le format HD 720p de 1280 pixels x 720 lignes en mode progressif, soit un peu moins de 1 Mpx (mégapixels) ;
- le format Full HD de 1920 pixels x 1080 lignes horizontales en mode entrelacé. Ce format a un total d’environ 2 Mpx (soit deux mégapixels).

³ Les roues de Formule 1 tournent très vite, parfois plus de 1000 tours par minute. La caméra, elle, capture des images à une fréquence fixe, par exemple 25 images par seconde. Si la roue a fait un peu moins qu’un tour complet entre deux images, alors les images successives semblent restituer un mouvement à l’envers. Cet effet est plus visible quand la vitesse de la roue est proche d’un multiple de la fréquence de la caméra.

Ces deux formats sont des standards mondiaux équivalents en termes de perception (du fait d'un balayage progressif pour l'un et entrelacé pour l'autre) sont utilisés à la fois dans les anciens pays NTSC et les anciens pays PAL ».

Le format UHD (Ultra Haute Définition) est un format de vidéo numérique qui offre une image plus détaillée que le format HD (Haute Définition). Il existe deux types de UHD :

- le 4K : 3840 pixels x 2160 lignes : quatre cadres Full HD de 1920 x 1080 réunis » ;
- le 8K : 7680 pixels x 4320 lignes ;

qui ont respectivement 4 fois et 16 fois plus de pixels que le HD.

Un des avantages d'enregistrer en 4K et d'éditer en HD, est la possibilité de pouvoir recadrer et zoomer dans l'image sans perte de qualité. Cela nous permet aussi de réaliser certains effets de mouvement de caméra en post production comme les "pans", sans perte de qualité et de pouvoir générer des graphismes en finesse.

De nombreuses salles de cinéma sont équipées de projecteurs numériques qui peuvent afficher une image de 2K (2048 x 1080), 4K (4096 x 2160) ou même 8K (7680 x 4320).

Mais il reste de la marge : la limite de définition de l'œil humain est de 576 millions de pixels ! Si on compare avec les 12 mégapixels de l'iPhone 11... l'œil humain a une définition 48 fois supérieure.

2.4.2 De la résolution de l'image :

La résolution de l'image de télévision est le nombre de pixels affichés à l'écran.

- Les **DPI** (dots per inch, soit « points par pouce » en anglais) indiquent le nombre de points se trouvant sur une ligne d'un pouce (environ 2,5cm) lors d'un scan ou d'une impression.
- Pour les moniteurs, et les écrans, on parle de **PPI** (pixels per inch, soit pixels par pouce).

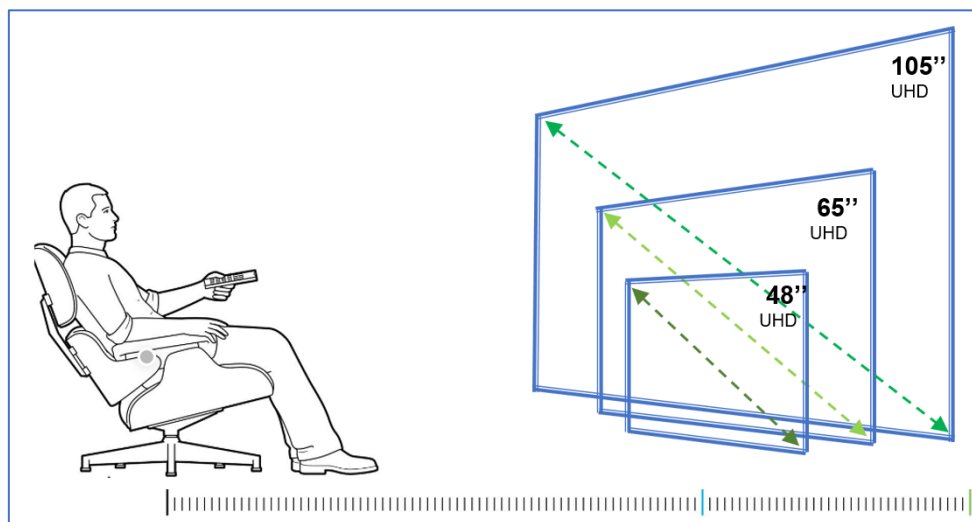


Figure 2-5 : La taille de l'écran vs le recul du spectateur

Plus la résolution est élevée, plus l'image est nette et détaillée. Il faut aussi tenir compte de la taille de l'écran, de la distance de visionnage, du format d'image et du contenu.

La résolution d'écran permet d'augmenter la taille du téléviseur sans augmenter le recul du spectateur. En effet, plus la résolution est élevée, moins les pixels sont visibles et plus l'image est fluide. Un téléviseur UHD 4K ne nécessite pas plus de recul qu'un Téléviseur HD⁴.

2.4.3 De l'UHD à Ultra HD !

L'ère de la Full HD 1080p est en train de laisser place progressivement à la 4K pour le cinéma (4096 x 2160 pixels et un ratio de 17/9) et l'UHD (Ultra HD : 3840 x 2160 pixels) qui offre une définition et un niveau de détails nettement supérieur.

La définition d'un écran n'est pas une valeur absolue en soi, elle n'a de sens qu'en combinaison avec la taille de l'image et la luminosité de la salle TV. Il est évident qu'il faut plus de points sur un téléviseur qui affiche 1 mètre de diagonale pour obtenir la même densité que sur un smartphone de 15 centimètres de diagonale. CQFD. Plus l'écran est grand, plus un point sera gros à définition égale et plus il se verra.

Reste à se poser la question suivante : à partir de quelle taille le passage à l'Ultra HD fait sens ? Généralement, on situe cette taille autour de 80 cm à 1 mètre de diagonale pour un téléviseur. Or les tailles d'écran ne font qu'augmenter depuis l'avènement du LCD et du LED qui permet de créer des modèles de plus en plus fins et quasiment sans bord, sans parler d'un prix au centimètre, en chute libre aussi. D'où la tentation de passer à des grandes tailles dans le salon, 140 cm ou plus, qui elles nécessitent à coup sûr l'Ultra HD.

L'Ultra HD permet de s'approcher sans être gêné. Si en Full HD, un écran de 1 mètre exige un recul de 2,5 mètres environ, en Ultra HD on peut passer à 140 cm à la même distance. Evidemment, plus on est proche et plus la sensation d'immersion va augmenter, comme au cinéma.

Tous les films et toutes les séries un peu spectaculaires par leurs mouvements de caméra, l'action, les paysages... prennent une autre dimension quand les bords de l'image s'effacent du champ de vision. Désormais avec un grand écran et l'Ultra HD, l'immersion à la maison est proche de celle au cinéma. L'exemple le plus flagrant sera toute fois le jeu vidéo. Avec une console qui diffuse en Ultra HD et un grand écran idoine pas trop loin, on pilote vraiment une Formule 1.

Pour profiter pleinement de l'effet « woow », il faut que le film ou la série ait été tourné et surtout diffusé en Ultra HD. C'est seulement le cas via Internet avec Netflix ou avec un lecteur Blu-ray UHD.

2.4.4 De la profondeur de champs :

La profondeur de champ au cinéma est la zone de l'image où les éléments au centre de l'image sont nets, tandis que l'avant-plan et l'arrière-plan sont flous. La profondeur de champ est définie par la focale de l'objectif, l'ouverture du diaphragme, la distance de mise au point et la définition de l'image. La profondeur de champ est une technique d'écriture utilisée dans la dramaturgie pour créer des effets esthétiques ou dramatiques : une faible profondeur de champ peut isoler un personnage ou un objet du décor, tandis qu'une grande profondeur de champ montrer l'espace et les relations entre les éléments filmés.

⁴ Il existe des formules pour calculer la distance de recul idéale en fonction de la taille et de la résolution de l'écran. Par exemple, pour un téléviseur UHD 4K de 55 pouces (140 cm), la distance de recul recommandée est de 1,5 mètre environ.

Aujourd'hui la profondeur de l'image est utilisée par les cinéastes qui usent abondamment de très longues focales, qui aplatissent la perspective et privilégient un seul objet ou personnage central. Celui-ci est paradoxalement mis en évidence par le flou de l'arrière-plan.



Figure 2-6 : La profondeur de champ

Il existe quelques techniques très efficaces, et pas forcément si compliquées pour donner plus de profondeur de champ à un plan :

- l'éclairage : soit utiliser le contre-jour pour détacher l'objet central de l'arrière-plan ou au contraire renforcer l'éclairage de l'arrière-plan ;
- le diaphragme : une faible profondeur de champ s'obtient avec une grande ouverture du diaphragme, et une grande profondeur de champ s'obtient avec une petite ouverture du diaphragme ;
- la mise au point qui peut varier en fonction du script ;
- les effets visuels, notamment les CGI (Computer Generated Imagery) et les fonds verts, permettent de renforcer la profondeur de champ en post-production.

3 La transformation des essences : Codage – Transcodage – Décodage :

3.1 La numérisation :

La **numérisation** est un processus qui transforme des données textuelles/graphiques, iconiques, sonores/musicales, audiovisuelles, 3D, ...) ou un signal électrique en données numériques (fichiers/dossiers) que des dispositifs informatiques numériques pourront traiter, rendre accessible et exploiter en ligne ! La numérisation traite exclusivement l'aspect « représentation » des documents.

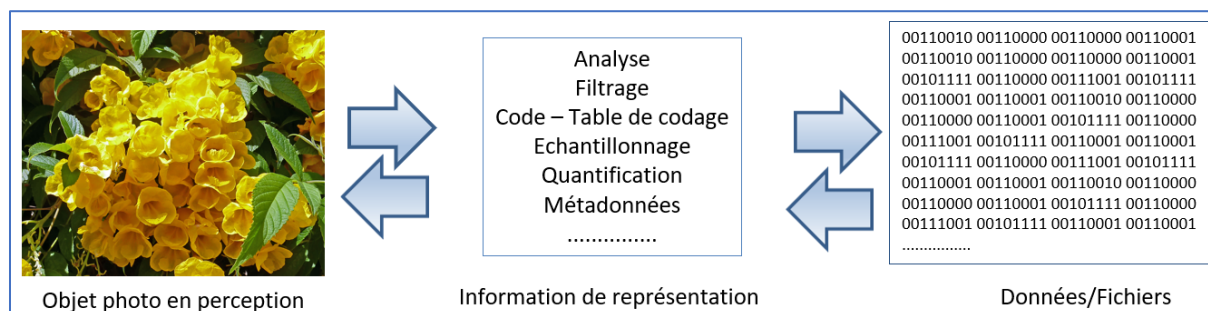


Figure 3-1 : codage - décodage

La numérisation implique l'usage d'un **[codec]** (COdeur-DECodeur qui transforme le signal en fichier et le fichier en signal en vue de représenter l'objet en perception et de codes de correction d'erreur en vue de pouvoir recréer l'objet avec des bits manquant, un fichier corrompu ou corriger des erreurs de transmission (redondance de 15% pour un CD audio, 50% pour le format BETA NUM de chez SONY).

Enfin la numérisation génère le "code génétique" d'un document. La connaissance de cette suite de valeurs, permet de reproduire un document à l'identique, le concept d'original s'évanouit à condition de ne pas coder avec pertes ([compression numérique]).

Le Codage : Un codage des caractères définit une manière de représenter les caractères (lettres, chiffres, symboles) dans un système informatique. Il existe différents types de codage :

- Le codage de source, qui permet de faire de la compression de données.
- le codage de canal, qui permet une représentation des données de façon à être résistant aux erreurs de transmission.
- Le codage visuel, qui permet une représentation des données en schémas 2D : code-barres ou QRcode par exemple.

Le Codage peut être défini :

- sans perte (ZIP, TAR, WAV, RAW, NEF, DPX, ...)
- Ou encore avec perte (MP3, JPEG, MPEG, HEVC, ...) mais en fonction de paramètres appliqués avec ou sans perception de cette perte d'information.

Le Transcodage : la traduction d'une information dans un code différent. Il ne s'agit pas d'un codage au sens strict du terme car le plus souvent le transcodage génère des pertes lors du recodage.

Le Décodage : le processus inverse s'appelle décodage afin de remettre l'objet codé préalablement dans une forme techniquement comparable à l'original. (Exemple : la carte son qui reconstitue dans un mode synchrone la forme d'onde d'origine pour l'envoyer vers les haut-parleurs en tant que signal analogique).

3.2 Codage – transcodage – décodage :

Tous les processus de communication partagent trois éléments de base: la source (émetteur), un support d'émission (fichier - canal de messagerie) et la cible (récepteur). Le processus ci-dessous code le signal dans un format fichier, ce dernier sert alors de référence au processus de décodage qui extrait les données en fonction du périphérique d'affichage :

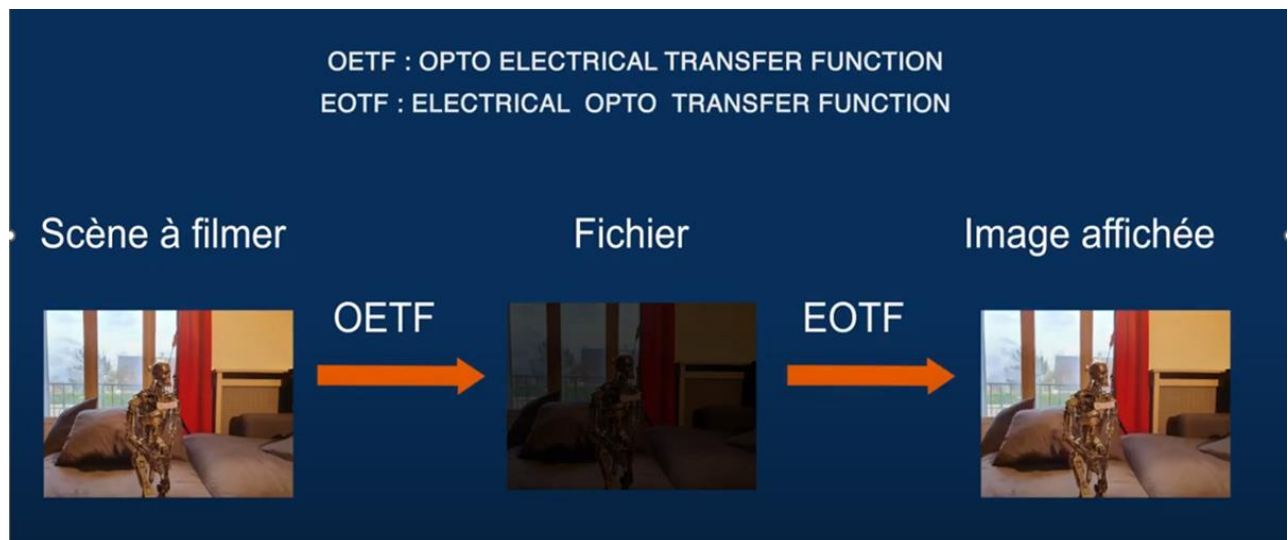


Figure 3-2 : opto/électrique (camera) vs électrique/opto (vers écran)

- **OETF** (fonction de transfert **optoélectronique**) est la fonction de transfert ayant la lumière de la scène en entrée et la convertissant en fichier image ou vidéo en sortie. Cela se fait généralement dans une caméra.
- **EOTF** (fonction de transfert **électro-optique**) est la fonction de transfert ayant le fichier image ou vidéo en entrée et le convertissant en sortie de lumière linéaire de l'écran. Cela se fait dans le dispositif d'affichage.

Le transcodage est souvent synonyme de pertes (notamment destructeur pour les métadonnées ancillaires). Un transcodage d'un fichier généré avec perte ne peut régénérer le format d'origine.

3.3 L'encapsulation des données :

Les applications informatiques sont des programmes qui permettent de créer, modifier et stocker des fichiers et des dossiers sur un ordinateur :

Les fichiers peuvent être de différents types tels que des fichiers exécutables, compressés, texte, documents, images, audio ou vidéo.

Les dossiers sont utilisés pour organiser les fichiers en catégories et sous-catégories.

L'encapsulation de données est le processus qui ajoute aux données des informations d'en-tête de protocole supplémentaires avant leur transmission. Pour les encapsulations, on utilise en général des termes comme wrapping ou boxing qui décrivent le concept "d'emballage" d'un élément relativement "primitif".

Un [Wrapper] (ou container) est un format conteneur qui peut contenir divers types de données (et parfois compressées à l'aide de codecs normalisés).

Les données enveloppées à chaque couche portent le nom de PDU (Protocol Data Unit) et contiennent deux choses : la donnée en elle-même et l'en-tête spécifique à cette couche ainsi que les

en-têtes des couches précédentes. L'encapsulation permet de définir des droits d'accès aux données : publique, privée ou protégée.

Le [ZIP] est un format de fichier « archive » qui peut contenir un ou plusieurs fichiers ou répertoires qui prend en charge la compression de données sans perte. Ce format de fichier peut poser des problèmes d'interopérabilité au niveau des noms de fichiers (caractères autres que des caractères latins non accentués - ASCII).

3.3.1 La définition du fichier :

Un fichier informatique est l'unité de stockage de données communes à tout ordinateur. Toutes les applications génèrent des données structurées dans un fichier fermé, et l'OS de l'ordinateur écrit, lit, stocke, importe et exporte des fichiers numériques. Un fichier informatique constitue donc un ensemble de données sous format binaire numériques réunies sous un même label, enregistrées sur un support de stockage permanent (mémoire de masse : un disque dur, un CD-ROM, une mémoire flash ou une bande magnétique), et manipulées comme un tout.

La syntaxe des titres des fichiers comporte un label (généré par l'utilisateur) et une "extension" courte figurant après le point séparateur :

- le titre (label) proprement dit : les lettres de A à Z et les chiffres de 0 à 9
- l'extension facultative (en général 3 caractères) qui synthétise l'application utilisée.

Les différents formats de fichiers :

Les formats de fichiers décrivent l'organisation et dépendent des processus de codage/décodage à appliquer aux données du dit fichier pour la relecture.

Les fichiers sont souvent désignés en fonction du codage (et donc du décodage). Par exemple, *.jpg est une image JPEG, *.doc est un fichier texte codé par l'application traitement de texte « Word » de Microsoft et windows7.exe est une application exécutable dans l'OS Windows.

3.3.2 La définition du dossier :

Chaque outil dispose d'une zone de stockage, intitulée [Dossier], dans laquelle sont rangés les différents Fichiers que l'on va trouver sur un périphérique numérique. Ces dossiers, ou "répertoires", sont créés sur le disque dur lorsque le système d'exploitation et les applications sont installés. Pour ce faire, il existe un dossier racine qui est à la source de tous les autres (arborescence), symbolisé par le signe \ (backslash sous Windows ou / (slash) sous Linux, et automatiquement créé lors du formatage du disque dur de stockage.

Les dossiers sont également parfois appelés "répertoires" et sont créés sur le disque dur lorsque le système d'exploitation et les applications sont installés. Par exemple, le bureau de l'utilisateur est en réalité un dossier. Certains fichiers sont vraiment des dossiers. Bien qu'ils ne soient pas identifiés comme tels, un seul fichier peut en réalité être un dossier.

Un dossier contient un ou plusieurs fichiers (ou peut être vide) avec juste un titre. Les dossiers fournissent une méthode pour classer, organiser et trier les fichiers similaires à un dossier Manila (une chemise cartonnée) contenant des documents papier dans un classeur.

L'encapsulation des données en mode fichier/dossier structure l'interface graphique d'un PC et organise l'enregistrement des données. Il existe d'autres formes d'encapsulation comme le ZIP (qui diminue la taille), METS, MXF, ... et bien sûr les bases de données.

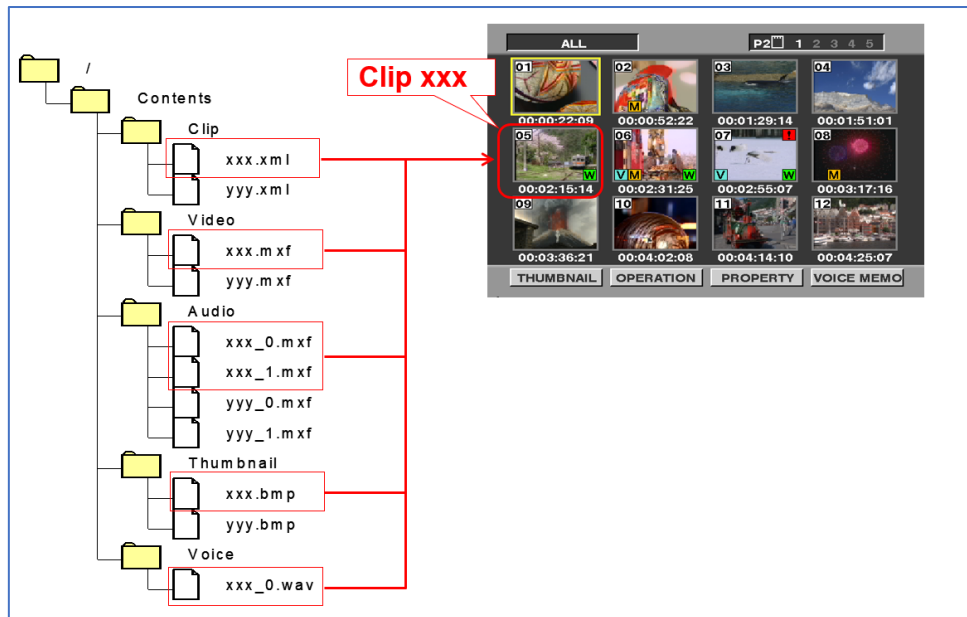


Figure 3-3 : le MXF (Material eXchange Format) OP ATOM

Un exemple d'encapsulation de « Fichiers/Dossiers » issu du monde de l'audiovisuel : le MXF (Material eXchange Format) OP ATOM est un conteneur défini par la SMPTE et utilisé par l'industrie audiovisuelle pour structurer les données (fichiers vidéo, audio, images, ...). D'autres OP « modèles opérationnels » peuvent contenir ou référencer plusieurs matériaux, tout comme une EdL (Edit list : chronologie d'un programme de montage vidéo).

4 La transformation des essences (image – son-vidéo – texte):

La transformation de l'audio, de la photographie ou de la vidéo comporte en général au moins trois phases :

- une **phase d'échantillonnage** où des caractéristiques d'un document (par exemple, le temps pour l'audio, l'espace pour la photo, ...) sont explorées à intervalles temps régulières ;
- une **phase de quantification** qui attribue une valeur représentant l'objet aux points d'échantillonnage (résolution pour le son, pixel pour l'image) ;
- une **phase de numérisation** qui confère à chacune de ces valeurs une valeur numérique (par exemple sous forme binaire) similaire au système de codage du texte !

L'ensemble de ces traitements effectués par un appareil de transformation numérique (Numériseur ou digitaliseur) constitue un fichier de données qui va stocker l'ensemble des données numérisées relative à un document, un objet, un son, une vidéo (image + son), ... sur la base d'un [titre de fichier].

La transformation binaire (transformation des signes interprétables par l'ordinateur) est la primitive du langage de l'ordinateur et le format de fichier (typé e.g. *.txt, *.docx, *.Jpg, *.mp4, *.Wav, ...) une forme d'organisation du fichier en fonction de codage d'un objet (texte/ASCII, image/JPEG, video/MPEG,...).

Le premier scanner numérique date de 1957 ! Mais il faut attendre la fin des années 80 et le développement du CCD (Charged Coupled Device) afin de disposer d'un capteur photographique. Ce composant électronique photosensible convertit un flux lumineux en un signal électrique analogique. Ce signal est ensuite amplifié, puis échantillonné et quantifié et enfin numérisé par un convertisseur analogique-numérique en vue d'obtenir une représentation numérique de l'ensemble des points (Pixels) de l'image et une quantification.

Le signal de base issu du CCD est fragmenté en points appelés « pixels », abréviation de « picture » (image) et de « el(ement) » (« élément »). L'ensemble des pixels, reliés entre eux, constituent la définition de l'image. Plus il y a de pixels, en horizontal et vertical, meilleure est la résolution de l'image en fonction de la taille de l'écran.

Un pixel est calculé par l'assemblage électronique de trois "points": un rouge + un vert + un bleu. La couleur est produite par l'addition de ces 3 couleurs à des intensités variables, ce qui permet de produire toute une palette de couleurs différentes.

Le premier appareil photo numérique aurait été développé par KODAK (mais qui n'a jamais développé le concept)⁵

Le premier appareil photo numérique Model 1 a été présenté en 1991 par la société Dycam lors du CeBIT à Hanovre. A l'époque les prises de vue étaient en noir et blanc avec une résolution minimaliste de 376x284, mais cela générait déjà un volume de ... par photo.

⁵ Voir : https://fr.wikipedia.org/wiki/Appareil_photographique_numerique

4.1 La transformation des images le pixel (picture x element)

Si l'œil humain est capable de distinguer de nombreuses nuances pour une couleur donnée, globalement il faut tout de même une différence significative entre deux nuances d'une même couleur pour que, sans ambiguïté, cette personne puisse affirmer qu'il a bien perçue deux couleurs différentes !

Sur cette base d'un nombre de 200 nuances perçues par couleur (soit : 200 x 200 x 200) cela fait 8 millions de couleurs différentes ! Dans la réalité on en est fort éloigné car les experts estiment que 300.000 nuances peuvent suffire pour des petits écrans !

La numérisation d'une image découpe la résolution en une suite de blocs et d'instructions permettant de la coder et de la reconstituer. Ces informations de représentations sont soit des données (basée sur sensations visuelles) pour chaque élément (pixels définis pour chaque couleur fondamentale en RGB, codée en YUV), soit avoir une organisation plus complexe (images comprimées, images vectorielles,).

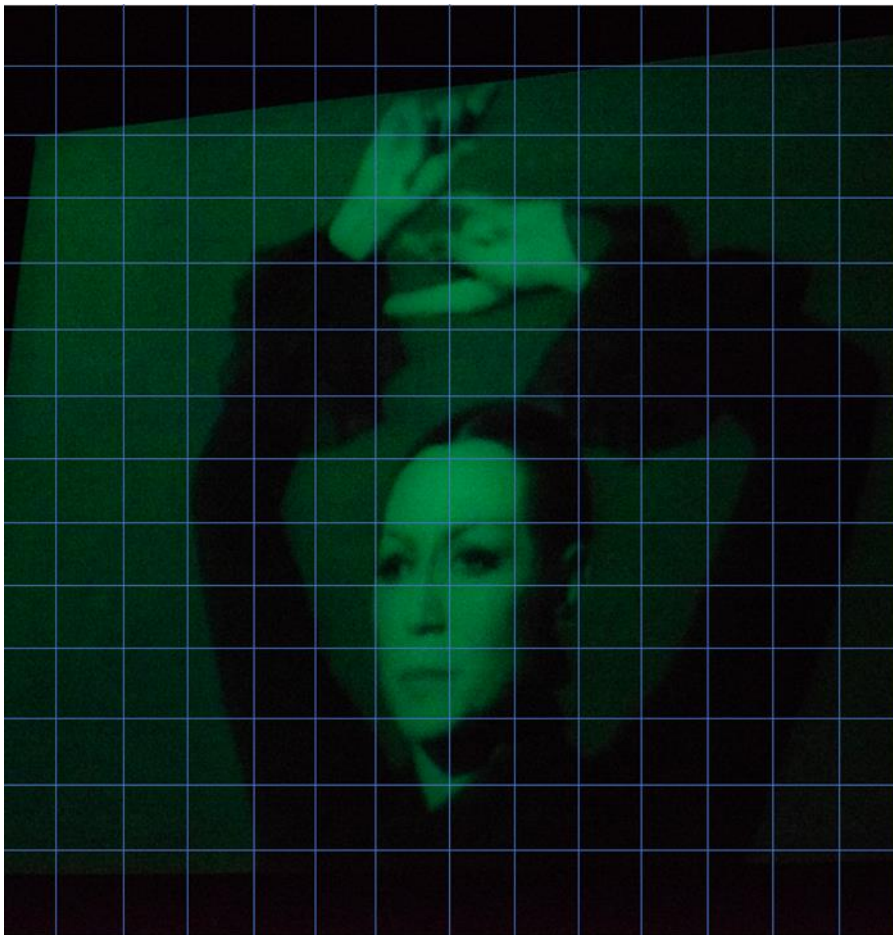


Figure 4-1 : le découpage image en "pixel"

4.2 Le codage "informatique" d'un pixel dans une couleur :

Chaque couleur est décrite sur la base d'une valeur (de 0 à 256), toujours avec comme objectif de reproduire le fonctionnement de l'œil humain. Ainsi une couleur aura comme valeur :

- R 158, V 128, B 84 - en RVB additif
- et C - 80%, M - 64%, J - 21%, N - 12% - en soustractif.

La définition d'une image représente sa « dimension informatique ». Une image avec 1024 pixels/ligne et 768 en hauteur aura une définition de 1024 pixels par 768 (proche d'une résolution d'écrans).

La résolution détermine le nombre de DPI (points par unité de surface, soit en anglais Dots Per Inch).

Une **table de codage** attribue à chaque pixel une valeur en fonction de la palette (N/B, 256 niveaux de gris, RGB ou YUV sur 8 bits, ...°). Il est également possible de rajouter une information de transparence (avec un codage en 32 bits). Une image couleur de 1024 pixels par 768, aura un volume de : $1024 \times 768 \times 3 \times 256 = 603.979.976$ bits, soit 75.400 octets !

Avec le temps des quantifications sur 10 ou 12 bits sont devenues courantes !

Le **codage d'une image noir / blanc** est aisé ! Pour scanner un texte, le pixel ne peut uniquement prendre que les valeurs noir ou blanc et donc 0 et 1.

Le **codage de la couleur** est réalisé sur trois octets, afin de représenter la gradation de la valeur d'une composante couleur de 0 à 256. Ces trois valeurs codent généralement la couleur dans l'espace RVB.

Le nombre de couleurs différentes pouvant être ainsi représenté est de $256 \times 256 \times 256$ possibilités, soit environ 16,7 millions de couleurs.

Valeur de transparence : dans les logiciels d'édition d'image, dans l'incrustation vidéo, il est intéressant de séparer les images et les fonds. C'est également important pour pouvoir représenter des images en réflexions ou en transmission lumineuse. Pour que ce soit possible, il faut indiquer, à chaque pixel, le degré de transparence.

L'attribution d'une valeur de transparence, sur un octet, permet tous les mélanges. Cette valeur est généralement appelée α (alpha). Le rendu d'un pixel s'obtient par multiplication et addition, récursive, des valeurs des couches des codages sur 24 bits (Paint : 256 valeurs $\times$ 4) ou sur 32 bits (pour les logiciels Gimp, Photoshop).

Un fichier **RAW** (de l'anglais RAW signifiant 'brut') est un fichier généré par un appareil numérique et qui contient l'ensemble des données générées par le capteur, brutes de traitement, plus des informations complémentaires selon la marque et le modèle de l'appareil.

Le logiciel de développement offre le choix des paramètres à appliquer aux données brutes issues du capteur CCD/CMOS qui assure transformation des signaux lumineux en signaux électriques (luminosité, contraste, couleurs, ...) avant une exportation sous format fichier *.tiff, *.jpeg, *.png, ...

5 Les systèmes de codage des couleurs (RVB – YUV)

RVB (Rouge, Vert, Bleu) désigne un système de traitement optique, d'affichage électronique ou d'un codage de signal vidéo analogique ou un codage informatique des couleurs.

Il existe deux systèmes de codage des couleurs : primaires RVB (Red Green Blue en anglais) et YUV basé sur les trois longueurs d'ondes auxquelles l'œil humain est réceptif : surtout le vert, puis le rouge, et enfin le bleu.

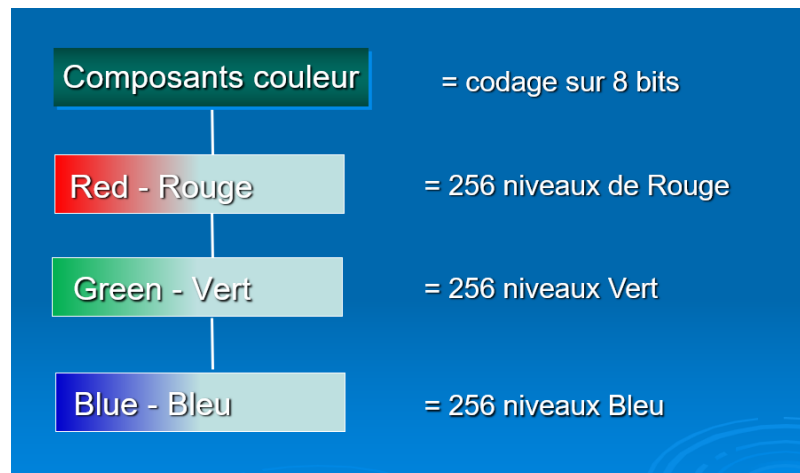


Figure 5-1 : le codage R V B sur 8 bits

RVB : Dans le système RVB, on utilise la synthèse additive des trois couleurs : l'addition des trois couleurs donne du blanc, le rouge et le vert donnent du jaune, le vert et le bleu du cyan et le bleu et le rouge du magenta. La couleur résultante dépend de l'intensité de chaque couleur primaire. Ce système RVB est utilisé par les capteurs CCD et Cmos qui équipent les appareils photo et les caméscopes numériques.

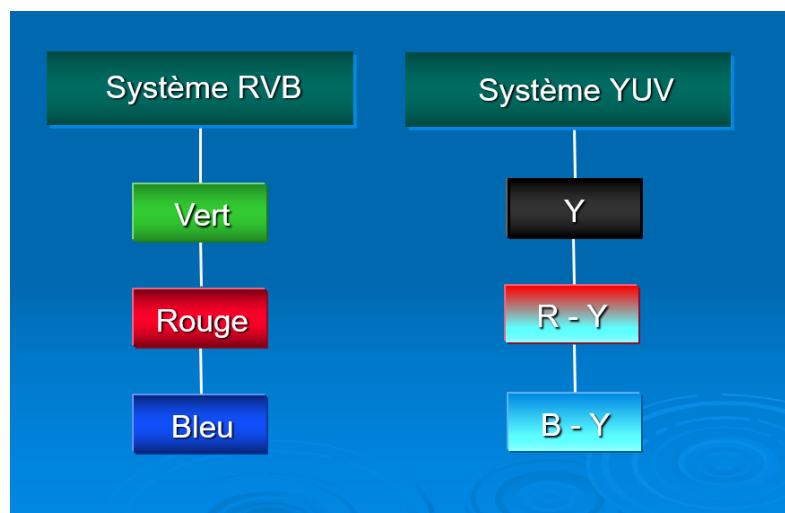


Figure 5-2 : le système RVB vs le système YUV

YUV : le système de codage YUV est créé depuis une source RVB. Il est codé en trois composantes : la première (Y) représente la luminance (informations de luminosité) tandis que les deux autres (U et V) sont des données de chrominance (informations de couleur). Chaque composant est formaté en fonction d'une table de codage. Exemple : la luminance (Y) est égale à l'addition de $0,59 V + 0,30 R + 0,11 B$ (soit 60 % de vert).

A l'origine ce codage YUV permettait d'économiser de la bande passante par rapport au codage RVB (qui transportait les informations de luminance pour chaque canal de couleur) ainsi que d'assurer la compatibilité entre les différents récepteurs analogiques de la télévision. Durant les recherches sur les technologies numériques il est apparu que ce modèle colorimétrique permettait en plus de réduire la taille d'une image via une compression. Comme l'œil humain est plus sensible à la luminance (définition) qu'à la chrominance, ce codage permet de dégrader plus fortement la chrominance d'une image tout en assurant un niveau de qualité.

6 Les différents modes d'échantillonnage des images (4.2.2 – 4.1.1 – 4.2.0)

L'unité de codage de base d'une image est le macrobloc. Les macroblocs de chaque image sont codés successivement, de gauche à droite, et de haut en bas.

Chaque macrobloc comporte 8 x 8 pixels soit en R, en V, ou en B. Chaque macrobloc est converti :

- un bloc de luminance Y
- un bloc de chrominance R - Y
- et un bloc de chrominance B - Y

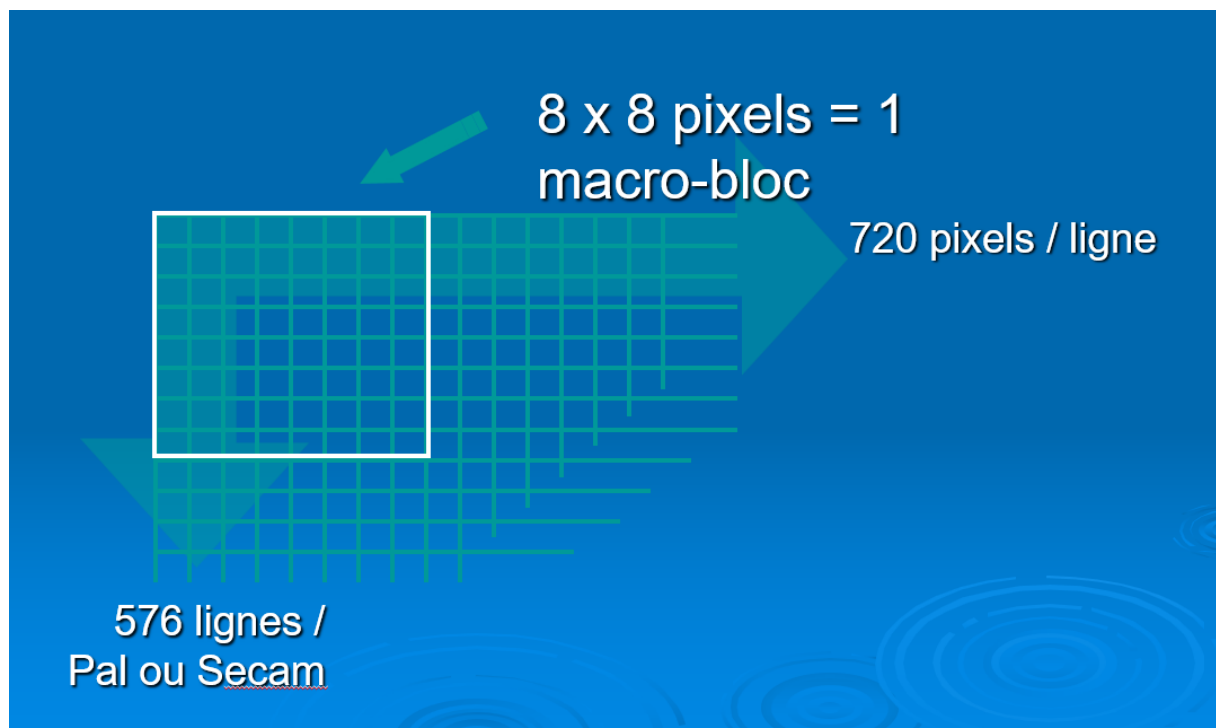


Figure 6-1 : le découpage du signal en macroblocs

Une Image Vidéo numérique TV SD = 6.480 Macro-Blocs

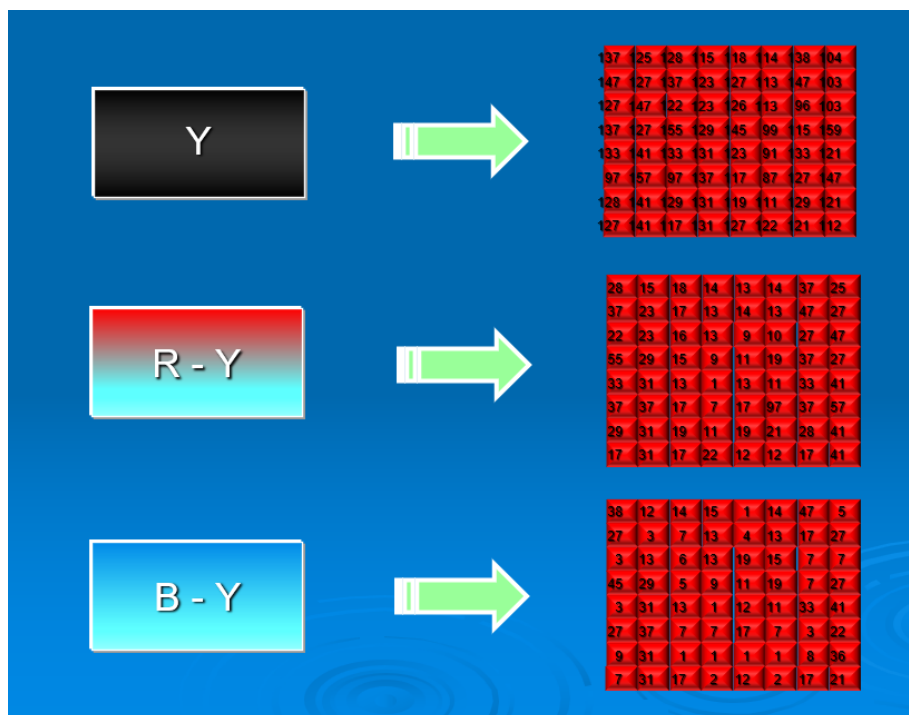


Figure 6-2 : échantillonnage en 4.4.4

L'échantillonnage 4.4.4 est principalement adapté pour générateurs graphiques, les projections cinéma en numérique.

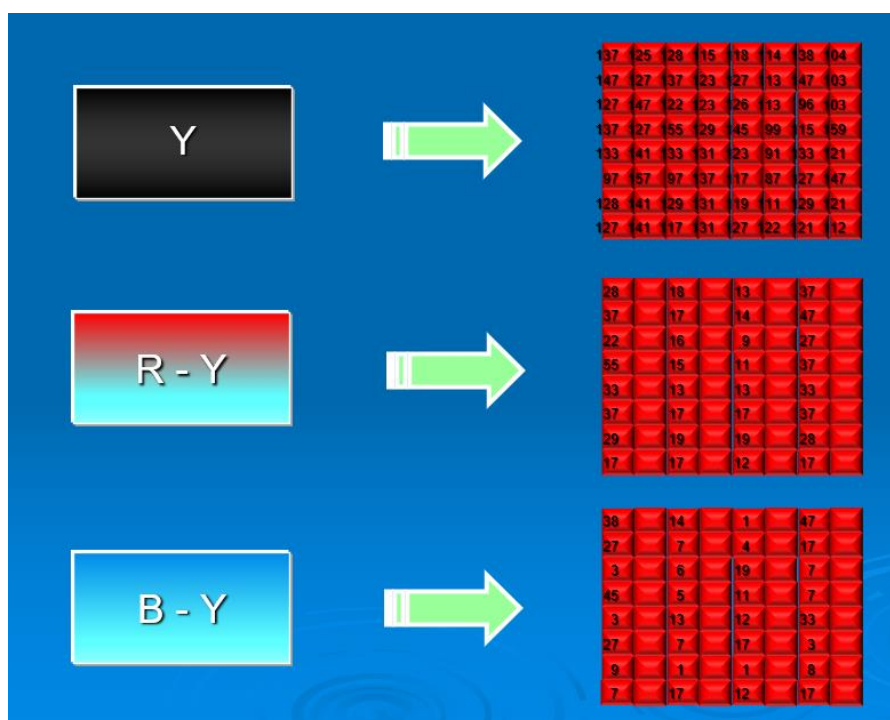


Figure 6-3 : échantillonnage en 4.2.2

4.2.2. : sur un échantillon de 4 pixels : 4 échantillons luminance, mais seulement 2 échantillons chrominance (B – Y) et R - Y) tous les deux pixels. C'est la norme mondiale de production pour des images de télévision : le format D1, Betacam numérique, DVC-Pro 50,

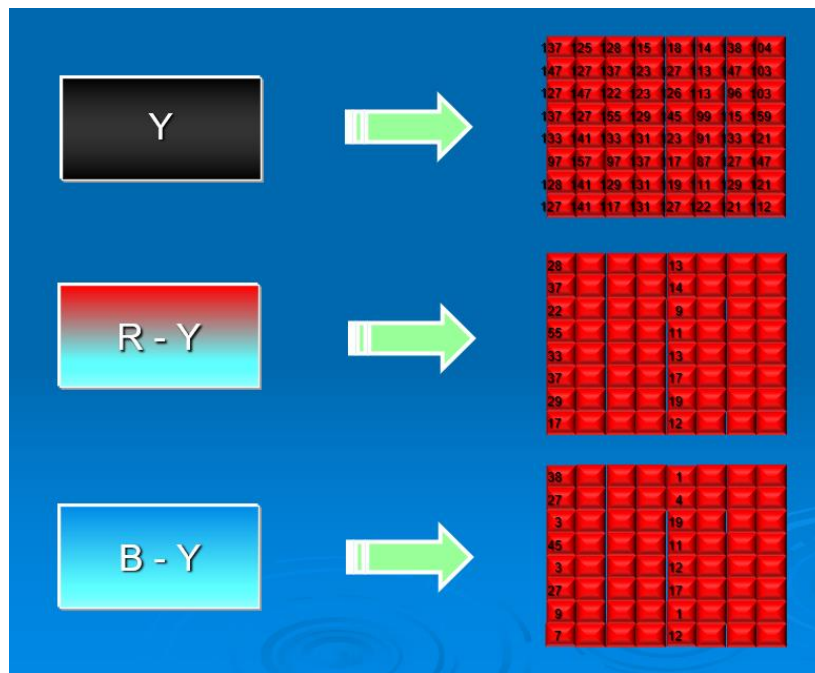


Figure 6-4 : échantillonnage en 4.1.1

4.1.1. : sur un échantillon de 4 pixels : 4 échantillons luminance, mais seulement 2 échantillons chrominance (B – Y et R – Y) tous les quatre pixels. C'est le format DVC-Pro 25, ...

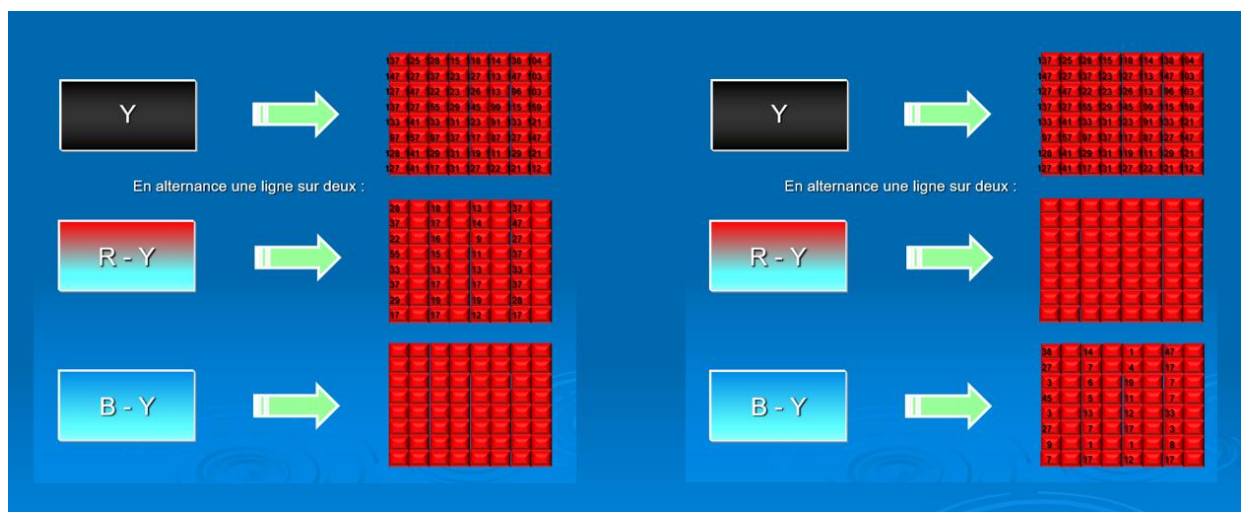


Figure 6-5 : échantillonnage en 4.2.0 – une image sur deux

4.2.0. : sur un échantillon de 4 pixels : 4 échantillons luminance, mais seulement 2 échantillons chrominance (B – Y ou R – Y en fonction de la ligne) tous les deux pixels. C'est le format du SECAM, du Dv, du Dv-Cam, le codec H264, de toutes les vidéos en streaming, sur les téléphones mobiles, en usage "courant".

Il y a donc deux fois plus d'informations de chrominance sur une image en 4.4.4 qu'en 4.2.2.

En 4.2.2 l'image perd de l'information colorimétrique et du détail dans les nuances cependant l'image conserve une très haute qualité visuelle et un poids plus léger.

7 La compression numérique (DCT – DWT – fractales) :

La vidéo regroupe l'ensemble des techniques, technologie, permettant l'enregistrement ainsi que la restitution d'images animées, accompagnées ou non d'un son (synchrone ou non synchrone), sur différents types de support et de formats.

Les données d'un signal vidéo peuvent être divisées en trois parties :

- les éléments redondants (répétitifs),
- les éléments irrelevantes (visuellement subjectifs),
- les éléments de base (éléments essentiels).

La quantité d'éléments redondants et irrelevantes dans un signal vidéo dépend de la complexité de la scène. La complexité signifie :

- le nombre de détails fins (information spatiale),
- le type et la vitesse du mouvement (information temporelle).

Le pourcentage d'éléments redondants et irrelevantes est donc fonction du contenu (complexité des images et du temps (nombre d'images par secondes) :

- traitement Image par Image (Intra image) : redondance spatiale,
- traitement inter images : redondance temporelle.

Objectif de la [compression numérique] : réduire le volume/débit en comprimant fortement le poids des essences à stocker ou à transmettre. Exploitation de :

- redondances spatiales : intra image
- redondances temporelles : inter images
- redondances subjectives : vision de détails
- redondances statistiques : codages entropiques
- Base de connaissances

But : réduire le débit et conserver "autant que possible" la qualité des essences (textes, images, sons, images animées) originales.

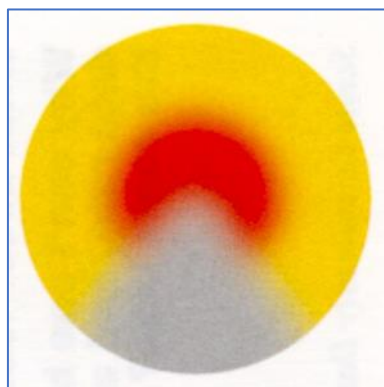


Figure 7-1 : les principes de la compression numérique

La compression (avec ou sans pertes) vise à réduire les éléments redondants et non pertinents du signal vidéo numérique. Afin de « contrôler » le débit la quantification applique des coefficients afin de réduire les écarts de niveaux dans les hautes fréquences où l'œil est moins sensible.

Cette réorganisation s'effectue dans une matrice par ordre croissant, les éléments « clés » en rouge au centre, les éléments jugés redondant en jaune et les détails les plus fins étant situés en bas (en gris).

7.1 Le codage des images : la compression numérique :

A titre d'exemple, pour réduire le volume de données à stocker ou à transporter, le codage numérique des images JPEG utilise le principe de la transformation en cosinus discrète (DCT).

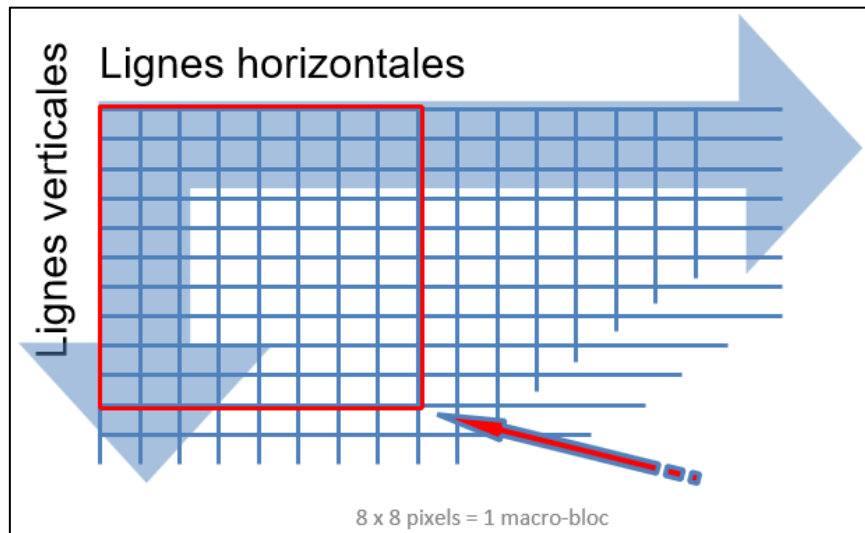


Figure 7-2 : Le concept de macro-bloc : découpe en blocs de 8 pixels sur 8 pixels.

Afin de mettre en évidence des zones à forte redondance spatiale, l'image est découpée en blocs de 8 pixels sur 8 pixels.

Ce concept de macro-bloc est la base de calcul pour la DCT.

La taille de 8 X 8 pixels était un bon compromis dans les années 90 car elle faisait appel à des puissances de calcul raisonnables.

7.2 Compression des images : la DCT

Afin de procéder à une étude de la redondance spatiale (la similarité entre les pixels adjacents aussi bien en horizontal qu'en vertical) la transformée en cosinus discrète DCT permet de passer d'une représentation spatiale de l'information (chaque nombre représentant la valeur de l'intensité d'un pixel] à une représentation fréquentielle de l'information (un coefficient représentant des variations d'intensité lumineuses ou colorimétriques (via la table de codage Y'CbCr qui est utilisé pour les images JPEG).

À chacun de ces macro-blocs est ensuite appliquée cette transformation qui remplace les données des 8 x 8 pixels par 64 autres coefficients ! On s'aperçoit ainsi que les 64 pixels de base ne portent pas tous une information de même importance :

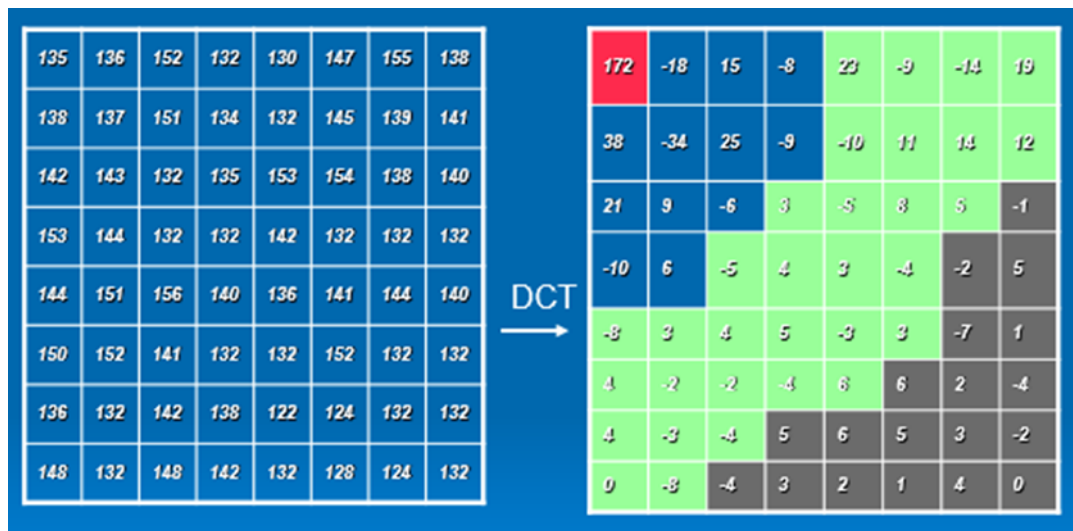


Figure 7-3 : DCT : Réduction de niveaux dans les hautes fréquences

La DCT applique des coefficients afin de réduire les écarts de niveaux dans les hautes fréquences où l'œil est moins sensible.

Cette réorganisation s'effectue par ordre croissant, les détails les plus fins étant situés en bas et à droite de la matrice, la première valeur en haut à gauche, représentant la valeur calculée de la matrice.

Les pixels de base en haut à gauche sont essentiels, et ne pas les prendre en compte se traduirait par une nette différence entre un macro-bloc source et un macro-bloc reconstitué. En revanche, négliger les valeurs qui figurent en bas à droite n'influe pas énormément sur l'aspect du macro-bloc reconstitué.

La compression JPEG ne sélectionne qu'une partie bien choisie de l'information traduite par les coefficients de la DCT. Ce sont des critères d'efficacité qui déterminent le fait que certains coefficients sont en partie ou complètement négligés et qui génèrent une compression numérique avec plus ou moins de pertes (surtout en chrominance).

7.3 Compression des images : la DWT (Discrete Wavelet Transform)

D'autres normes de codage sont apparues qui elles sont intra uniquement (JPEG2K ou JPEG-XS) basés sur une autre technologie DWT = décorrélation intra-composant : Concentre l'énergie de l'image de façon hiérarchique :

- Pas d'artefacts de type DCT avec des taux de compression élevés
- Permet des représentation multi-résolution de l'image

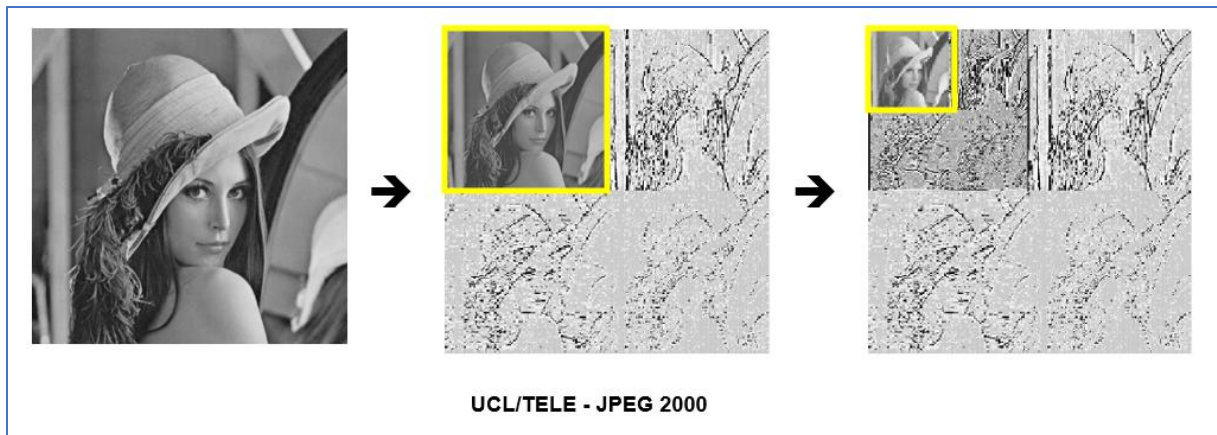


Figure 7-4 : DWT : une transformée en ondelettes discrètes

Un exemple de la transformée en ondelettes discrètes 2D utilisée dans JPEG2000. L'image d'origine est filtrée passe-haut, produisant les trois grandes images, chacune décrivant les changements locaux de luminosité (détails) dans l'image d'origine. Il est ensuite filtré passe-bas et réduit, donnant une image d'approximation; cette image est filtrée passe-haut pour produire les trois images de détail plus petites et filtrées passe-bas pour produire l'image d'approximation finale dans le coin supérieur gauche.

En analyse numérique et en analyse fonctionnelle , une transformée en ondelettes discrètes (DWT) est toute transformée en ondelettes pour laquelle les ondelettes sont échantillonnées de manière discrète. Un avantage clé qu'elle présente par rapport aux transformées de Fourier est la résolution temporelle : elle capture à la fois les informations de fréquence et de localisation (emplacement dans le temps).

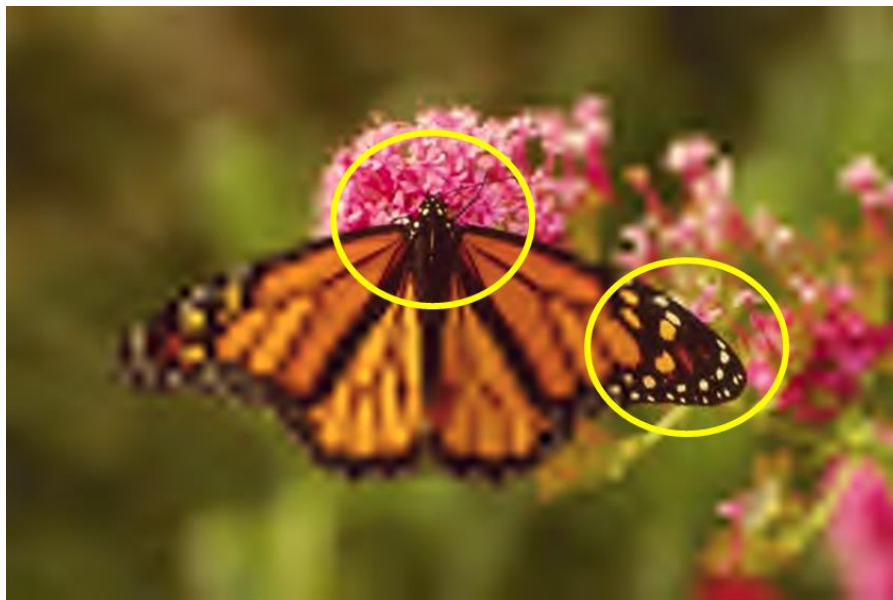


Figure 7-5 : DWT : les zones d'intérêts

Le codec cible « des zones d'intérêt » qui offrent plusieurs mécanismes pour prendre en charge l'accès aléatoire spatial ou l'accès à une zone d'intérêt à divers degrés de granularité.

Ces normes présentent une caractéristique qui permet un décodage hiérarchique (en fonction de l'application sélectionnée on n'est pas obligé de tout décoder par exemple si la qualité de reproduction n'est pas importante).

7.4 Compression des images : les fractales

La compression fractale est une méthode de compression avec perte pour les images numériques, basée sur les fractales. La méthode est la mieux adaptée aux textures et aux images naturelles, en s'appuyant sur le fait que des parties d'une image ressemblent souvent à d'autres parties de la même image. La compression fractale consiste donc à remplacer l'image tout entière par une série de formules mathématiques qui permettent de la recomposer dans son entièreté. Le succès de l'opération de compression est par conséquent proportionnel à l'importance des propriétés géométriques de l'image.

Les algorithmes fractals convertissent ces parties en données mathématiques appelées "codes fractals" qui sont utilisées pour recréer l'image codée.

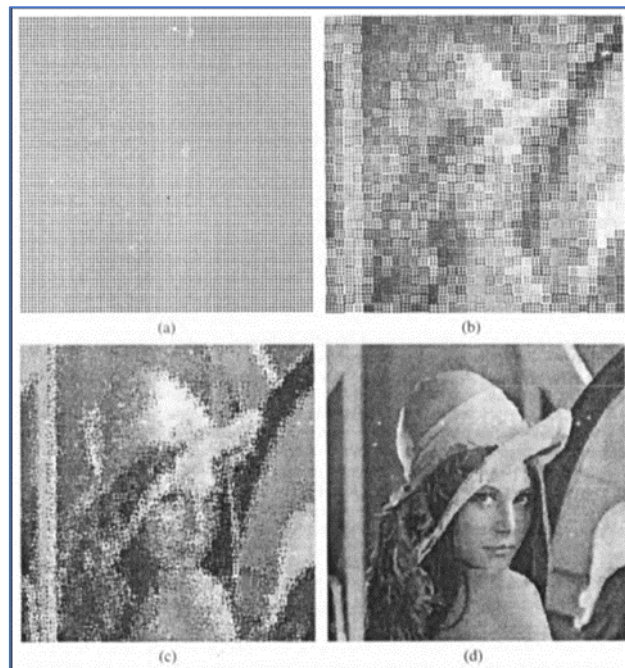


Figure 7-6 : une transformée en fractales

La compression fractale consiste tout d'abord à réaliser deux segmentations (appelées aussi pavages, ou partitionnements) sur une image : une segmentation de figures Sources et une segmentation de figures Destinations. Il s'agit alors de trouver pour chaque figure Source, quel est le meilleur couple (figure source, figure destination) minimisant une erreur. Cette erreur est généralement calculée en soustrayant les deux figures. Pour réaliser l'opération de soustraction, il est nécessaire d'opérer une transformation de la figure source aux dimensions (et à la géométrie) de la figure destination.

8 Le codage des images fixes : JPEG (fruit du Joint Photographic Expert Group)

Les premières applications visant à réduire la quantité d'informations nécessaire au codage d'images fixes ont vu le jour il y a une dizaine d'années, avec comme principal objectif de réduire de façon significative la taille des fichiers graphiques, puis de photographies, en vue de leur stockage ou de leur transmission.

L'ISO (International Standard Organization) a créé en 1990 un groupe de travail international baptisé JPEG (Joint Photographic Experts Group) pour élaborer une norme de compression d'images photographiques de résolutions variables. Le format utilise une représentation YCbCr des couleurs, qui permet de séparer la luminance (Y) et les deux composantes de chrominance (Cb et Cr) pour chaque pixel.

Cette norme, publiée en 1993 sous la référence ISO/IEC 10918, est très ouverte et peut être considérée comme une « boîte à outils » pour la compression d'images fixes.

La compression JPEG ne sélectionne qu'une partie bien choisie de l'information traduite par les coefficients de la DCT. Ce sont des critères d'efficacité qui déterminent le fait que certains coefficients sont en partie ou complètement négligés et qui génèrent une compression numérique avec plus ou moins de pertes (surtout en chrominance).

JPEG définit deux classes de processus de compression :

- sans pertes ou compression réversible : il n'y a pas de pertes d'information et il est donc possible de revenir aux valeurs originales de l'image. Elle fait l'objet d'une norme spécifique appelée JPEG-LS ;
- avec pertes ou compression irréversible : c'est le JPEG « classique » qui permet des taux de compression de 3 à 100 (réduction du débit binaire à la sortie du codage et donc du volume de données à stocker). Ce genre de codage utilise de la DCT (Discrete Cosine Transform) sur des macro-blocs afin de permettre la suppression des coefficients nuls ou proches de zéro (et donc réduire le volume de données à conserver).

Le format de fichier .jpeg :

- ✓ C'est l'un des formats de fichier image le plus couramment utilisé par les APN (appareils photo numériques) et autres dispositifs de capture d'images.
- ✓ Supporte 16.777.216 couleurs, qui sont générées en utilisant un codage sur 8 bits pour chaque couleur dans le modèle RGB.
- ✓ Peut afficher plus de 16 millions de couleurs à la fois, ce qui rend l'agencement des couleurs et la résolution du contraste presque parfaits.
- ✓ Supporte une taille maximale de 65.535 x 65.535 pixels.
- ✓ Réduit la taille d'une image d'environ 50 à 75% (en raison de la compression avec perte) lors de l'enregistrement.

Cependant ce type de format de fichier n'est pas idéal pour les images avec des bords nets car les couleurs se mélangent davantage que si elles étaient enregistrées en format .png (qui affiche les pixels individuels en combinaison).

Il existe d'autres formats de fichiers d'images que le JPEG. Les deux types principaux sont les images matricielles et les images vectorielles :

- les fichiers d'images matricielles stockent des informations sur les pixels dans une grille ou tableau,

- tandis que les fichiers vectoriels stockent des données sous forme d'opérations géométriques avec des points, des lignes et des courbes.

Voici une liste de formats de fichiers d'images populaires :

PNG : Format de fichier d'image matricielle sans perte qui prend en charge la transparence et la compression sans perte.

GIF : Format de fichier d'image matricielle qui prend en charge l'animation et la transparence. Il utilise une palette de couleurs limitée, ce qui le rend idéal pour les images simples telles que les icônes.

WebP : Format de fichier d'image matricielle développé par Google qui prend en charge la compression avec perte et sans perte. Il offre une meilleure qualité d'image à une taille de fichier plus petite que JPEG ou PNG.

TIFF : Format de fichier d'image matricielle sans perte couramment utilisé pour l'impression professionnelle. Il prend en charge la profondeur de couleur élevée et la compression sans perte.

BMP : Format de fichier d'image matricielle non compressé couramment utilisé pour les images Windows. Il prend en charge la profondeur de couleur élevée, mais produit des fichiers volumineux.

HEIF : Format de fichier d'image matricielle développé par MPEG qui prend en charge la compression avec perte et sans perte. Il offre une meilleure qualité d'image à une taille de fichier plus petite que JPEG ou PNG.

SVG : Format de fichier d'image vectorielle qui stocke des données sous forme d'opérations géométriques avec des points, des lignes et des courbes. Il est évolutif, ce qui signifie qu'il peut être redimensionné sans perdre sa qualité.

9 La transformation des images animées

9.1 Codage vidéo : l'échantillonnage des couleurs (rappel)

Si d'une manière générale la vision humaine présente une sensibilité moindre aux couleurs qu'à la définition, c'est évidemment encore plus vrai dans le cadre d'images animées ! En vidéo il est donc possible de conserver moins de données de chrominance que de luminance. Il est donc possible de sous-échantillonner la chrominance, ce qui correspond à un besoin de réduire le poids du signal pour l'enregistrement sur disque ou carte compact flash.

137	125	128	115	118	114	138	104
147	127	137	123	127	113	147	103
127	147	122	123	126	113	96	103
137	127	155	129	145	99	115	159
133	141	133	131	123	91	133	121
97	157	97	137	117	87	127	147
128	141	129	131	119	111	129	121
127	141	117	131	127	122	121	112

Pour les besoins de l'échantillonnage, le signal est découpé en tranches temporelles de 4 pixels de référence.

Dans ce cas l'échantillonnage est réputé 4.4.4, soit une information complète en luminance et en chrominance.

Le sous-échantillonnage de la chrominance est une méthode de réduction de volume de données contenues dans des images numériques.

28	18	13	37
37	17	14	47
22	16	9	27
55	15	11	37
33	13	13	33
37	17	17	37
29	19	19	28
17	17	12	17

Dans tous les cas, l'information de luminance sera entièrement conservée car elle est indispensable pour une bonne perception de l'image enregistrée par notre œil et comporte donc les 4 valeurs captées.

38	14	1	47
27	7	4	17
3	6	19	7
45	5	11	7
3	13	12	33
27	7	17	3
9	1	1	8
7	17	12	17

Pour l'échantillonnage de la chrominance, seule la moitié des valeurs seront conservées (en mode 4.2.2 : ci-contre) ou le quart (4.1.1 ou 4.2.0.).

Figure 9-1 : un échantillonnage en 4.2.2

9.2 Codage vidéo : Compression intra – inter image⁶

La compression vidéo est une méthode de compression de données, qui consiste à réduire le volume de la même image (codage inter-image identique au JPEG) ou des pixels d'images voisines (utilisant des techniques intra-image).

La compression vidéo inter-image (prédiction spatiale) traite chaque image séparément. L'image est divisée en zones (macros-blocks), qui sont en général de 8 pixels sur 8 pixels pour un traitement DCT.

⁶ <https://www.gypsevideo.fr/tout-savoir-sur-la-video/les-techniques/codecs>

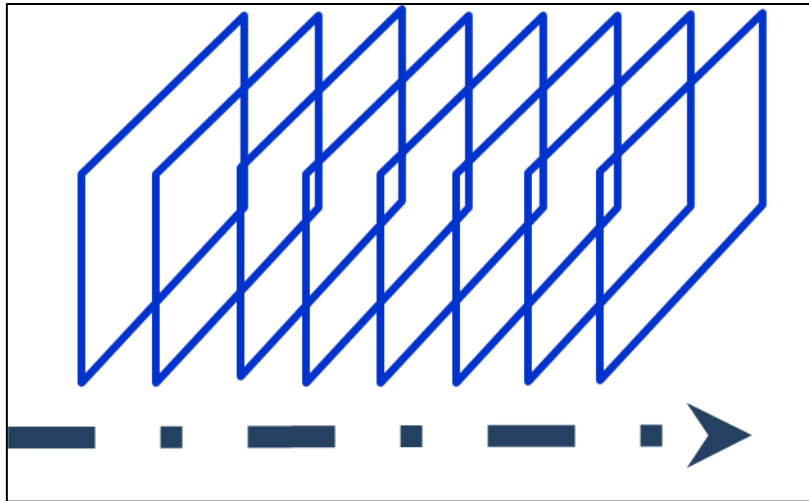


Figure 9-2 : La compression vidéo inter-image (prédiction spatiale)

Avantage :

Pas de temps de traitement : la DCT travaille chaque bloc de 8 X 8 pixels indépendamment.

Désavantage :

- La réduction de débit est faible.
- Formats utilisant la compression inter-image : DV-AVI, Mjpeg, AVC-intra, DVCPROHD.
- La compression temporelle (prédictions inter-images) s'effectue par GOP (Group of Pictures).

9.3 Codage vidéo : le traitement de l'AV : le MPEG 1- 2 (Moving Picture Expert Group)

Les réunions à l'ISO du groupe MPEG (Moving Picture Experts Group) : des spécialistes provenant de l'industrie de l'électronique des composants, de l'informatique et des télécommunications ont démarré en 1988 dans le but de développer une première norme ISO **MPEG-1** pour des applications de stockage audio/vidéo du type Vidéo CD. La caractéristique de la famille MPEG est basée sur l'usage du **GOP** (Group of Pictures). Cette caractéristique engendre une latence au codage/décodage (plusieurs secondes dépendant de la longueur de GOP).

En novembre 1994, l'ISO a publié la norme **MPEG-2** faisant suite à MPEG-1.

MPEG-2 définit tous les aspects du codage source de l'image et du son et ainsi que du transport (TS : Transport stream) à travers des réseaux de télévision numérique ou de la production de DVD (PS Packet Stream) pour des usages liés à la télévision numérique (diffusion DVB et DVD).

Fin des années 90, l'ISO a publié **MPEG-4**, également appelé ISO/CEI 14496, une norme de codage d'objets audiovisuels. Cette norme MPEG-4 spécifie des techniques pour gérer le contenu de scènes comprenant un ou plusieurs objets audio-vidéo. Contrairement au MPEG-2, le MPEG-4 englobe 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 film, la vidéo combinent des dimensions à la fois spatiales (images) qui évoluent dans le temps (un nombre d'images par seconde reproduisant un mouvement).

Afin d'optimiser le codage des images, le codec découpe le flux en séquences de l'ordre de 12 à 16 images (ou plus Gop : Group of Pictures) en déterminant une image mère (I = I Frame codée en mode intra) et des images P (prédites ; calculée par différence avec les images précédentes) ou B (calculée en

fonction des images qui précèdent ou suivent) afin de compresser au maximum le débit à transmettre ou à stocker.

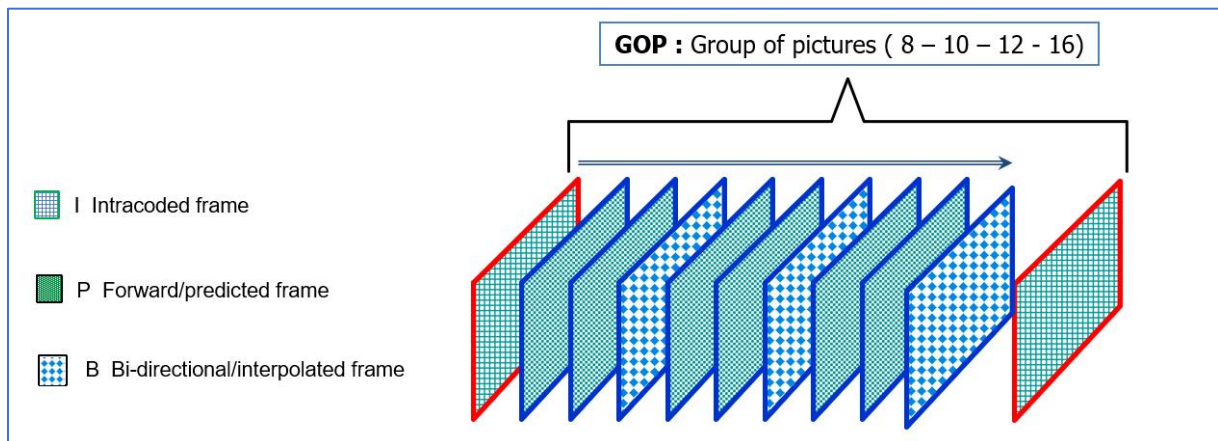


Figure 9-3 : le codec découpe le flux en séquences de l'ordre de 12 à 16 images

- Temps de traitement (longueur du GOP = latence)
- La DCT travaille chaque macro-bloc de 8 X 8 pixels indépendamment (MPEG 2) ou par regroupement (MPEG 4)
- La réduction de débit est forte
- Ce système de compression a été développé pour la diffusion broadcast

9.4 Le codage vidéo : la prédiction du mouvement

Les différents macro-blocs constituent la base pour le calcul de l'estimation du mouvement déterminant le vecteur de mouvement.

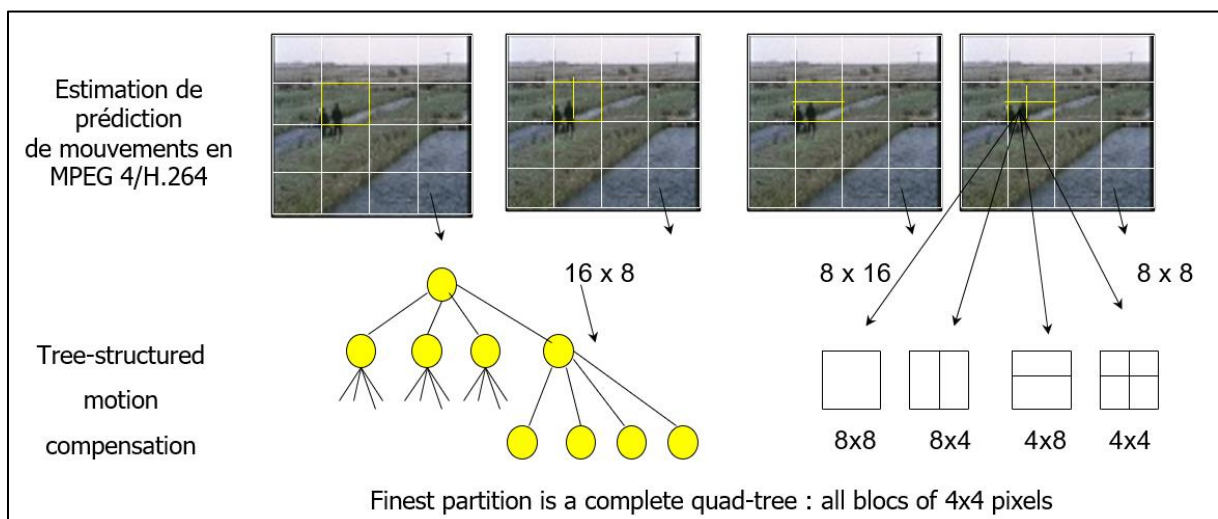


Figure 9-4 : MPEG 4 : décomposition en blocs pour l'estimation de mouvement

En MPEG-4 (qui permet de coder des objets), il existe trois types de VOP (Vidéo Object Plane) :

- des I-VOP (Intra-coded) sans compensation de mouvement,
- des P-VOP I (Predicted) prédits à partir des derniers I-VOP ou P-VOP,
- et des B-VOP (Bidirectionnaly predicted) prédits à partir des plus récents I-VOP ou P-VOP antérieur et/ou postérieur qui contribuent efficacement à la réduction de données à stocker et à transmettre.

10 Les normes de format et de compression numériques :

10.1 du MPEG-2 au MPEG-4

Dans l'évolution MPEG-4 AVC, du fait de puissances de calcul disponibles, des macro-blocs de tailles plus petites (4 X 4) différentes et plus grandes (16 X 16) ont été introduites.

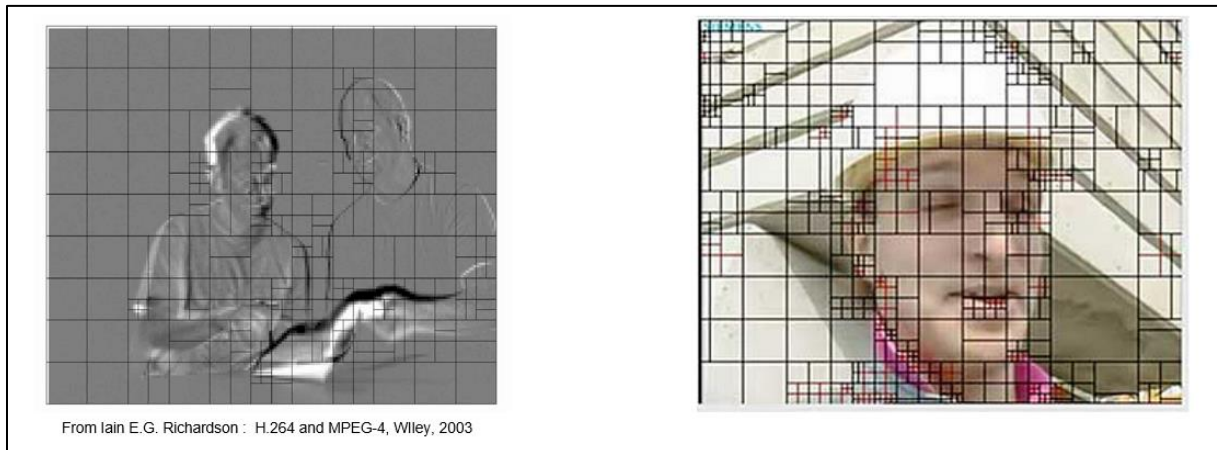


Figure 10-1 : Segmentation des macro-blocs (plus petits ou plus grands)

Par rapport aux macro-blocs de taille fixe, cette méthode présente l'intérêt de mieux s'adapter au contenu, en traitant très différemment les aplats des zones les plus détaillées. Elle est en outre nettement plus propice au traitement parallélisé, tel qu'avec un Graphics Processing Unit [GPU] et offre plus de possibilité pour des extractions sémantiques.

10.2 Autre format : le JPEG2K :

JPEG 2000 décompose l'image en une représentation à résolution multiple au cours de son processus de compression. Cette représentation pyramidale peut être utilisée à d'autres fins de présentation d'image au-delà de la compression.

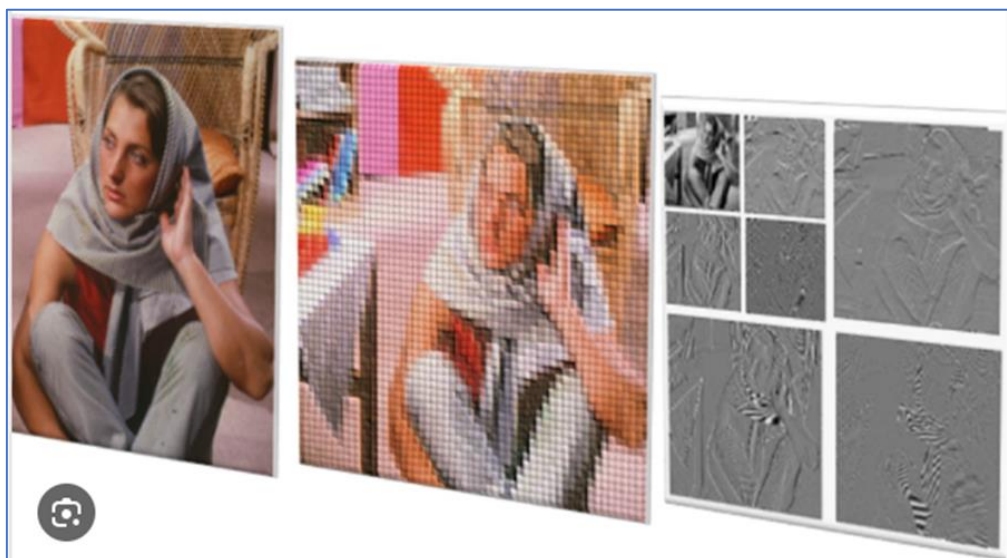


Figure 10-2 : un exemple de codage JPEG2K : Transformée en Ondelettes Discrète

Ces fonctionnalités sont plus communément appelées décodage progressif et évolutivité du rapport signal sur bruit (SNR). JPEG 2000 fournit des organisations de flux de code efficaces qui sont progressives par précision de pixel et par résolution d'image (ou par taille d'image). De cette façon,

après avoir reçu une plus petite partie du fichier entier, le spectateur peut voir une version de qualité inférieure de l'image finale. La qualité s'améliore ensuite progressivement en téléchargeant davantage de bits de données à partir de la source.

Comme la norme JPEG, la norme JPEG 2000 fournit à la fois une compression sans perte et avec perte dans une architecture de compression unique. La compression sans perte est fournie par l'utilisation d'une transformée en ondelettes entière réversible dans JPEG 2000.

11 En développement : l'Ultra HD – HDR (High Dynamic Range)

Aujourd'hui, les téléviseurs 4K (Ultra HD) sont les « futurs » nouveaux rois de la TV haut de gamme avec une résolution d'image numérique de 3840 × 2160 pixels, soit environ 8 millions de pixels. Cela signifie que l'image est plus riche en détails, ... et offre ainsi une expérience immersive sans devoir augmenter la distance entre l'écran et le spectateur.

Depuis quelques années, la télévision japonaise (NHK) présente les évolutions d'un projet dénommé « Super Hi Vision », qui propose d'augmenter la résolution à 7680 pixels sur 4320 lignes, soit une définition 16 fois supérieure à celle de la télévision Full HD. Conçue pour être visionnée à une distance de 0,75 de la hauteur de l'écran et pour remplir 110° du champ visuel, cette expérience procure un réalisme incroyable !

Par ailleurs, la BBC a démontré récemment que le système de balayage progressif à 50 images/sec n'était pas suffisant pour éliminer tous les artéfacts dans les mouvements rapides. Elle propose d'y remédier en augmentant la fréquence d'image à 300 images progressives/sec.

Place donc à la norme H.265/HEVC (High Efficiency Video Coding), capable de supporter jusqu'à une définition 8K (8.192 x 4.320 pixels). Grâce à un processeur plus puissant et des algorithmes complexes, le HEVC inaugure une nouvelle méthode de découpage des trames «Coded Tree Unit» variable (trois tailles) et sur trois niveaux avec des gains de compression deux fois supérieurs.

Bref, ce n'est qu'un début, continuons à contribuer à améliorer « l'expérience du spectateur » !

11.1 UHD vs HDR :

La véritable différence entre UHD et HDR réside dans le fait que l'UHD fait référence à la résolution, tandis que le HDR est né pour améliorer les couleurs et les contrastes d'un écran UHD. High Dynamic Range". En français, il est traduit par "Haute Gamme Dynamique". Il s'agit d'une technologie qui va améliorer la qualité visuelle de l'image sur votre téléviseurs, grâce à un fin jeu de couleurs.

Le but de la HDR est donc de se servir des contrastes et de l'intensité lumineuses pour en améliorer la perception par notre œil, tout en proposant un rendu réaliste, avec une colorimétrie beaucoup plus naturelle. Cela joue donc surtout sur les couleurs, plus que sur la qualité de l'image en elle-même.

Le High Dynamic Range n'est pas tout à fait nouveau, puisqu'il est employé dans l'univers de la photographie. Il désigne alors la combinaison logicielle de plusieurs clichés à l'exposition différente afin d'obtenir des résultats à la fois plus lumineux et plus contrastés, sans zones bouchées ni surexposées.

Dans l'univers de la télévision, le HDR, même si son résultat est similaire, s'appuie sur des principes techniques bien différents : le HDR (High Dynamic Range) propose d'étendre la plage dynamique de l'image. Tout est une histoire de luminosité, ou plutôt de luminance, laquelle s'exprime en NITS ou en Candelas par mètre carré (cd/m²).

Les téléviseurs HDR sont ainsi capables d'offrir une plage de luminosité plus étendue que leurs homologues SDR, et de monter bien plus haut dans la luminance et de descendre très bas dans le noir.

La coalition « UHD Alliance » recommande ainsi un pic de luminosité d'au moins 1000 nits et un niveau de noir à moins de 0,05 nit, ou d'au moins 540 nits, avec un niveau de noir à moins de 0,0005. De quoi espérer que les scènes les plus éclairées soient particulièrement lumineuses, que les zones sombres ne soient pas délavées, et qu'entre les deux, les détails de l'image ne se perdent pas dans des zones bouchées et/ou surexposées. Car c'est bien l'idée derrière le HDR : il est conçu pour laisser voir les détails de l'image quelle qu'en soit l'exposition, et pour offrir une image la plus proche possible de ce que voit l'œil humain dans la réalité.

Une bonne gestion des différents spectres colorimétriques est également garante d'un HDR de qualité.

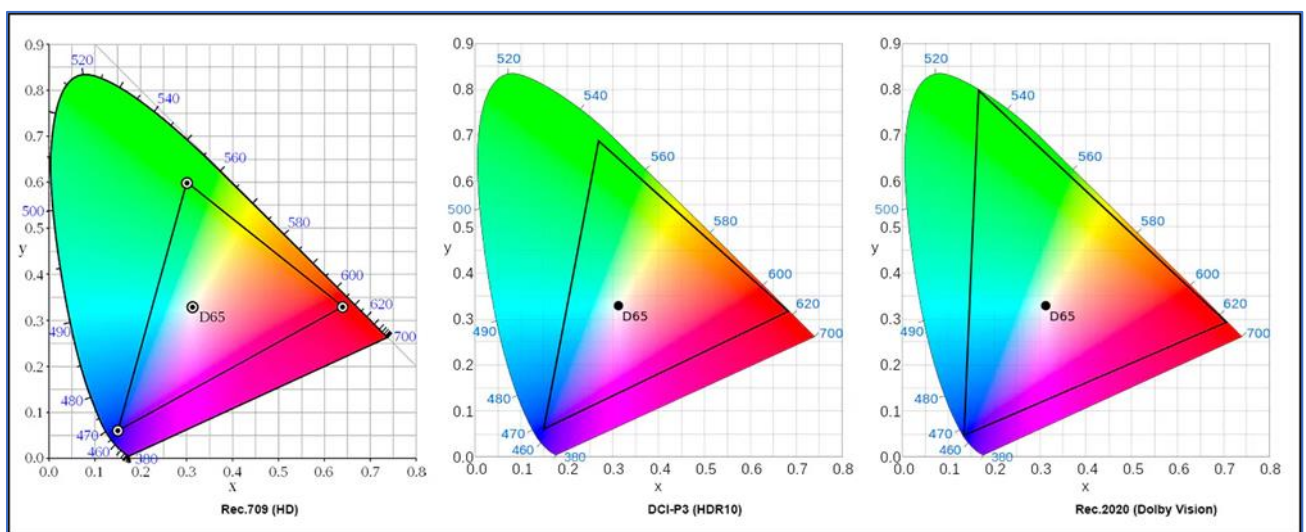


Figure 11-1 : HDR : comparaison de la plage de luminosité de différents standards

C'est là qu'intervient la deuxième recommandation de l'UHD Alliance : non seulement les appareils doivent disposer d'une entrée compatible avec l'espace colorimétrique Rec.2020 (ou BT2020⁷), mais aussi de l'affichage d'au moins 90 % de l'espace colorimétrique DCI-P3⁸. Les téléviseurs doivent être capables d'afficher l'intégralité des couleurs situées à l'intérieur d'un triangle (gamut) plus étendu de 25 % que le traditionnel sRGB, notamment dans les rouges et les verts. Les constructeurs peuvent évidemment aller plus loin et afficher davantage de couleurs, mais ils doivent conserver l'équilibre entre chaque teinte pour éviter les dérives, ce qui n'est pas toujours une mince affaire. Il faut également noter que la norme P3 est utilisée pour la projection cinématographique.

⁷ Voir : https://fr.wikipedia.org/wiki/Rec._2020

⁸ DCI-P3 est un espace colorimétrique RVB défini pour la première fois en 2005 dans le cadre de la Digital Cinema Initiative, à utiliser pour la distribution cinématographique numérique en salle[1] (DCDM[2]). Display P3 est une variante développée par Apple Inc. pour les écrans à large gamme.

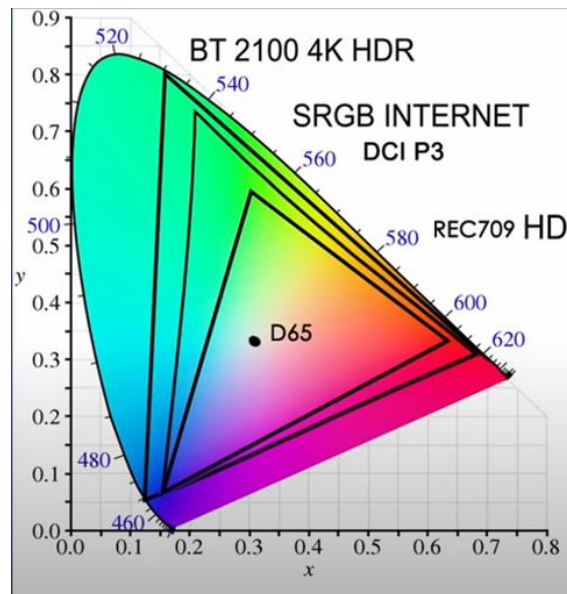


Figure 11-2 : espace couleur : le gamut en HDR

Le gamut de l'œil "moyen" est très étendu : c'est le fameux espace $L^*a^*b^*$. Cet espace qui a été décrit par un consortium international (CIE) en 1976; englobe toute la lumière visible par un être humain avec un œil « standard ». Il permet de définir une couleur de manière absolue. Pour chaque code L^*a^*b correspond une couleur absolue, précise, une longueur d'onde précise (chiffres en bleu au pourtour du fer à cheval). Si on indique une couleur précise avec une longueur d'onde précise elle aura un unique code L^*a^*b . C'est le mètre étalon de la couleur.

- ✓ D65 : la lumière du jour naturelle.
- ✓ La recommandation UIT-R BT 709 ou plus communément Rec.709 est un standard de l'industrie audiovisuelle pour la SD, la HD et la télévision 4K UHD.
- ✓ DCI P3 : la nouvelle norme de gamme de couleurs pour le cinéma.
- ✓ SRGB Internet.

BT 2100 4K HDR : BT.2100 permet aux producteurs de télévision de choisir parmi trois résolutions : Full HD' (1920 x 1080), UHD TV 4K' (3840 x 2160) et UHD TV 8K' (7680 x 4320)

Equipement grand-public : parmi les boîtiers adaptés au HDR,

- les Shield de Nvidia, dont la dernière mouture date de 2019 (Dolby Vision et HDR10),
- mais aussi le tout récent boîtier Chromecast avec Google TV (HDR10, HDR10+, Dolby Vision).

Il est à noter que Netflix et Amazon Prime Video ne sont plus les seuls à fournir des contenus en HDR : comme eux, le plus récent Disney+ est compatible avec les normes HDR10 et Dolby Vision. Apple TV+ est lui aussi compatible HDR et Dolby Vision.

11.1.1 Le HDR10- 10+- HDR10+ Adaptive

- Le **HDR10** est un standard qui s'appuie sur une plateforme ouverte. Il est ainsi nommé, car son échelle de quantification est basée sur 10 bits au sein de l'espace colorimétrique BT.2020 . Chaque constructeur de matériel peut s'en emparer pour rendre ses appareils compatibles et ainsi offrir aux utilisateurs les bénéfices d'une telle technologie sur leurs écrans. Il permet d'afficher une plus large palette de couleurs et de nuances par rapport à des images classiques encodées sur 8 bits dans l'espace colorimétrique Rec.709. Il utilise la norme SMPTE ST 2084 EOTF

qui se base sur la réelle sensibilité de l'œil humain concernant le contraste et permettant de couvrir une luminance maximale de 10 000 cd/m².

- Le **HDR10+** est une évolution du HDR10. Il a été développé pour mieux concurrencer le format Dolby Vision. Comme lui, le HDR10+ exploite les métadonnées dynamiques qui sont intégrées au sein du flux vidéo afin de pouvoir moduler la réponse selon les scènes et ainsi en optimiser le rendu. Ce standard est soutenu par plusieurs acteurs majeurs de l'industrie de l'électronique dont la 20Th Century Fox, Panasonic et Samsung et ce, depuis 2018. En août 2021, le format HDR10+ revendiquait 120 partenaires, dont des fabricants de télévisions, mais également des marques de smartphones, tablettes tactiles et lecteurs Blu-ray. À cette date, 2000 contenus prennent en charge ce format.
- Le **HDR10+ Adaptative** est une évolution du HDR10+. Lancé depuis 2021, il prend non seulement en charge les métadonnées dynamiques du flux vidéo, mais ajuste encore un peu plus les images afin de s'adapter aux conditions d'éclairage de la pièce. Pour ce faire, le diffuseur utilise un capteur de luminosité afin d'optimiser les scènes comme l'aurait souhaité le réalisateur en limitant les pertes de détails et de contraste selon la lumière ambiante. Le HDR10+ Adaptative vient concurrencer le format Dolby Vision IQ qui utilise le même principe d'adaptativité selon les conditions lumineuses de visionnage.

11.1.2 Le format HDR Dolby Vision

Le format Dolby Vision est un format HDR. Mais à la différence du HDR10, il n'est pas libre de droits. Il est la propriété des laboratoires Dolby. Pour en profiter, il faut que les appareils de lecture et de diffusion soient compatibles Dolby Vision. Le caractère HDR ne suffit pas.

Le flux vidéo Dolby Vision est encodé sur 12 bits.

Comme le format HDR10+, il utilise des métadonnées dynamiques pour optimiser les images en termes de luminosité, de contraste et de colorimétrie selon les scènes de la vidéo. Par rapport au HDR10+, il profite d'un suivi et d'un contrôle particulièrement rigoureux de l'étalonnage depuis les studios de production jusqu'à la télévision du consommateur final.

Il s'agit d'une évolution du format Dolby Vision. On peut le comparer au HDR10+ Adaptative dans le sens où il utilise un capteur de luminosité installé sur le diffuseur pour mesurer la quantité de lumière dans la pièce.

Ainsi, travaillant de concert avec le processeur du diffuseur, il permet d'ajuster les propriétés de l'image pour tenter d'obtenir le meilleur rendu possible depuis une source Dolby Vision. Cette option peut être désactivée sur les appareils qui le proposent, le cas échéant.

12 Conclusion :

À partir du XXe siècle avec l'apparition des technologies de radiodiffusion et la télévision, différentes normes et standards voient le jour dans de multiples parties du monde. Pour faciliter les échanges nécessaires entre ces différentes zones géographiques, des dispositifs sont élaborés pour transformer les sources produites dans des formats ou standards locaux et de les rendre compatibles avec la zone qui les reçoit. Ce sont des appareils ou équipements électroniques qui effectuent les traitements (soit en tête de réseau ou dans le périphérique) visant à transcoder des signaux « NTSC », « PAL », « SECAM », ... dans un format accessible par le spectateur.

L'émergence du numérique dans la dernière décennie du XXe siècle révolutionne ces traitements. Comme les signaux sont échantillonnés, analysés, traités et convertis en fichiers de nouveaux standards font leur apparition !

Les normes et standards des formats de codage employés pour spécifier un média listent l'ensemble paramètres propres à la norme ou au format :

En photographie

- l'affichage : la dimension de la photo en représentation) ;
- le type d'affichage (en transparence – en réflexion),
- la définition (le nombre de points ou pixels significatifs de la photo ligne horizontale / nb de lignes verticales) ;
- la résolution (pixels par centimètre carré) ;
- la profondeur de couleur (sur 8 ou 10 bits) ;
-

Pour la vidéo, toutes les spécifications de la photographie, avec en plus :

- le rapport hauteur/largeur de l'image (4:3, 16:9 par exemple) ;
- la fréquence images par seconde (25, 30, 50, 60, ...) ;
- le balayage entrelacé ou progressif ;
-

Et d'une manière plus générale, le type d'OS (operating system) utilisé, le taux de compression, le débit numérique et les algorithmes exploités. Bien que performants, les codecs « propriétaires », peuvent produire des médias inutilisables dans un proche avenir, ou s'avérer inexploitable sur d'autres plateformes matérielles ou logicielles.

Enfin, il faut bien distinguer ce que le numérique a rendu compatible en lecture, d'un véritable transcodage de format (en vue de finaliser un montage d'un sujet vidéo par exemple).

En acquisition/production/postproduction, l'étape du transcodage peut se révéler ardue et parfois douloureuse (dégradations, pertes importantes de temps comme le nombre de calculs à effectuer pour restituer l'intégralité d'un média, de qualité). Par exemple, l'affichage d'un film HD codé avec en H264 demande d'importantes capacités de processeur, ou des circuits intégrés spécifiques qui ne sont pas toujours disponibles localement). Afin d'éviter ces déboires, les sources audio et vidéo captées doivent dans la mesure du possible être enregistrées dans la représentation des données, qui seront utilisées pour la post-production.

Dans le cas de la post-production, il est recommandé de privilégier l'usage d'un format proxy (peu coûteux en matière de calculs pour la lecture et le traitement des essences).

Un des problèmes importants rencontrés lors d'un transcodage est le nombre d'images vidéo générées par seconde ainsi que les balayages respectifs à chaque norme.

- Le cinéma utilise toujours des formats non compatibles avec la télévision : 24 images secondes – format 2K ou 4K.
- Dans le cas du PAL, nous avons 25 images complètes par seconde composées de 2 frames entrelacées transmises successivement (lignes paires, puis impaires).
- Le NTSC, lui, utilise 29,97 images par seconde ; soit une différence au niveau du flux d'images d'environ 5 images existantes dans une des normes, et manquantes pour l'autre.

Pour un transcodage du format NTSC à PAL, on ne peut pas simplement ôter une image toutes les 5 images car cela provoque un effet stroboscopique et une désynchronisation du son. Pour le transcodage du format PAL vers le NTSC, la démarche de dupliquer une image toutes les 5 images pour « combler » les vides amène aux mêmes piètres résultats. La numérisation du média offre aujourd'hui des traitements numériques d'interpolation, effectués de manière logicielle, qui améliore le résultat.

Balayage entrelacé vs progressif :

Le balayage progressif a toujours été une méthode plus simple que le balayage entrelacé conçu à l'époque en raison des goulots d'étranglement techniques du passé. La bande passante de transmission ne pose plus de problème aujourd'hui et les vidéos et écrans progressifs dominent le marché.

Les seules vidéos entrelacées restantes sont souvent des archives. Comme la plupart des écrans d'aujourd'hui sont progressifs, ces anciennes vidéos doivent passer par le processus de désentrelacement avant d'être diffusées sur des écrans progressifs.

Le transcodage d'une image entrelacée en progressif peut causer des problèmes :

- comme le peignage - où les lignes paires de la trame A et de la trame impaire B constituant de l'image sont synchronisées sur un écran, le résultat affiche une image qui est la somme des deux trames. Dans le cas d'un événement sportif, la vitesse de déplacement des objets fait que les deux trames enregistrent une modification visible de la position dans l'espace !
- lors de l'affichage de textes, les lignes de la structure des lignes deviennent plus visibles, ce qui affecte la qualité de l'incorporation des détails de l'image et l'apparence de l'image sur les écrans modernes qui sont physiquement beaucoup plus grands que les anciens.

Algorithmes de désentrelacement : Il existe un tas d'algorithmes différents qui peuvent faire cela :

- désentrelacement de combinaison de champs - fusionne les champs pairs et impairs en une seule image. Par exemple, la conversion de 50i ou 60i en 25p ou 30p réduit de moitié la fréquence d'images perçue.
- désentrelacement d'extension de champ - étire chaque champ (avec seulement la moitié des lignes) sur tout l'écran pour former une image complète. Il préserve le taux de champ initial tout en réduisant de moitié la résolution verticale de l'image (50i ou 60i est converti en 50p ou 60p).
- désentrelacement de compensation de mouvement - utilise des algorithmes de plus en plus sophistiqués pour trouver le mouvement à travers les champs, en modifiant les techniques si nécessaire. Bien qu'il utilise le plus de puissance de traitement, cela produit le produit de la plus haute qualité.

Pour l'archivage, il ne faut pas prendre en compte les capacités matérielles de codage/décodage des ordinateurs actuels. Si le média a perdu de son intégrité à cause de plusieurs transcodages ou d'un transcodage trop destructif, il sera impossible de revenir en arrière.

Il est donc impératif de choisir un transcodage le moins destructif possible pour le média original et son stockage (conservation du format de l'original). On conserve ainsi la possibilité d'effectuer un nouveau transcodage avec de nouveaux codecs aux goûts du jour et/ou plus performants.

Ce n'est qu'au prix de cette interpolation que l'on arrive à transcoder des flux de formats incompatibles entre eux pour atteindre une qualité satisfaisante !

La migration informatique est un autre problème générique ! Le passage d'un état existant d'un système d'information ou d'une application vers un autre constitue un vrai défi. Elle peut prendre la forme d'un transfert, d'une transformation ou d'une conversion de données d'un espace de stockage, d'un format ou d'un type de système à un autre. Le processus de migration informatique peut impliquer un ou plusieurs types de déplacements : une migration de données, d'applications, entre systèmes d'exploitation et vers le cloud.

L'incompatibilité des systèmes d'exploitation est une réalité, elle peut être due à plusieurs raisons.

Par exemple :

- les programmes peuvent être conçus pour fonctionner avec une version spécifique d'un système d'exploitation et ne pas offrir de rétrocompatibilité avec une version plus ancienne ou future. C'est le cas par exemple des formats LTO (Linear Tape-Open (LTO)), un format « ouvert » de stockage sur bande. Les différentes générations de LTO sont rétro-compatibles en écriture jusqu'à la génération n-1 et en lecture jusqu'à la génération n-2 ;
- les applications ou programmes plus anciens de Microsoft peuvent ne pas être compatibles avec les dernières versions de Windows 10 ou Windows 11.

D'autres raisons peuvent inclure des erreurs qui endommagent les clés de Registre sur les systèmes d'exploitation 64 bits ou qui empêchent l'installation de nouveaux programmes.

Il est également important de vérifier si un PC est compatible avec un OS par exemple Windows 11 avant de l'installer. Pour ce faire, il faut vérifier la compatibilité de tous les éléments essentiels : processeur, mémoire vive, Bios, TPM, etc...

Comment éviter les problèmes d'incompatibilité des systèmes d'exploitation :

- vérifiez les exigences du système pour les programmes que vous installez et assurez-vous que votre système d'exploitation répond à ces exigences ;
- exécutez l'utilitaire de résolution des problèmes de compatibilité pour les programmes qui ne fonctionnent pas correctement ;
- si vous utilisez plusieurs antivirus, assurez-vous qu'ils sont compatibles entre eux.

Si vous rencontrez des problèmes de compatibilité avec les applications, vous devez les classer par ordre de priorité, puis leur affecter une solution.

Il est vrai que les systèmes d'exploitation Apple, Windows et Android sont différents et peuvent être incompatibles entre eux. Cependant, il existe des « solutions » pour résoudre ce problème en utilisant des applications tierces pour transférer des fichiers entre les différents systèmes d'exploitation.

Cependant, la migration informatique peut entraîner des pertes d'informations difficiles à identifier en raison de la multitude des interactions entre logiciels. De plus, la migration est dans une certaine mesure irréversible du fait de la séparation des données de leur contexte technologique de création.

ALI OUMAROU - PRESENTATION SUR LES FORMATS

DU CODAGE AU FORMAT DE FICHIER

1. 1 - Joint Photographic Experts Group (.jpeg ou .jpg)
2. 2 - Portable Network Graphics (.png)
3. 3 - Graphics Interchange Format (.gif)
4. 4 - Tagged Image File (.tif)
5. 5 - Photoshop Document (.psd)
6. 6 - Portable Document Format (.pdf)
7. 7 - Encapsulated Postscript (.eps)



Définition et résolution : quelle différence ?

On pense souvent, à tort, que parler de résolution et de définition revient à la même chose. C'est faux et c'est encore une fois la faute aux Anglais ! Une bonne occasion de revenir aussi sur les différentes définitions en photo et en vidéo.

Définition

La définition d'une image est la quantité d'information contenue dans celle-ci. En photographie numérique, une image est composée de points : les pixels. La définition est une valeur absolue qui permet de quantifier une image et, souvent, le capteur qui a servi à sa création. Vous pouvez ainsi avoir des images en 2, 8, 16 ou 100 millions de pixels, ou Mpx.

Notez que cette "écriture" de définition ne permet pas de connaître les dimensions de l'image : nous savons juste combien de pixels elle est composée. En effet, un pixel n'a pas de taille "physique" en soi : c'est une unité d'affichage ou d'impression.

La taille physique d'un pixel peut ainsi très bien aller de plusieurs millimètres de côté, sur un écran géant par exemple, à quelques μm (ex. sur une dalle de smartphone).

Résolution

Toutefois, dès que cette image est affichée ou imprimée, il faut prendre en compte un autre paramètre : la surface d'affichage, donc la résolution. En effet, une fois que vous avez le support d'affichage (un écran, du papier...), il faut faire rentrer toutes les informations de l'image sur cette surface. Plus cette surface est petite, plus la résolution augmente.

Pour résumer, la résolution est le nombre de pixels par unité de mesure. Les Anglo-Saxons ont une nouvelle fois gagné la bataille, puisque le système métrique n'est pas utilisé. La résolution s'exprime en nombre de pixels par pouce — les fameux ppp (pixel per inch) — ou de points par pouce — les non moins fameux dpp (dot per inch). On pourrait parfaitement parler de pixels par cm, mais c'est rarement le cas.

Publicité, votre contenu continue ci-dessous

Quel terme pour quel usage ?

Ces notions sont donc bien différentes. Mais alors d'où vient cette confusion ? Elle provient du fait que les Anglo-Saxons, pour évoquer la définition, utilisent le terme resolution. Le raccourci est donc facile et... tenace : vous trouvez encore des sites marchands qui, pour parler de la définition du [5D Mark IV de Canon](#), utilisent (fautivement) à la place le terme de résolution.

La notion qu'il faut retenir est que plus il y a de pixels, plus vous pourrez exploiter facilement votre image pour une impression, une visualisation sur écran ou un recadrage, lorsqu'il faudra faire intervenir une dimension finale de votre image.

Exemple d'un recadrage – sévère – dans une image en 2 Mpx (gauche) et 20 Mpx (droite) pour une dimension finale de 450 px de large.

En photo, la définition est donc une notion assez simple, mais qui n'apporte finalement pas beaucoup d'information. En vidéo, par contre, les choses se compliquent sérieusement...

Les cas de la vidéo et du cinéma

En vidéo, les conventions d'écriture varient, mais la plupart du temps, ce sont les dimensions de l'image qui sont utilisées pour définir... la définition.

Ainsi, l'affichage HDTV 1080 correspond à une image en 1 920 x 1 080 px, soit 2 073 600 pixels (environ 2 Mpx). C'est donc plus précis qu'en photographie, avec la possibilité de connaître les dimensions, donc le ratio d'une image.

Le problème de ce type de nomenclature, c'est qu'elle multiplie les références parfois peu explicites. Ainsi, en vidéo ou en cinéma numérique, on parle également des définitions UHD, 2K, 4K ou 8K. Pas simple de s'y retrouver.

Le format UHD correspond à une définition de 3 840 x 2 160 px. Pour les formats 2K, 4K...xK, c'est à peine plus complexe...

Le k ou K est une abréviation provenant du préfixe kilo, qui représente 1 000. On utilise cette expression comme abréviation, couramment dans le système monétaire : les milieux financiers ont pris l'habitude de parler de kilo-euros (k€) pour des milliers d'euros.