

# Qualité et Formats d'image - (2 chapitres de Photographie Numérique)

Henri Maître

17 février 2014

# Table des matières

<b>1 Qualité d'image</b>	<b>3</b>
1.1 Les attributs de la qualité . . . . .	3
1.1.1 La résolution . . . . .	3
1.1.2 Le piqué . . . . .	4
1.1.3 L'acuitance . . . . .	5
1.1.4 Le masquage flou . . . . .	6
1.1.5 La quantité d'information . . . . .	7
1.2 Diffraction incohérente . . . . .	7
1.2.1 Cohérence - Incohérence . . . . .	7
1.2.2 Définitions et Notations . . . . .	8
1.2.3 Pour une longueur d'onde unique . . . . .	8
1.2.4 Diaphragme circulaire . . . . .	9
1.2.5 Discussion . . . . .	11
1.2.6 Cas d'un spectre large . . . . .	11
1.2.7 Pouvoir de séparation . . . . .	13
<b>2 Le codage des images et leurs formats</b>	<b>14</b>
2.1 Format RAW : brut ou natif . . . . .	14
2.1.1 Vers une normalisation du RAW . . . . .	15
2.1.2 Que contient un format RAW ? . . . . .	15
2.1.3 Intérêt du RAW . . . . .	16
2.1.4 Les limites du RAW . . . . .	16
2.1.5 Format RAW via TIFF/EP . . . . .	16
2.2 Le format d'entête Exif . . . . .	17
2.2.1 Les données du fichier Exif . . . . .	17
2.2.2 Limitations du fichier Exif . . . . .	18
2.3 Les formats comprimés sans perte . . . . .	18
2.3.1 Les algorithmes généraux de codage sans perte . . . . .	19
2.3.2 Les codages sans perte de JPEG . . . . .	19

2.4	Les formats d'image issus des arts graphiques . . . . .	20
2.4.1	Le format PNG . . . . .	20
2.4.2	Le format TIFF . . . . .	21
2.4.3	Le format GIF . . . . .	22
2.5	Les formats de compression d'images avec pertes . . . . .	22
2.5.1	JPEG . . . . .	23
2.5.2	JPEG 2000 . . . . .	25
2.6	Codage compressé ( <i>Compressed sensing</i> ) . . . . .	29
2.7	Le codage vidéo . . . . .	29
2.7.1	Codage MPEG . . . . .	30

# Chapitre 1

## Qualité d'image

### 1.1 Les attributs de la qualité

L'étude de la qualité des images produites par un appareil photo est complexe. Elle prend en compte de nombreux aspects techniques mesurables, des considérations perceptuelles qui ont été plus ou moins normalisées et enfin des critères esthétiques souvent culturels et subjectifs liés tant à la scène imagée, qu'aux conditions de son examen et à l'observateur. Nous n'examinerons en détail que les premiers et les seconds, laisserons de côté les derniers.

#### 1.1.1 La résolution

Rappelons que l'on définit la résolution comme la distance minimale entre 2 points de l'image qui pourront être distingués (*cf.* paragraphe 1.2.7).

C'est un aspect relativement simple de la qualité de l'image. Tout d'abord on dispose toujours des spécificités du capteur : sa taille ( $l_1 \times l_2$ ) et son nombre de pixels ( $N_1 \times N_2$ ). De ces éléments, on déduit une résolution maximale théorique selon les 2 directions horizontales et verticales :  $u_1 = l_1/N_1$ ,  $u_2 = l_2/N_2$ , indépendante de l'optique d'une part et de l'électronique d'autre part.

Prenant en compte toute la chaîne de construction de l'image, on peut expérimentalement déterminer la Fonction de Transfert de Modulation, FTM<sup>1</sup>, à partir d'une mire sinusoïdale de fréquence croissante. Son image contient précisément la FTM, c'est-à-dire, pour chaque fréquence spatiale le coefficient de réduction de cette fréquence :

$$FTM(u) = \frac{\text{contraste en sortie a la fréquence } u}{\text{contraste en entrée a la fréquence } u} \quad (1.1)$$

Cette FTM s'observe également qualitativement ; lorsque le contraste disparaît on en déduit la fréquence ultime transmise par la photo. La perte progressive de contraste illustre bien l'évolution de la FTM.

Cette méthode utilisant une mire à pas variable, quoique beaucoup utilisée en pratique, est cependant trop globale, ce que l'on constate bien en déplaçant la mire dans l'image. La fréquence d'extinction du contraste est généralement plus élevée au centre qu'au bord de l'image, exprimant que la FTM n'est pas la même en tout point du champ de l'image.

Afin d'obtenir une mesure plus locale permettant d'indiquer la limite de résolution en tout point du champ de l'image, on préfère revenir à la définition de la FTM qui la rattache à la réponse

---

1. en anglais : *Modulation Transfer Function*, *MTF*.

impulsionnelle, RI<sup>2</sup>. En effet :

$$FTM(u, v) = |\mathcal{TF}(RI(x, y))|^2 \quad (1.2)$$

Pour cela une première idée pourrait être d'analyser une image d'une scène constituée de points très fins. L'image de chaque point étant par définition la réponse impulsionnelle, par transformée de Fourier, chacune permettrait de connaître la FTM autour du point de mesure. Malheureusement cette méthode est très sensible au bruit et donne des résultats très médiocres. On préfère utiliser la formule 1.2 dans sa version intégrale et monodimensionnelle, c'est-à-dire à partir de l'image d'une marche d'escalier. C'est ce que l'on appelle la *Line Spread Function*. De nombreuses versions de cette technique sont disponibles dans la littérature, faisant ou non l'hypothèse de l'isotropie de la FTM. Il suffit dans le premier cas d'une seule orientation de contours, dans le second il faut utiliser des images contenant une variété d'orientations, souvent réduite à ... 2 ou 3.

La méthode de la *Line Spread Function* a le mérite de pouvoir être assez aisément utilisée sur des images naturelles qui contiennent souvent des lignes et ceci est très précieux lorsque l'on ne dispose pas du capteur pour l'étalonner (ça peut être le cas en robotique mobile, c'est le cas en imagerie satellitaire), car beaucoup d'images naturelles possèdent de telles lignes : bord de route, bâtiment, ou pont.

La FTM a donné naissance à des indicateurs plus concis pour caractériser la résolution d'une caméra. Le plus utilisé est probablement MTF50 (ou MTF50P) qui réduit la résolution à la fréquence spatiale pour laquelle la modulation est réduite à 50 % (ou à 50 % de sa valeur maximale dans MTF50P (*peak = P*)). C'est la valeur qui est adoptée par la norme ISO 12233 [ISO, 2000].

Généralisant la démarche de mesure de la réponse impulsionnelle à partir de lignes, plusieurs techniques l'ont améliorée, en raffinant la mire pour avoir une meilleure estimation de la FTM (certaines sont même normalisées dans la norme ISO citée ci-dessus). C'est ainsi que la mire de Joshi utilise des profils circulaires de façon à mieux rendre compte de l'anisotropie potentielle du capteur.

D'autres travaux se sont proposés également d'inverser l'équation de formation de l'image sur une image connue, mais en s'appuyant sur des motifs aléatoires, qui, s'ils sont bien définis, donnent théoriquement un spectre plat. Il suffit alors de normaliser le spectre de l'image obtenue pour en déduire la FTM recherchée. Ainsi Delbracio [Delbracio-Bentancor, 2013] a démontré que toute mesure de FTM était bornée dans sa précision par une borne liée au capteur (essentiellement due au bruit affectant le capteur). Il a ainsi proposé une mire et un protocole qui permettent de s'approcher au mieux de cette borne (*cf. figure 1.1*).

Par une tout autre voie, il a été proposé d'utiliser pour mire un schéma aléatoire de feuilles mortes, présentant l'avantage d'offrir une invariance non seulement en translation, rotation et changement de dynamique, mais aussi en changeant d'échelle grâce à ses propriétés fractales [Gousseau et Roueff, 2007, Cao et al., 2010, McElvain et al., 2010].

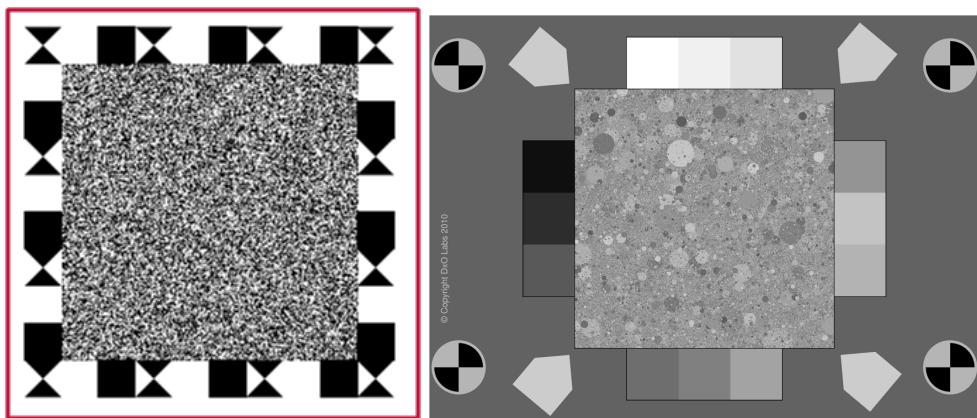
Est-ce que le problème de la résolution est résolu avec la connaissance d'une FTM en divers points du champ de l'image pour chaque distance focale, chaque ouverture et chaque distance de mise au point ? Oui, car il y a là toute la matière pour prédire une image quelle que soit la scène observée. Non car cette information très volumineuse et complexe ne permet pas de prédire simplement comment l'observateur "verra" cette image. Vaut-il mieux privilégier une bonne reproduction des fréquences basses au détriment des hautes fréquences ou l'inverse ?

### 1.1.2 Le piqué

Les photographes usent d'un terme particulier pour exprimer la qualité du contenu fréquentiel satisfaisant pour l'expert, c'est le piqué de l'image qui exprime un bon rendu des détails fins accompagné d'une bonne résolution de l'image. Malheureusement ce terme ne se traduit pas en une

---

2. en anglais *Impulse Response* ou plus souvent *Point Spread Function*, PSF.



**Figure 1.1 – Deux exemples de mires utilisées pour mesurer la résolution d'un système d'imagerie. A gauche la mire aléatoire de M. Delbracio-Bentancor, à droite la mire feuilles-mortes de DxO, issue des travaux [Alvarez et al., 1999], permettant de caractériser la préservation des textures.**

mesure que l'on peut faire sur l'image et reste affaire de spécialiste qui permet généralement de classer plusieurs photos d'une même scène prises avec des paramètres différents. On juge ainsi généralement qu'une image prise avec une ouverture intermédiaire (par exemple 5,6 ou 8) offre souvent un meilleur piqué que celles prises avec une ouverture plus forte (donc à priori avec moins de flou de diffraction) ou une ouverture plus faible (donc avec plus de profondeur de champ). Le piqué est donc une donnée subjective qui se traduit en pratique par des hautes fréquences mieux préservées (moins atténueées) dans les zones d'intérêt.

A partir d'une image de FTM donnée, il est assez simple de modifier les fréquences spatiales qui portent ces détails jugés utiles et ainsi de forcer le piqué de l'image au delà de ce qu'un système idéal pourrait donner. Ces opérations sont très classiques en photographie et ont été mises en œuvre avant même l'arrivée du traitement numérique des images.

### 1.1.3 L'acutance

C'est une grandeur peu utilisée en France, mais qui a le mérite de donner une base solide au critère de piqué dès lors que l'on fixe les conditions exactes de l'observation. Dans cette situation précise où l'on fixe non seulement la distance d'observation, mais aussi la luminosité de l'image et l'éclairage ambiant, on connaît bien la sensibilité du système visuel humain moyen au contenu fréquentiel d'une image. Elle est résumée dans la Fonction de Sensibilité au Contraste Spatial (FSCS) [Le Grand, 1972]. L'acutance prend en compte cette FSCS ainsi que la FTM pour qualifier par un chiffre unique la capacité de l'image à transmettre des détails fins utiles à la formation de l'image pour un observateur dans des conditions précises d'observation :

$$A = \int_O^{u_{max}} FTM(u) FSCS(u) du \quad (1.3)$$

Ainsi, pour une résolution fixée, l'acutance donnera plus de poids aux fréquences dans la zone de forte sensibilité de l'œil en fonction des conditions précises d'observation. L'acutance d'une même image ne sera pas identique si elle est examinée sur un écran de cinéma, de télévision ou d'ordinateur, ce qui correspond bien à l'expérience que nous avons de la qualité de l'image.

Certains ont fait remarquer que l'acutance, déterminée à partir d'une seule FTM, ne caractérisait que le centre de l'image et ont proposé d'utiliser plutôt une combinaison linéaire d'acutances en divers points de l'image. Ce raffinement ne corrige cependant pas les objections faites à l'acutance pour décrire la qualité, objections que nous allons voir ici.

#### 1.1.4 Le masquage flou

Pourquoi ne généralise t'on pas la notion d'acutance qui semble disposer à la fois d'une base expérimentale solide (la FTM mesurée) et d'une bonne prise en compte du système visuel (la FSCS) ? Parce que l'acutance est une mesure très fragile, aisément modifiée artificiellement, elle ne reflète pas toujours la qualité perçue de l'image.



**Figure 1.2 – Accroissement de l'acutance d'une image par masquage flou.** L'image à gauche est l'originale, les deux images à droite ont la même résolution, mais elles ont subi une amplification croissante des fréquences moyennes.

Pour accroître l'acutance, on exploite la grande sensibilité du piqué d'une image au contenu de certaines fréquences. Il est facile d'amener toute image, à une acutance très grande, au prix d'une forte dégradation du rapport signal à bruit. Quoique l'acutance croîsse, la qualité subjective s'en trouve donc très affectée

L'acutance est pourtant une composante précieuse de la qualité, et l'expérience montre souvent un gain de qualité lorsqu'on augmente un peu l'acutance. Ceci est illustré sur les figures 1.2. Cette propriété était d'ailleurs connue depuis fort longtemps en photographie sur film et des méthodes habiles – et complexes<sup>3</sup> – ont été alors développées pour y parvenir, nommément le masquage flou (plus connu sous son nom anglo-saxon de *Unsharp Masking*).

Le masquage flou réalise, le plus souvent par une combinaison directe des pixels autour d'un point donné, une amplification modérée des hautes fréquences de l'image<sup>4</sup>. Le masquage flou renforce les contours de l'image, rend plus apparents les détails, améliore incontestablement sa lisibilité, mais au détriment de la fidélité des rendus. Par ailleurs, il accroît le bruit, ce qui fait apparaître des détails étrangers à l'image, généralement interprétés comme du piqué, mais parfois intempestifs.

Le masquage flou est disponible dans toutes les boîtes à outils de traitement d'image.

3. en photographie, le masquage flou est obtenu en exposant le tirage à travers un négatif de l'original, atténué et légèrement flou, d'où le nom du procédé.

4. Dans ses formes les plus simples, on retire du pixel une petite fraction de l'image filtrée par une gaussienne. La proportion retirée et l'étalement de la gaussienne sont les deux paramètres du masquage flou

### 1.1.5 La quantité d'information

## 1.2 Diffraction incohérente

### 1.2.1 Cohérence - Incohérence

Rappelons tout d'abord quelques éléments de Physique concernant la formation des images et la cohérence des sources optiques. Une image est formée par la distribution, sur le plan image, des photons émis par les divers points des sources constituant l'objet. Chaque photon, dans l'espace libre, suit bien sûr les lois de propagation que l'on retrouve dans les équations de diffraction de Fresnel et de Fraunhofer [Pérez, 1991]. L'image est formée par la sommation des contributions des divers photons émis. Par définition, la cohérence est la qualité qui permet à ces photons d'interférer, c'est-à-dire de se combiner en amplitude complexe. Elle se compose de deux éléments :

1. une cohérence temporelle qui assure le maintien dans le temps des relations de phase des sources pendant le temps de la mesure ;
2. une cohérence spatiale qui assure que les divers points sources maintiennent la même relation de phase.

La cohérence temporelle est une propriété des sources monochromatiques qui se mesure par la longueur des photons émis (la longueur de cohérence) : plus les sources sont monochromatiques, plus leur bande fréquentielle est étroite et plus la durée de vie des photons est grande. La cohérence spatiale est régie par la capacité des diverses sources à émettre des ondes en phase et à maintenir cette phase au long de la prise de vue.

En photographie, en lumière naturelle comme en lumière artificielle, l'éclairage est très généralement incohérent quelles que soient les sources car ces sources ne sont cohérentes ni spatialement (les photons sont très brefs) ni temporellement (les diverses sources sont indépendantes). Les exceptions concernent :

1. l'imagerie radar car on sait faire des sources très cohérentes en micro-ondes, avec des longueurs de cohérence de plusieurs centaines de mètres,
2. l'éclairage par laser qui procure, dans le domaine visible, des ondes stimulées, très monochromatiques et des fronts d'onde plans parfaitement en phase,
3. la microscopie, car on étudie des phénomènes très petits qui peuvent rester dans la longueur de cohérence des éclairages utilisés et ces éclairages sont souvent ponctuels et uniques.

Nous ne traiterons pas ici des sources cohérentes (radar, laser), ni des sources partiellement cohérentes de la microscopie.

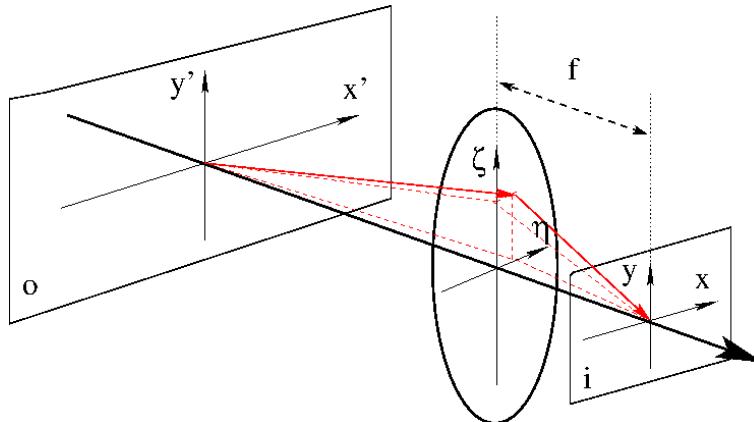
Dans le cas des sources incohérentes de la photographie ordinaire, l'image est portée par des photons dans toutes les longueurs d'ondes du domaine visible. Comme on l'a dit, ces photons ont généralement une durée de vie très courte devant le temps de pose. Ils sont émis de façon spontanée par les sources (le soleil, une lampe, ...), sans aucune relation de phase entre photons successifs et voisins. Si l'on s'intéresse à un domaine étroit de longueurs d'onde (domaine dans lequel les photons seraient susceptibles de se combiner), alors on montre que les contributions de photons sans relation de phase sont en moyenne nulles et que seules interviennent dans le bilan énergétique les interactions de chaque photon avec lui-même [Blanc-Lapierre et Dumontet, 1955].

Le cas le plus simple à étudier est celui de la photographie d'un objet situé à l'infini, dont l'image se forme au plan focal de la lentille, cas qui s'étend sans difficulté à la situation générale de formation d'image entre plans conjugués, tant que l'objet est "loin" de l'optique<sup>5</sup>.

5. Dans le cas d'un objet à distance  $p$  grande devant  $f$  (et non à l'infini), il convient de prendre en compte des termes de phase pour revenir au cas à l'infini et le résultat est essentiellement équivalent. Dans le cas où l'objet est très proche, comme en microscopie, on ne peut plus passer par la diffraction de Fresnel (à l'infini). Il faut modéliser la propagation par diffraction de Fraunhofer, la situation est alors nettement plus complexe et doit être associée au fait que la source est fréquemment partiellement cohérente. On ne peut plus appliquer les relations 1.7 et 1.8 que nous allons voir (*cf.* [Born et Wolf, 1970] p. 418).

### 1.2.2 Définitions et Notations

Nous nous plaçons dans les conditions représentées en figure 1.3. Nous réduisons l'objectif à une lentille simple et l'ouverture à son diaphragme. Nous considérons un objet plan décrit par son intensité lumineuse  $o$  (carré du module du champ électromagnétique) et une mise au point parfaite (absence d'aberrations ou d'autre défaut). Afin d'alléger les notations, nous ferons tous les calculs dans le plan image (décrit par ses coordonnées  $(x, y)$ ). Pour cela nous transporterons virtuellement l'objet (donc sans diffraction) dans ce plan. L'objet, dans son plan, serait décrit par  $o(x', y')$ . Porté dans le plan image, pour notre calcul, nous utiliserons  $o(x, y)$ , les variables  $(x, y)$  se déduisant de  $(x', y')$  par l'intermédiaire du grandissement transversal  $G$  (relation ??,  $G$  est négatif, image et objet sont inversés) :  $x = Gx'$ ,  $y = Gy'$ . Nous utiliserons abondamment la représentation du filtrage optique par transformée de Fourier introduite par [Duffieux et Lansraux, 1945, Dumontet, 1955].



**Figure 1.3 – Le montage utilisé pour calculer l'influence de la diffraction. En fait nous ferons tous les calculs dans l'espace image  $(x, y)$  afin de ne pas transporter les termes de grandissement.**

Dans le plan de la lentille, nous utilisons les variables d'espace  $(\eta, \zeta)$ . La diffraction pourra s'exprimer de façon équivalente entre l'objet et la lentille (diffraction à l'infini) ou entre la lentille et l'image (diffraction à l'infini ramenée à la distance focale  $f$  de la lentille).

L'imagerie serait donc parfaite s'il n'y avait le diaphragme qui diffracte. Notons  $K(\eta, \zeta)$  ce diaphragme (ce sera souvent la fonction cercle de diamètre  $D$ ).

Nous allons maintenant décomposer l'image en nous intéressant à chaque longueur d'onde  $\lambda$ . On peut alors appliquer à la composante  $o_\lambda$ , la théorie de la diffraction incohérente ([Born et Wolf, 1970] p. 484).

Pour la longueur d'onde  $\lambda$ , on démontre que la formation de l'image en éclairage incohérent est régie par une relation linéaire (relation de convolution) entre les intensités lumineuses dans le plan objet et dans le plan image. Cette convolution met en œuvre 2 fonctions :

1. la réponse impulsionnelle  $h_\lambda(x, y)$  qui décrit l'image d'un point objet infinitement fin<sup>6</sup>,
2. la fonction de transfert  $H_\lambda(u, v)$  qui exprime l'atténuation de la fréquence spatiale  $(u, v)$ . On la dénote souvent par FTM (fonction de transfert de modulation).

### 1.2.3 Pour une longueur d'onde unique

La transformée de Fourier  $O_\lambda(u, v)$  de l'objet s'exprime en fonction des fréquences spatiales  $(u, v)$ , par :

$$O_\lambda(u, v) = \mathcal{TF}[o_\lambda(x, y)] = \iint_{\text{objet}} o_\lambda(x, y) \exp \{-2j\pi(ux + vy)\} dx dy \quad (1.4)$$

6. la réponse impulsionnelle s'appelle en anglais *point spread function* et de ce fait est souvent désignée par PSF.

ou, en faisant apparaître les angles de diffraction  $(\eta x/\lambda f, \zeta y/\lambda f)$ , :

$$O_\lambda(\eta, \zeta) = \iint_{objet} o_\lambda(x, y) \exp \left\{ \frac{-2j\pi(\eta x + \zeta y)}{\lambda f} \right\} dx dy \quad (1.5)$$

Les fréquences spatiales  $(u, v)$  s'expriment naturellement en fonction des variables d'espace  $(\eta, \zeta)$  sur la lentille par les relations :

$$u = \frac{\eta}{\lambda f} \quad \text{et} \quad v = \frac{\zeta}{\lambda f} \quad (1.6)$$

Par filtrage incohérent par le diaphragme l'intensité de l'image est obtenue comme la convolution de  $o_\lambda$  par la réponse impulsionale incohérente  $h_\lambda$  du diaphragme :

$$i_\lambda(x, y) = o_\lambda(x, y) * h_\lambda(x, y) = \iint_{objet} o_\lambda(x_o, y_o) h_\lambda(x - x_o, y - y_o) dx_o dy_o \quad (1.7)$$

où la réponse impulsionale incohérente,  $h_\lambda$ , est le carré du module de la réponse impulsionale cohérente [Born et Wolf, 1970] :

$$h_\lambda(x, y) = |k_\lambda(x, y)|^2 \quad (1.8)$$

$k_\lambda(x, y)$  est la transformée de Fourier inverse de la pupille :

$$k_\lambda(x, y) = \mathcal{T}\mathcal{F}^{-1}(K(u, v)) = \iint_{lentille} K(\eta, \zeta) \exp \left\{ \frac{2j\pi(\eta x + \zeta y)}{\lambda f} \right\} d\eta d\zeta \quad (1.9)$$

La relation 1.7 s'exprime également dans le domaine de Fourier :

$$I_\lambda(u, v) = O_\lambda(u, v) \cdot H_\lambda(u, v) \quad (1.10)$$

de même que la relation 1.8 selon les fréquences spatiales :

$$H_\lambda(u, v) = K(u, v) * K^*(-u, -v) \quad (1.11)$$

ou selon les variables d'espace :

$$H_\lambda(\eta, \zeta) = K(\eta, \zeta) * K^*(-\eta, -\zeta) \quad (1.12)$$

La FTM est une caractéristique importante du système optique : elle mesure comment une fréquence spatiale pure  $(u, v)$  est atténuée par le système :

$$H_\lambda(u, v) = \frac{I_\lambda(u, v)}{O_\lambda(u, v)} \quad (1.13)$$

Elle peut se mesurer en photographiant des mires sinusoïdales de pas variable (donc des fréquences  $(u, v)$  pures) et de contraste fixé et en mesurant le contraste obtenu dans l'image.

#### 1.2.4 Diaphragme circulaire

Examinons le cas pratique le plus courant en photographie, celui d'un diaphragme circulaire. On a dans ce cas tout à fait intérêt à passer en coordonnées polaires puisque le système optique a la symétrie de révolution ; posons :  $\rho^2 = \eta^2 + \zeta^2$ ,  $w^2 = u^2 + v^2$  et  $r^2 = x^2 + y^2$  :

$$\begin{array}{ccc} K(\eta, \zeta) & \longrightarrow & K(\rho) = \text{ cercle}(2\rho/D) \\ h_\lambda(x, y) & \longrightarrow & h_\lambda(r) \end{array}$$

Partant de l'équation 1.8, et connaissant la TF de la fonction cercle, on trouve que :

$$h_\lambda(x, y) = \left| \mathcal{TF}^{-1}(\text{cercle}(2\rho/D)) \right|^2 \quad \rightarrow \quad h_\lambda(r) = \left[ \frac{2J_1(\pi r D / \lambda f)}{\pi r D / \lambda f} \right]^2 \quad (1.14)$$

où  $J_1$  est la fonction de Bessel de première espèce.

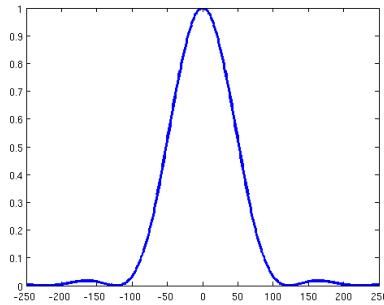
Le réponse impulsionnelle  $h_\lambda$  est appelée tache d'Airy (cf. figure 1.4). Elle s'annule pour  $r_o D / \lambda f = 0,61$  soit :

$$r_o = 0,61 \frac{\lambda f}{D} = 0,61\lambda N \quad (1.15)$$

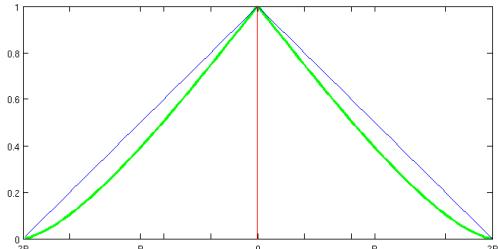
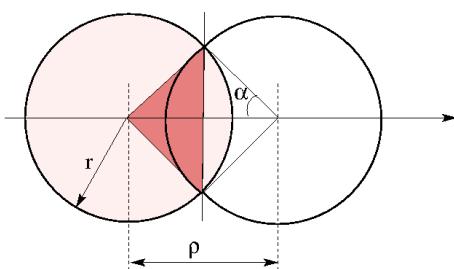
où  $N$  est le nombre d'ouverture de l'objectif. Le calcul de la fonction de transfert  $H_\lambda(u, v)$  se fait directement à partir de la relation 1.11 et à l'aide de la figure 1.5 à gauche. On obtient<sup>7</sup> :

$$H_\lambda(\rho) = \frac{D^2}{4} \left[ 2\text{Arc cos}\left(\frac{\rho}{D}\right) - \sin\left(2\text{Arc cos}\left(\frac{\rho}{D}\right)\right) \right] \quad (1.16)$$

C'est une fonction à symétrie de révolution, régulièrement décroissante selon une loi presque linéaire et qui s'annule pour  $\rho = \sqrt{(\eta^2 + \zeta^2)} = D$  (cf. figure 1.5 à droite).



**Figure 1.4 –** A gauche : diffraction incohérente monochromatique par un diaphragme circulaire, ou tache d'Airy. Le zéro de cette fonction se trouve à la position  $x' = 0,61\lambda f/D = 0,61N\lambda$ ,  $\lambda$  étant la longueur d'onde ( $\lambda = 0,5\mu\text{m}$  pour le vert),  $f$  la longueur focale de la lentille,  $D$  l'ouverture du diaphragme et  $N$  le nombre d'ouverture. Pour un objectif de 50 mm ouvert à  $f/2$ , la figure de diffraction a donc un diamètre de l'ordre de 1,2 microns. A droite : pouvoir de séparation. C'est la distance limite permettant de séparer 2 objets.



**Figure 1.5 –** A gauche : calcul de la fonction de transfert incohérente du diaphragme circulaire. La valeur de  $H$  est égale à l'aire de la lunule entre les 2 cercles. A droite cette fonction de transfert, en vert (on a tracé en bleu la décroissance linéaire d'un cône de diamètre  $4R$ ). En éclairage cohérent, cette fonction de transfert aurait vallu 1 entre  $-R$  et  $R$ .

7. merci Saïd !

### 1.2.5 Discussion

Les équations 1.8 et 1.11, illustrées dans le cas d'un diaphragme circulaire par les équations 1.14 et 1.16, méritent d'être discutées car dans leur simplicité, elles cachent des subtilités.

- nous avons exprimé  $H_\lambda(\eta, \zeta)$ , de façon un peu incorrecte, en fonction des variables d'espace mesurées directement sur l'objectif  $(\eta, \zeta)$ , de façon à bien montrer que la fonction de transfert est directement liée au diaphragme, puisqu'elle en est la fonction d'autocorrélation. Il suffirait donc de remplacer le cercle de la figure 1.5 pour l'adapter exactement à un diaphragme, à 9 lames curvilignes par exemple, si on le souhaitait pour obtenir une FTM plus adaptée à un appareil photo donné ;
- en toute rigueur,  $H_\lambda$  est en fait une fonction des fréquences spatiales  $(u, v)$  qui décrivent l'objet  $o(x, y)$  ou l'image  $i(x, y)$ . C'est par l'intermédiaire des fréquences spatiales qu'intervient la longueur d'onde  $\lambda$ .  $\lambda$  n'apparaît pas dans la formule 1.16, par exemple, mais on peut néanmoins la faire réapparaître en la réécrivant sous la forme :

$$H_\lambda(\lambda fw) = \frac{D^2}{4} \left[ 2\text{Arc cos}\left(\frac{\lambda fw}{D}\right) - \sin(2\text{Arc cos}\left(\frac{\lambda fw}{D}\right)) \right]. \quad (1.17)$$

- les fréquences spatiales ne sont pas un artifice de calcul mathématique introduit par la transformée de Fourier, mais des grandeurs physiques, directement liées à la propagation de l'onde électromagnétique lors de la formation de l'image, grandeurs étroitement liées à l'inclinaison des rayons optiques en interaction avec le diaphragme.
- ainsi s'explique que si la fonction d'autocorrélation du diaphragme  $K(\eta, \zeta)$  est indépendante de la longueur d'onde, la FTM qui en est issue lors du passage de la lumière à travers l'objectif  $H_\lambda$ , elle, en dépend, et donc la réponse impulsionale  $h_\lambda$ . Plus  $\lambda$  est grand, plus la réponse impulsionale est large (et donc l'image floue).
- Remarquons que l'expression 1.7 est symétrique de celle que l'on aurait obtenue en éclairage cohérent [Pérez, 1991], mais où cette fois les grandeurs  $o$  et  $i$  seraient les amplitudes complexes de l'onde. Dans ce cas cohérent  $h(x, y) = k(x, y)$ , les calculs se font en amplitude complexe, et l'on ne repasse en intensité qu'au moment de la détection : le carré du module de  $i(x, y)$  :  $|i(x, y)|^2$  permet de former l'image à partir de l'onde complexe.

On retiendra que la fonction de transfert du système en optique incohérente étant l'autocorrélation de la fonction de transfert cohérente, elle est 2 fois plus étalée que celle de l'optique cohérente. Peut-on en déduire qu'elle sera plus riche en détails ? Oui et non. Oui, car sa fréquence de coupure sera 2 fois plus élevée et l'image incohérente aura donc des hautes fréquences qui ne seront pas contenues dans l'image cohérente. Non, car son module sera généralement inférieur à celui de la FTM cohérente (en vertu du théorème de Schwartz, et compte tenu du fait que la FTM a un module strictement inférieur à 1). Les fréquences différentes du continu seront donc généralement plus atténuerées dans le cas incohérent<sup>8</sup>

### 1.2.6 Cas d'un spectre large

Reprendons ici la construction d'une image en lumière blanche, c'est à dire avec un spectre large de longueurs d'onde.

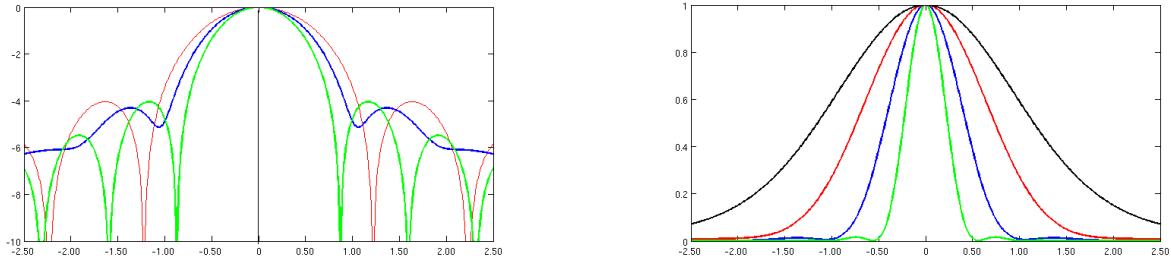
Chaque composante X ( $X = R, V$  ou  $B$ ), contribuant aux 3 canaux de l'image, est elle-même composée d'un ensemble de longueurs d'ondes qui subiront toutes un filtrage un peu différent. En prenant en compte la sensibilité  $\sigma(x, y, \lambda)$  du capteur (par exemple à l'aide des courbes telles que celles présentées en figure ??), on obtient une composante X de l'image donnée par :

$$i_X(x, y) = \int_{\lambda_{min-X}}^{\lambda_{max-X}} o_\lambda(x, y) * h_\lambda(x, y) \sigma(x, y, \lambda) d\lambda \quad (1.18)$$

---

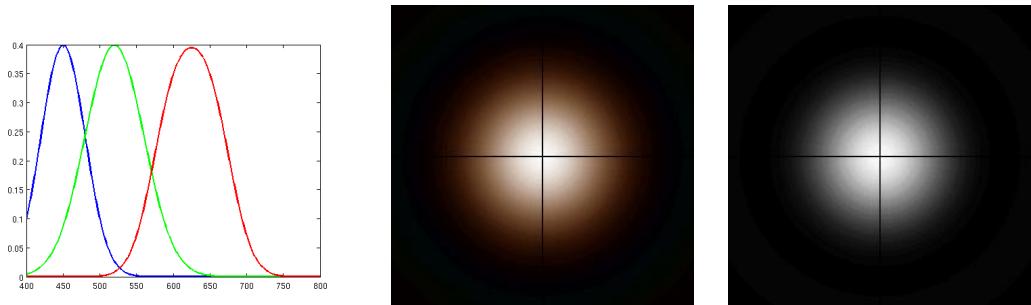
8. Il faut aussi souligner que l'imagerie cohérente souffre généralement d'un défaut plus important : elle est affectée d'un bruit de granularité (le *speckle*) dû aux interférences des ondes cohérentes en raison des micro-irrégularités des surfaces. Il est en pratique très difficile de s'affranchir de cette granularité qui affecte considérablement les très hautes fréquences de l'image (cf. [Goodman, 1976]).

On présente sur la figure 1.6 à gauche la réponse impulsionnelle à 2 longueurs d'onde limites dans le cas d'un diaphragme circulaire :  $\lambda = 0,7\mu m$  (rouge) et  $\lambda = 0,5\mu m$  (vert), ainsi que la psf moyenne d'un spectre qui serait uniforme entre ces deux valeurs. On note que la réponse impulsionnelle moyenne, à la différence des réponses impulsionnelles monochromatiques ne s'annule plus. Enfin on présente dans ces mêmes conditions, en échelle linéaire, les psf globales à diverses ouvertures (figure 1.6 à droite).



**Figure 1.6 – Diffraction (à gauche en échelle logarithmique) d'une lumière uniformément répartie entre le rouge ( $0,7 \mu m$ ) et le vert ( $0,5 \mu m$ ). L'objectif circulaire est ouvert à  $f/2,8$ . A droite, diffraction d'une lumière uniformément répartie (en échelle linéaire) pour divers nombres d'ouverture : vert :  $N=1,4$ , Bleu :  $N=2,8$ , Rouge :  $N=5,6$  et Noir :  $N=11$ . L'échelle des abscisses est en microns.**

On présente sur la figure 1.7 l'exemple d'une psf en lumière blanche pour un spectre plat entre  $0,4 \mu m$  et  $0,8 \mu m$ . A gauche : les 3 courbes colorimétriques des capteurs ; au centre, la psf en lumière blanche ; à droite la psf telle qu'elle serait si toutes les longueurs d'onde avaient la même psf (une psf moyenne à  $540 nm$ ). On note à droite, un figure uniformément monochrome, tandis qu'au centre, si la tache de diffraction est majoritairement blanche, les longueurs d'onde rouge, qui décroissent moins vite sont dominantes au bord de la psf, donnant une teinte rougeâtre au pourtour de la tache. Notons que l'anneau du second maximum n'est pas visible avec une dynamique linéaire.



**Figure 1.7 – Psf obtenue avec une lumière blanche. Le capteur dispose de 3 CCD sensibles selon les courbes de la figure gauche. Le spectre de la lumière est uniforme sur la bande visible de  $0,4$  à  $0,8$  microns. Au centre la psf obtenue par sommation des contributions de toutes les longueurs d'onde : la tache est dominée par les longueurs d'onde rouges en dehors de l'axe. A droite hypothèse d'une diffraction moyenne identique pour toutes les longueurs d'ondes. La tache serait monochrome.**

Afin de conduire une étude exacte de l'effet de la diffraction sur la formation de l'image, il faudrait enfin prendre en compte le filtrage par les filtres de la matrice de Bayer, ainsi éventuellement que par les filtres infra-rouge et ultra-violet.

Il n'y a donc pas véritablement de réponse impulsionnelle globale d'une optique en éclairage polychromatique, mais un mélange de réponses impulsionnelles, mélange qui change avec le contenu spectral de l'objet imagé.

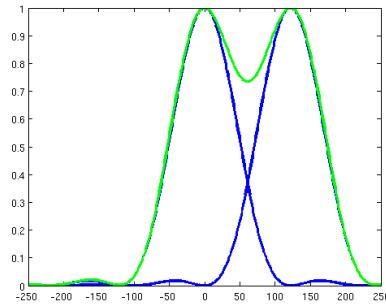
La formule 1.14 utilisée pour les longueurs d'onde vertes ( $\lambda$  de l'ordre de  $0,5\mu m$ ) donne cependant une bonne valeur moyenne de la réponse impulsionnelle due au diaphragme (ce que l'on nomme

généralement « la diffraction par le diaphragme »).

### 1.2.7 Pouvoir de séparation

Le pouvoir de séparation ou de résolution d'un système optique est la distance minimale qui doit séparer deux points images pour qu'ils soient vus séparément. Le pouvoir de résolution, en pratique, prend en compte tous les défauts de l'optique. Mais comme il a pour limite théorique la diffraction par le diaphragme, le pouvoir de séparation d'une ouverture circulaire a fait l'objet d'une étude particulière.

Par convention, on situe la limite de résolution d'une optique circulaire à l'étendue de la tache d'Airy. Deux points images seront dits séparés dès que la position de l'un sera au delà du premier zéro de la tache d'Airy de l'autre, (*cf.* figure 1.8). Notons alors que le minimum est environ à 70 % du maximum.



**Figure 1.8 – Pouvoir de séparation d'un système optique limité par un diaphragme circulaire.** Par définition, c'est la distance limite permettant de séparer 2 objets et par convention, lorsque la psf est une tache d'Airy, c'est la distance au premier minimum de la Bessel.

## Chapitre 2

# Le codage des images et leurs formats

La diversité des formats de présentation des images constitue l'un des freins à leur usage au sein de la société. Si le format JPEG a acquis aujourd'hui pour l'usager de la photographie, quotidien, ou plus encore occasionnel, un statut particulier qui lui donne, de fait, l'apparence d'une norme, de nombreux autres formats sont disponibles - et parfois préférables.

Le recours à ces formats est dicté par deux objectifs différents et bien-sûr contradictoires : diminuer le volume de l'image ou mieux respecter ses propriétés : dynamique, palette chromatique, géométrie des lignes, qualité des aplats, etc. Les applications concernées par l'image sont très nombreuses et elles sont aussi très diversifiées certaines imposent des échanges de documents très rapides tandis que d'autres s'attachent à leur qualité qui pourra être jugée avec des critères différents : respect des nuances ou précision des détails, finesse des tracés, reproductibilité des couleurs, lisibilité des textes, capacité à entrer dans une composition complexe, etc.

Rappelons qu'une photo ordinaire est aujourd'hui constituée d'une dizaine de millions de pixels, chacun codé souvent sur 3 octets, un par canal chromatique, rouge, vert ou bleu. Si l'on ne prend pas de mesure pour compresser le signal, on est donc amené à stocker, aussi bien dans l'appareil que dans l'ordinateur quelques dizaines de millions d'octets par image. Par ailleurs il est devenu très simple d'enregistrer des centaines, voire des milliers d'images à l'occasion d'un évènement. Un amateur passionné ou un professionnel maintiennent donc des téra-octets d'images. Au delà du coût de l'archivage, la consultation et le transfert des données exigent des volumes aussi réduits que possible. Les nombreux formats développés reflètent ces besoins [Barlaud et Labit, 2003].

### 2.1 Format RAW : brut ou natif

Le format RAW est celui qui respecte le mieux le signal mesuré par le photodétecteur. Il est constitué d'un part des signaux issus des convertisseurs analogiques-numériques et d'autre part des paramètres de la prise de vue qui permettent de reconstruire le signal en connaissance des conditions de leur acquisition.

Quoiqu'il existe une norme RAW standardisée par l'ISO, chaque fabricant maintient sa propre version et la fait évoluer au rythme de ses besoins. Les divers formats disposent d'une extension spécifique dans le nom du fichier (au nombre d'une centaine aujourd'hui). Ces formats ne sont donc généralement pas lisibles avec les outils ordinaires d'affichage, mais beaucoup de bibliothèques de traitement d'images disposent des formats les plus courants. La transformation d'un fichier RAW en un autre format est assurée par le fabricant de l'appareil photo lors du transfert de la mémoire de l'appareil à un ordinateur. C'est souvent la meilleure conversion disponible. Les bibliothèques

proposent parfois des algorithmes différents de ceux du fabricant, qui conduisent alors à des reconstructions légèrement différentes : elles peuvent être meilleures car des algorithmes plus complexes peuvent-être mis en œuvre, par exemple pour la colorimétrie ou pour le démosaïcage. Par ailleurs, ces bibliothèques peuvent intégrer des effets que la caméra ne prend pas en compte : vignettage optique (voir paragraphe ??), vignettage de pixel (voir paragraphe ??), aberrations géométriques, etc. La reconstruction peut alors être supérieure à celle du fabricant.

Le format RAW est bien sûr très volumineux puisqu'il contient non-seulement tous les signaux d'image, mais aussi toutes les informations pour reconstruire une image de qualité. Il est généralement constitué des éléments suivants :

1. une entête comportant l'ordre des octets (petit-indien ou grand indien), le nom de l'image, l'adresse des données brutes ;
2. les métadonnées décrivant le capteur : nombre de pixels, type de matriçage couleur et profile colorimétrique,
3. les métadonnées décrivant l'image : date, type d'appareil, d'objectif, temps de pose, ouverture, focale, réglages, etc.
4. une vignette permettant de consulter la mémoire, et éventuellement une image réduite JPEG permettant de l'afficher en une taille raisonnable,
5. les données d'image généralement présentées dans la géométrie du capteur (entrelacement RVB de Bayer, éventuellement quinconce des pixels).

### 2.1.1 Vers une normalisation du RAW

Le format normalisé par l'ISO (ISO 12234-2 [ISO, 2001]) est dénommé TIFF/EP (pour *Tagged Image File Format - Electronic Photography*) et dérive du format TIFF 6.0 (voir paragraphe 2.1.5) plus répandu mais propriété de Adobe. Il autorise des représentations sans perte ou avec perte (en particulier avec codage JPEG / DCT (voir paragraphe 2.5.1)). Il se voudrait donc un sur-ensemble des formats d'image.

La normalisation ISO n'est pas suivie par la plupart des fabricants qui s'en tiennent donc à des formats propriétaires. Néanmoins, les formats adoptés sont souvent dérivés de la norme TIFF/EP qu'ils étendent.

La partie traitant du codage sans perte de TIFF/EP est compatible en particulier avec un format proposé par Adobe, mais ouvert et libre d'accès : le format DNG (*Digital Negative*) qui dispose de bibliothèques de traitement publiques.

### 2.1.2 Que contient un format RAW ?

Le format RAW doit refléter les éléments suivants qui font la diversité des capteurs :

- la disposition des photosites : sur maille carrée, rectangulaire ou hexagonale, ou leur configuration spécifique pour des besoins particuliers (voir paragraphe ??),
- la disposition des filtres chromatiques, selon une matrice de Bayer, ou selon d'autres matrices (paragraphe ??), ou en configuration trichrome vraie (paragraphe ??),
- la composition précise des filtres chromatiques
- la balance des blancs telle qu'utilisée par le capteur,
- la dynamique du signal sur 8 bits, 10 bits ou 12 bits,
- la position du pixel afin de tenir compte des effets d'aberration et de vignettage,
- les paramètres de la chaîne électronique pour déterminer le niveau de bruit.

Ces éléments peuvent également se retrouver dans un fichier de métadonnées disponible pour les images d'autres format (par exemple le fichier EXIF, voir paragraphe 2.2), mais dans le fichier RAW se trouvent également des informations plus spécifiques au capteur (en particulier sur les paramètres fins qui régissent l'électronique : bruit thermique, taux de correction du vignettage pixellique) que le

fabricant ne souhaite pas toujours rendre public. Pour cela les métadonnées des fichiers RAW sont parfois cryptés de façon à n'être pas lus par des logiciels que ne contrôle pas le fabricant.

Enfin certains constructeurs ont choisi des formats RAW dont les fichiers images subissent un codage sans perte (type Liv-Zampel) permettant de gagner quelques pourcents de volume sur les données (ceci est possible car les images codées sur 16 bits utilisent de façon très inégale les 65 536 niveaux de gris disponibles).

Il est assez aisément de déterminer la taille qu'occupera un fichier RAW. Tout d'abord, le fichier d'en-tête est de quelques milliers d'octets et donc généralement de taille négligeable devant le fichier image. Si l'image a  $n$  pixels, chaque pixel délivre un seul signal dans un des canaux R, V ou B, signal généralement codé sur 2 octets ou moins. Sans compression la taille du fichier sera alors au plus de l'ordre de  $2n$  octets. Avec un algorithme de compression sans perte, on atteindra un rapport de compression de l'ordre de 2 et le fichier sera environ de taille 20 Mo. Cette compression, si elle a été retenue par le fabricant s'appliquera d'office.

### 2.1.3 Intérêt du RAW

Comme nous l'avons dit, le format RAW est celui qui préserve le mieux la qualité du signal capturé par l'appareil. Par ailleurs, il fournit dans les métadonnées tous les éléments indispensables à un traitement efficace des images.

A partir de ces éléments, il sera possible de procéder aux opérations de démosaïcage, de correction des pixels aveugles, de suppression du bruit électronique, de correction chromatique, de correction des distorsions, de rattrapage des défauts d'exposition, de changement d'illuminant et d'espace de couleur, de compression de dynamique pour repasser à un signal sur 8 bits et éventuellement de codage avec perte (JPEG).

Ces éléments peuvent être parfois accomplis dans la caméra, mais plus sûrement par des logiciels déportés dans un ordinateur, soit propriétaires, soit spécialisés dans le post-traitement des images (PhotoShop, LightRoom, Gimp, DxO, ...).

### 2.1.4 Les limites du RAW

Si le format RAW garantit la meilleure fidélité de rendu de l'image, on doit lui opposer les défauts suivants :

- le volume des images est souvent prohibitif; pour un usage grand public, il multiplie par 10 la capacité de stockage nécessaire;
- la taille de l'image réduit la cadence de prise de vue car le bus qui conduit du capteur à la carte mémoire a un débit limité;
- les temps d'affichage ou de navigation dans les bases d'images est considérablement accru ;
- le manque de standardisation du format RAW laisse planer un doute sur la capacité d'accéder dans le futur aux images stockées de cette façon. Cette menace est d'autant plus sérieuse que la restructuration très rapide des fabricants d'appareils photos fait disparaître des acteurs majeurs qui avaient protégé leurs logiciels de lecture et d'écriture par des codages cryptés.

### 2.1.5 Format RAW via TIFF/EP

Nous venons de voir qu'Adobe portait le format TIFF (pour *Tagged Image File Format*) dont une version TIFF/EP (*electronic photography*) est librement accessible pour les utilisateurs des fichiers RAW. Le format TIFF couvre cependant très largement le champ des documents graphiques et sera plus en détail ci-dessous.

Voyons comment cette version TIFF/EP permet de transporter des données RAW de façon plus interchangeable.

Dans l'exemple cité plus haut d'un fichier RAW issu d'un capteur de Bayer, reconstruit puis transcodé en TIFF. Son codage se fera dans TIFF/EP sur 2 octets par canal et par pixel, soit 6 octets par pixel et donc  $6n$  octets pour l'image (120 Mo pour une image de 20 Mpixels). L'application d'une compression sans perte est alors souhaitable pour ramener cette taille à des nombres plus raisonnables : on peut alors espérer une compression d'un rapport 2 ou 3.

Trois types de codage sans perte sont disponibles : un codage de Hufmann par plages, un codage (de type Hufmann) par blocs, le codage Liv-Zampel depuis qu'il est du domaine public. En compression avec perte, seul JPEG/DCT est disponible.

Le format TIFF est capable de supporter une représentation d'image par tuiles, facilité très précieuse pour naviguer sur les très gros fichiers (comme en imagerie aérienne ou satellitaire). Les tuiles sont alors des sous-images, toutes de même taille, chacune pointant sur ses voisines pour permettre une navigation fluide.

## 2.2 Le format d'entête Exif

Le fichier Exif (*exchangeable image file format*) est initialement un format d'échange d'images et de sons, adapté pour l'audiovisuel et composé de 2 parties : une partie compression des données et un partie méta-données. Il est surtout utilisé aujourd'hui pour le fichier des méta données qui accompagne une image. Pour le signal on lui préfère JPEG, TIFF ou PNG (qu'il incorpore de fait dans ses spécifications). Pour les métadonnées, on retrouve au contraire les éléments du fichier Exif au cœur de nombreux autres standards : TIFF, JPEG, PNG, etc.

Mis en place par les industriels japonais à la fin des années 90, il s'est naturellement retrouvé dans tous les matériels et même s'il n'est pas normalisé ni maintenu par une organisation, il persiste comme standard de fait. Malheureusement il a tendance à dériver et perdre son universalité.

Exif s'inspire beaucoup de la structure des fichiers TIFF. C'est un fichier qui s'inscrit à l'intérieur du fichier image (JPEG ou TIFF). Malheureusement, tous les programmes de manipulation d'images, s'ils reconnaissent le fichier Exif au sein d'une image, ne le transmettent pas toujours dans les images dérivées, ou le transmettent de façon partielle seulement entraînant une dilution des informations Exif.

### 2.2.1 Les données du fichier Exif

Les données inscrites dans le fichier Exif sont relatives à plusieurs domaines :

1. les caractéristiques du matériel
  - boîtier : fabricant, modèle
  - objectif : fabricant, modèle
  - numéro de référence de l'appareil
  - logiciel : version
  - version d'Exif
  - flash : fabricant, modèle
  - version de FlashPix (logiciel de tuilage)
  - dimension du commentaire de l'utilisateur
2. les conditions de la prise de vue
  - date, heure et lieu (si les images sont géocodées)
  - orientation de l'appareil (format paysage ou portrait, éventuellement orientation par rapport au Nord magnétique)
  - position respective des canaux chromatiques
  - configuration des composants
  - type de compression et paramètres associés si JPEG
  - résolution en  $x$ , en  $y$  et unité employée (mm ou inch)

- distance focale, distance de mise au point
  - nombre de pixels en  $x$  et en  $y$
  - choix de la mesure de lumière
  - nombre d'ISO de la sensibilité
  - choix de la mesure de distance
3. les réglages internes de l'appareil
    - espace de couleur : CMYK ou sRGB
    - paramètres de balance des blancs
    - correction d'exposition
    - temps de pose, ouverture
    - contrôle du gain, de la saturation, du contraste, de la netteté, du zoom numérique
    - indice d'interopérabilité et version
    - programme choisi
    - traitement personnalisé
  4. les données propres au logiciel de traitement
    - date du traitement
    - logiciel
    - fonction et paramètres
    - traitement personnalisé

### 2.2.2 Limitations du fichier Exif

Le fichier Exif est un élément précieux pour une utilisation rationnelle et automatisée des images. Il pèche néanmoins par plusieurs défauts :

- Exif ne considère que des composantes chromatiques codées sur 8 bits, ce qui est insuffisant pour les appareils modernes.
- Exif n'a pas prévu de traiter des fichiers RAW, poussant les fabricants à utiliser des formats de repli : TIFF ou DNG.
- Exif expose son utilisateur à une perte de confidentialité et de vie privée car s'il ne prend pas des dispositions explicites, ses photos comporteront des éléments précis sur les conditions de prise de vue : référence unique à l'appareil photo, date de prise de vue et de plus en plus souvent géolocalisation. XKeyScore par exemple est un logiciel de traitement des fichiers Exif associés aux images postées sur le web ou transitant par mail.
- Exif n'étant pas maintenu par une instance officielle, il évolue de façon anarchique sous la pression des diverses innovations qui en adaptent des morceaux sans souci de compatibilité avec les autres applications.

## 2.3 Les formats comprimés sans perte

La compression sans perte est une opération inversible qui permet de diminuer le volume occupé par les images sans porter atteinte à leur qualité. Ils recherchent les redondances présentes dans l'image et s'appuient sur la théorie de l'information de Shannon pour viser la limite entropique de volume [Guillemot et Pateux, 2003, Prost, 2003].

Les images en Noir et Blanc ne présentent pas beaucoup de redondance si on les examine pixel par pixel. Leur entropie est souvent supérieure à 6 bits/pixel indiquant un gain possible inférieur à 2 bits par pixel. Les images en couleur présentent généralement une plus forte redondance et l'on peut souvent atteindre un rapport de 2. Mais on peut également rechercher les redondances entre pixels en codant des groupes de pixels (plages ou blocs). Il est possible alors d'exploiter la corrélation spatiale qui est souvent forte. Cela se fait au coût d'une plus grande complexité du codage, qui s'accompagne de temps importants tant au codage qu'au décodage, et à la nécessité de disposer de mémoires importantes pour garder les mots de code. Ces tables doivent aussi être transmises ou

conserver et leur coût fait chuter le bénéfice de la compression si les images ne sont pas de très grande taille.

Les algorithmes employés pour le codage sans perte sont tout d'abord ceux dont disposent tous les informaticiens, ce sont ensuite des formats dédiés à l'image ou plus généralement au graphisme.

### 2.3.1 Les algorithmes généraux de codage sans perte

Ce sont des techniques de codage qui exploitent la redondance de séquences de bits, sans considérer l'origine de ces séquences. Elles sont disponibles de façon banalisée dans les systèmes d'exploitation des ordinateurs [Guillemot et Pateux, 2003].

- le codage de Huffmann, code à longueur variable, qui associe des mots de code très courts à des messages très fréquents (le message est ici la valeur d'un pixel ou d'un ensemble de pixels). Le code de Huffmann est optimal pour des messages indépendants les uns des autres. Lorsque les messages sont corrélés, Huffmann ne tire profit de cette corrélation qu'en associant les messages dans des super-messages. Il n'y a donc pas de modélisation des dépendances et ce n'est pas très efficaces pour les images. C'est par contre très efficace pour le codage des graphiques et des dessins (avec donc des aplats sur des plages ou des blocs et ce codage permet alors des gains importants pour une complexité informatique modique. On n'est pas surpris de voir donc le codage de Huffmann par plages (*run-length*) ou par zones dans de nombreuses normes (dont TIFF).
- le codage arithmétique ne sépare pas l'image en pixels mais code tout le flot de bits de l'image par une représentation arithmétique unique adaptée aux divers états du signal. Le codage arithmétique n'est pas employé en traitement d'image car il n'y a pas de bon modèle d'apparition des symboles.
- le codage de Liv-Zempel (codes LZ77 et LZ78) qui utilise une fenêtre glissante balayant l'image et un dictionnaire qui garde trace des chaînes élémentaires la première fois qu'il les rencontre et qui remplace les occurrences ultérieures par l'adresse de la chaîne dans le dictionnaire. Le codage de LZ est maintenant utilisé en routine pour coder les très gros fichiers à forte redondance comme les fichiers TIFF.
- des codages hybrides utilisant des références arrière à la façon de Liv-Zempel, conjuguées avec des représentations plus courtes pour les séquences plus fréquentes, représentations construites à l'aide d'arbres de préfixes. De tels codes sont devenus très populaires sous des noms génériques : deflate, zip, ...
- des codages plus spécifiques à l'image ont été développés, tirant profit explicitement des dépendances bidimensionnelles et hiérarchiques des images, on les trouvera détaillés dans [Prost, 2003]. Ils ont conduit à des codeurs comme Ebcot qui sert à la norme JPEG 2000 et sera examiné au paragraphe 2.5.2.

Tous ces codages sont asymptotiquement très proches de la limite entropique de Shannon sur des séquences aléatoires. Ils font payer le poids de leur entête pour des messages courts. Mais les images ne sont pas des messages courts. S'ils ne sont pas très efficaces sur les images, c'est que la redondance statistique y est en général assez faible dans une représentation par flot de pixels. Elle est cachée dans la structure même de l'image.

### 2.3.2 Les codages sans perte de JPEG

Des algorithmes spécifiques pour le codage d'images ont donc été développés qui s'appuient sur les propriétés particulières bidimensionnelles de l'image. Ils ont un certain succès dans des domaines comme l'imagerie médicales où l'on exige des compressions sans perte. Ils sont beaucoup utilisés dans le domaine des arts graphiques où l'on manipule conjointement images par balayage, textes et dessins [Guillois et al., 2003].

Les approches les plus souvent employées s'appuient sur la prédiction d'un pixel  $X$  à la position  $(i, j)$  par ses voisins immédiats déjà transmis, donc connus du récepteur, en particulier ceux que l'on

désigne par  $A(i, j-1)$ ,  $B(i-1, j-1)$ ,  $C(i, j-1)$  et  $D(i, j+1)$  (figure 2.1). La différence entre la valeur prédite et la valeur vraie est souvent très faible et suit une loi de distribution souvent gaussienne que l'on peut coder avec profit avec un codage de Huffmann.

Ainsi, la norme JPEG sans perte de 1993 (JPEG *lossless*), propose t'elle un algorithme de prédiction de  $X$  conditionnel à la présence d'un contour horizontal ou vertical. Huit cas de figures sont répertoriés qui conduisent à huit décisions différentes. Le codage du résidu se fait par Huffmann. Des taux de compression pouvant aller jusqu'à 3 sont obtenus en imagerie médicale, mais ils sont souvent plus faibles en imagerie naturelle où JPEG sans perte n'a jamais eu de grand succès.

Un second codeur pour JPEG a ensuite été proposé : JPEG-LS qui s'appuie sur un schéma type de prédiction LOCO-1 qui dénombre les configurations à partir d'un prédicteur dénommé *median edge detector*. 365 configurations sont recensées et la prédiction est faite sous l'hypothèse de loi de distribution de Laplace de l'erreur. Un codage de Golomb-Rice est alors utilisé pour comprimer la séquence transmise (c'est un codage adapté à des messages pour lesquels les faibles valeurs sont très fortes. Il est optimal pour les séquences géométriques (loi de Laplace)). Il a des performances supérieures à JPEG lossless et peut atteindre 3 sur des images naturelles. Il a fait l'objet d'une normalisation ISO-14495-1.

Avec la norme JPEG 2000 (noté parfois JPEG 2000<sub>R</sub> pour reversible), un codage sans perte a ensuite été développé, qui utilise un codage par ondelettes bi-orthogonales 5/3 (ondelettes de Le Gall avec un filtre passe-bas à 5 coefficients rationnels et un filtre passe-haut à 3 coefficients rationnels). Ce troisième codage, JPEG sans perte, a le mérite d'être scalable progressif et compatible avec les facilités offertes par JPEG 2000 pour manipuler les images (voir paragraphe 2.5.2). A la différence de JPEG-LS, il utilise le même principe de codage avec ou sans perte mais il est plus lent et de mise en œuvre plus complexe. Il n'a pas de meilleures performances de compression que JPEG-LS.

Pour les images en couleur, JPEG 2000<sub>R</sub> propose la représentation suivante :

$$Y = \lfloor \frac{R + 2V + B}{4} \rfloor \quad C_B = B - V \quad C_R = R - V \quad (2.1)$$

(où  $\lfloor x \rfloor$  exprime l'arrondi à l'entier immédiatement inférieur à  $x$ ), qui peut être reconstruit par :

$$V = Y - \lfloor \frac{C_B + C_R}{4} \rfloor \quad R = C_R + V \quad B = C_B + V \quad (2.2)$$

## 2.4 Les formats d'image issus des arts graphiques

### 2.4.1 Le format PNG

Le format PNG (pour *portable networks graphics*) a été conçu comme codage de graphiques en mode balayage (et non vectoriel) et compression sans perte. Il a été développé en concurrence du format GIF (*graphics interchange format* de CompuServe qui n'était pas d'usage libre pendant de nombreuses années (les brevets de GIF sont maintenant dans le domaine public)).

Le format PNG supporte complètement les images couleurs codées sur 24 bits RVB (au contraire de GIF qui ne code les images que sur 8 bits), ainsi que les graphiques codés par des palettes sur 24 ou 32 bits. Il supporte également les informations de transparence associées à une image (le canal  $\alpha$ , qui varie de 0 à 1 : 0= opaque, 1 = totalement transparent). Mais PNG ne supporte pas les graphiques animés (au contraire de GIF).

PNG a été normalisé par l'ISO (ISO-15948 :2004).

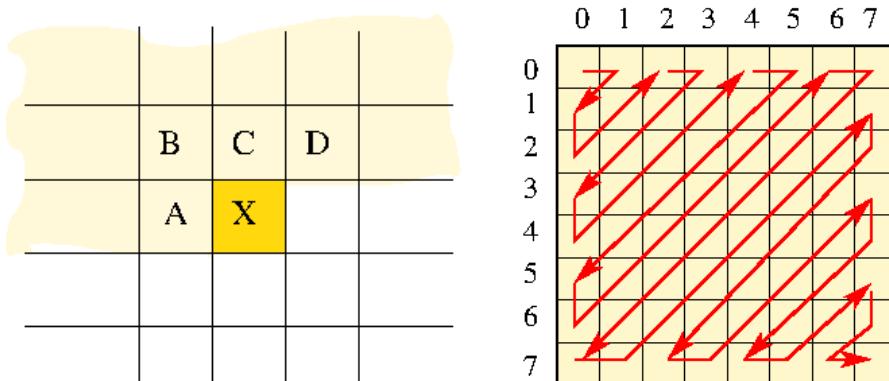
Après une entête de signalisation, PNG comporte un certain nombre de fragments (*chunks*), certains obligatoires, d'autres facultatifs, permettant de lire l'image. les fragments sont identifiés par un code à 4 symboles et protégés par un code correcteur d'erreur. Sont obligatoires :

- le fragment IHDR comportant la largeur de l'image, son nombre de ligne et le nombre de bits par pixel,

- le fragment PLTE comportant la palette de couleurs lorsque l'image est graphique,
- le fragment IDAT contenant l'image elle-même, en un ou plusieurs morceaux,
- le fragment IEND portant la marque de fin.

A l'aide des fragments facultatifs, on peut préciser de nombreuses compléments d'usage concernant le fond de l'image, l'espace de couleur utilisé, des corrections sur le blanc, le gamma, le facteur d'homothétie lors de l'affichage et même des informations en cas de visualisation stéréoscopique. PNG a été progressivement étendu jusqu'à permettre des opérations de gestion de couleur (*color management*) en post-production, c'est-à-dire le contrôle de la chaîne de reproduction de l'image jusqu'au produit final.

PNG est donc très utilisé pour les arts graphiques. Il permet de travailler de façon très transparente et économique avec des graphiques ou textes binaires, des graphiques définis jusqu'à 16 bits par canal et 4 canaux (images true color + canal  $\alpha$ ). Ce n'est pas alors ses capacités de compression qui sont ses atouts premiers. Néanmoins, une image PNG est compressée par la conjonction d'un codage prédictif (exploitant particulièrement la redondance sur les aplats) et un codeur Deflate. Le codeur prédictif est choisi parmi 5 (les valeurs possibles sont : 0 (pas de prédition), A, B,  $(A + B)/2$ , A ou B ou C selon la valeur la plus proche de  $A + B - C$ , voir le schéma de la figure 2.1) par l'utilisateur.



**Figure 2.1** – A gauche, lors d'un codage prédictif, le pixel X est codé à partir de ses voisins déjà transmis : A, B, C, D. A droite, ordre de balayage d'une fenêtre de DCT lors d'un codage JPEG.

#### 2.4.2 Le format TIFF

Nous avons eu l'occasion de rencontrer plus haut le format TIFF à travers son extension EP permettant d'accueillir un format RAW. Revenons maintenant sur ce format. Le format TIFF (*tagged image file format*), maintenant propriété de Adobe est un format spécifique pour documents scannés dédiés aux arts graphiques qui a progressivement évolué des documents binaires aux documents couleur en haute définition. Conçu comme un conteneur capable de conserver des documents de nature diverse (graphique vectoriel, texte, image par balayage codée avec ou sans perte, etc.), c'est par excellence un format d'archivage et d'échange qui permet d'assurer la qualité des documents source utilisés conjointement dans la construction d'un document final.

Il est constitué de blocs de codes, identifiés par des étiquettes (*tags* qui ont donné leur nom au code) chaque bloc spécifiant le contenu des données. Les étiquettes sont codées sur 16 bits. À partir de 32 000, elles sont attribuées à des entreprises ou organisations pour des besoins collectifs. Au-delà de 65 000 et jusqu'à 65 535 les étiquettes peuvent être spécifiées par l'usager à titre privé pour ses applications. TIFF a donc donné naissance à de nombreuses variantes parfois peu compatibles. En 2014, la version de base 6.0, établie il y a plus de 20 ans, indique les points essentiels permettant d'assurer une compatibilité entre systèmes. TIFF permet de contenir des structures très variées dans une seule image : des sous-fichiers, des bandes ou des tuiles qui chacune dispose de caractéristiques spécifiques qui expliquent leur coexistence.

TIFF doit obligatoirement pouvoir supporter le codage sans perte de Huffmann par plage ainsi que JPEG par DCT, mais éventuellement beaucoup d'autres schémas de compression (comme LZW et JPEG ondelettes) et pour cela des étiquettes spécifiques sont consacrées à ces compressions.

Les images couleur reconnues par TIFF sont de type RVB, YCbCr ou CIE  $la^*b^*$ , codées sur 24 bits ou de type CMYK codées sur 32 bits. Des extensions ont été faites pour des images à plus de 8 bits par pixels, ainsi que pour les très grosses images.

En règle général, le fichier TIFF associé à une image est assez volumineux. Dans le cas d'une image de  $n$  pixels en sRGB (3 octets par pixel), le fichier TIFF qui utilise exactement 1 octet par canal, aura pour taille  $3n$  octets sans compression, probablement la moitié environ après compression sans perte, mais c'est une décision de l'utilisateur ou non d'y faire appel et cela ralentit son emploi.

#### 2.4.3 Le format GIF

Indiquons ici quelques mots du format GIF, également réputé sans perte. C'est également un format propriétaire, maintenu par CompuServe, principalement destiné au graphique, mais permettant cependant de transporter des images sous certaines conditions.

GIF est un codage sur 8 bits de graphiques codés à travers une table de couleurs (ou palette) utilisant le codage de Lempel-Ziv suivant un codage par plage. Une image en noir et blanc est donc transmise sans beaucoup d'altération par une table transparente (quoique quelques niveaux de gris soient souvent réservés pour une représentation graphique minimale), mais sans grande compression. Une image en couleur devra être sévèrement quantifiée (il n'y a plus alors copie sans perte), ou devra être partagée en morceaux, chacune disposant de sa propre table de couleur limitée à 256. Les performances de compression sont alors plus faibles. Une autre solution utilise la technique du tramage (*dithering*) qui remplace un petit bloc uniforme de pixels par un mélange de pixels issus de la palette.

### 2.5 Les formats de compression d'images avec pertes

Nous avons eu l'occasion de dire que la compression sans perte d'images est assez limitée. Elle permet typiquement de gagner un facteur 2 avec une image en noir et blanc, un facteur 3 avec une image couleur et un facteur de 5 à 10 pour une image codée sur  $3 \times 16$  bits.

On s'est cependant très vite rendu compte qu'il était possible de modifier légèrement l'image par un codage qui devenait irréversible, sans altérer notablement son apparence, en jouant sur la tolérance du système visuel à des infidélités de reproduction chaque fois que le signal est assez complexe (effet de masquage), c'est-à-dire, dans les zones texturées et près des contours.

De cette remarque sont nés les premiers codages avec perte, utilisant les schémas que nous avons déjà signalés : codages prédictifs ou codages par longueur de plage ou de zones, mais pour lesquels on s'autorise une tolérance pour approcher le signal par des chaînes plus longues. Malgré la sophistication de ces approches, elles n'ont pas permis d'obtenir des taux suffisants et des méthodes nouvelles sont apparues où l'on tire profit de la forte redondance spatiale des images sur des blocs relativement larges en projetant les images sur des bases de fonctions adaptées pour regrouper l'énergie en un petit nombre de coefficients. Ces bases sont d'une part la transformation de Fourier (et surtout sa variante la DCT) et la transformation en ondelettes [Barlaud et Labit, 2003].

Dans tout codage avec perte, l'utilisateur dispose d'un paramètre  $\alpha$  lui permettant d'imposer le compromis qualité/compression qu'il souhaite atteindre. Ce compromis est très approximatif et dépend beaucoup de l'image qu'il veut comprimer. Il existe des critères objectifs de l'effet d'un codage. Si l'on désigne par  $I(i, j)$  l'image avant codage et par  $I'(i, j)$  l'image après codage :

1. le taux de compression  $\tau(\alpha) = \frac{\text{nombre de bits de } I}{\text{nombre de bits de } I'}$ ,
2. l'erreur quadratique moyenne de codage  $e = \sum_{i,j \in I} (I(i, j) - I'(i, j))^2$ ,

3. le rapport signal à bruit :  $\sigma(\alpha) = 20 \log_{10}(e/I^2)$ .

### 2.5.1 JPEG

C'est à partir des années 1980 qu'un groupe de chercheurs s'est penché sur l'élaboration d'un standard efficace de compression avec perte. Ils ont abouti dans les années 1990 et le format JPEG (pour *joint photographic expert group*) a été normalisé en 1992 sous le nom ISO/CEI 10918-1. C'est une norme qui laisse de nombreux espaces de liberté pour les développeurs de logiciel et il existe donc des variantes du logiciel de codage. Seule contrainte, tous les logiciels de décodage doivent pouvoir relire un ensemble d'images test déposées [Guillois et al., 2003].

#### Les principes du codage JPEG

Le codage JPEG est constitué de 6 étapes :

1. un changement d'espace de couleur pour ramener le signal, généralement présenté dans l'espace RVB, dans un espace luminance/chrominance (généralement YCrCb),
2. un sous-échantillonnage des composantes de chrominance par un facteur variable. Il peut atteindre 4 en  $x$  et en  $y$ , mais est le plus souvent de 2 (cette étape tire profit de la plus faible discrimination de l'oeil aux fluctuations de couleur qu'aux fluctuations de gris),
3. un découpage de l'image en blocs de taille fixe :  $8 \times 8$  pour les images de luminance,  $16 \times 16$  pixels pour les images de chrominance,
4. la transformation de ces blocs par un opérateur qui « concentre » l'énergie du bloc dans un petit nombre de coefficients, en l'occurrence la transformation en cosinus discrets DCT (*discrete cosine transform*),
5. la quantification de ces coefficients de façon à réduire leur entropie,
6. la transmission de ces coefficients réordonnés par un balayage en zig-zag de la fenêtre, puis un codage de type Lempel-Ziv de la chaîne de bits.

Le décodage inversera chaque étape hormis celles qui ont entraîné des pertes irréversibles d'information : les étapes 2 et 5.

La transformée en cosinus discret est une variante de la transformée de Fourier, mieux adaptée aux signaux réels et positifs comme les images que la transformée de Fourier [Ahmed et al., 1974]. C'est une transformation bidimensionnelle, exacte et inversible d'un signal scalaire qui crée un signal réel ayant autant d'échantillons que le signal d'entrée.

Soit  $J(i, j)$  une fenêtre de  $N \times N$  pixels (par exemple  $N = 8$ ) du canal à transmettre (canal de luminance ou de chrominance de l'image d'origine). La DCT de  $J$  est donnée par :

$$K_J(k, l) = \frac{2}{N} C(i) C(j) \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} J(i, j) \cos \left[ \frac{(2i+1)k\pi}{2N} \right] \cos \left[ \frac{(2j+1)l\pi}{2N} \right] \quad (2.3)$$

avec  $C(0) = \sqrt{2}/2$  et  $C(i, i \neq 0) = 1$ .

La transformation inverse :

$$J(i, j) = \frac{2}{N} \sum_{k=0}^{N-1} \sum_{l=0}^{N-1} C(k) C(l) K_J(k, l) \cos \left[ \frac{(2i+1)k\pi}{2N} \right] \cos \left[ \frac{(2j+1)l\pi}{2N} \right] \quad (2.4)$$

permet de retrouver exactement  $J(i, j)$  tant que l'on n'a pas procédé à la quantification de  $K_J(k, l)$ .

Les variables  $k$  et  $l$  sont des fréquences spatiales comme pour une transformée de Fourier. Le couple  $(k = 0, l = 0)$  porte la fréquence nulle (c'est à dire, au facteur  $N$  près, la valeur moyenne de  $J(i, j)$  sur la fenêtre). Lorsque  $k$  croît, les fréquences horizontales croissent. Lorsque  $l$  croît, ce sont

	0	1	2	3	4	5	6	7
0	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000
1	0,981	0,831	0,556	0,195	-0,195	-0,556	-0,831	-0,981
2	0,924	0,383	-0,383	-0,924	-0,924	-0,383	0,383	0,924
3	0,831	-0,195	-0,981	-0,556	0,556	0,981	0,195	-0,831
4	0,707	-0,707	-0,707	0,707	0,707	-0,707	-0,707	0,707
5	0,556	-0,981	0,195	0,831	-0,831	-0,195	0,981	-0,556
6	0,383	-0,924	0,924	-0,383	-0,383	0,924	-0,924	0,383
7	0,195	-0,556	0,831	-0,981	0,981	-0,831	0,556	-0,195

**Table 2.1 – Coefficients des filtres DCT pour  $N = 16$  i en ligne et  $k$  en colonne (équation 2.3).**

les fréquences horizontales et toute valeur de couple  $(k, l)$  où  $k \neq 0, l \neq 0$  correspond à une fréquence inclinée par rapport aux axes  $0x$  et  $0y$ .

Le terme  $K_J(0, 0)$ , appelé composante continue, est toujours beaucoup plus grand que les autres. Il est traité à part dans la DCT, souvent transmis sans altération de façon à donner une information exacte, mais imprécise, de la valeur du bloc.

Les autres termes feront l'objet au contraire d'une quantification d'autant plus importante que l'on souhaitera des taux de compression élevés. On définit un quantificateur  $q(k, l)$  et on procède à l'approximation du coefficient  $K_J(k, l)$  par  $K'_J(k, l)$  tel que :

$$K'_J(k, l) = \lfloor \frac{K_J(k, l) + \lfloor q(k, l)/2 \rfloor}{q(k, l)} \rfloor \quad (2.5)$$

où  $\lfloor x \rfloor$  désigne l'arrondi entier par défaut de  $x$ .

L'ensemble des coefficients  $q(k, l)$  forme ainsi une matrice de quantification  $\mathcal{Q}$  qui sera appliquée à toute l'image. Elle sera transmise avec les coefficients pour permettre de reconstruire l'image. Les divers codeurs se distinguent par la stratégie d'emploi du sous-échantillonnage de chrominance d'une part et le choix des matrices de quantification d'autre part.

On peut montrer que les termes de haute fréquence contiennent généralement très peu d'énergie et que ces termes sont moins importants pour la qualité perçue de l'image. On fera en sorte qu'ils soient ramenés à zéro par la quantification.

On procèdera alors à la transmission des coefficients  $K'_J(k, l)$  par un balayage en zigzag qui place en premier, juste après la composante continue les fréquences dans un ordre bidimensionnel (autant qu'on le peut) croissant (voir figure 2.1 à droite). Le train de coefficient ne comporte rapidement plus que des zéros ce qu'un codage de Lampel-Zif ou de Huffmann par plage comprimeront très bien.

Lors de la décompression des images JPEG, on procède à l'inversion de toutes les étapes ci-dessus. La quantification des coefficients peut faire apparaître des défauts de bloc, caractéristiques de JPEG. Certains algorithmes de reconstruction comportent donc une étape de filtrage passe bas qui gomme ces défauts. Le sous-échantillonnage des composantes chromatiques peut aussi nécessiter une étape d'interpolation qui pourra à son tour dégrader l'image.

## JPEG en pratique

La compression par JPEG est aujourd'hui la façon ordinaire de transmettre des images de qualité acceptable sous un volume lui aussi acceptable. Pour des images en couleur, on a l'habitude de considérer qu'un taux de compression de 10 s'accompagne d'une très faible dégradation, pratiquement imperceptible à l'oeil. C'est d'autant plus vrai que les images sont très grandes et observées sous un faible grossissement.

Des taux de compression de 100 sont couramment utilisés, en particulier sur le web ou pour des



**Figure 2.2 – Agrandissement d'une zone d'image codée par JPEG. Les facteurs de compression (pour l'image entière) valent en haut de gauche à droite : 6, 30 et 65, puis, en dessous : 100 et 150.**

échanges par mail. La qualité des images en est très fortement affectée et il n'est guère possible alors d'en faire un autre usage qu'une visualisation rapide sans être géné par les artéfacts.

Notons que tous les codeurs disposent d'un paramètre souvent unique qui contrôle le taux de compression. Cet indicateur n'est cependant formellement lié ni au gain lors de la compression, ni à la qualité de l'image. Un même indice pour des images différentes pourra donner des résultats très différents quoique l'indice, lorsqu'il décroît et pour une même image s'accompagne toujours de dégradations plus fortes et de taux de compression supérieurs. Cet indice vaut généralement 100 pour un codage sans perte et décroît jusqu'à la valeur minimale de 1.

Rappelons que JPEG sans perte n'utilise cependant pas du tout le principe général de la DCD pour coder l'image puisqu'il fait un codage prédictif.

Il est parfois possible d'utiliser JPEG de façon progressive lors d'une transmission, les termes de très basse fréquence sont transmis les premiers et l'image est raffinée ensuite lors de la transmission des fréquences plus élevées.

L'image constituée des seules composantes continues, 64 fois plus petite que l'image originale, est souvent utilisée pour des fonctions d'affichage ou de recherche rapide dans les bases de données.

### 2.5.2 JPEG 2000

JPEG 2000 est une norme reconnue par l'ISO sous le nom : ISO/CEI 15444-1. C'est une norme assez générale d'échange d'images avec des fonctionnalités variées (archivage, internet, vidéo). Nous nous attacherons surtout aux aspects liés au format d'image.

JPEG 2000 exploite la propriété des décompositions en ondelettes de mieux comprimer les images que la DCT, c'est-à-dire de fournir, pour un même taux de compression, une image de meilleure qualité, ou, pour une qualité fixée un taux de compression plus fort [Barlaud et Antonini, 2003].

Elle permet aussi d'assurer des fonctions de scalabilité. La scalabilité est la propriété d'un système de s'adapter aux canaux de transmission ou aux systèmes d'affichage pour ne transmettre que la partie de l'information qui sera utile pour l'usager (par exemple en n'envoyant qu'un signal réduit pour un téléphone mobile tandis qu'un téléviseur recevra le flux complet).

JPEG 2000 utilise également un codeur arithmétique très évolué EBCOT (*embedded block coding with optimized truncation*) qui tire profit de la structure en arbre bidimensionnel de la décomposition en ondelettes [Taubman, 2000].

Enfin JPEG 2000 offre une grande possibilité de décrire à un niveau supérieur le contenu des images pour faciliter, d'une part l'utilisation intelligente des collections d'images, d'autre part leur transmission optimisée dans les réseaux de nouvelle génération. Elle exploite pour cela les possibilités de transmission progressive données par la scalabilité des ondelettes à travers une structure de transmission appelée codestream ou la possibilité de définir des régions d'intérêt (*ROI region of interest*) qui seront transmises en priorité et avec une meilleure qualité. Elle prévoit explicitement la présence de texte ou de graphique dans une structure appelée *compound*. Elle est faite enfin pour des images très grandes (jusqu'à  $2^{32}$  pixels par ligne!) ou dotées de très nombreux canaux (jusqu'à  $2^{14}$ ) et avec des pixels à très forte dynamique (jusqu'à 38 bits par échantillon).

Comme beaucoup d'autres normes, JPEG 2000 est ouverte et ne précise pas explicitement le mode de compression, laissant au développeur la possibilité d'apporter des améliorations au cours du temps. N'est imposée que la capacité à décoder des fichiers types qui sont joints à la norme.

JPEG 2000 est une norme de compression très adaptée aux besoins professionnels. Elle a plus de mal à trouver sa place dans le grand public car elle réclame une puissance de calcul supérieure à JPEG et des logiciels de décodage assez lourds.

## Principe de JPEG 2000

Comme JPEG, JPEG 2000 peut transformer les images en couleur de l'espace RVB dans un espace luminance/chrominance (YUV ou  $YC_R C_B$ ) pour cela deux transformations sont disponibles, l'une réversible, l'autre irréversible (mais plus efficace en termes de compacité de représentation). Les composantes chromatiques sont ensuite traitées séparément.

JPEG 2000 travaille ensuite, soit sur l'image complète, soit sur des tuiles (*tiles*) qui sont des fragments d'image tous de même taille (à l'exception de ceux des bords). Ces tuiles sont décomposées de façon hiérarchique par un banc de filtres en ondelettes bi-dimensionnelles et bi-orthogonales [Barlaud et Antonini, 2003].

Deux types de filtres sont autorisés par la norme<sup>1</sup> :

1. les filtres à coefficients rationnels (5,3) de Le Gall que nous avons décrits au paragraphe 2.3 ils ne travaillent que sur des entiers sans arrondis et permettent un codage avec ou sans perte, selon que l'on transmet tous les coefficients non nuls ou pas (voir schéma EBCOT ci-dessous),
2. les filtres à coefficients réels (9,7) de Daubechies<sup>2</sup> (voir tableau 2.2), qui ne permettent qu'un codage avec perte, car les calculs se font en réels avec une précision finie. Ces ondelettes donnent de meilleures performances que les ondelettes rationnelles pour le codage avec pertes.

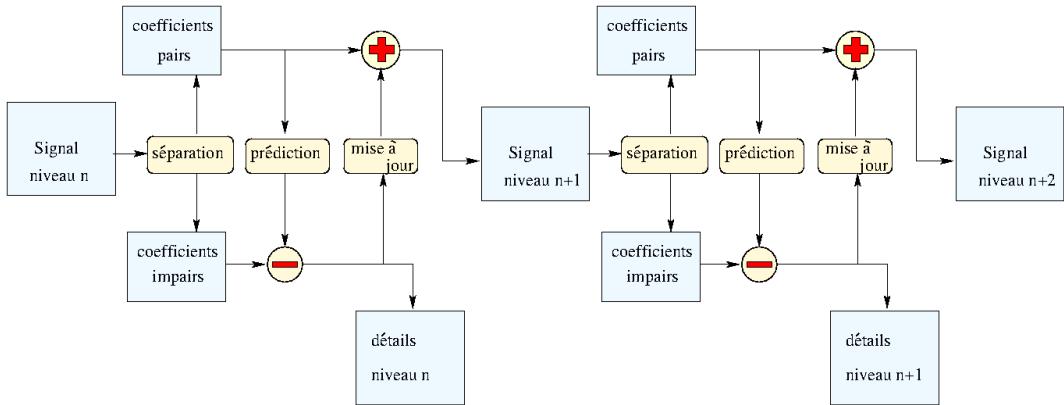
Ces filtres sont des filtres bi-orthogonaux<sup>3</sup> qui sont définis par un filtre passe-bas servant à calculer des images à résolution de plus en plus faible depuis le niveau original de la tuile, et des filtres passe-haut qui permettent de définir à chaque niveau le contenu des détails, caractéristiques du niveau considéré. Ces détails sont portés par des coefficients de la décomposition en ondelettes :  $a_n(i, j)$ .

---

1. ces filtres n'ont malheureusement pas de forme analytique simple. Ils s'expriment soit par la liste de leurs coefficients soit par leur équation dans l'espace polyphase où ils sont définis.

2. ou plus exactement, ondelettes de Cohen-Daubechies-Feauveau [Cohen et al., 1992]

3. Les filtres sont orthogonaux si le produit scalaire de deux filtres de rang différent est nul. Si l'on ne peut tenir cette contrainte très difficile à réaliser, on peut imposer la bi-orthogonalité, c'est-à-dire l'orthogonalité entre les filtres passe-bas et passe-haut.



**Figure 2.3 – Schéma de lifting entre 3 niveaux de  $n$  et  $n + 2$  de décomposition en ondelettes : on sépare les échantillons pairs et impairs. On définit les détails du niveau  $n$  comme la différence entre les échantillons impairs et leur prédiction faite à partir des échantillons pairs. On prend en compte le risque de repliement de spectre en filtrant le signal de détail et en le réinjectant dans les échantillons pairs qui forment alors le signal du niveau  $n + 1$  auquel peut de nouveau appliquer le schéma de lifting, et ainsi de suite jusqu'à la profondeur souhaitée.**

Le filtrage peut se faire de deux façons : soit comme d'habitude en traitement du signal, par convolution, soit, et c'est l'un des mérites de l'approche hiérarchique par ondelettes, par application itérative de 2 filtres l'un passe-bas, l'autre passe-haut, ces filtres étant identiques à tous les étages de la pyramide. C'est ce que l'on appelle le schéma de *lifting*. Dans ce schéma, on exprime les coefficients impairs d'une décomposition à partir de la valeur des coefficients pairs affectée d'un terme correctif. Ce terme correctif constitue l'image de détails propre au niveau considéré [Sweldens, 1996] (voir figure 2.3).

La transformée en ondelettes de base est monodimensionnelle. On l'applique donc une fois en lignes et une fois en colonnes sur le domaine d'intérêt. Il en résulte 4 images (voir la figure 2.4) :

- une image de hautes fréquences en lignes et en colonne,
- une image de hautes fréquences en lignes et de basses fréquences en colonnes,
- une image de hautes fréquences en colonnes et de basses fréquences en lignes,
- une image de basses fréquences en lignes et en colonnes.

On montre que, pour des bases d'ondelettes bien choisies, ces coefficients présentent des propriétés statistiques très intéressantes : ils sont très concentrés et peuvent être très bien modélisés à l'aide d'une loi gaussienne généralisée qui permet donc de les quantifier de façon très efficace :

$$p(a_n) = \alpha \exp(-|bx|^d) \quad (2.6)$$

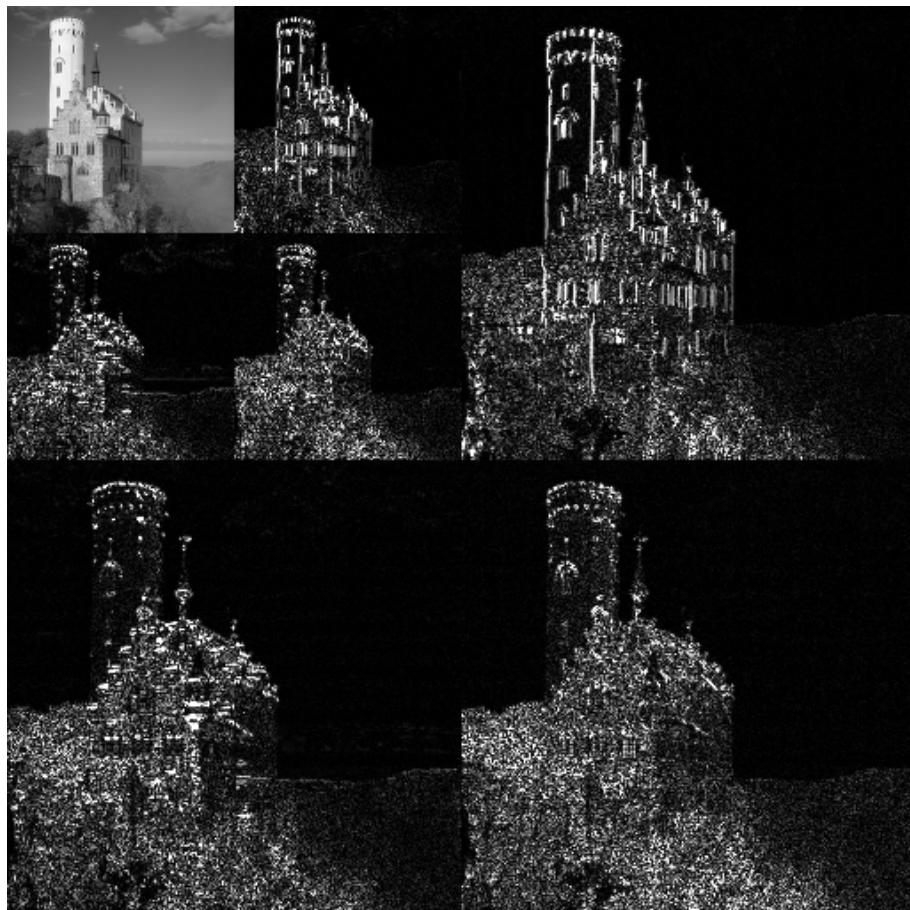
Les tables de quantifications sont définies pour chaque niveau de résolution. Elles sont calculées pour fournir une grande diversité de compression.

Ce sont ces coefficients quantifiés, transformés en plans de bits, regroupés par sous-bande qui sont présentés au codeur. Comme nous l'avons évoqué plus haut, le codeur adopté, EBCOT, utilise un schéma très performant de codage contextuel dans l'arbre des coefficients d'ondelettes prédisant un bit à partir du bit du niveau de résolution supérieur et de ses voisins immédiats, en commençant par les bits de poids fort. Le codeur travail sur des sous-blocs dont la taille est laissée libre mais assez réduite (maximum :  $64 \times 64$ )

La puissance du codage repose sur l'importance de l'analyse hiérarchique fournie par l'analyse en ondelettes. On s'intéresse à des voisinages en 6 ou 27 connexité dans la pyramide des niveaux de résolution. Les pixels majoritaires (identifiés par le bit du niveau supérieur) ne sont pas codés, les autres, sont codés par leur signe et leur écart au niveau supérieur s'ils sont significatifs. Enfin on examine les situations où il n'y a pas de configuration dominante. L'arbre ainsi créé est codée par un code algébrique.

ordre du filtre	0	$\pm 1$	$\pm 2$	$\pm 3$	$\pm 4$
analyse : filtre passe-bas	0,602949	0,266864	-0,078223	-0,016864	0,026749
analyse : filtre passe-haut	0,557543	0,295636	-0,028772	-0,045636	0
synthèse : filtre passe-bas	0,602949	-0,266864	-0,078223	0,016864	0,026749
synthèse : filtre passe-haut	0,557543	-0,295636	-0,028772	0,045636	0

**Table 2.2 – Coefficients des filtres des ondelettes de Daubechies 9/7 pour l’analyse et la reconstruction de l’image (d’après [Barlaud et Antonini, 2003]). Notez que l’on trouve parfois des formules différent d’un facteur  $\sqrt{2}$ . Les filtres d’analyse et de reconstruction doivent être symétriques par rapport à l’ordre 0 de façon à ne pas déplacer les contours des images.**



**Figure 2.4 – Codage d’une image par JPEG 2000. En haut à gauche, image de niveau 3, entourée des 3 images de détails du niveau 2 et des 3 images de détails du niveau 1 (issu de Wikipedia).**

### les performances de JPEG2000

JPEG 2000 permet des transmissions d’images fixes avec des taux de compression supérieurs à JPEG de l’ordre de 20 % à qualité fixée. Les défauts de JPEG 2000 sont subjectivement moins gênants que ceux de JPEG car il n’y a pas d’effet de bloc systématique. Les grandes images sont particulièrement bien codées par JPEG, ainsi que les images contenant peu de très hautes fréquences. JPEG 2000 donne des résultats de qualité satisfaisante à très fort taux de compression (de l’ordre de 100). Il apparaît également comme une norme bien adaptée pour les images de télédétection pour

sa capacité à traiter de grandes images dans un mode compatible avec le tuilage nécessaire à une navigation à haute résolution et avec la possibilité de spécifier des zones d'intérêt qui recevront une attention plus élevée.

## 2.6 Codage compressé (*Compressed sensing*)

On ne peut terminer un chapitre sur le codage et la compression des images sans aborder même brièvement cette technique qui porte aujourd'hui (2014) beaucoup d'espoirs pour réduire le volume occupé par un signal d'image au delà des valeurs fournies par JPEG 2000. Néanmoins, le codage compressé est bien loin de constituer une norme ou un format d'image et ne saurait être recommandé aujourd'hui à quiconque en concurrence des standards proposés ci-dessus.

L'idée sous-jacente au codage compressé est de reconstruire une image à partir d'un nombre d'échantillons très faibles, notamment moins nombreux que ce que recommande le théorème de Shannon. Les pistes proposées consistent à faire quelques mesures pertinentes et très complémentaires permettant de reconstruire l'image par des techniques de reconstruction sophistiquées qui exploitent la parcimonie des signaux d'images, essentiellement par des reconstructions en norme  $\mathcal{L}_0$  ou en norme  $\mathcal{L}_1$ . La parcimonie est la propriété d'un signal de pouvoir être décrit dans une base appropriée par un petit nombre de coordonnées non nulles seulement. Que les images soient parcimonieuses n'est ni démontré ni trivial. Il existe de bons arguments et des expériences convaincantes pour soutenir cette hypothèse qui mérite cependant d'être affermée et définie.

On voit que le codage compressé place d'enjeu l'acquisition de l'image au cœur du problème de compression (ce que les termes anglo-saxons de *compressed sensing* font mieux sentir) et que l'on est bien loin du schéma que nous développons dans cet ouvrage d'un capteur qui échantillonne de façon régulière et dense toute image avec le même pas. Néanmoins, cet échantillonnage « à la Shannon » n'est pas à priori incompatible avec un futur codage comprimé. Il laisse cependant au processeur embarqué le soin de calculer la forme compressée qu'un capteur plus malin pourrait peut-être mesurer directement (mais cela reste à démontrer pour les photographies de la vie courante).

Si l'on suppose qu'une image se décompose avec un tout petit nombre de coefficients sur une base appropriée, quelle stratégie peut-on adopter pour découvrir quels coefficients sont actifs ? Candès a montré [Candès et al., 2006, Candès et Wakin, 2008] qu'il fallait analyser l'image avec une base la plus « orthogonale » possible des fonctions de bases en question de façon que même un petit nombre de points d'analyse soient assurés de rendre compte pour partie des contributions des fonctions de base. Comme on ignore quelles fonctions de bases sont actives, Candès a suggéré d'utiliser une famille d'échantillons aléatoires et indépendants. La reconstruction, si l'on se donne une base, se fait alors par un gros problème d'optimisation sous contrainte de parcimonie, optimisation conduite généralement par des techniques de projection. Quelle base doit-on choisir ? Nous ignorons quelle base sera optimale, mais nous disposons de nombreux candidats : bases de Fourier, DCT, ondelettes comme celles utilisées dans JPEG, et paquets d'ondelettes utilisés en traitement des images (*curvelets*, *bandelets*, etc.).

Les travaux conduits dans ces directions n'ont pas encore apporté de résultats intéressants le domaine de la photographie en général. Ils ont cependant conduits à d'intéressantes solutions pour des problèmes complexes de mesure tridimensionnelle en imagerie médicale (tomographie axiale, résonnance magnétique [See Lustig et al., 2008]), en microscopie pour la biologie [Le Montagner et al., 2012], en holographie pour l'instrumentation physique [Denis et al., 2009], etc.

## 2.7 Le codage vidéo

Le codage vidéo déborde très largement du cadre de cet ouvrage, néanmoins de plus en plus nombreux sont les appareils de photo qui permettent d'enregistrer soit des images fixes, soit des séquences animées. Ils comportent alors obligatoirement un codeur vidéo qui comprime le flot d'images. Ces

codeurs exploitent, comme les codeurs d'images fixes, la redondance au sein de chaque image (c'est ce que l'on appelle le codage intra-image), mais peuvent également réduire la redondance entre images successives (c'est le codage inter-image).

Les travaux de normalisation ont débuté au sein de l'UIT<sup>4</sup>, avec des recommandations appelées H.120, puis H.261 (ou  $p$  fois 64 kbit/s), H.263 qui le premier fait intervenir un mode inter-images et H.264 (à la base de MPEG-4).

### 2.7.1 Codage MPEG

Ces recommandations ont été reprises dans un cadre moins dédié aux télécommunications pour des applications d'archivage, de jeux vidéo, d'usage privatif. C'est le groupe de travail MPEG *moving picture expert group*, groupe joint de l'ISO et de la CEI, qui s'est chargé de construire des solutions de consensus aptes à passer en normalisation [Nicolas et Guillemot, 2003].

La première recommandation MPEG-1 visait des transmissions à bas débit pour les transmissions sur fil avant l'ADSL (par pas de 64 kbit/s, elle permettait de monter jusqu'à 1,5 Mbit/s environ). Ce très faible débit permettait la transmission vidéo d'images de taille réduite ( $352 \times 288$  pixels par image, à 25 image/s). MPEG-1 a beaucoup été utilisée dans le cadre de l'archivage sur vidéo-CD.

MPEG-2 est une norme adaptée à la télévision. Elle vise des transmissions de très bonne qualité (codage dit de contribution) et convient particulièrement à la transmission hertzienne, sur câble ou sur ADSL. Elle autorise des transmissions à des débits de 2 à 6 Mbit/s pour l'image vidéo ordinaire et des débits de 15 à 20 Mbit/s pour des vidéo en TVHD<sup>5</sup>.

MPEG-4 est une norme de transmission à bas débit (autour de 2 Mbit/s) utilisant des mécanismes de compression plus élaborés que ceux de MPEG-2 et aptes à s'adapter à de très fortes compressions (avec bien-sûr une perte de qualité importante). Elle tire profit des propriétés de scalabilité présentes dans la recommandation H.263. Elle propose de nombreuses passerelles avec le monde de l'information et dispose en particulier de désignation de zones d'intérêt (ROI) et de mécanismes de gestion de la propriété.

MPEG-7, MPEG-21 et MPEG-x (x prenant les valeurs de A à V) sont des développements qui visent les applications multimédias des séquences vidéo et définissent des fonctionnalités spécifiques à l'archivage, la recherche dans les bases de données, la production (en particulier par intégration de sources variées), l'interaction, la diffusion et la protection des contenus.

### Les mécanismes mis en œuvre dans MPEG

MPEG-2 et MPEG-4 utilisent des mécanismes de compression assez semblables tant que la notion de zone d'intérêt n'est pas utilisée [Nicolas et Guillemot, 2003, Pereira et Ebrahimi, 2002].

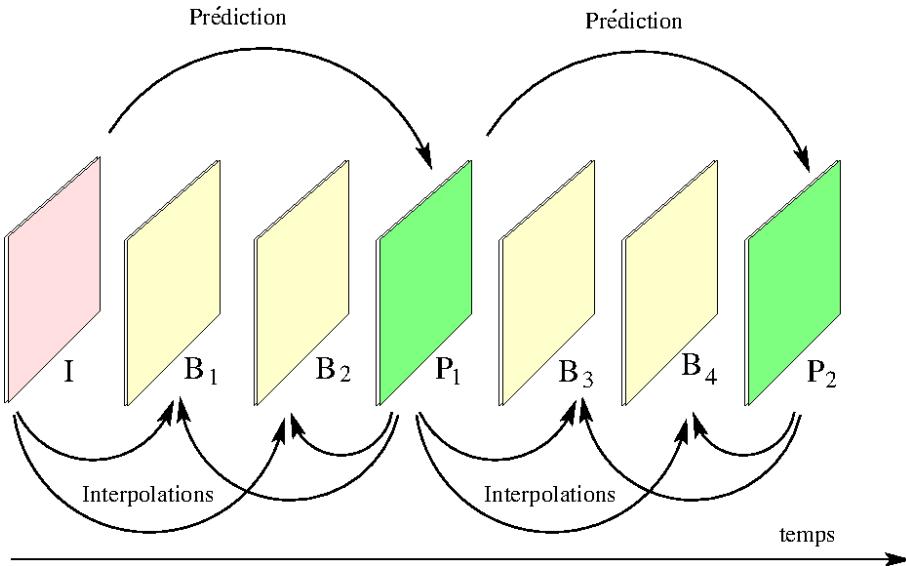
L'image en couleur, généralement acquise dans un espace RVB, est convertie dans un espace luminance/chrominance et les canaux de chrominance sont généralement sous-échantillonnés d'un facteur 2 en ligne et en colonne.

Le codage intra-image se fait par subdivision de l'image en groupes de blocs (GOB), construits à partir de macroblocs MB. Un macrobloc est lui-même constitué de 4 blocs de  $8 \times 8$  pixels de luminance et généralement 1 bloc de chrominance (codage à chrominance sous-échantillonnée dit 4 : 2 : 0) ou 4 blocs (cas du codage 4 : 4 : 4 sans sous-échantillonnage). Le train binaire sera donc constitué d'une couche image (I), constituée des couches de groupes de blocs (GOB), elles-mêmes constituées de la couche de macroblocs (MOB). Cette couche de macroblocs comportera typiquement les 4 blocs de luminance suivis de 2 blocs de chrominance, l'un pour la dominante R, l'autre pour la dominante B.

4. UIT = Union Internationale des Télécommunications, ici le groupe intéressé est l'UIT-T, chargé de la normalisation

5. TVHD = Télévision haute définition =  $1920 \times 1080$  pixels par image à 50 images par seconde.

Les blocs sont codés par transformation en cosinus discret (DCT), quantifiée et suivie d'un codage entropique (codeur par longueur de plage, ou mieux, codage arithmétique).



**Figure 2.5 – Codage d'un flot d'images par MPEG :** I est l'image de référence codée en intra-image. P<sub>1</sub> et P<sub>2</sub> sont les images codées par prédition de mouvement à partir de I d'une part et de P<sub>1</sub> d'autre part. B<sub>1</sub> et B<sub>2</sub> sont interpolées à partir de I et P<sub>1</sub>, B<sub>3</sub> et B<sub>4</sub> à partir de P<sub>1</sub> et P<sub>2</sub>.

Le codage inter-images se fait selon 2 modes :

1. le mode prédictif (mode P) : une image du flot est prise comme référence et transmise avec un codage intra-image (image I). La nouvelle image (P) à transmettre à un temps postérieur est comparée à I de façon à déterminer le déplacement que chaque pixel a subi. C'est une étape de prédition de mouvement qui fonctionne par corrélation de blocs (*block matching*) et détermine un vecteur de déplacement ( $\Delta x$  en  $x$  et  $\Delta y$  en  $y$ ) pour chaque macrobloc. La différence entre le macrobloc de l'image P et celui de l'image I, déplacé de  $\{\Delta x, \Delta y\}$ , constitue un macro bloc d'erreur qui est codé par DCT. La transmission se fera par l'envoi de l'adresse du macrobloc de référence dans I, du vecteur de déplacement  $\{\Delta x, \Delta y\}$ , puis du bloc d'erreur quantifié et compressé entropiquement.
2. le mode interpolé bidirectionnellement (mode B) : ce mode permet d'interpoler le flot d'images en utilisant à la fois une image antérieure I et une image postérieure P pour reconstruire des images intermédiaires notées B (voir la figure 2.5) à l'aide d'une très petite surcharge de transmission (on n'ajoute que les éléments indispensables au positionnement du macrobloc de B s'il ne se déduit pas par simple interpolation de celui de P). Ce mode de fonctionnement présente néanmoins l'inconvénient d'introduire un délai dans la reconstruction de l'image. Il est réservé à ce titre aux archivages sur supports magnétiques qui n'imposent pas de temps réel.

Ce mode de codage fait apparaître d'importants effets de blocs pour des taux de compression élevés. De nombreuses améliorations ont été proposées pour les éviter : multiplication des vecteurs de déplacement pour un macro-bloc, superposition de blocs lors de la recherche du déplacement, filtrage à postériori, etc.

MPEG-4 se distingue de MPEG-2 par une amélioration de la plupart des fonctionnalités ci-dessus : meilleure prédition des coefficients DCT, meilleure prédition de mouvement, resynchronisation du flux des données, codage à longueur variable entropique, avec correction d'erreur, etc. Il introduit surtout un objet nouveau : le VOP (*video object plane*), c'est-à-dire la boîte englobant un objet d'intérêt. Un VOP pourra recevoir un mode particulier de compression, différent du reste de l'image. Un VOP est un nombre entier de macro-blocs. Il sera identifié dans le flux des données. Par exemple, un speaker parlant sur un fond fixe pourra être situé dans un VOP qui évoluera à haute définition

et haute fréquence de raffraîchissement tandis que le fond du studio sera codé à très bas débit. Les VOP, qui sont généralement décrits à la main par l'utilisateur, ne sont pas aujourd'hui considérés par les appareils photographiques, même lorsqu'ils disposent d'une compatibilité MPEG-4.

# Bibliographie

- [Ahmed et al., 1974] Ahmed, N., Natarajan, T., and Rao, K., 1974. Discrete cosine transform. *IEEE trans on Computers* C-23-(1), 90–93.
- [Alvarez et al., 1999] Alvarez, L., Gousseau, Y., and Morel, J., 1999. The size of objects in natural and artificial images. *A. i. I. & Electron Physics* 111, 167.–242.
- [Barlaud et Antonini, 2003] Barlaud, M. and Antonini, M., 2003. Transformées en ondelettes pour la compression d'images. In Barlaud, M. and Labit, C., editors, *Compression et codage des images et des vidéos* chapter 3, pages 73–95. Hermes-Lavoisier.
- [Barlaud et Labit, 2003] Barlaud, M. and Labit, C., 2003. *Compression et codage des images et des vidéos*. Hermes-Lavoisier.
- [Blanc-Lapierre et Dumontet, 1955] Blanc-Lapierre, A. and Dumontet, P., 1955. La notion de cohérence en optique. *Revue de Physique Appliquée* 34, 1–21.
- [Born et Wolf, 1970] Born, M. and Wolf, E., 1970. *Principles of Optics*. Pergamon Press (4th Edition).
- [Candès et al., 2006] Candès, E., Romberg, J., and Tao, T., 2006. Stable signal recovery from incomplete and inaccurate measurements. *Communications on Pure and Applied Mathematics* 58 (8), 1207–1222.
- [Candès et Wakin, 2008] Candès, E. and Wakin, M., 2008. An introduction to compressive sampling. *IEEE Signal Processing Magazine* 25 (2), 21–30.
- [Cao et al., 2010] Cao, F., Guichard, F., and Hornung, H., 2010. Dead leaves model for measuring texture quality on a digital camera. In Proc. SPIE, Vol. 7537.
- [Cohen et al., 1992] Cohen, A., Daubechies, I., and J.C., F., 1992. Biorthogonal bases of compactly supported wavelets. *Communications on Pure and Applied Mathematics* 45-5, 485–560.
- [Delbracio-Bentancor, 2013] Delbracio-Bentancor, M., 2013. Two problems of digital image formation. PhD thesis Ecole Normale Supérieure de Cachan, France.
- [Denis et al., 2009] Denis, L., Lorentz, D., Thiébaut, E., Fournier, C., and Trede, E., 2009. Inline hologram reconstruction with sparsity constraint. *Optics Letters* 34(22), 3475–3477.
- [Duffieux et Lansraux, 1945] Duffieux, P. and Lansraux, G., 1945. L'intégrale de Fourier et ses applications à l'optique. *Revue d'Optique* 24, 65, 151, 215.
- [Dumontet, 1955] Dumontet, P., 1955. Sur la correspondance objet-image en optique. *Optica Acta* 2, 53–63.
- [Goodman, 1976] Goodman, J., 1976. Some fundamental properties of speckle. *Journal Optical Society of America* 66, 1145–1150.
- [Gousseau et Roueff, 2007] Gousseau, Y. and Roueff, F., 2007. Modeling occlusion and scaling in natural images. *SIAM J. of Multiscale Modeling and Simulation* 6(1), 105–134.
- [Guillemot et Pateux, 2003] Guillemot, C. and Pateux, S., 2003. Eléments de théorie de l'information et de communication. In Barlaud, M. and C.Labit, editors, *Compression et codage des images et des vidéos* chapter 1, pages 21–41. Hermes-Lavoisier.
- [Guillois et al., 2003] Guillois, J., Charrier, M., Lambert, C., and Paucard, B., 2003. standards de compression d'images fixes. In Barlaud, M. and Labit, C., editors, *Compression et codage des images et des vidéos* chapter 4, pages 97–130. Hermes-Lavoisier.

- [ISO, 2000] ISO, 2000. Norme ISO 12233 - photography – electronic still picture cameras – resolution measurements. norme ISO.
- [ISO, 2001] ISO, 2001. ISO 12234-2 :2001 electronic still-picture imaging – removable memory – page 2 : TIFF/EP image data format. ISO Norm.
- [Le Grand, 1972] Le Grand, Y., 1972. Optique physiologique, tome 2 : lumière et couleur. Masson (Paris).
- [Le Montagner et al., 2012] Le Montagner, Y., Angelini, E., and Olivo-Marin, J., 2012. Video reconstruction using compressed sensing measurements and 3D total variation regularization for bio-imaging applications. In IEEE Int. Conference on Image Processing ICIP-12 pages 917–920 Orlando (Florida).
- [McElvain et al., 2010] McElvain, J., Campbell, S., and Miller, J., 2010. Texture-based measurement of spatial frequency response using the dead leaves target : extensions, and application to real camera systems. In Electronic Imaging Proc SPIE 7537, 75370D.
- [Nicolas et Guillemot, 2003] Nicolas, H. and Guillemot, C., 2003. Normes de compression vidéo. In Barlaud, M. and Labit, C., editors, Compression et codage des images et des vidéos chapter 6, pages 177–208. Hermes-Lavoisier.
- [Pereira et Ebrahimi, 2002] Pereira, F. and Ebrahimi, T., 2002. The MPEG-4 book. Prentice-Hall Upper Saddle River (NJ, USA).
- [Pérez, 1991] Pérez, J.-P., 1991. Optique géométrique, ondulatoire et polarisation. Masson (Paris).
- [Prost, 2003] Prost, R., 2003. Compression sans perte. In Barlaud, M. and Labit, C., editors, Compression et codage des images et des vidéos chapter 9, pages 285–305. Hermes-Lavoisier.
- [See Lustig et al., 2008] See Lustig, M., Donoho, D., Santos, J., and Pauly, J., 2008. Compressed sensing MRI. IEEE Signal Processing Magazine 25(2), 78–82.
- [Sweldens, 1996] Sweldens, W., 1996. The lifting scheme : a custom-design construction of biorthogonal wavelets. Applied and Computational Harmonic Analysis 3-2, 186–200.
- [Taubman, 2000] Taubman, D., 2000. High performance scalable image compression with EBCOT. IEEE trans on Image Processing 9 - 7.

# Index

- acutance, 4
- codage compressé, 28
- compressed sensing, 28
- compression sans perte, 17
- DCT (transformation), 21
- dithering, 21
- DNG (format), 14
- EBCOT (codeur), 24
- Exif (fichier), 16
- FSCS = fonction de sensibilité au contraste spatial, 4
- FTM = fonction de transfert de modulation, 2, 7
- GIF (format), 21
- granularité, 10
- Huffmann (code), 18
- image vidéo, 28
- JPEG (codage), 22
- JPEG 2000 (codage), 24
- JPEG sans-perte (codage), 18
- Liv-Zempel (code), 18
- lumière cohérente, 6
- lumière incohérente, 6
- MPEG (codage), 29
- palette, 21
- piqué, 3
- PNG (format), 19
- pouvoir de séparation, 12
- PSF = point spread function, 3, 7
- réponse impulsionale, 2, 7
- Résolution, 2
- RAW (format), 13
- scalabilité, 24
- speckle, 10
- TIFF (format), 20
- TIFF/EP (format), 14
- tramage, 21
- TVHD = Télévision haute définition, 29
- UIT = Union Internationale des Télécommunications, 28