

Traitement de l'image : de l'équation de la chaleur aux ondelettes

Jean-Pierre Antoine* et Laurent Jacques[§]

*professeur, resp. [§]assistant, à l'Université Catholique de Louvain

I. Généralités sur le traitement numérique des images

Avant de discuter du traitement d'images, il convient de préciser l'objet de notre étude. Une *image numérisée* est un ensemble de pixels (picture elements), à chacun desquels est attribué un niveau de gris, de 0 (noir) à 255 (blanc). Le nombre $256 = 2^8$ traduit le codage sur 8 bits en binaire, p.ex. 10010101. Une image numérisée en *niveaux de gris* n'est donc rien d'autre qu'une fonction définie sur une grille discrète et prenant des valeurs discrètes, c'est-à-dire un tableau (matrice) de nombres (binaires). Une telle image contient donc une quantité d'information énorme (se chiffrant facilement en MB!). Toutefois, de par sa définition même, elle va présenter des difficultés géométriques : courbes et droites sont mal reproduites, il y a des effets de pixellisation (lignes en escalier, p.ex.). La Figure 1 illustre le principe de numérisation et la nature de ces effets indésirables, tandis que la Figure 2 présente un exemple concret.

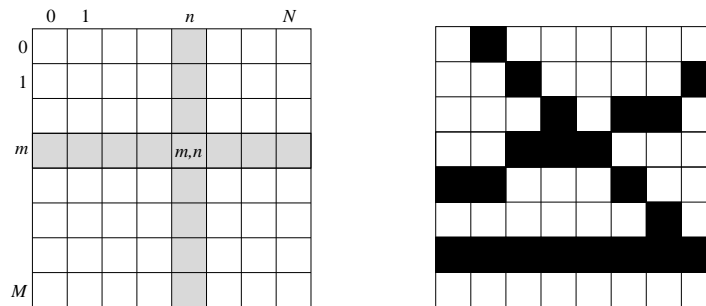


Figure 1: (a) Schéma d'une image numérisée. (b) Difficultés liées à la pixellisation.

Ceci dit, que veut-on faire d'une image numérisée ? Plusieurs types d'actions peuvent être envisagés:

- (1) Analyser : détecter, mesurer, compter des "objets" particuliers
 - . photos aériennes : routes, bâtiments, ...
 - . images du ciel : galaxies
 - . fluide, trafic automobile : "objets" en mouvement
 - . photos : reconnaître des personnages ("image retrieval")
 - . texte : lecture optique
 - . empreintes digitales (FBI)
- (2) Modifier :
 - . nettoyer l'image : enlever le "bruit", des défauts, ...
 - . transmettre et reconstruire une image (télévision!), ce qui implique compresser sans perte d'information intolérable (variable!)
 - . imprimer une signature invisible dans l'image (tatouage, "watermarking")

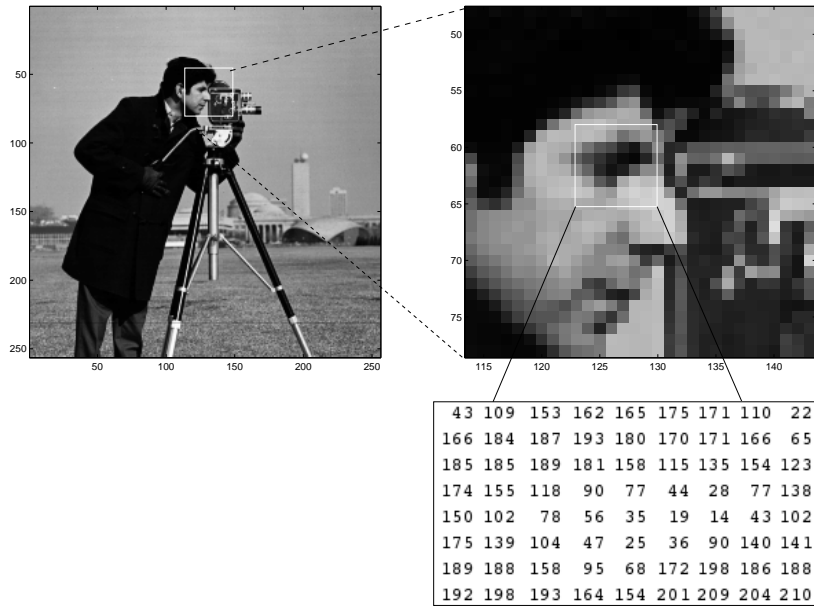


Figure 2: Un exemple concret d'image numérique.

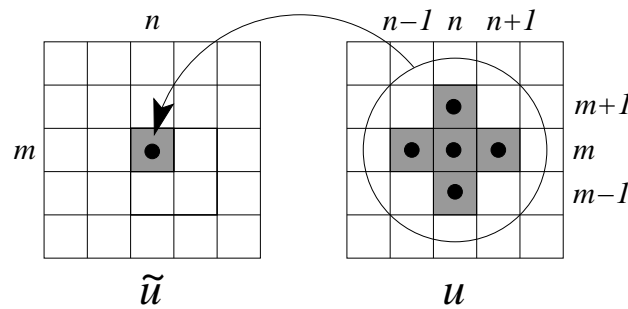


Figure 3: Lissage d'une image par moyenne.

Pour toutes ces opérations, il faut rendre le signal plus "propre", plus petit, mais toujours utilisable, sans perte intolérable d'information. Les techniques utilisées pour arriver à ce but constituent le *traitement numérique des images*. Dans tous les cas, l'on devra transformer, manipuler l'image, c'est-à-dire, agir sur les nombres qui la constituent.

Un exemple typique est le *lissage*. Sous la forme la plus simple, cette opération consiste à remplacer chaque pixel par la moyenne de celui-ci et de ses 4 plus proches voisins (Figure 3). Soit l'image $u(m, n) \in [0, 255]$, avec (m, n) la position d'un pixel. L'image lissée de cette manière est alors

$$\tilde{u}(m, n) = \frac{1}{5}(u(m, n) + u(m + 1, n) + u(m - 1, n) + u(m, n + 1) + u(m, n - 1)) \quad (1.1)$$

La Figure 4 montre le résultat de cette opération sur l'image cameraman présentée à la Figure 2. Ce procédé peut se réécrire sous forme d'une *convolution* :

$$\tilde{u}(m, n) = \sum_{i,j} u(i, j) b(m - i, n - j) = (u * b)(m, n), \quad (1.2)$$



Figure 4: Lissage par moyenne de l'image cameraman.

avec $b(0, 0) = b(1, 0) = b(-1, 0) = b(0, 1) = b(0, -1) = \frac{1}{5}$ et $b = 0$ ailleurs. Sous cette forme, on peut généraliser la méthode en prenant une fonction b plus compliquée, p. ex. une gaussienne g_l de taille $l > 0$:

$$g_l(m, n) = \exp\left(-\frac{1}{l^2}(m^2 + n^2)\right), \quad (1.3)$$

Appliquant ce dernier lissage sur l'image cameraman, pour $l = 3$ et $l = 10$, on obtient les résultats indiqués à la Figure 5.

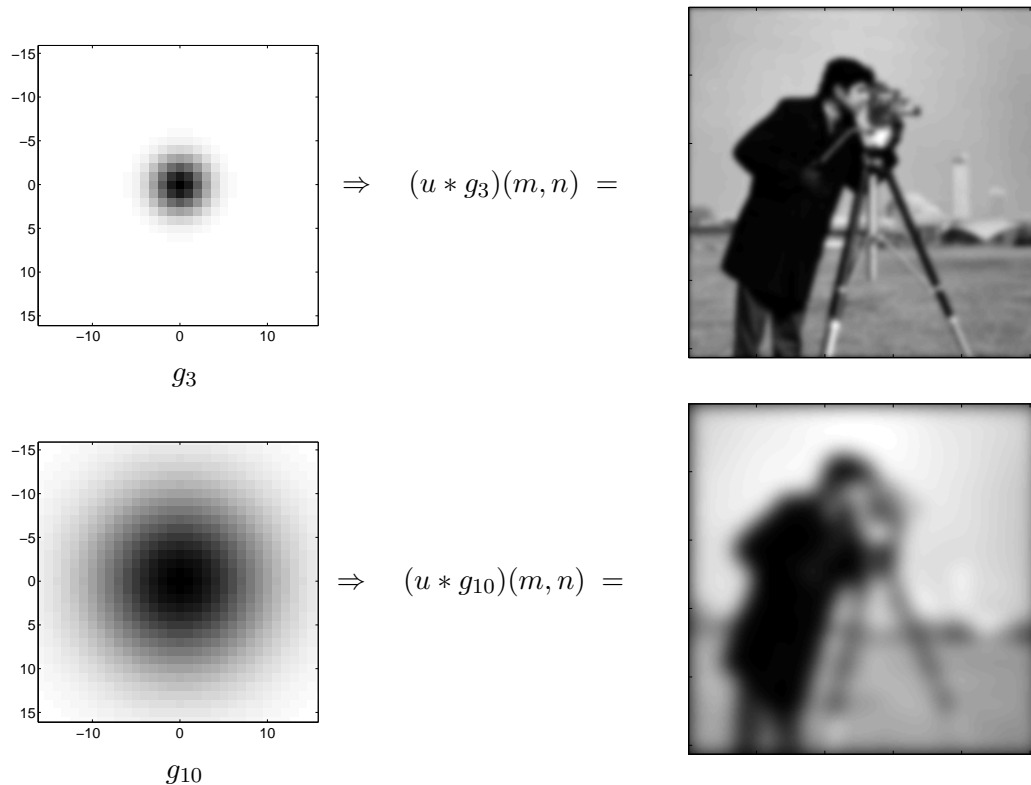


Figure 5: Lissage de l'image cameraman par des gaussiennes de taille 3 et 10.

Parmi les méthodes de traitement numérique des images, on peut distinguer deux grandes classes.

(a) Méthodes de type spectral (Fourier)

La technique consiste à décomposer l'image en *blocs*, dont la taille ($4 \times 4 \text{ pixels}$, $8 \times 8 \text{ pixels}$, ...) varie en fonction de celle des détails qui s'y trouvent (Figure 6), puis à représenter chaque bloc par des fonctions trigonométriques (DCT : Discrete Cosine Transform) :

$$f_k(x, y) = \sum_{m,n} c_{mn}^k \cos mx \cos ny \quad (k^{\text{ième}} \text{ bloc}). \quad (1.4)$$

Dans cette approche, l'image est donc décrite par l'ensemble des coefficients c_{mn}^k .

Quoique très largement utilisée, cette méthode présente de sérieux inconvénients :

- Elle ne contient aucune notion de *résolution*, les indices m, n correspondent seulement à une fréquence d'oscillation.
- Il y a des difficultés de continuité aux frontières entre blocs, ce qui crée des défauts ("blocking effects").
- Il y a aussi des difficultés pour représenter des lignes droites obliques et des courbes ("escaliers"); la raison en est évidemment que cette approche est intrinsèquement liée à la géométrie cartésienne "x,y".



Figure 6: Décomposition en blocs d'une image.

(b) Méthodes multirésolution

Le principe d'une approche multirésolution est de considérer qu'une approximation donnée (résolution fixée) s'obtient en combinant l'approximation deux fois plus grossière avec des détails additionnels, d'où l'importance de la notion d'*échelle*.

Une réalisation intéressante de cette approche est la *transformée en ondelettes*. Les propriétés essentielles de celle-ci peuvent se résumer comme suit :

- L'image est représentée par des fonctions mieux adaptées, de moyenne nulle, plutôt que par des fonctions trigonométriques, identiques dans tous les cas :

$$f(x, y) = \sum_j c_j \psi_j(x, y), \quad (1.5)$$

où j représente collectivement les paramètres de ψ_j , tels que la position, l'échelle ou l'orientation. Il s'ensuit que le développement (1.5) nécessite beaucoup moins de termes que celui de la DCT (1.4) pour une même qualité d'approximation.

- La méthode est économique, puisque les ψ_j sont bien localisés dans l'espace et ont une moyenne nulle : si une portion de l'image ne contient aucune information (région unie), il n'y a rien à calculer, la transformée est identiquement nulle dans cette région.
- La transformée possède une structure hiérarchique en échelle : l'indice j correspondant entre autres à la résolution, la fonction ψ_j peut s'interpréter comme une lentille *locale*, de plus en plus forte (zoom) à mesure que j augmente.
- La transformée en ondelettes détecte surtout les *discontinuités*, p.ex. les contours dans une image.
- La méthode répond bien à la question initiale, par un traitement en trois étapes : décomposition, compression par "seuillage des coefficients", reconstruction.

Comme exemples spectaculaires d'application de la transformée en ondelettes bidimensionnelle, on peut citer la lecture optique (la détection du contour des lettres mène univoquement à leur identification), le débruitage des images en astrophysique, ou encore la segmentation en imagerie médicale.

Avec le recul, on peut affirmer que le succès de la méthode des ondelettes repose sur la collaboration entre ingénieurs (possédant les techniques de traitement du signal), les physiciens théoriciens (qui ont apporté des idées venant des théories quantiques), et des mathématiciens (qui ont démontré les théorèmes fondamentaux).

II. Détection des contours ('edge detection') et multirésolution

Le principe de la détection des contours dans une image est de détecter et de représenter les changements (brusques) d'intensité. Cette opération est souvent nécessaire; c'est le cas pour la segmentation d'une image (décomposition en "cartoon + texture", détermination de région d'intérêt (ROI) en imagerie médicale), le débruitage d'une image, etc.

En fait, il s'agit d'une opération analogue à une différentiation. Il faut donc lisser (= filtrer) l'image, de façon à éliminer le bruit de haute fréquence, comme discuté à la Section I. De plus, les changements d'intensité peuvent se produire à des échelles différentes, il faut donc lisser l'image à des résolutions différentes.

Il est bien connu, depuis D. Gabor, que le filtre optimal pour lisser une image est la gaussienne, déjà présentée en (1.3), car elle présente le compromis idéal entre localisation spatiale et localisation fréquentielle. Pour une largeur σ , nous la normaliserons comme suit :

$$G_\sigma(x, y) = \frac{1}{\sigma^2} \exp\left(-\frac{|\vec{x}|^2}{2\sigma^2}\right), \quad \vec{x} = (x, y). \quad (2.1)$$

Dès lors, l'image lissée à la résolution σ est donnée par la formule

$$(G_\sigma * I)(\vec{x}) = \int d\vec{y} I(\vec{y}) \frac{1}{\sigma^2} \exp\left(-\frac{|\vec{x} - \vec{y}|^2}{2\sigma^2}\right). \quad (2.2)$$

Nous retrouvons donc ici le noyau de la chaleur, décrivant un phénomène de diffusion.

Par ailleurs, un contour, c'est-à-dire un changement brutal d'intensité, se traduit par le passage par zéro ("zero-crossing") de $D^2(G_\sigma * I) = D^2 G_\sigma * I$, où D représente une dérivée directionnelle. Ceci est facile à comprendre intuitivement. Considérons une image I à la résolution σ , c'est-à-dire $G_\sigma * I$. Les contours contenus dans cette image sont les points (pixels) où le gradient est maximal, plus précisément les points où $|D(G_\sigma * I)|$ est maximal dans la direction de $D(G_\sigma * I)$, c'est-à-dire

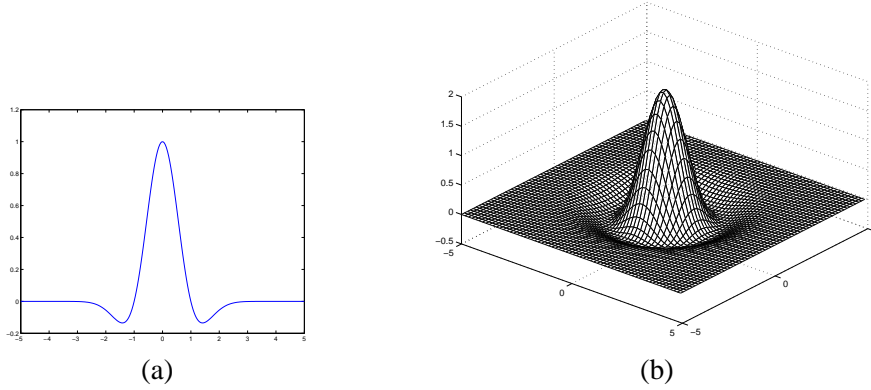


Figure 7: Ondelette “chapeau mexicain” : (a) en 1-D, $-G''_\sigma$; (b) en 2-D, $-\Delta G_\sigma$.

les points où $D^2(G_\sigma * I) = 0$.¹ S’il n’y a aucune direction privilégiée, on prendra $[1, 2] D^2 \Rightarrow \Delta$, c’est-à-dire l’opérateur différentiel isotrope le plus simple :

$$-\Delta G_\sigma(\vec{x}) \equiv \psi_\sigma(\vec{x}) = \frac{2}{\sigma^4} \left(1 - \frac{|\vec{x}|^2}{2\sigma^2}\right) \exp\left(-\frac{|\vec{x}|^2}{2\sigma^2}\right)$$

Mais ceci n’est rien d’autre que l’ondelette “chapeau mexicain” bidimensionnelle, représentée à la Figure 7 (b). La Figure 8 donne un exemple de contours à des résolutions de plus en plus fines.

Il est clair sur ces relations que fixer σ revient choisir un niveau de résolution. Ceci nous amène à l’idée de *multirésolution*, selon laquelle le principe de la vision consiste à partir d’une image grossière et y ajouter des détails de plus en plus fins, aux résolutions $\sigma_1, \sigma_2, \sigma_3, \dots$. Le problème qui se pose est alors de vérifier la cohérence entre ces différentes échelles.

Précisons le rôle de la dilatation. On a (en normalisation L^1) :

$$(D_a f)(\vec{x}) = \frac{1}{a^2} f\left(\frac{\vec{x}}{a}\right)$$

et, dès lors,

$$(D_a \psi_\sigma)(\vec{x}) = \psi_{a\sigma}(\vec{x}).$$

Autrement dit, varier σ équivaut à dilater. La fonction dont il faut détecter les passages à zéro s’écrit donc, à l’échelle σ :

$$\begin{aligned} (-\Delta G_\sigma * I)(\vec{x}) &= (\psi_\sigma * I)(\vec{x}) \\ &= \int d\vec{y} I(\vec{y}) \psi_\sigma(\vec{x} - \vec{y}) \\ &= \int d\vec{y} I(\vec{y}) \frac{1}{\sigma^2} \psi_1\left(\frac{\vec{x} - \vec{y}}{\sigma}\right). \end{aligned} \quad (2.3)$$

Cette quantité est précisément la *transformée en ondelettes* de l’image I avec l’ondelette $\tilde{\psi}_1(\vec{x}) = \psi_1(-\vec{x})$.

Afin de mettre en oeuvre le principe de multirésolution, Marr et Hildreth [1] proposent la stratégie suivante. A partir d’une image donnée, on construit des versions de plus en plus dégradées (lissées),

¹Pour une fonction d’une variable f , un bord peut être défini comme un point d’inflexion, c.-à-d. un point où $f'' = 0$.

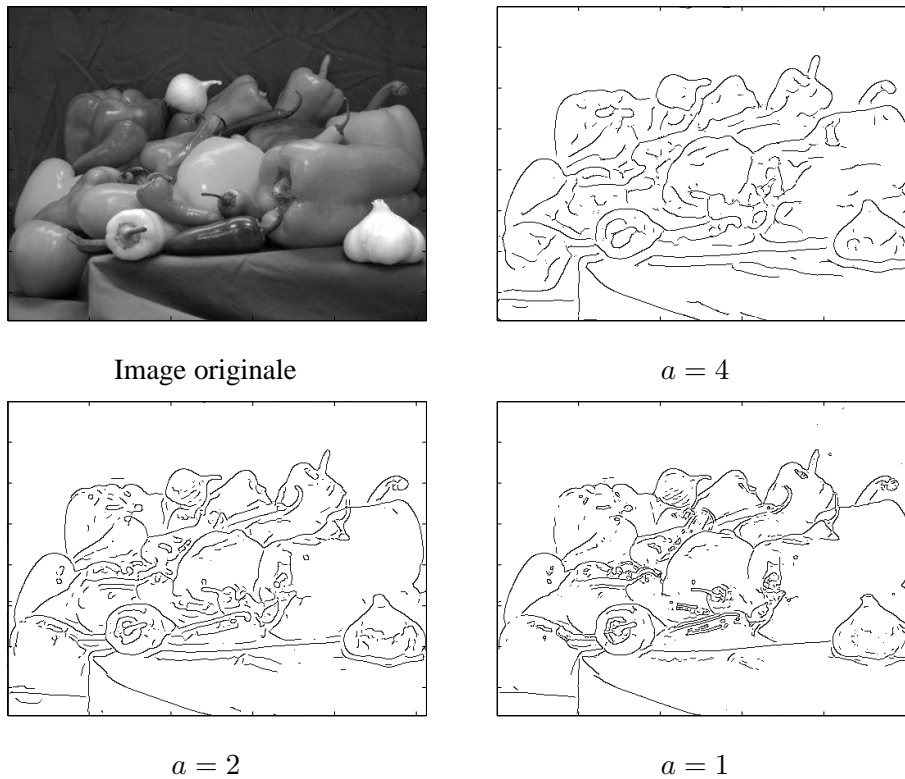


Figure 8: Contours contenus dans une image à des résolutions de plus en plus fines.

jusqu'à arriver au "Primal sketch" (Marr), constitué des "zero crossings" ou passages à zéro de la transformée (en ondelettes) (2.3). Des exemples frappants sont présentés dans la référence [1] (voir aussi la Figure 8). Fait remarquable, cette procédure est en fait optimale. En effet, Mallat et Zhong [3] ont démontré que l'ensemble des "zero crossings" (de la transformée en ondelettes) suffit à caractériser entièrement l'image.

III. Structure des images

A. Le "scale space"

Appliquant la stratégie de Marr et Hildreth, on obtient une série d'images lissées à des échelles de plus en plus grossières. Un problème évident est de faire le lien entre ces différentes échelles. Une solution a été proposée par Witkin [4] et Koenderink [5], utilisant une représentation tridimensionnelle, appelée "scale space" et consistant à ajouter aux coordonnées x, y du plan de l'image une coordonnée verticale z , représentant la résolution ou l'échelle, suivant le schéma indiqué à la Figure 9(a).

Le principe de base peut s'énoncer comme suit : afin que les points représentant un détail particulier de l'image originale $u_o(\vec{x})$ puissent être identifiés dans les images aux différentes résolutions, il faut plonger celles-ci dans une famille à un paramètre d'images dérivées $u(\vec{x}, z)$, $z > 0$, avec $u(\vec{x}, 0) = u_o(\vec{x})$. Dans cette paramétrisation, z représente une échelle "interne" ou la résolution, décroissante pour z croissant, comme indiqué sur la Figure 9(b).

Ceci dit, il faut identifier les contraintes à imposer à $u(\vec{x}, z)$ de façon à assurer la cohérence entre les différentes échelles. Selon Koenderink [5], cela revient à établir le lien entre $u(\vec{x}, z)$ et $\partial_z u(\vec{x}, z)$.

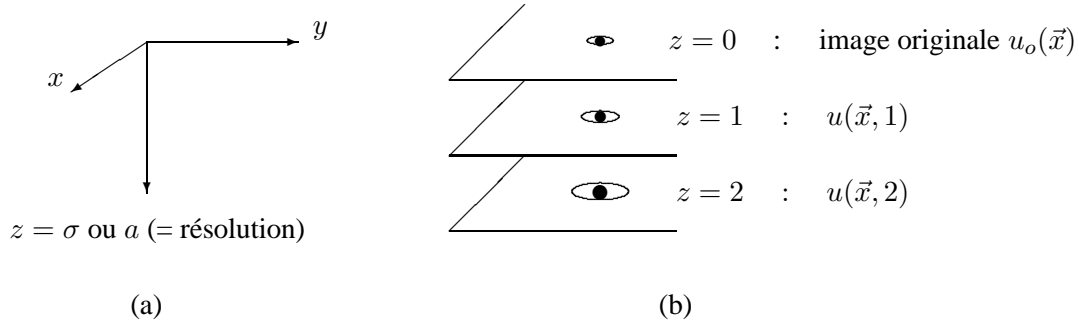


Figure 9: Utilisation du “scale space” : (a) Représentation schématique; (b) Principe de cohérence entre échelles différentes.

On peut procéder en trois étapes:

- (1) *Principe de causalité* : $u(\vec{x}, z_1)$ contient plus d’information que $u(\vec{x}, z_2)$ si $z_1 < z_2$:
 - l’information se dégrade quand z croît
 - chaque détail à z donné possède un prédécesseur à une résolution plus fine
 - propriété de semi-groupe : $u(\vec{x}, z_2)$ peut se déduire de $u(\vec{x}, z_1)$ sans connaître l’image originale $u(\vec{x}, 0)$ (mais pas l’inverse : il s’agit d’un semi-groupe, pas d’un groupe !)
- (2) *Considérations géométriques* dans le “scale space” (\vec{x}, z) , portant essentiellement sur la courbure, qui mènent à la relation

$$\Delta u = \alpha^2(\vec{x}, z) \frac{\partial u}{\partial z}.$$

- (3) *Homogénéité et isotropie* à chaque échelle : $\alpha^2 = \alpha^2(z)$. Introduisant la reparamétrisation : $z \mapsto t = \varphi(z)$, $u = u(\vec{x}, t)$, on arrive finalement à l’équation de la chaleur

$$\frac{\partial u}{\partial t} = \Delta u. \quad (3.1)$$

La solution générale s’écrit dès lors :

$$u(\vec{x}, t) = \frac{1}{4\pi t} \int u_o(\vec{y}) \exp\left(-\frac{|\vec{x} - \vec{y}|^2}{4t}\right) d\vec{y}. \quad (3.2)$$

On peut imposer en outre le principe de *causalité forte*, à savoir, la conservation des arêtes = “edges” :

$$\{\text{arêtes à l'échelle } t_2\} \subset \{\text{arêtes à l'échelle } t_1\} \text{ si } t_2 > t_1$$

Ce principe est vérifié pour la méthode de Witkin, mais pas vraiment dans l’algorithme de Canny [6], basé sur un filtre $\simeq \partial G_\sigma$ (cet algorithme est néanmoins standard en traitement d’images et considéré par beaucoup comme le meilleur détecteur d’arêtes).

Bien entendu, la résolution (3.2) de l’équation de la chaleur (3.1) se fera numériquement, en discrétisant aussi bien le temps ($t \mapsto t_k$ avec $t_{k+1} > t_k$ et $k \in \mathbb{N}$) que la position ($u(\vec{x}, t) \mapsto u_k(m, n)$)



Figure 10: Diffusion anisotrope par la méthode de Perona et Malik; k est croissant vers la droite.

avec $u_0(m, n)$ l'image initiale). On obtient ainsi la solution en itérant l'opération de lissage, sous la forme

$$u_{k+1} = u_k + (t_{k+1} - t_k)\Delta u_k,$$

où Δu_k représente l'accroissement obtenu par convolution, comme indiqué en (1.2). Ceci représente bien le phénomène de diffusion dans l'image, traité ici de façon isotrope.

B. Formulation variationnelle

La méthode décrite ci-dessus possède une formulation variationnelle, qui présente des avantages numériques indéniables. On définit l'énergie associée, qui donne la quantité d'information contenue dans u , comme :

$$E(u) = \int |\nabla u|^2 d\vec{x}.$$

Dès lors, résoudre (numériquement) l'équation de la chaleur revient à minimiser itérativement la fonctionnelle (pour $t = \lambda^2$ petit)

$$E_\lambda(u) = \lambda^2 \int |\nabla u|^2 d\vec{x} + \int (u - u_o)^2 d\vec{x}.$$

Cette formulation se prête bien aux généralisations que nous allons introduire ci-dessous.

IV. Généralisations

La théorie linéaire de Marr-Hildreth-Witkin n'est pas suffisante. En effet, elle est numériquement très lourde et, en outre, les arêtes à basse résolution donnent une vue incorrecte des bords des différentes régions. Différentes généralisations ont donc été proposées dans la littérature, et elles sont toutes non linéaires [9].

(1) Méthode de diffusion de Perona et Malik

L'idée est d'introduire une partie de la détection d'arêtes dans le filtrage, en permettant dès le début une interaction entre les échelles [7]. Ceci conduit à l'équation

$$\frac{\partial u}{\partial t} = \text{div} (f(|\nabla u|) \nabla u), \quad u(0) = u_o, \quad (4.1)$$

dans laquelle f est une fonction régulière, décroissante, positive, telle que $f(0) = 1$ et $f(s) \rightarrow 0$ pour $s \rightarrow \infty$. L'exemple-type est $f(s) = \frac{1}{s}$ ou encore $f(s) = \frac{1}{1+s}$.

Cette technique mène à un filtrage conditionnel :

- pour $\nabla u(\vec{x})$ grand, la diffusion est faible, les arêtes restent stables;
- pour $\nabla u(\vec{x})$ petit, la diffusion lisse davantage autour de \vec{x} .

Le résultat est une bien meilleure localisation des arêtes (et donc une bonne réalisation du principe de causalité forte), mais la méthode présente des difficultés en présence de bruit important (car dans ce cas $\nabla u(\vec{x})$ oscille rapidement). Notons qu'ici aussi, il y a une formulation variationnelle [9].

Bien sûr, l'équation (4.1) sera aussi résolue par discrétisation, ce qui donne :

$$u_{k+1} = u_k + \nabla \cdot (f(|\nabla u_k|) \nabla u_k), \quad (4.2)$$

avec f une fonction décroissante, p. ex. $f(s) = \frac{1}{1+s}$. Il résulte de (4.2) que la diffusion est faible sur les zones à fort gradient, c.-à-d. les contours. La figure 10 montre le résultat sur l'image cameraman. Les contours sont manifestement mieux préservés au cours du lissage que dans le cas de la diffusion isotrope utilisé dans les Figures 4 et 5.

(2) Méthode de la variation totale

Un cas particulier de la méthode de Perona-Malik, introduit par Osher et Rudin [8], consiste à prendre $f(s) = \frac{1}{s}$. Ceci donne

$$\frac{\partial u}{\partial t} = \operatorname{div} \left(\frac{\nabla u}{|\nabla u|} \right), \quad u(0) = u_o,$$

correspondant à la minimisation de la fonctionnelle

$$E_\lambda(u) = \lambda^2 \int |\nabla u| d\vec{x} + \int (u - u_o)^2 d\vec{x}.$$

(3) Diffusion anisotrope

Le point de départ de cette généralisation est le constat que les arêtes sont régulières dans la direction de la tangente, le saut brusque est dans la direction perpendiculaire. Il faut dès lors diffuser davantage dans la direction tangentielle, ce qui mène à l'équation

$$\frac{\partial u}{\partial t} = \Delta u - \frac{D^2 u(\nabla u, \nabla u)}{|\nabla u|^2}, \quad u(0) = u_o.$$

A noter que l'on retrouve exactement la même idée dans la définition des "ridgelets" et des "curvelets" généralisant les ondelettes 2-D, que nous étudierons dans la section V.E.

(4) Image sur une surface S

Dans ce cas, l'équation de la chaleur devient

$$\frac{\partial u}{\partial t} = \Delta_S u,$$

où Δ_S est l'opérateur de Laplace-Beltrami de S (avec $\Delta_S = \Delta$ si S est plane). Un cas particulièrement intéressant est celui de la sphère.

En conclusion de cette section, nous pouvons affirmer que le traitement des images par des méthodes mathématiques (EDP, méthodes variationnelles) est un domaine de recherche très actif, avec des applications telles que la vision robotique ou les techniques de déformation des images 2-D et 3-D ("morphing"). Pour tous ces développements, nous renverrons le lecteur à l'ouvrage de Morel et Solimini [9].

V. Transformée en ondelettes bidimensionnelle

A. Dérivation intuitive

Nous représenterons une *image* par une fonction définie sur le plan et de carré intégrable (“signal d’énergie finie”), $I \in L^2(\mathbb{R}^2, d^2\vec{x})$:

$$|I|^2 = \int_{\mathbb{R}^2} d^2\vec{x} |I(\vec{x})|^2 < \infty$$

Pratiquement, une image en noir et blanc sera représentée par une fonction positive bornée, prenant des valeurs discrètes entre 0 et 255, correspondant au niveau de gris de chaque pixel :

$$0 \leq I(\vec{x}) \leq M, \forall \vec{x} \in \mathbb{R}^2 \quad (M > 0)$$

Le principe de l’analyse en ondelettes est d’analyser l’image point par point à l’aide d’une *sonde* ψ :

$$\langle \psi | I \rangle = \int_{\mathbb{R}^2} d^2\vec{x} \overline{\psi(\vec{x})} I(\vec{x})$$

Nous appellerons *ondelette* une sonde admissible, c’est-à-dire une fonction de carré sommable, bien localisée en position *et* en fréquence et vérifiant la condition d’admissibilité (faible) :

$$\int_{\mathbb{R}^2} d^2\vec{x} \psi(\vec{x}) = 0. \quad (5.1)$$

Cette condition implique que l’ondelette est nécessairement une fonction oscillante.

L’étape suivante est de déterminer les opérations géométriques dans le plan que l’on veut utiliser, à savoir :

- (i) une translation par $\vec{b} \in \mathbb{R}^2$: $\vec{x} \mapsto \vec{x}' = \vec{x} + \vec{b}$
- (ii) une dilatation (zoom) par un facteur $a > 0$: $\vec{x} \mapsto \vec{x}' = a\vec{x}$
- (iii) une rotation d’angle θ : $\vec{x} \mapsto \vec{x}' = r_\theta(\vec{x})$ où r_θ est la matrice de rotation 2×2 habituelle

$$r_\theta \equiv \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}, \quad 0 \leq \theta < 2\pi.$$

Appliquées à un signal (image ou sonde), c’est-à-dire une fonction $s \in L^2(\mathbb{R}^2, d^2\vec{x})$, ces opérations sont réalisées par les opérateurs unitaires suivants :

- (i) translation : $(T_{\vec{b}}s)(\vec{x}) = s(\vec{x} - \vec{b})$, $\vec{b} \in \mathbb{R}^2$
- (ii) dilatation : $(D_a s)(\vec{x}) = a^{-1} s(a^{-1}\vec{x})$, $a > 0$
- (iii) rotation : $(R_\theta s)(\vec{x}) = s(r_{-\theta}(\vec{x}))$, $\theta \in [0, 2\pi)$.

Combinant ces trois opérations, on obtient l’action suivante (similitude) :

$$U(\vec{b}, a, \theta) = T_{\vec{b}} D_a R_\theta.$$

La sonde transformée s’écrit donc :

$$\psi_{\vec{b}, a, \theta}(\vec{x}) \equiv [U(\vec{b}, a, \theta)\psi](\vec{x}) = a^{-1}\psi(a^{-1}r_{-\theta}(\vec{x} - \vec{b})).$$

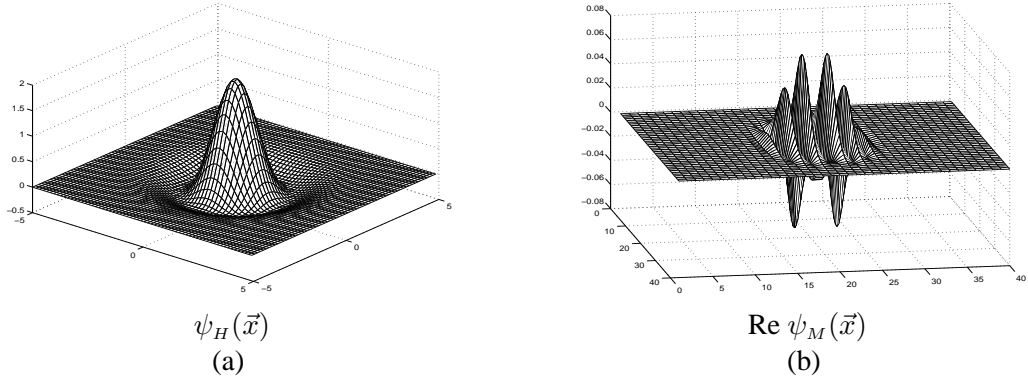


Figure 11: Deux ondelettes usuelles : (a) Chapeau mexicain isotrope; (b) Ondelette de Morlet.

Nous appellerons enfin *transformée en ondelettes de l'image* l'analyse de celle-ci à l'aide de la sonde transformée :

$$\begin{aligned}
 W_\psi I(\vec{b}, a, \theta) &= \langle \psi_{\vec{b}, a, \theta} | I \rangle \\
 &= a^{-1} \int_{\mathbb{R}^2} d^2 \vec{x} \overline{\psi(a^{-1} r_{-\theta}(\vec{x} - \vec{b}))} I(\vec{x})
 \end{aligned}$$

Les ondelettes les plus utilisées sont les suivantes (Figure 11):

- *Chapeau mexicain isotrope* (ondelette de Marr), déjà présenté à la Figure 7 :

$$\psi_H(\vec{x}) = -\Delta G(\vec{x}) = (2 - |\vec{x}|^2) \exp(-\frac{1}{2}|\vec{x}|^2).$$

On retrouve exactement le contexte de l'équation de la chaleur !

- *Ondelette de Morlet* :

$$\psi_M(\vec{x}) = \exp(i\vec{k}_o \cdot \vec{x}) \exp(-\frac{1}{2}|\vec{x}|^2) + \text{terme correctif}$$

Dans cette définition, le terme correctif garantit que l'ondelette vérifie la condition d'admissibilité (5.1). On peut montrer toutefois que ce terme est numériquement négligeable dès que $|\vec{k}_o| > 5.5$.

B. Propriétés essentielles

On peut résumer comme suit les propriétés essentielles de la transformée en ondelettes 2-D (T.O.). Pour plus de détails, nous renvoyons le lecteur à l'ouvrage classique d'I. Daubechies [10], et surtout au tout récent volume [11], spécifiquement consacré à l'analyse d'images par ondelettes.

- (1) La T.O. est une convolution avec une fonction oscillante de moyenne nulle. Dès lors, la T.O. prend des valeurs significatives uniquement là où l'ondelette et l'image se ressemblent. Autrement dit, la T.O. réalise un *filtrage* en position, en échelle et en direction.
- (2) La T.O. est *inversible* si et seulement si l'ondelette est admissible. Ceci mène à la *formule de reconstruction*

$$\begin{aligned}
 I(\vec{x}) &= c_\psi^{-1} \iiint_G d^2 \vec{b} \frac{da}{a^3} d\theta \psi_{\vec{b}, a, \theta}(\vec{x}) W_\psi I(\vec{b}, a, \theta) \\
 G &\equiv \{\vec{b} \in \mathbb{R}^2, a > 0, \theta \in [0, 2\pi)\}
 \end{aligned} \tag{5.2}$$

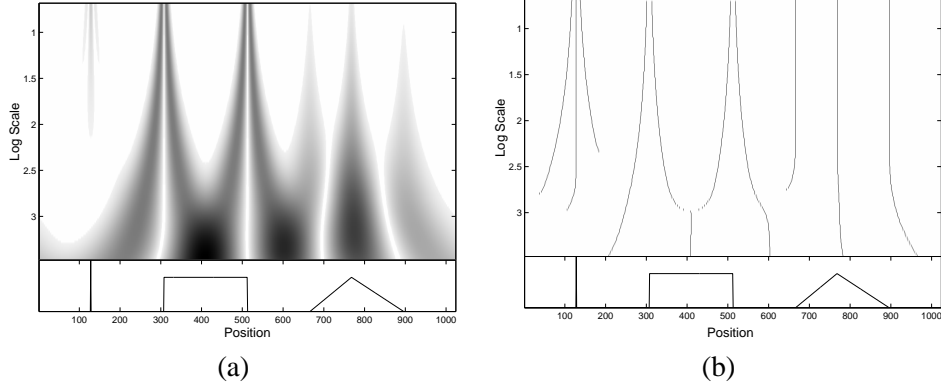


Figure 12: Arêtes en 1-D : Signal discontinu (représenté dans le panneau inférieur), analysé avec un chapeau mexicain : (a) T.O. du signal; (b) Squelette de la T.O.

où ψ est (fortement) admissible si et seulement si $c_\psi < \infty$ (cette condition implique la condition d'admissibilité faible (5.1), qui n'est que nécessaire, mais en général suffisante en pratique).

La formule (5.2) peut s'interpréter comme la représentation de l'image $I(\vec{x})$ par une combinaison linéaire des ondelettes $\psi_{\vec{b},a,\theta}(\vec{x})$: autrement dit, la famille $\{\psi_{\vec{b},a,\theta}\}$ joue le rôle d'une base continue.

(3) *Propriétés de support* en fonction de l'échelle :

- Pour $a \gg 1$, $\psi_{\vec{b},a,\theta}$ est une fenêtre large, centrée autour d'une basse fréquence : cette ondelette est sensible aux *basses* fréquences.
- Pour $a \ll 1$, $\psi_{\vec{b},a,\theta}$ est une fenêtre étroite, centrée autour d'une haute fréquence : cette ondelette est sensible aux *hautes* fréquences.
- Il en résulte que la T.O. est la plus efficace à haute fréquence, c'est-à-dire à petite échelle, et donc la T.O. est un *détecteur de singularités*. En particulier, la T.O. est efficace pour détecter les arêtes : elle sera donc un outil privilégié pour l'opération de "edge detection" !

(4) *Redondance* de la T.O. :

$f \in L^2(G)$ est la T.O. d'un certain signal si et seulement si f vérifie la propriété de *reproduction* :

$$f(\vec{b}', a', \theta') = c_\psi^{-1} \iiint_G d^2\vec{b} \frac{da}{a^3} d\theta \langle \psi_{\vec{b}',a',\theta'} | \psi_{\vec{b},a,\theta} \rangle f(\vec{b}, a, \theta).$$

Ceci signifie que la valeur de f au point (\vec{b}', a', θ') est déterminée par les valeurs prises aux points (\vec{b}, a, θ) dans un voisinage dont la taille est contrôlée par la fonction $\langle \psi_{\vec{b}',a',\theta'} | \psi_{\vec{b},a,\theta} \rangle$ (appelée noyau reproduisant). Il s'ensuit que la T.O. est extrêmement *redondante* !

C. Discrétisation de la T.O. : bases et repères

La propriété de redondance de la T.O. signifie que toute l'information est contenue dans un (petit) sous-ensemble de la T.O. ! Deux possibilités sont utilisées en pratique, les arêtes et les ensembles discrets.

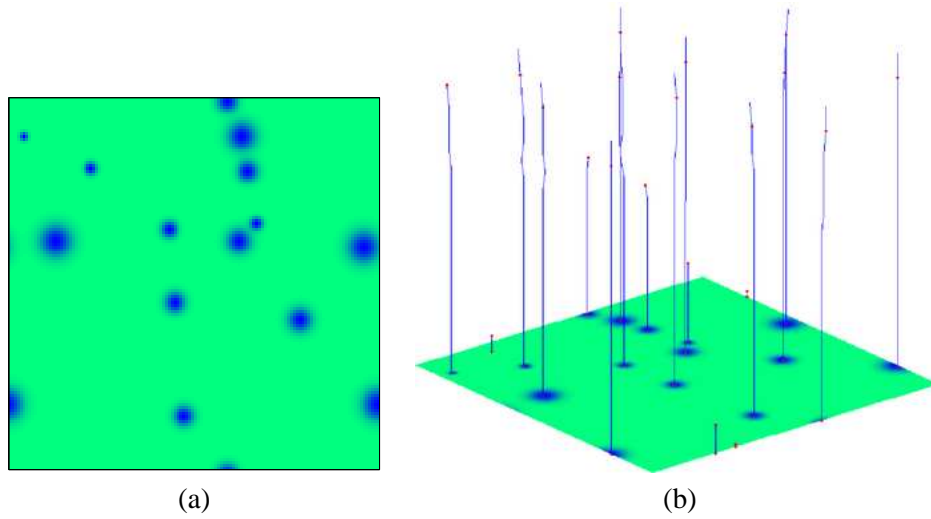


Figure 13: Arêtes en 2-D : Simulation de points brillants à la surface du Soleil : (a) T.O. du signal; (b) Squelette de la T.O.

(1) *Arêtes (ridges) et squelette*

Les *arêtes* de la T.O. sont essentiellement des lignes de maxima locaux et le *squelette* de la T.O. est sa restriction aux arêtes. L'intérêt de cette notion est que le squelette contient essentiellement toute l'information de la T.O. . Il est clair sur les exemples (voir Figures 12 et 13) que l'utilisation du squelette à la place de la T.O. elle-même réduit considérablement le coût numérique du calcul !

(2) *Sous-ensemble discret, réseau*

Si on se limite à un ensemble discret de points, l'intégrale dans la formule de reconstruction (5.2) est remplacée par une somme discrète

$$\tilde{I}(\vec{x}) = \sum_{mjl} \psi_{\vec{b}_m, a_j, \theta_l}(\vec{x}) W_{\psi} I(\vec{b}_m, a_j, \theta_l), \quad (5.3)$$

où l'on espère \tilde{I} aussi proche que possible de I . Devant une telle formule, il s'agit d'abord de vérifier la convergence rapide de la somme, ensuite de tronquer celle-ci après un petit nombre de termes et estimer l'erreur ainsi commise.

Le problème est que la famille $\mathcal{I} = \{\psi_{\vec{b}_m, a_j, \theta_l}\}$ n'est en général pas une base orthonormale (la représentation est donc redondante !). Le meilleur cas possible est celui où la famille \mathcal{I} est un *repère (frame)*, ce qui signifie

$$A \|f\|^2 \leq \sum_{mjl} |\langle \psi_{\vec{b}_m, a_j, \theta_l} | f \rangle|^2 \leq B \|f\|^2, \quad \forall f \in L^2, \quad 0 < A \leq B < \infty \quad (5.4)$$

Alors :

- . si $A = B > 1$, on dit que \mathcal{I} est un repère strict (tight frame);
- . si $A = B = 1$ et $\|\psi\| = 1$, \mathcal{I} est une base orthonormale.

Il faut noter que le développement (5.3) est en fait un développement en puissances de $|B/A - 1|$. Dès lors, un bon repère est un repère où $|B/A - 1| \ll 1$, en particulier, un repère strict. Ainsi les

deux ondelettes usuelles, le chapeau mexicain et l'ondelette de Morlet fournissent toutes deux de bons repères, mais non stricts [10]. Dans le cas d'un bon repère, on obtient, au premier ordre, $\tilde{I} = \frac{2}{A+B}I$, et bien sûr $\tilde{I} = I$ pour une base orthonormale

D. Solution alternative : la T.O. discrète

Une autre solution — et plus populaire — est d'utiliser la *transformée en ondelettes discrète*. Mais celle-ci procède d'une philosophie diamétralement opposée. En effet, on exige d'emblée une *base orthonormale* d'ondelettes, mais le prix à payer est que l'on perd le choix de l'ondelette-mère, car elle est dérivée indirectement (théorème d'existence) et on ne peut pas toujours la visualiser. Commençons par le cas unidimensionnel.

1. T.O. discrète en 1-D

La construction se fait à partir d'une *analyse multirésolution*, c.-à-d. une suite croissante de sous-espaces fermés de $L^2(\mathbb{R})$,

$$\dots \subset V_{-2} \subset V_{-1} \subset V_0 \subset V_1 \subset V_2 \subset \dots \quad (5.5)$$

où $\bigcap_{j \in \mathbb{Z}} V_j = \{0\}$ et $\bigcup_{j \in \mathbb{Z}} V_j$ est dense in $L^2(\mathbb{R})$, et telle que

- (1) $f(x) \in V_j \Leftrightarrow f(2x) \in V_{j+1}$
- (2) Il existe une fonction $\phi \in V_0$ (fonction d'échelle), telle que la famille $\{\phi(x - k), k \in \mathbb{Z}\}$ soit une base orthonormale de V_0 .

Dans la suite (5.5), V_j est interprété comme un espace d'approximation à la résolution 2^j . La condition (1) signifie qu'aucune échelle n'est privilégiée. D'autre part, les détails additionnels nécessaires pour passer de 2^j à 2^{j+1} seront décrits par le sous-espace W_j , complément orthogonal de V_j dans V_{j+1} :

$$V_j \oplus W_j = V_{j+1}.$$

On a donc :

$$L^2(\mathbb{R}) = \bigoplus_{j \in \mathbb{Z}} W_j = V_{j_0} \oplus \left(\bigoplus_{j=j_0}^{\infty} W_j \right) \quad (j_0 = \text{résolution minimale}) \quad (5.6)$$

Le théorème central de la théorie affirme alors que l'ondelette-mère ψ peut être dérivée à partir de la fonction d'échelle ϕ (on n'a donc aucun contrôle sur ψ !). Il en résulte que la famille $\{\psi_{jk}(x) \equiv 2^{j/2} \psi(2^j x - k), j, k \in \mathbb{Z}\}$ est base orthonormale de $L^2(\mathbb{R})$: ce sont les ondelettes orthogonales.

L'exemple le plus simple est la base de Haar (Figure 14):

- . fonction d'échelle : $\phi(x) = 1$ pour $0 \leq x < 1$, et 0 sinon
- . ondelette associée : $\psi_{\text{Haar}}(x)$

$$\psi_{\text{Haar}}(x) = \begin{cases} 1, & \text{if } 0 \leq x < 1/2 \\ -1, & \text{if } 1/2 \leq x < 1 \\ 0, & \text{sinon} \end{cases}$$

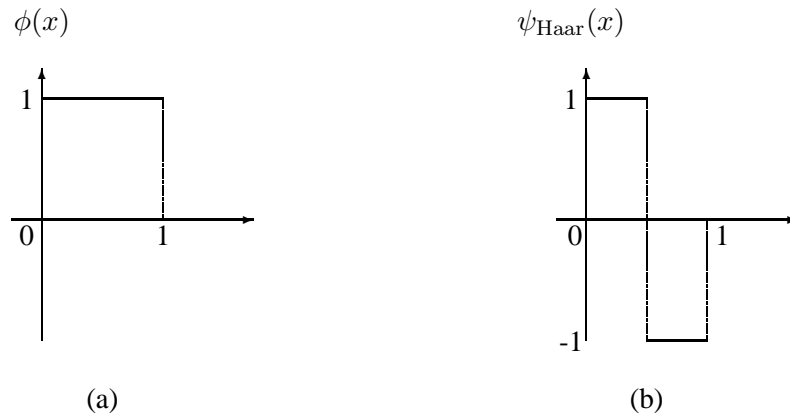


Figure 14: Base de Haar: (a) Fonction d'échelle $\phi(x)$; (b) Ondelette $\psi_{\text{Haar}}(x)$.

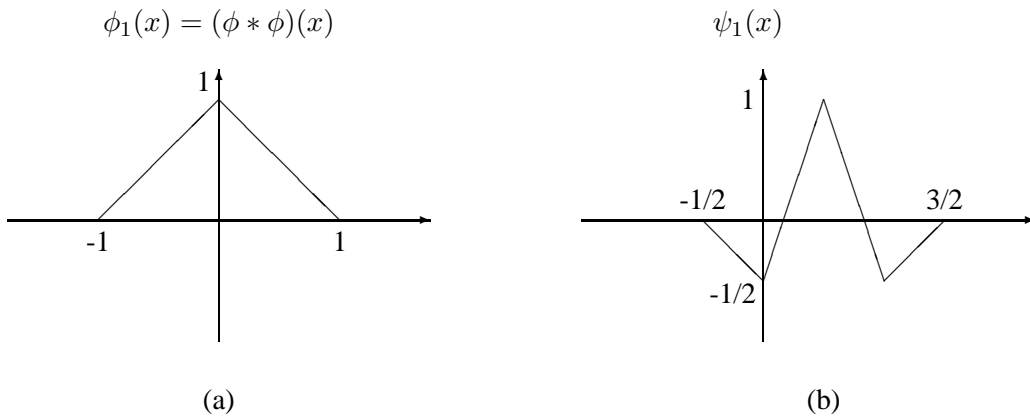


Figure 15: Ondelettes splines d'ordre 1 : (a) Fonction d'échelle $\phi_1(x)$; (b) Ondelette $\psi_1(x)$

A partir de la base de Haar, on peut construire les bases d'ondelettes splines d'ordres successifs, correspondant aux fonctions d'échelle

$$\begin{aligned}\phi_1 &= \phi * \phi \\ \phi_n &= \phi * \phi_{n-1}\end{aligned}$$

La fonction d'échelle ϕ_1 et l'ondelette-mère associée ψ_1 sont représentées à la Figure 15.

En pratique, le signal échantillonné est pris dans un certain V_J , on aura donc une représentation finie, au lieu de (5.6) :

$$V_J = V_{j_0} \oplus \left(\bigoplus_{j=j_0}^{J-1} W_j \right) \quad (5.7)$$

Un exemple explicite d'une telle représentation est donné à la Figure 16, qui montre la décomposition sur 6 niveaux d'un signal discontinu.

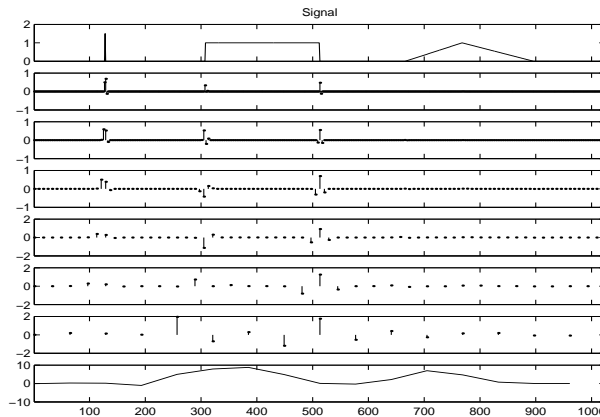


Figure 16: Décomposition sur 6 niveaux du signal discontinu.

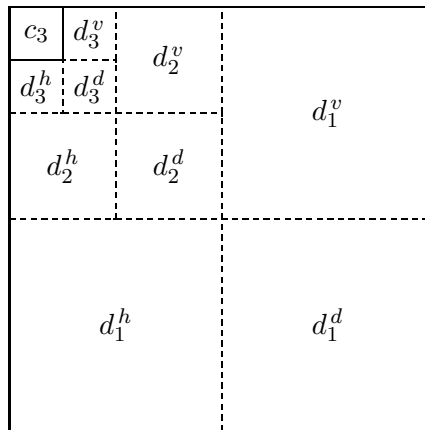


Figure 17: Schéma typique : décomposition en détails horizontaux, verticaux et obliques

2. T.O. discrète en 2-D :

Il n'y a pas grand chose à ajouter pour passer d'une à deux dimensions dans le cas de la T.O. discrète, l'extension se faisant par un simple produit tensoriel : $2\text{-D} = 1\text{-D} \otimes 1\text{-D}$. Concrètement, si $\{V_j, j \in \mathbb{Z}\}$ est une analyse multirésolution de $L^2(\mathbb{R})$, la famille $\{(2)V_j = V_j \otimes V_j, j \in \mathbb{Z}\}$ sera une analyse multirésolution de $L^2(\mathbb{R}^2)$. Dans cette construction, il faut une seule fonction d'échelle $\Phi(x, y) = \phi(x) \phi(y)$, mais 3 ondelettes :

$$\Psi^h(x, y) = \phi(x) \psi(y), \quad \Psi^v(x, y) = \psi(x) \phi(y), \quad \Psi^d(x, y) = \psi(x) \psi(y)$$

Mettant cette construction en oeuvre, on obtient le schéma de base, qui consiste à décomposer une image en une approximation de basse résolution, à laquelle on ajoute successivement trois types de détails, horizontaux, verticaux et obliques (Figure 17). Un exemple bien connu est présenté à la Figure 18, à savoir la décomposition sur trois niveaux de l'image lena.

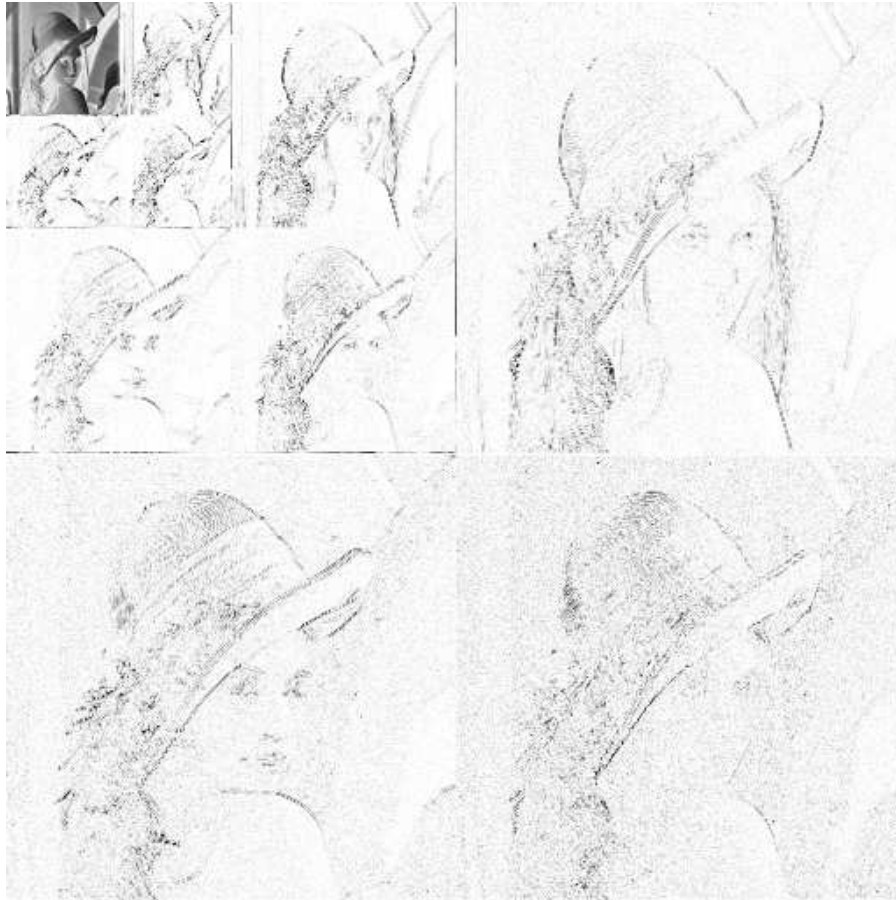


Figure 18: Décomposition typique de l'image Lena sur 3 niveaux.

3. Comparaison T.O. continue \Leftrightarrow T.O. discrète

Les deux méthodes ont des philosophies diamétralement opposées et, dès lors, elles ont des propriétés différentes et sont utiles pour des applications différentes :

- . T.O. continue : analyse des images, détection d'éléments caractéristiques (d'où son utilité pour le monitoring de divers processus);
- . T.O. discrète : compression en vue de la transmission (TV(HD), satellites), débruitage d'images.

A titre d'exemple, la Figure 19 montre un exemple de débruitage d'une image par ondelettes (qui dans ce cas possèdent un certain caractère directionnel).

En termes de redondance :

- . la T.O. continue est très redondante, ce qui lui confère une meilleure stabilité et une bonne résistance au bruit;
- . la T.O. discrète n'est pas redondante, ce qui implique une indépendance statistique optimale des coefficients, et lui permet donc d'atteindre d'excellents taux de compression.

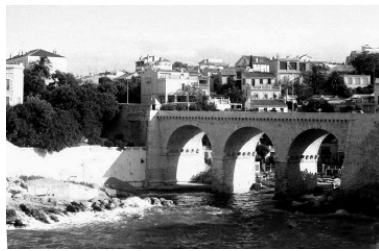


Image nette

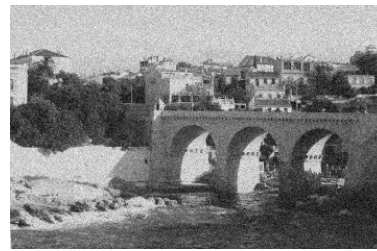


Image bruitée



Image débruitée

Figure 19: Débruitage d'une image.

E. Généralisation de la T.O.

Par définition, la T.O. discrète 2-D repose sur le pavage du plan en carrés dyadiques, de côté 2^{-j} , de plus en plus petits quand j augmente : ceci est un raffinement *isotrope*, parfaitement adapté pour détecter des singularités *ponctuelles*. Si on définit la complexité comme le nombre de carrés nécessaires pour recouvrir un élément donné de l'image, on constate que celle-ci augmente trop vite avec j , et donc la convergence perd beaucoup de sa rapidité ("curse of dimensionality"). Mais, en réalité, une courbe régulière est un objet très anisotrope : elle est régulière dans la direction parallèle, singulière dans la direction perpendiculaire, c'est donc pratiquement un objet 1-D. Dès lors, utiliser des ondelettes orthogonales 2-D pour représenter un tel objet est un gaspillage énorme !

La solution sera de construire des "sondes" plus adaptées à la *géométrie* des objets à représenter :

- les *ridgelets*, qui sont des fonctions de type ondelette dans une direction et constante dans la direction perpendiculaire. Les ridgelets fournissent une bonne représentation des lignes droites (Figure 20 (a)). Remarquons que l'on retrouve exactement l'idée de la diffusion anisotrope (lissage adaptatif) en traitement d'image !



Figure 20: Deux généralisations de la T.O. en 2-D : (a) Ridgelets; (b) Curvelets.

- les *curvelets* sont obtenues en utilisant une dilatation parabolique $a \times a^2$ et un paramètre de rotation : on obtient ainsi des “cellules” allongées, bien adaptées aux courbes (Figure 20 (b)).

Différentes autres techniques généralisant les ondelettes ont été proposées récemment, ce qui atteste bien l’actualité des problèmes du traitement d’image et l’intérêt de l’approche multirésolution sous-jacente à la méthode des ondelettes !

References

- [1] D. Marr and E. Hildreth, Theory of edge detection, *Proc. R. Soc. London B* **207**, 187–207 (1980)
- [2] D. Marr, *Vision*, Freeman, New York, 1982
- [3] S.G. Mallat and S. Zhong, Characterization of signals from multiscale edges, *IEEE Trans. Pattern Anal. and Mach. Intell.*, **14**: 710–732, (1992)
- [4] A.P. Witkin, Scale-space filtering, *Proc. 8th Int. Joint Conf. on Artif. Intell.*, Aug. 1983, Karlsruhe
- [5] J.J. Koenderink, The structure of images, *Biol. Cybern.* **50**, 363–370 (1984)
- [6] J. Canny, A computational approach to edge detection, *IEEE Trans. Pattern Anal. Machine Intell.*, **8**, 679–698 (1986)
- [7] P. Perona and J. Malik, A scale space and edge detection using anisotropic diffusion, *Proc. IEEE Computer Soc. Workshop on Computer Vision*, 1987
- [8] S. Osher and L. Rudin, Feature-oriented image enhancement using shock filters, *SIAM J. on Numerical Analysis*, **27**, 919–940 (1990); Shocks and other nonlinear filtering applied to image processing, *Proc. SPIE Appl. of Digital Image Processing XIV*, vol. 1567, 414–430 (1991)
- [9] J-M. Morel et S. Solimini, *Variational Methods in Image Segmentation*, Progress in Nonlinear Differential Equations and Their Applications, vol. 14, Birkhäuser, Boston, 1995
- [10] I. Daubechies, *Ten Lectures on Wavelets*, SIAM, Philadelphia, PA, 1992
- [11] J-P. Antoine, R. Murenzi, P. Vandergheynst and S.T. Ali, *Two-dimensional Wavelets and Their Relatives*, Cambridge University Press, Cambridge (UK), 2004