

Algorithmes optimaux pour la génération de pyramides d'images passe-bas et laplaciennes

Optimal Algorithms for Low-Pass and Laplacian Image Pyramids Computation



Alain CHÉHIKIAN

Laboratoire de Traitement
d'Images et Reconnaissance
de Formes, LTIRF/INPG,
46 avenue Félix Viallet,
38031 Grenoble Cedex

Professeur à l'Université Joseph Fourier de Grenoble, directeur du laboratoire LTIRF. Domaines de recherches : description d'images, algorithmes rapides, vision par ordinateur, architectures de machines de vision.

RÉSUMÉ

La représentation multi-échelle des images est un outil très utilisé en vision artificielle, car elle procure un moyen naturel de décrire une image par une hiérarchie de structures. Cependant, de par le processus de décimation mis en œuvre, ces algorithmes génèrent des images bruitées par le repliement de spectre, phénomène d'autant plus gênant que la recherche de structures dans ces images passe par une opération préalable de dérivation. Le choix du noyau de filtrage est donc essentiel et paradoxalement, le problème de la création du bruit de recouvrement de spectre n'a pas été, ou mal été, pris en compte pour sa définition. Nous étudions dans cet article deux algorithmes très utilisés pour générer une

représentation pyramidale d'image, et ce par rapport au bruit de repliement de spectre qu'ils créent. Nous définissons un noyau de filtrage optimal qui : minimise le bruit de repliement de spectre, ne crée pas de structures nouvelles, a de bonnes propriétés de symétrie en rotation et de coût de calcul réduit. Nous proposons un algorithme rapide pour la génération de pyramides d'octave passe-bas et laplacien.

MOTS CLÉS

Représentation pyramidale d'images. Pyramide gaussienne. Pyramide de Laplaciens. Repliement de spectre.

ABSTRACT

Multiple resolution representations are often used in computer vision, as they provide a natural way for describing an image by a hierarchy of structures. However, because of the resampling process, the images they provide are corrupted by an aliasing noise which makes difficult the detection of structures, specially when the detection process implies the computation of derivatives. Choosing the filtering kernel is thus essential. Paradoxically, the importance of proper signal-to-noise analysis have been widely neglected by the vision community. In this paper, we study two commonly used algorithms from the point of view of the aliasing noise

they create. Then we propose an optimum filtering kernel which : minimizes the aliasing noise, does not create new structures, has interesting properties of rotational symmetry and reduced computation cost. We also propose a fast algorithm for the computation of low-pass and laplacian octave-spaced pyramids.

KEY WORDS

Pyramid representation. Gaussian pyramids. Laplacian pyramids. Aliasing noise.

1. Introduction

La représentation multi-échelle des images est un outil très utilisé en vision artificielle, car elle procure un moyen naturel de décrire une image par une hiérarchie de structures. Elle est également un moyen privilégié de réduire la charge de calcul des algorithmes mettant en œuvre des procédures de mise en correspondance, tels que les algorithmes de mesure de disparité ou les algorithmes de mesure de champs de déplacement (« coarse-to-fine matching »). Pour ces applications, trois qualités essentielles sont exigées d'une telle représentation : i) ne pas créer des structures non existantes dans l'image d'origine, ii) traiter les structures présentes dans l'image indépendamment de leur orientation, iii) ne pas annihiler par la charge de calcul de l'algorithme de représentation multi-échelle le bénéfice escompté par l'utilisation de cette représentation. Les deux premiers critères sont habituellement satisfaits par le choix d'un noyau de filtrage de type gaussien qui en outre, grâce à sa propriété de séparabilité, conduit à une réduction de la charge de calcul. Cependant, de par le processus de décimation mis en œuvre, ces algorithmes génèrent des images bruitées, phénomène d'autant plus gênant que la recherche de structures dans ces images passe par une opération préalable de dérivation : recherche de contours, coins, taches, pics, crêtes, etc... Le choix du noyau de filtrage est donc essentiel et paradoxalement, le problème de la création du bruit de recouvrement de spectre n'a pas été, ou mal été, pris en compte dans la définition de ce noyau. Nous nous proposons, dans ce qui suit, de montrer qu'on peut définir un noyau de filtrage optimal tant du point de vue des propriétés de la représentation multi-échelle, que du point de vue du bruit de repliement de spectre et du coût de calcul. Les propriétés d'une représentation multi-échelle dépendant à la fois des propriétés du noyau de filtrage et de la procédure de décimation mise en œuvre, nous étudierons les conséquences de l'utilisation de ces noyaux dans les algorithmes de Burt [1], [2] d'une part, de Crowley [3] d'autre part. Ces algorithmes ont été choisis car ils sont les plus cités et utilisés dans la littérature de la vision artificielle. Nous comparerons les résultats ainsi obtenus.

Une version préliminaire de cet article a été publiée [4], nous donnons ici une version étendue. L'article sera organisé comme suit. En section Deux, nous rappellerons les algorithmes de Burt et de Crowley qui nous serviront de référence. En section Trois, nous définirons une famille de noyaux de filtrage pour la représentation échelle-espace, nous étudierons les propriétés des algorithmes de Burt et de Crowley lorsque les noyaux de filtrage appartiennent à cette famille, nous critiquerons ces algorithmes et proposerons un algorithme qui possède leurs avantages mais corrige leurs inconvénients. En section Quatre, après avoir rappelé les techniques utilisées par ces mêmes auteurs pour construire une pyramide de laplaciennes, nous proposerons un algorithme qui tout en réduisant le coût de calcul donne des images de laplaciennes de meilleure qualité. Nous concluons en section Cinq.

2. Construction d'une pyramide d'images passe-bas

Cette représentation des images tire son nom du processus de création qui résulte d'une suite de filtrages passe-bas et de re-échantillonnage produisant un ensemble d'images de taille $N \times N$, $N/2 \times N/2$, $N/4 \times N/4$,..., où N est le nombre de lignes et de colonnes de l'image initiale. Nous décrivons dans ce qui suit les algorithmes de Burt d'une part et de Crowley d'autre part, qui nous serviront de référence.

2.1. L'ALGORITHME DE BURT

L'algorithme de Burt utilise un noyau de filtrage unique B pour produire une suite d'images d'indice $k \geq 0$ caractéristique du nombre de re-échantillonnages. La figure 1 décrit cet algorithme sous la forme d'un organigramme. L'image initiale I est filtrée par convolution avec ce noyau avant re-échantillonnage, et ce processus est répété pour chacun des niveaux k de la pyramide. A noter qu'à chaque niveau k de la pyramide, deux images de même bande passante peuvent être utilisées : les images I_k obtenues après filtrage mais avant re-échantillonnage et les images g_{k+1} obtenues après re-échantillonnage (en fait celles utilisées par Burt). Le coût de calcul des images g_{k+1} sera très inférieur à celui des images I_k mais comme nous le verrons ultérieurement (cf. fig. 2), les images g_{k+1} étant fortement bruitées par le repliement de spectre, seules nous intéresseront par la suite les images I_k .

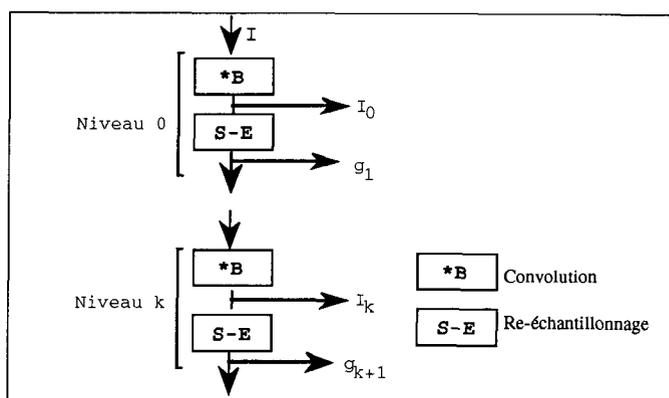


Figure 1. — L'algorithme de Burt.

Burt définit cinq contraintes que doit satisfaire le noyau de filtrage :

1. Séparabilité : vise à réduire le coût de calcul.
2. Normalisation : réalise un filtrage passe-bas.
3. Symétrie : la réponse en fréquence qui s'exprime par une somme de $\cos(n\omega)$ est réelle.
4. Réponse impulsionnelle unimodale : évite la création de faux contours.
5. Iso-contribution : chaque pixel de l'image initiale contribue avec le même poids à une image quelconque de la pyramide.

Burt propose des noyaux de taille 4×4 et 5×5 , nous ne nous intéresserons qu'à ces derniers (5×5), seuls à ne pas introduire de déphasage. L'élément unidimensionnel du noyau est alors :

$$Bx = [b_2 \ b_1 \ b_0 \ b_1 \ b_2]$$

Avec les contraintes définies précédemment :

$$b_1 = 0,25$$

$$0,25 \leq b_0 \leq 0,5$$

$$b_2 = 0,25 - 0,5 \ b_0.$$

La réponse en fréquence du noyau bidimensionnel est alors :

$$B(u, v) = [b_0 + 2 \ b_1 \cos(u) + 2 \ b_2 \cos(2u)] \times [b_0 + 2 \ b_1 \cos(v) + 2 \ b_2 \cos(2v)].$$

La réponse en fréquence associée à un niveau k de la pyramide est :

$$R_k = \prod_{m=0}^k F(2^m u, 2^m v).$$

La fréquence de Nyquist associé à ce même niveau k , étant :

$$u_{Nk} = v_{Nk} = \pi/2^k \quad \text{pour les images } I_k,$$

$$u_{Nk} = v_{Nk} = \pi/2^{k+1} \quad \text{pour les images } g_{k+1}.$$

Pour ces images, en supposant que le spectre de l'image d'origine est constant dans $[-\pi, +\pi]$, le rapport signal-à-bruit dû au bruit de repliement de spectre est :

$$S/B_k = \sqrt{\frac{\int_0^{u_{Nk}} \int_0^{v_{Nk}} R_k^2 dv du}{\int_{u_{Nk}}^{2u_{Nk}} \int_{v_{Nk}}^{2v_{Nk}} R_k^2 dv du}}$$

Nous avons calculé numériquement l'évolution de ce rapport signal-à-bruit lorsque b_0 varie de 0,2 à 0,5 déformant l'intervalle permis par les contraintes de Burt. La figure 2 montre les résultats obtenus avec les images g_{k+1} ($k = 0, 1, 2$ courbes de gauche), et avec les images I_k ($k = 1, 2$ courbes de droite).

On constate que dans les images g_{k+1} le repliement de spectre dégrade le rapport signal-à-bruit. Nous avons repéré sur ces courbes les rapports signal-à-bruit 36 db et 48 db correspondant respectivement au rapport signal-à-bruit d'une image fournie par une caméra de bonne qualité (64 niveaux de gris) et au rapport signal-à-bruit maximum possible lors d'une quantification des pixels sur 256 niveaux (on suppose que le signal est équiréparti dans $[0, 2^N - 1]$, le bruit de quantification étant équiréparti dans $[-0,5, +0,5]$). Ce phénomène est particulièrement sensible au niveau $k = 0$ de la pyramide, il s'atténue ensuite grâce au filtrage effectué au niveau initial. Par contre,

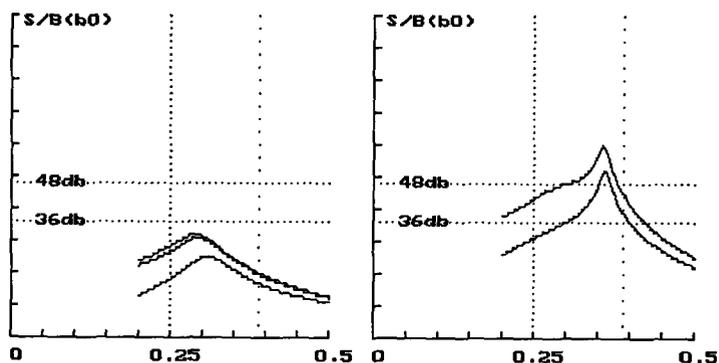


Figure 2. — Évolution du rapport signal-à-bruit en fonction de b_0 dans les images de la pyramide de Burt. A gauche : images g_{k+1} ($k = 0, 1, 2$ de bas en haut). A droite : images I_k ($k = 1, 2$ de bas en haut).

grâce à ce même filtrage initial, les images I_k sont nettement moins bruitées par le repliement de spectre. Par la suite, nous considérerons que la pyramide de Burt sera constituée par les images I_k .

On constate également que le rapport signal-à-bruit est fortement dépendant de b_0 . Burt préconise $b_0 \cong 0,39$ (repéré sur les courbes) pour obtenir la meilleure approximation d'une pyramide gaussienne, on notera que cette valeur est loin d'optimiser le rapport signal-à-bruit.

Un autre inconvénient de l'algorithme de Burt concerne la bande passante aux différents niveaux de la pyramide. Les noyaux de filtrage étant symétriques et normalisés, la variance associée à une cascade de filtres est la somme des variances de chacun des filtres de la cascade. σ_0^2 étant la variance associée au noyau, compte tenu du coefficient d'échelle introduit par le re-échantillonnage, la variance associée au filtre générant l'image I_k est :

$$\sigma_k^2 = \sigma_0^2 \sum_{m=0}^k 2^{2m} = \sigma_0^2, 5 \sigma_0^2, 21 \sigma_0^2, 85 \sigma_0^2, \dots$$

tandis que la bande passante est proportionnelle à $1/\sigma_k$. Ainsi la pyramide générée est proche d'une pyramide d'octave, cependant les images ne sont pas filtrées de manière homogène par rapport à la fréquence d'échantillonnage, ce qui serait le cas si σ_k^2 était proportionnel à 2^{2k} .

2.2. L'ALGORITHME DE CROWLEY

Crowley utilise deux noyaux F et $F^{*2} = F * F$ pour produire une suite d'images $I_{k\ell}$, où $\ell \in \{0, 1, 2\}$ caractérise le niveau de filtrage à un niveau de re-échantillonnage donné. L'image initiale subit une première convolution par le noyau F . L'image ainsi obtenue est convoluée par F puis par F^{*2} pour générer les images de niveaux $\ell = 1$ et $\ell = 2$, avant le re-échantillonnage. Le processus est ensuite répété pour les divers niveaux de re-échantillonnage de la pyramide. La figure 3 décrit cet algorithme sous la forme d'un organigramme. On notera que l'image obtenue par re-échantillonnage est effectivement utilisée, nous le verrons par la suite, pour créer une pyramide de laplaciens par différence de passe-bas. Crowley utilise donc des noyaux

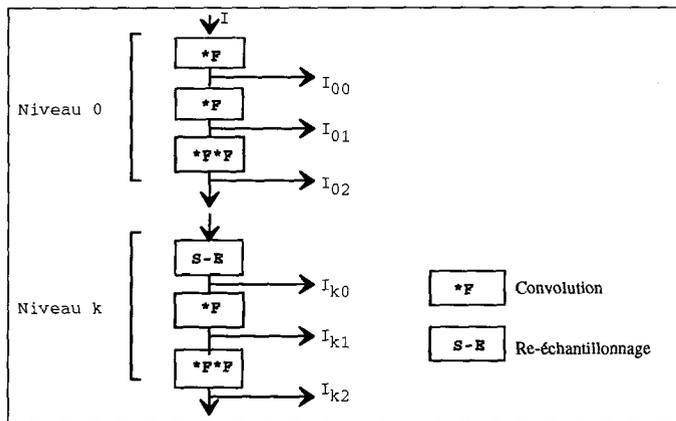


Figure 3. — L'algorithme de Crowley.

de filtrage de plus grande taille pour obtenir la réduction de bande passante souhaitable.

Dans la pyramide de Crowley, le noyau de filtrage F est un noyau binomial :

$$F(x, y) = \frac{1}{16^3} \begin{pmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{pmatrix}^{*3}$$

choisi pour approximer un noyau gaussien. La notation « *3 » en exposant signifie que le noyau est convolué 2 fois par lui-même, le noyau résultant approche une distribution gaussienne.

A noter que ce noyau satisfait aux contraintes de Burt.

Par ailleurs, la variance associée au filtre générant $I_{k,\ell}$ étant

$$\sigma_{k,\ell}^2 = 2^{2k+\ell} \sigma_0^2 = 2^r \sigma_0^2 = \sigma_0^2, 2 \sigma_0^2, 4 \sigma_0^2, 8 \sigma_0^2, \dots$$

$\ell \in \{1, 2\}$

la pyramide de Crowley fournit donc une représentation multi-échelle d'une 1/2 octave de résolution en fréquence.

3. Un algorithme optimal pour une représentation pyramidale d'image

Burt et Crowley utilisent des noyaux satisfaisant les mêmes contraintes, ils cherchent tous deux à approximer un noyau gaussien, mais les noyaux qu'ils définissent sont très différents tant par la taille que par leurs coefficients. Dans la démarche de Burt, les contraintes priment et les coefficients du noyau sont choisis en fonction de ces contraintes et pour réaliser un filtrage gaussien. Dans celle de Crowley, le caractère gaussien est recherché a priori, la taille du noyau est fixée de telle sorte qu'à la nouvelle fréquence de Nyquist définie par le re-échantillonnage, la réponse en fréquence soit suffisamment réduite pour que le bruit de repliement de spectre soit négligeable. Nous montrons dans ce qui suit qu'on peut définir une famille de

noyaux de filtrage de manière analytique en posant un certain nombre de contraintes que ces noyaux devront satisfaire. Cependant, le choix d'un noyau particulier doit tenir compte du processus de filtrage et décimation qui contribue à définir la bande passante de l'image générée à un niveau donné de la pyramide. Nous critiquerons donc en fonction de ces contraintes les choix faits par Burt et par Crowley et nous nous appuierons sur cette critique pour proposer un algorithme qui présente les avantages de ces algorithmes et en corrige les inconvénients.

3.1. UNE FAMILLE DE NOYAUX DE FILTRAGE POUR UNE REPRÉSENTATION ÉCHELLE-ESPACE

Les noyaux de filtrage appartenant à une telle famille doivent répondre aux contraintes :

Contrainte 1 : Séparabilité.

Contrainte 2 : Normalisation.

Contrainte 3 : Symétrie et dimension impaire.

Contrainte 4 : Une contrainte importante que doivent satisfaire ces noyaux est de ne pas créer au cours du processus de filtrage, des structures nouvelles dans l'image filtrée. Lindeberg [5] a montré que, pour que cette propriété soit acquise, la réponse impulsionnelle correspondante doit être unimodale et non négative, et la réponse en fréquence doit être unimodale et non négative dans le domaine de Nyquist défini par la fréquence d'échantillonnage courante. Ces contraintes sont nécessaires et suffisantes dans le cas unidimensionnel, elles sont nécessaires mais il n'est cependant pas démontré qu'elles soient suffisantes dans le cas bidimensionnel. Cette contrainte est une extension de la contrainte 4 de Burt.

Contrainte 5 : Pour ne pas introduire un bruit de repliement de spectre, la réponse en fréquence associée à ces noyaux doit être nulle en dehors du domaine de Nyquist. Cette contrainte ne pourra être satisfaite car elle conduirait à définir des noyaux à réponse impulsionnelle infinie, et en contradiction avec la contrainte précédente. Par contre on peut énoncer une contrainte que l'on pourra satisfaire : pour ne pas dégrader le rapport signal-à-bruit de l'image générée, le bruit résultant du repliement de spectre doit être négligeable par rapport au bruit de l'image (bruit du capteur et bruit résultant de la quantification). Cette contrainte est une extension de la contrainte 5 de Burt.

Considérons alors la famille de noyaux de la forme :

$$F(x, y, n) = f(x) * f(y)$$

avec :

$$f(x) = [w_1 w_0 w_1]^{*n}$$

La notation « *n » en exposant signifiant que $[w_1 w_0 w_1]$ est convolué $(n - 1)$ fois par lui-même.

La transformée de Fourier de $F(x, y, n)$ est :

$$F(u, v, n) = [w_0 + 2 w_1 \cdot \cos(u)]^n [w_0 + 2 w_1 \cdot \cos(v)]^n$$

Un tel noyau satisfait les contraintes 1 et 3. Il satisfait la contrainte 2 si

$$w_0 + 2 w_1 = 1,$$

il satisfait la contrainte 4 si :

$$0,5 \leq w_0 < 1,$$

enfin la variance associée à ce noyau est :

$$\sigma_0^2 = 2 w_1 n = (1 - w_0) n.$$

A ce stade nous ne pouvons mieux définir ni w_0 , ni n . Nous pouvons juste remarquer que w_0 contrôle la forme de la réponse en fréquence, tandis que n contrôle la taille du noyau et par conséquent la bande passante du filtre. Ces paramètres seront définis au moyen de la contrainte 5 mais il faudra pour cela considérer un processus particulier de filtrage-décimation.

3.2. APPLICATION A LA PYRAMIDE DE BURT

Burt utilisant un noyau de taille 5×5 , cela revient à imposer $n = 2$ et le noyau unidimensionnel correspondant aura pour coefficients :

$$[b_2 \ b_1 \ b_0 \ b_1 \ b_2] = [w_1 \ w_0 \ w_1]^* [w_1 \ w_0 \ w_1]$$

soit :

$$b_0 = w_0^2 + 2 w_1^2 = \frac{3 w_0^2 + 2 w_0 + 1}{2}$$

$$b_1 = 2 w_0 w_1 = w_0(1 - w_0)$$

$$b_2 = w_1^2 = \frac{(1 - w_0)^2}{4}$$

avec :

$$0,5 \leq w_0 < 1.$$

Comme dans le noyau de Burt, il n'existe qu'un paramètre libre w_0 , les coefficients du noyau : b_0 , b_1 , b_2 , s'en déduisent par les relations précédentes.

Nous avons calculé numériquement l'évolution du rapport signal-à-bruit lorsque w_0 varie de 0,3 à 0,7 débordant l'intervalle permis selon nos contraintes. La figure 4 montre les résultats obtenus pour l'image I_1 .

On constate sur cette courbe qu'il existe une valeur de w_0 qui maximise le rapport signal-à-bruit dû au repliement de spectre. En choisissant pour w_0 une valeur voisine de cet optimum : $w_0 = 0,5 - \epsilon$, la valeur du coefficient b_1 qui en résulte est $b_1 = 0,25 - \epsilon^2$. Burt en imposant $b_1 = 0,25$ (résultant du respect de la contrainte 5) se met donc d'emblée dans les conditions qui assureront un rapport signal-à-bruit convenable. Examinons alors les cas suivants :

* $w_0 = 0,475$ maximise le rapport signal-à-bruit, soit 50 db. Alors $b_0 = 0,363$ et $b_1 = 0,249$, cependant cette valeur ne respecte pas la contrainte d'unimodalité de la réponse en fréquence. Elle satisfait cependant le critère défini par Burt tel que : $0,25 \leq b_0 \leq 0,5$.

* $w_0 = 0,5$ permet d'atteindre le meilleur rapport signal-à-bruit possible tout en respectant la contrainte d'unimodalité en fréquence : 44 db, soit un bruit de repliement de spectre inférieur de 12 db au bruit d'image et de 4 db au bruit de

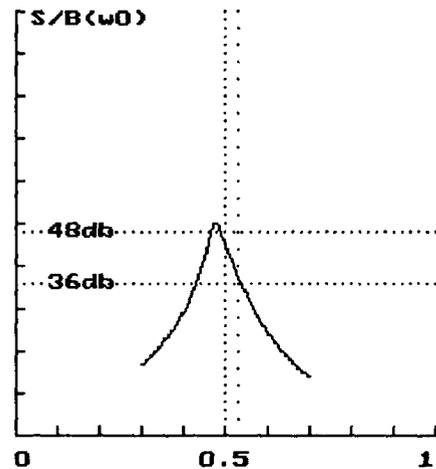


Figure 4. — Pyramide de Burt. Évolution du rapport signal-à-bruit en fonction de w_0 .

quantification. Alors le noyau optimal correspondant se met sous la forme :

$$B(x, y) = \frac{1}{16^2} \begin{pmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{pmatrix}^{*2}.$$

Ce noyau correspond aux coefficients de Burt :

$$b_0 = 0,375, \quad b_1 = 0,25, \quad b_2 = 0,0625.$$

* $w_0 = 0,528$. Alors $b_0 = 0,39$ et $b_1 = 0,249$, ces valeurs sont celles préconisées par Burt pour avoir la meilleure approximation d'une gaussienne. Cependant, ce choix de w_0 conduit à une dégradation notable du rapport signal-à-bruit qui est alors de 37 db.

Nous avons calculé pour les différentes images de la pyramide : le rapport signal-à-bruit, la dissymétrie en rotation caractérisée par la distance moyenne entre les sections à 0° et 45° de la réponse en fréquence, ainsi que la ressemblance avec une gaussienne caractérisée par la distance moyenne entre la réponse en fréquence et celle d'une gaussienne de même variance. Ces résultats sont résumés par la table I.

3.3. APPLICATION A LA PYRAMIDE DE CROWLEY

Crowley utilise un noyau F où les coefficients w_0 , w_1 sont fixés ($w_0 = 0,5$, $w_1 = 0,25$) ainsi que le paramètre $n = 3$.

La réponse en fréquence associée à une image $I_{k, \ell}$ de la pyramide est :

$$R_{k, \ell} = [F(2^k u, 2^k v)]^{(2^\ell - 1)} \prod_{m=0}^{k-1} R_{m, 2}, \quad k > 0$$

avec

$$R_{0, \ell} = [F(u, v)]^{(2^\ell - 1)} F(u, v)$$

où

$$F(u, v) = [w_0 + 2 w_1 \cos(u)]^n [w_0 + 2 w_1 \cos(v)]^n$$

Table 1.

Niveau k	0			1			2			3		
Algorithme	Burt	Crow.	Prop.	Burt	Crow.	Prop.	Burt	Crow.	Prop.	Burt	Crow.	Prop.
σ_k^2	1	3	1	5	12	4	21	48	16	85	192	64
S/B	—	—	—	45 db	174	58 db	51 db	171	57 db	52 db	171	57 db
$\sigma_k^2/2^{2k}$	1		1	1,25	0,33	1	1,31	0,33	1	1,33	0,33	1
Dissymétrie $\times 10^{-2}$	1,5	ϵ	1,5	1	ϵ	1	0,9	ϵ	1	0,9	ϵ	1
Ressemb. Gauss. $\times 10^{-2}$	2,9	ϵ	2,9	1,9	ϵ	2	1,7	ϵ	1,9	1,7	ϵ	1,9

tandis que le rapport signal-à-bruit dû au repliement de spectre est :

$$S/B_{k, \ell} = \sqrt{\frac{\int_0^{u_{Nk}} \int_0^{v_{Nk}} R_{k, \ell}^2 dv du}{\int_{u_{Nk}}^{2u_{Nk}} \int_{v_{Nk}}^{2v_{Nk}} R_{k, \ell}^2 dv du}}$$

Nous avons calculé l'évolution du rapport signal-à-bruit des images $I_{k\ell}$ lorsque w_0 varie de 0,4 à 0,9, débordant l'intervalle permis par nos contraintes, ainsi que les valeurs fixées par Crowley, et pour des valeurs de n allant de 1 à 3. La figure 5 montre les résultats obtenus au niveau $k = 1$.

On constate que la valeur $w_0 = 0,5$ préconisée par Crowley est celle qui maximise le rapport signal-à-bruit tout en respectant la contrainte d'unimodalité en fréquence. Cependant, le choix $n = 3$ conduit à un rapport signal-à-bruit inutilement élevé compte tenu du bruit de quantification, et en conséquence à un filtrage excessif par rapport à la fréquence d'échantillonnage.

3.4. UN ALGORITHME OPTIMAL POUR LA REPRÉSENTATION MULTI-ÉCHELLE

L'étude précédente menée sur les algorithmes de Burt et Crowley avec les noyaux de filtrage tels que nous les avons définis, nous conduit à proposer un algorithme rapide de représentation multi-échelle qui allie les avantages des deux algorithmes et corrige leurs inconvénients. Nous retiendrons de l'algorithme de Crowley la séquence de filtres re-échantillonnage car elle permet d'obtenir une pyramide d'octave exacte : à chaque niveau d'échantillonnage une séquence de filtres par les noyaux F et F^*F , excepté le niveau initial où la séquence est précédée d'un filtrage par le noyau F . Cependant, dans la mesure où nous voulons construire une pyramide d'octave, nous retiendrons parmi les trois sorties possibles à chaque niveau k celle qui, pour le plus petit noyau de filtrage, assure le rapport signal-à-bruit compatible avec le bruit d'image et le bruit de quantification. On constate sur la figure 4 que le niveau $\ell = 1$ est le plus favorable. En effet, à ce niveau, la valeur $w_0 = 0,5$ est très proche de la valeur qui optimise le rapport signal-à-bruit et elle correspond aux contraintes que nous nous sommes définies. En outre, pour $n = 1$ le rapport signal-à-bruit est de 58 db ; le bruit de

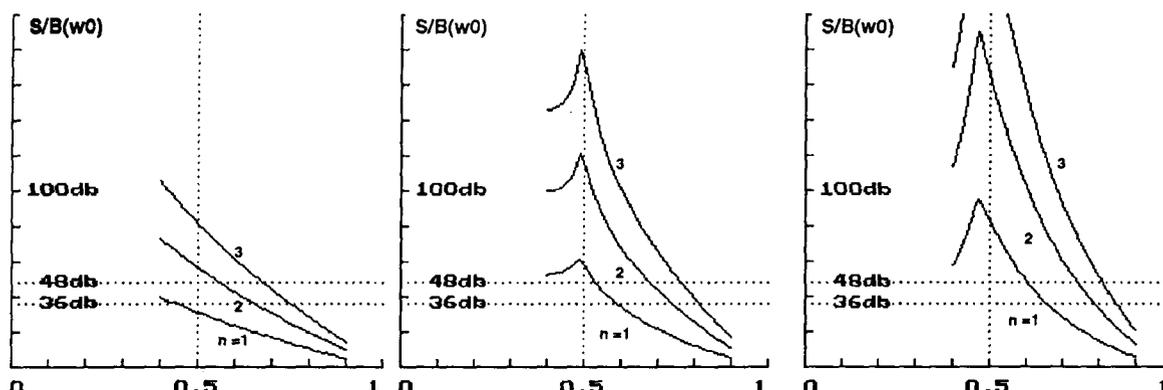


Figure 5. — Pyramide de Crowley. Évolution du rapport signal-à-bruit en fonction de w_0 et n .
Gauche : image I_{10} . Centre : image I_{11} . Droite : image I_{12} .

repliement de spectre demeure très inférieur (de 22 db) au bruit propre de l'image et très inférieur (de 10 db) au bruit de quantification. Il est donc inutile de donner à n une valeur plus élevée qui ne se traduirait que par un accroissement du coût de calcul d'une part, et par une perte d'information due au filtrage excessif d'autre part.

Nous choisissons donc les noyaux de filtrage :

$$F(x, y) = \frac{1}{16} \begin{pmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{pmatrix} = \frac{1}{16} \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix} * (1 \ 2 \ 1)$$

et

$$F(x, y)^{*2} = \frac{1}{16^2} \begin{pmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{pmatrix} * 2 = [F(x) * F(x)] * = = \frac{1}{16} \begin{pmatrix} 1 \\ 4 \\ 6 \\ 4 \\ 1 \end{pmatrix} * \frac{1}{16} (1 \ 4 \ 6 \ 4 \ 1).$$

Compte tenu que nous ne construisons qu'une pyramide d'octave, il est possible de réduire le coût de calcul en n'effectuant la convolution par F^{*2} que pour ceux des pixels qui seront retenus par le re-échantillonnage, conformément à la figure 6.

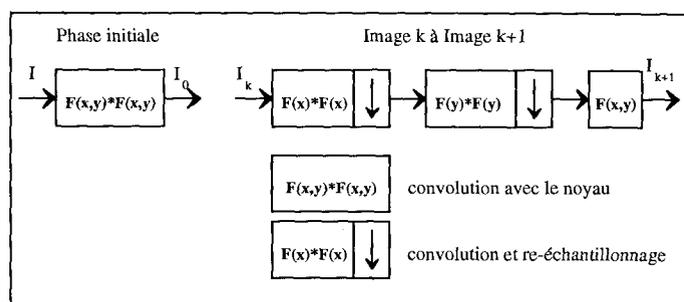


Figure 6. — Notre algorithme de représentation multi-échelle.

La première image I_0 de la pyramide est obtenue par convolution avec $F^{*2}(x, y)$, la génération de l'image I_{k+1} à partir de l'image I_k s'obtient par une première convolution par le noyau $F^{*2}(x)$ restreint à un pixel sur deux par ligne, puis par une convolution par le noyau $F^{*2}(y)$ restreint à une ligne sur deux. L'image obtenue à ce stade est analogue à l'image $I_{k+1,0}$ de la pyramide de Crowley, l'image I_{k+1} est alors obtenue par convolution avec le noyau $F(x, y)$.

La figure 7 représente la réponse en fréquence associée aux images des niveaux $k = 0, 1, 2, 3$ (de droite à gauche) de la pyramide générée par cet algorithme.

Nous avons calculé : le rapport signal-à-bruit dû au bruit de repliement de spectre, la dissymétrie en rotation et la ressemblance avec une gaussienne, de la même manière que pour la pyramide de Burt. Ces résultats sont résumés sur la table I.

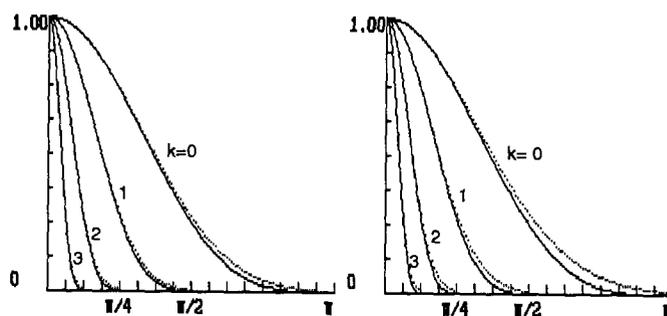


Figure 7. — Réponse en fréquence aux niveaux $k = 0, 1, 2, 3$ de la pyramide. Gauche : sections 0° (traits pleins) et 45° (pointillés). Droite : sections 0° (traits pleins), filtres gaussiens (pointillés).

3.5. DISCUSSION

En matière de représentation pyramidale d'image pour des applications de vision artificielle, les mérites d'un algorithme doivent être évalués sur les plans : i) du coût de calcul, ii) de l'adéquation entre la bande passante de l'image et la fréquence d'échantillonnage, iii) de la propriété de non-crédation de structures dans les images de la pyramide. Il est de plus éminemment souhaité que la réponse en fréquence associée aux images de la pyramide soit symétrique en rotation. Qu'il s'agisse de l'algorithme de Burt, de celui de Crowley ou de celui que nous proposons, le point trois est acquis grâce à un choix convenable du noyau de convolution. Le point deux est plus difficile à évaluer dans la mesure où la réponse en fréquence étant très éloignée de celle d'un filtre passe-bas idéal, l'interprétation du Théorème d'Echantillonnage est malaisée. On peut néanmoins convenir que la meilleure adéquation est obtenue lorsque le rapport signal-à-bruit dû au repliement de spectre est le plus élevé (bonne élimination des fréquences au-delà de la fréquence de Nyquist), et que par ailleurs la variance associée au filtre est la plus petite (bonne conservation des fréquences situées en deçà de la fréquence de Nyquist). On peut évaluer cette bonne adéquation par la variance normalisée par le coefficient d'échelle 2^{2k} . Nous avons reporté ces éléments dans la table I qui résume les propriétés des algorithmes de Burt et de Crowley, et de l'algorithme que nous proposons.

Sont indiqués en gras sur ce tableau ceux des éléments qui militent en faveur d'un algorithme. En ce qui concerne le bruit de repliement de spectre, la symétrie en rotation, ainsi que la ressemblance avec une gaussienne, l'algorithme de Crowley présente un avantage certain. Ceci est obtenu au détriment de l'adéquation de la bande passante au taux d'échantillonnage : l'information n'est pas dégradée par le repliement de spectre mais elle est partiellement perdue par le filtrage. A un moindre titre, l'algorithme de Burt souffre du même inconvénient. Notre algorithme présente la meilleure adéquation bande passante-taux d'échantillonnage, tout en générant un bruit de repliement de spectre négligeable par rapport au bruit d'image. Les figures 8 a et 8 b illustrent la pyramide d'images passe-bas obtenue par notre algorithme à partir de : d'une image de synthèse (8 a) comportant des anneaux alternativement noirs et blancs d'épaisseur 2, 4, 8, 16, 32 pixels, et d'une image réelle (8 b).

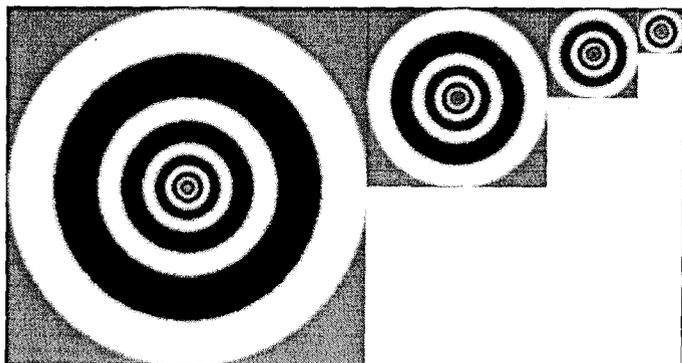


Figure 8 a. — Pyramide passe-bas : image synthétique.



Figure 8 b. — Pyramide passe-bas : image réelle.

En matière de vision artificielle, l'avantage d'un algorithme par rapport à un autre n'a réellement de sens que si le coût de calcul est, sinon moindre, du moins comparable. L'évaluation du coût de calcul doit se faire en prenant en compte le coût d'une convolution par pixel d'une part, le nombre de convolutions à effectuer d'autre part. Pour l'évaluation du coût d'une convolution nous considérerons que chaque convolution n'utilise que des opérations de décalage et addition, nous considérerons cette opération comme opération de base. On notera qu'un noyau de taille 3×3 n'exige que cette opération de base et que la convolution par un noyau de taille 5×5 peut être implémentée au moyen de cette même opération, ce qui est favorable au plan de la vitesse de traitement. Ainsi la convolution par F exigera 6 opérations de base, tandis que la convolution par F^{*2} en exigera 12.

Le coût de calcul de notre algorithme peut alors être évalué :

* Convolution initiale. N^2 pixels convolués ; coût $12 N^2$ opérations de base.

* Au niveau k il faut effectuer : $N^2/2^{2(k+1)}$ convolutions par un noyau (1×5) , puis $N^2/2^{2(k+1)}$ convolutions par un noyau (5×1) , enfin $N^2/2^{2(k+1)}$ convolutions par un noyau (3×3) . Soit $4 N^2/2^{2(k+1)}$ pixels à convoluer, chaque convolution impliquant 6 opérations de base.

Le coût de calcul est alors borné par : $12 N^2 + 24 N^2/3 = 20 N^2$ décalages et additions.

Le coût de calcul de l'algorithme de Burt peut être chiffré de la même façon : à chaque niveau le nombre de convolutions à effectuer est $N^2/2^{2k}$, chaque convolution impliquant 12 opérations de base. Le coût de calcul est alors borné par $12 \times 4 N^2/3 = 16 N^2$ décalages et additions.

Quant à l'algorithme de Crowley, en adoptant les noyaux de convolution que nous proposons, la qualité des images produites serait exactement identique, avec un coût de calcul : $6 \times N^2 + 18 \times 4 N^2/3 = 30 N^2$ décalages et additions. Ce coût de calcul serait de $90 N^2$ décalages et additions en adoptant les noyaux originaux de Crowley.

On constate ainsi que l'algorithme de Burt est environ 25 % plus rapide. Cependant cette caractéristique de vitesse est valable dans la mesure où le noyau de filtrage utilisé est bien celui que nous proposons. De plus, de notre point de vue, cet écart est largement compensé par la meilleure qualité (en terme de bruit de repliement de spectre et d'adéquation de la bande passante au taux d'échantillonnage) des images de la pyramide générée par notre algorithme. Quant à l'algorithme de Crowley, dans le meilleur des cas son coût de calcul est de 50 % supérieur, cependant avec l'objectif de construire une pyramide demi-octave, donc de meilleure résolution en échelle.

4. Construction d'une pyramide de laplaciens

La détection de structures dans une image passe généralement par le calcul du laplacien de l'image : le passage à zéro permet de détecter des contours, les maximums locaux mono ou pluri-directionnels à différentes échelles permettent de localiser et caractériser les structures présentes dans l'image. Dans une image numérique, ce sont en fait des différentielles que l'on calcule, assimilables à des dérivées si les fréquences d'échantillonnage en x et y sont identiques. Pour ce genre d'opérations, la symétrie en rotation est absolument recherchée et la solution consiste à utiliser une approximation du laplacien par une différence de gaussiennes [6].

Burt réalise cette opération par différence entre images I_k et I_{k+1} , cependant ces images n'étant pas échantillonnées à la même fréquence, il doit reconstituer la copie filtrée de l'image I_k en partant de I_{k+1} par une opération d'expansion et interpolation (cette dernière est réalisée par convolution avec le noyau B). Cependant, le rapport des variances associées à l'image I_k et à sa copie filtrée étant :

$$\frac{\sigma_{k+1}^2 + 2^{2k} \sigma_0^2}{\sigma_k^2} = 4 + \frac{(1 + 2^{2k}) \sigma_0^2}{\sigma_k^2}$$

le rapport des variances et par voie de conséquence le rapport des bandes de fréquence, n'est pas constant. Initialement égal à 6, il tend vers 4,75 aux grandes valeurs de k .

Crowley, disposant des images et de leur copie filtrée à chaque niveau de la pyramide, calcule le laplacien par

différence entre ces images. Le rapport des variances est dans ce cas :

$$\frac{\sigma_{k,\ell}^2}{\sigma_{k,\ell-1}^2} = 2 \quad \ell \in \{1, 2\} .$$

Il génère ainsi une pyramide de laplaciennes dont les bandes de fréquence sont espacées d'une demi-octave, ce qui facilite le suivi des structures à travers les niveaux de la pyramide.

Pour comparer les coûts de calcul entraînés par chacun de ces algorithmes, nous supposons que les noyaux de filtrage utilisés sont ceux que nous préconisons. Alors, les deux algorithmes deviennent comparables : dans celui de Burt, l'interpolation exige $4/3 N^2$ convolutions supplémentaires, soit $16 N^2$ opérations de base, et le coût de calcul devient sensiblement égal à celui de l'algorithme de Crowley avec un léger avantage pour ce dernier.

4.1. UN NOYAU POUR LE CALCUL DU LAPLACIEN D'UNE IMAGE

Calculer le laplacien d'une image revient à filtrer cette image par un filtre dont la réponse fréquentielle est une parabolicoïde de révolution. Dans la mesure où le spectre de l'image est limité, cette caractéristique doit être respectée dans le domaine de fréquence proche de l'origine, disons la moitié du domaine de Nyquist. Le noyau de filtrage utilisé doit donc :

1. Permettre de calculer le laplacien de l'image.
2. Avoir une réponse en fréquence symétrique en rotation.
3. Induire un coût de calcul acceptable.

La littérature offre des solutions multiples à ce problème qui toutes reviennent à utiliser un noyau de la forme :

$$L = \begin{pmatrix} c & b & c \\ b & a & b \\ c & b & c \end{pmatrix} .$$

En considérant le développement en série de Taylor appliqué au résultat de la convolution d'une image $I(x, y)$ par le noyau L , on déduit que les coefficients du noyau doivent respecter cette symétrie pour que les dérivées d'ordre impair soient éliminées. Il faut en outre :

$$a + 4b + 4c = 0$$

pour éliminer la composante continue.

Alors, toujours en supposant une fréquence d'échantillonnage identique dans les directions verticales et horizontales : $dx = dy = 1$, et en négligeant les dérivées d'ordre 6 et supérieures :

$$I(x, y) * L = (b + 2c) \nabla^2 I + c I_{x^2 y^2}^{(4)} + \frac{2(b + 2c)}{4!} (I_x^{(4)} + I_y^{(4)})$$

qui montre que la convolution par ce type de noyau permet effectivement de calculer le laplacien mais lui ajoute inévitablement des contributions de dérivées d'ordre 4 (et plus).

Ainsi les noyaux couramment proposés dans la littérature : Noyau L1 avec $a = -4, b = 1, c = 0$:

$$L1 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{pmatrix}$$

$$I(x, y) * L1 = \nabla^2 I + \frac{2}{4!} (I_x^{(4)} + I_y^{(4)}) .$$

Noyau L2 avec $a = -8, b = c = 1$:

$$L2 = \begin{pmatrix} 1 & 1 & 1 \\ 1 & -8 & 1 \\ 1 & 1 & 1 \end{pmatrix}$$

$$I(x, y) * L2 = 3 \nabla^2 I + I_{x^2 y^2}^{(4)} + \frac{6}{4!} (I_x^{(4)} + I_y^{(4)}) .$$

Noyau de Robinson L3 avec $a = -4, b = 2, c = -1$:

$$L3 = \begin{pmatrix} -1 & 2 & -1 \\ 2 & -4 & 2 \\ -1 & 2 & -1 \end{pmatrix}$$

$$I(x, y) * L3 = I_{x^2 y^2}^{(4)} .$$

Nous proposons quant à nous le masque Lp avec $a = -12, b = 2, c = 1$:

$$Lp = \begin{pmatrix} 1 & 2 & 1 \\ 2 & -12 & 2 \\ 1 & 2 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{pmatrix} - 16 = 16[-1]$$

où $F(x, y)$ est le noyau de base pour la construction de notre pyramide passe-bas. Alors :

$$I(x, y) * Lp = 16 [I(x, y) * F(x, y) - I(x, y)] = 4 \nabla^2 I + I_{x^2 y^2}^{(4)} + \frac{8}{4!} (I_x^{(4)} + I_y^{(4)}) .$$

Tous ces noyaux peuvent être représentés par des formes séparables, ce qui est favorable pour la réduction du coût de calcul. Tous à l'exception du masque de Robinson fournissent une estimation du laplacien au sens strict du terme. Par contre, leurs propriétés vis-à-vis de la symétrie en rotation ne peuvent aisément se déduire de cette formulation.

Pour caractériser ces propriétés nous considérons leurs réponses en fréquence d'une part, la réponse en fréquence associée à un niveau k de la pyramide d'autre part. Nous avons représenté, sur la figure 9, ces réponses en fréquence sous forme d'images où le gain nul est représenté par le noir, le gain maximum par le blanc. Pour rendre ces images plus lisibles, nous avons quantifié les réponses en fréquence sur 16 niveaux également répartis entre le niveau zéro et le niveau maximum. Ces réponses en fréquence ayant des maximums différents d'un noyau à l'autre, nous les avons normées par un coefficient d'échelle ad-hoc.

Nous avons représenté sur la figure 9a la réponse en fréquence propre au noyau et sur la figure 9b celle résultant de la convolution avec l'image I_0 de notre pyramide. Dans chacune de ces figures, nous avons

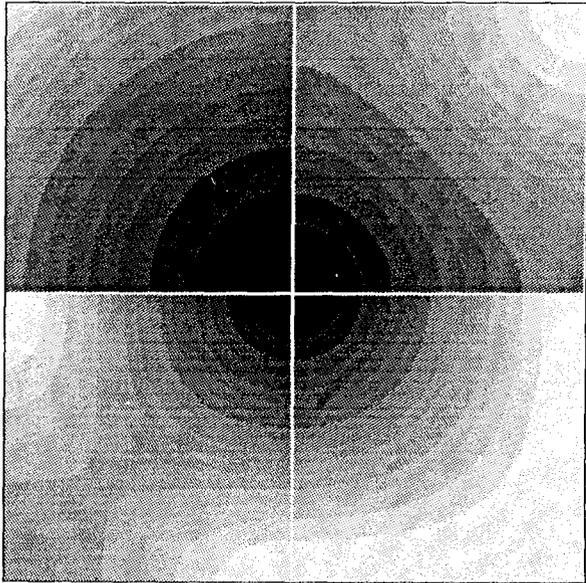


Figure 9 a. — Image des réponses en fréquence des noyaux de calcul du laplacien.

Quadrant Nord-Ouest : Opérateur Laplacien Idéal.
 Quadrant Nord-Est : Noyau L1.
 Quadrant Sud-Ouest : Noyau L2.
 Quadrant Sud-Est : Noyau Lp Proposé.

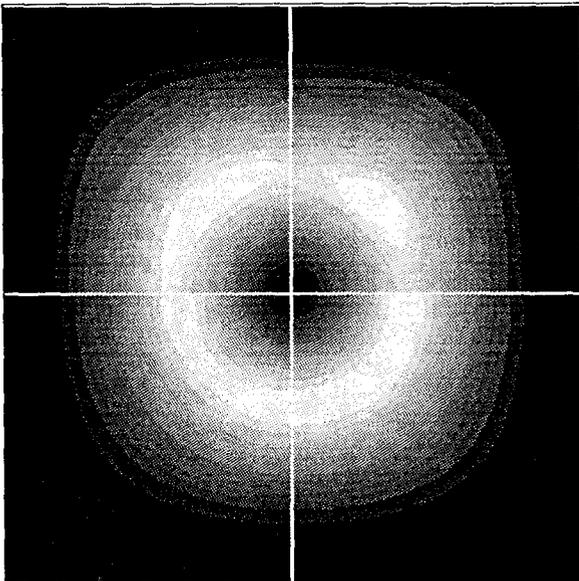


Figure 9 b. — Image des réponses en fréquence au niveau $k = 0$ de la pyramide de Laplacien.

Quadrant Nord-Ouest : Opérateur Laplacien Idéal.
 Quadrant Nord-Est : Noyau L1.
 Quadrant Sud-Ouest : Noyau L2.
 Quadrant Sud-Est : Noyau Lp Proposé.

également représenté, à titre de référence, la réponse en fréquence de l'opérateur Laplacien idéal. On constate sur la figure 9 a que le noyau L1, très fréquemment utilisé, a une réponse en fréquence pour le moins exotique ainsi que, mais à un degré moindre, le noyau L2. Des trois noyaux étudiés, celui que nous proposons est celui qui présente les

meilleures propriétés de symétrie en rotation. Si nous observons maintenant la réponse en fréquence associée au niveau $k = 0$ de la pyramide, figure 9 b, la symétrie en rotation est améliorée car le filtrage passe-bas élimine les contributions hautes fréquences les plus sujettes à dissymétrie. Mais là encore le noyau que nous proposons présente les meilleures qualités. Ceci peut s'expliquer simplement par le fait que convoluer une image par ce noyau revient à effectuer une différence d'images filtrées passe-bas par des noyaux proches de noyaux gaussiens.

4.2. PROPRIÉTÉS DE LA PYRAMIDE DE LAPLACIEN

Ces propriétés sont résumées sur la figure 10 d'une part et la table II d'autre part. Nous avons représenté sur la figure 10 la réponse en fréquence aux niveaux $k = 0, 1, 2, 3$ (de droite à gauche) de la pyramide ainsi les réponses en fréquence des laplaciens gaussiens de même variance.

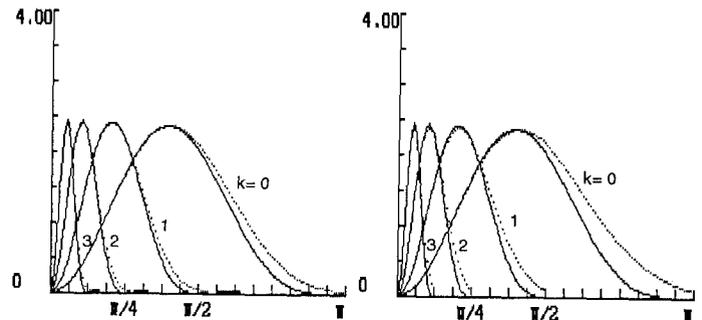


Figure 10. — Réponses en fréquence du laplacien aux niveaux $k = 0, 1, 2, 3$ de la pyramide. A gauche coupes à 0° (traits plein) et 45° (pointillés). A droite coupes à 0° (traits pleins) et différences de gaussiennes correspondantes (pointillés).

On constate, sur les courbes de gauche de cette figure, que la réponse en fréquence présente une bonne symétrie en rotation, les écarts entre coupes à 0° et 45° n'étant sensibles qu'aux limites du domaine de Nyquist où le gain devient très faible. Les courbes de droite permettent de comparer la réponse avec celle de la différence de gaussienne correspondante déterminée de la façon suivante : le rapport des variances associées à l'image I_k et à sa copie filtrée $I_k * F(x, y)$ est :

$$\frac{\sigma_k^2 + 2^{2k} \sigma_0^2}{\sigma_k^2} = \frac{2^{2k} + 2^{2k-1}}{2^{2k}} = 1,5.$$

La pyramide de laplacien que nous construisons ainsi est donc une bonne approximation d'une pyramide de différences de gaussiennes de variances 2^{2k} et $1,5(2^{2k})$. De même que dans la pyramide de Crowley, mais contrairement à celle de Burt, le rapport de ces variances est constant ce qui assure une réponse en fréquence identique, au coefficient d'échelle près, à tous les niveaux de la pyramide.

On constate par ailleurs sur la table II que la dissymétrie en rotation reste négligeable. Cependant, le rapport signal-à-bruit, bien que supérieur à celui de l'image, est diminué par rapport à ceux caractérisant les images passe-bas. Ceci est dû au fait que la puissance du signal est réduite par le

Table 2.

Niveau k	0	1	2	3
S/B	—	44 db	43 db	43 db
Dissymétrie	$3,1 \times 10^{-2}$	$2,1 \times 10^{-2}$	$2,1 \times 10^{-2}$	$2,1 \times 10^{-2}$

filtrage passe-bande sans que cela soit accompagné d'une réduction comparable du bruit de repliement de spectre. Ce phénomène est caractéristique de toutes les méthodes visant à estimer le laplacien.

Quant au coût de calcul additionnel, il est essentiellement dû au calcul de la copie filtrée des images passe-bas (nous négligerons par souci de simplicité le coût de la soustraction), soit $4/3 N^2$ convolutions nécessitant chacune 6 opérations de base. Le coût de calcul total est alors $28 N^2$ opérations de base, inférieur au coût de calcul de l'algorithme de Burt ($32 N^2$ opérations de base), inférieur également à celui de l'algorithme de Crowley ($30 N^2$ opérations de base). De plus, grâce au gain important mis en évidence par les courbes de la figure 10, notre noyau permet une estimation du laplacien avec une dynamique importante ainsi que l'illustrent les figures 11 a et 11 b représentant les pyramides d'images de laplacien obtenues par notre algorithme à partir des pyramides passe-bas des figures 8 a et 8 b.

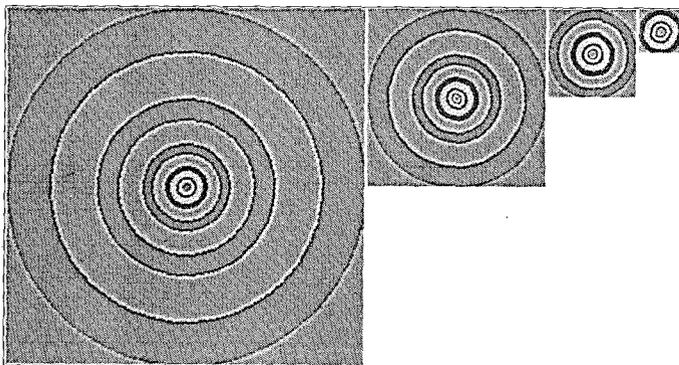


Figure 11 a. — Pyramide de Laplaciens : image synthétique.

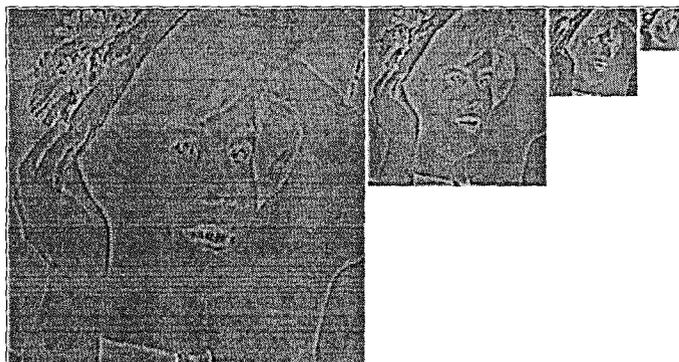


Figure 11 b. — Pyramide de Laplaciens : image réelle.

5. Conclusion

La représentation des images par des pyramides d'images passe-bas et de laplaciens sont des outils très utilisés en vision artificielle. Parmi tous les algorithmes proposés, ceux de Burt et de Crowley sont les plus couramment utilisés en raison principalement de leur coût de calcul réduit. Ce faible coût de calcul est obtenu grâce à l'utilisation de noyaux de filtrage de taille réduite et malgré cela, une réduction importante de la bande passante est acquise par une technique de re-échantillonnage avant filtrage qui étend le domaine de l'espace sur lequel la convolution est effectuée. Néanmoins, le bruit de repliement de spectre introduit par le processus de filtrage-décimation n'a pas été pris en compte par ces auteurs. En partant de contraintes sur la forme des réponses impulsionnelles et fréquentielles d'une part, et sur le bruit de repliement de spectre d'autre part, nous avons montré qu'il existe une solution optimale qui satisfait ces contraintes. Nous avons défini un algorithme de construction de pyramide demi-octave passe-bas, qui tout en induisant un coût de calcul comparable à celui de l'algorithme de Burt, produit des images de bien meilleure qualité, tant en ce qui concerne le bruit inévitable de repliement de spectre qu'en ce qui concerne l'adéquation entre filtrage et taux d'échantillonnage. Nous avons également proposé un algorithme de construction de pyramide de laplacien, qui produit des images de meilleure qualité que celles produites par l'algorithme de Burt et de plus grande dynamique que celles produites par l'algorithme de Crowley, tout en ayant un coût de calcul inférieur à celui de ces algorithmes.

Manuscrit reçu le 4 février 1992.

Remerciements

Ce travail a été effectué dans le cadre du projet ESPRIT-BRA 3038 « Vision as Process ». L'auteur tient à remercier les partenaires de ce projet pour les discussions fructueuses et plus particulièrement J. L. Crowley.

BIBLIOGRAPHIE

- [1] P. J. BURT, « Fast Filter Transform for Image Processing », *CGIP* **16**, 20-51, 1981.
- [2] P. J. BURT, « The Laplacien Pyramid as a Compact Image Code », *IEEE Trans. on Communications*, **COM 31**, No. 4, 532-540, 1984.
- [3] J. L. CROWLEY, R. M. STERN, « Fast computation of the difference of low-pass transform », *IEEE Trans. Pattern Anal. Mach. Intell.*, **6**, 212-222, 1984.
- [4] A. CHEHIKIAN, J. L. CROWLEY, « Fast Computation of Optimal Semi-Octave Pyramids », *Proceedings of the 7th Scandinavian Conference on Image Analysis*, 18-27, Aalborg Denmark, 1991.
- [5] T. LINDBERG, « Scale Space for Discrete Signals », *IEEE Transactions on P.A.M.I.*, **12**, No. 3, March 1990.
- [6] D. MARR, « Vision », Freeman Press, 1979.