

Chapitre 3

Espaces d'échelles

3.1 Introduction

Les objets réels, la réalité physique, par opposition aux objets mathématiques, existent sous forme d'entités bien différentes selon le niveau d'échelle auquel on les considère : Une feuille, un arbre, une forêt présentent ainsi des caractéristiques de forme, de couleur, de texture radicalement différentes. Etudier le comportement local d'un objet à partir de dérivées estimées par différences finies comme dans le chapitre précédent ne permettra d'appréhender qu'une petite partie de la réalité physique représentée par l'image : l'échelle la plus fine (celle du pixel). La théorie des espaces d'échelle est un modèle qui est apparu progressivement dans le domaine de la vision artificielle, principalement pour les aspects bas niveau (i.e. traitement d'images), pour prendre en compte cette nature foncièrement multiéchelles des données images.

Les capteurs d'images sont capables d'appréhender un certain domaine dans l'espace des échelles d'une scène réelle, correspondant à l'intervalle entre l'échelle interne, correspondant à la taille du plus petit objet discriminable dans l'image, lié à la résolution du capteur, et l'échelle externe, lié au champ du capteur. Or pour extraire une information d'une donnée image, on lui applique certaines transformations. La nature de l'information alors obtenue est fortement déterminée par le rapport entre la taille des structures concernées dans l'image et la taille (portée spatiale) des opérateurs appliqués.

A l'exception de certaines applications pour lesquelles on maîtrise - au moins en partie - l'environnement (contrôle industriel, imagerie microscopique), les échelles associées à une tâche de vision particulière ne sont en général pas connues à l'avance. C'est pourquoi il est essentiel que les opérateurs considèrent l'image à plusieurs niveaux d'échelles ; c'est à ce titre que les espaces d'échelles constituent un modèle formel adapté.

La section 3.2 présente les espaces d'échelles linéaires. Les sections suivantes montrent l'utilisation du raisonnement multi-échelles à travers deux exemples très importants : la détection de contours et la détection de points d'intérêt. La section 3.5 présente quelques exemples d'espaces d'échelles non linéaires.

3.2 Espace d'échelles linéaire

Les opérateurs de traitement d'images dont le support de calcul est réduit au pixel et à ses «plus proches» voisins - tels que les filtres 3×3 présentés au chapitre précédent pour approximer les grandeurs différentielles - permettent d'estimer des structures à l'échelle la plus fine : celle du pixel. Si l'on souhaite estimer, avec le même opérateur, une structure à différentes échelles, il faut opérer sur l'image à des résolutions différentes. C'est le principe de la *multi-résolution*, qui consiste à produire une suite d'images de résolution

décroissante, ou *pyramide*, et à appliquer des traitements locaux à différents niveaux de résolution.

Ainsi, sur la figure 3.1, la première ligne montre le calcul d'une dérivée partielle verticale par un opérateur local calculé à la résolution 256×256 (niveau 0 de la pyramide). La seconde ligne montre le même calcul réalisé sur la même image à la résolution 16×16 , obtenue en divisant par 16 le nombre d'échantillons spatiaux dans chaque dimension (niveau 4 de la pyramide dyadique, pour laquelle chaque dimension de l'image est divisée par 2 en passant d'un niveau au supérieur).

Mais il est essentiel de rappeler que, comme on l'a vu dans le premier chapitre, l'obtention de la pyramide ne doit pas se traduire par un simple sous-échantillonnage, car alors les artefacts dûs à l'aliasing seraient naturellement détectés par les opérateurs locaux dans les résolutions plus basses. La sous-résolution doit donc combiner un filtrage spatial en préalable au sous-échantillonnage. On a vu (Chapitre 1), que dans ces bonnes conditions, l'image sous-échantillonnée était finalement équivalente à l'image filtrée à pleine résolution.

C'est précisément sur cette équivalence que repose la représentation *multi-échelles*. Au contraire de la précédente, elle n'implique pas de sous-résolution des images, mais elle identifie l'indice d'échelle, significatif de la taille des structures concernées par l'opérateur, avec le paramètre de portée spatial d'un filtre caractérisant cet opérateur. Ainsi, sur la figure 3.1, la dernière ligne montre, à gauche l'image à l'échelle 16, au sens du filtre passe-bas idéal ($u_M = v_M = 8$). Du point de vue information, cette image est équivalente à l'image 16×16 au dessus. A droite, on peut voir l'estimation multi-échelle du gradient vertical, correspondant ici à une différence locale calculée sur l'image de gauche.

Dans l'axiomatique des espaces d'échelles, une image I est donc associée à un paramètre d'échelle s , lui même identifié à l'indice de portée spatial d'un opérateur P :

$$I_s = P_s(I) \quad (3.1)$$

Pour les espaces d'échelles linéaires, P sera supposé linéaire et invariant en translation, c'est-à-dire une convolution :

$$I_s = I \star h_s \quad (3.2)$$

L'axiomatique des espaces d'échelles linéaires est fondée sur un certain nombre de propriétés exprimant l'absence d'*a priori* quelconque sur les images :

- invariance en translation
- invariance en rotation
- invariance en échelle

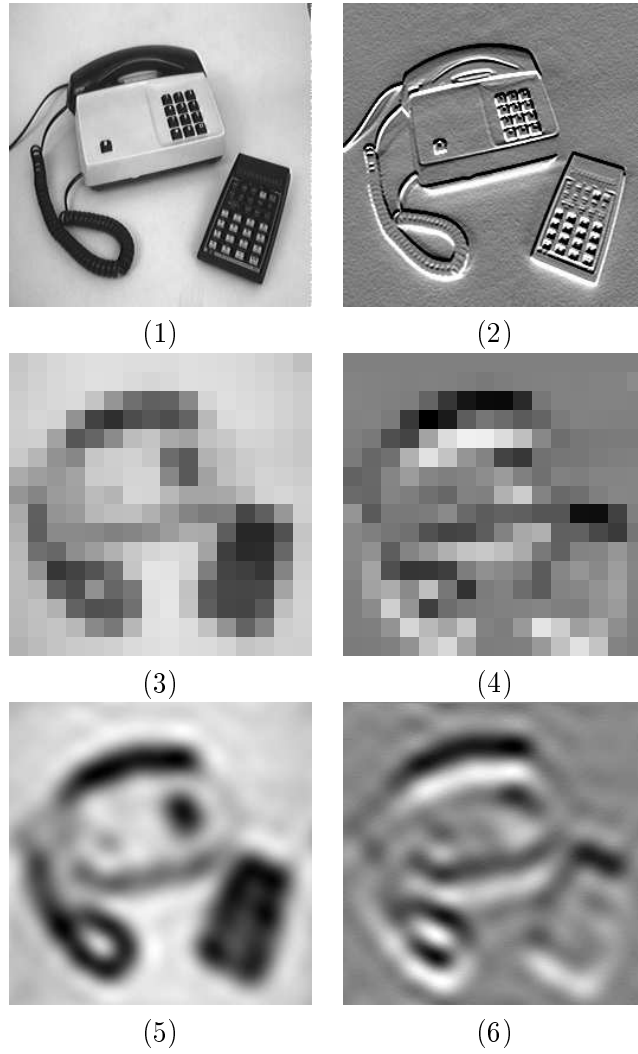


FIG. 3.1 – Multi-résolution ou Multi-échelle (1) et (2) : une image (256×256), et l'estimée de sa dérivée verticale à l'échelle du pixel (Sobel vertical). (3) et (4) : estimation de la dérivée verticale en multi-résolution, à l'échelle 16 (4ème niveau de la pyramide dyadique). (5) et (6) : estimation de la dérivée verticale en multi-échelle, à l'échelle 16.

Pour des raisons calculatoires, on impose aussi la condition que la composition de 2 opérateurs de changement d'échelle produise un changement d'échelle qui peut être réalisé avec un seul opérateur. Cela correspond à la propriété dite de semi-groupe, alléguant l'existence d'une loi \boxplus opérant sur les indices d'échelles, telle que :

$$h_{s_1} \star h_{s_2} = h_{s_1 \boxplus s_2} \quad (3.3)$$

Lindeberg et ter Haar Romeny ont montré que les propriétés d'invariances et de semi-groupe impliquent que l'opérateur de changement d'échelle h_s coïncide avec le noyau gaussien de variance $2s$:

$$h_s = \frac{1}{4\pi s} e^{\frac{-(x^2+y^2)}{4s}} \quad (3.4)$$

Les espaces d'échelles linéaires sont donc généralement synonymes d'espaces d'échelles gaussiens. La figure 3.2 montre ainsi la représentation d'un signal 1d dans l'espace d'échelle gaussien, sous la forme d'un faisceau de courbes, correspondant à l'application de filtres de convolution gaussien de variance croissante.

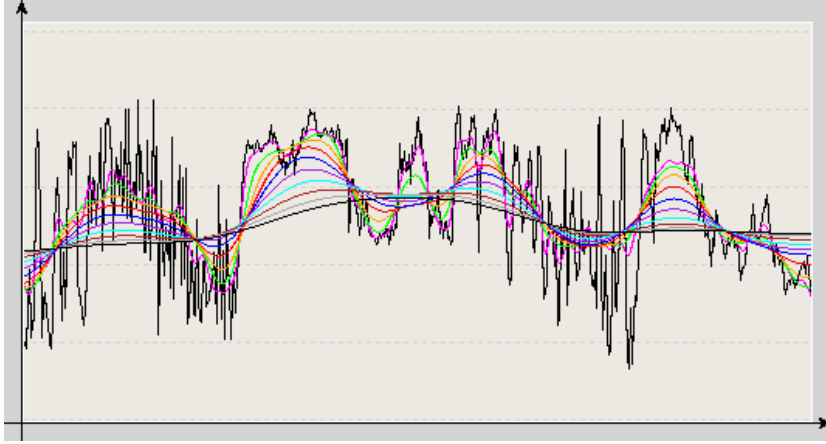


FIG. 3.2 – Représentation multi-échelles d'un signal 1d.

Cette figure illustre l'idée clef des espaces d'échelles, qui est le *principe de causalité*, selon lequel le signal ne doit pas faire apparaître de nouvelles structures lorsque l'échelle augmente. Pour un faisceau de courbes 1d, ce principe se traduit par :

- pas d'apparition de nouveaux extrema locaux.
- pas d'amplification des extrema existants (i.e. lorsque l'échelle augmente, la valeur d'un maximum diminue et la valeur d'un minimum augmente).

En 2d, cette propriété n'est plus vraie : Si l'on voit l'image comme un relief, où la hauteur correspond au niveau de gris, il suffit d'imaginer 2 cols de hauteurs différentes reliés par une corniche plus fine que le support spatial des deux cols. A l'échelle la plus fine, cette «image» ne comporte qu'un maximum local, qui est le sommet du plus grand col. Cependant, le lissage gaussien fait progressivement disparaître la corniche, si bien qu'à une certaine échelle, le deuxième col apparaît comme un nouveau maximum local.

Malgré cette limitation inhérente au filtrage linéaire en 2d, le noyau gaussien reste le meilleur support de calcul de l'espace d'échelle linéaire. La première ligne de la figure 3.3 montre une image à 3 échelles différentes.

Une des propriétés les plus importantes de l'espace d'échelle linéaire réside dans le fait que la convolution commute avec la dérivation :

$$\partial^n(I \star h_s) = I \star (\partial^n h_s) \quad (3.5)$$

Cela signifie qu'estimer une dérivée à l'échelle s revient à calculer une convolution avec la dérivée de la gaussienne. Ainsi, si l'on est capable de calculer efficacement une convolution par une dérivée de gaussienne $G_s^{(i,j)}$ avec :

$$G_s^{(i,j)} = \frac{\partial^{i+j}}{\partial x^i \partial y^j} G_s \quad (3.6)$$

et ce, quelques soient (s, i, j) , on dispose d'un outil permettant de calculer un grand nombre de caractéristiques multi-échelles très utiles. Les deux sections suivantes illustrent cela, dans le cas de la détection de contours, puis de la détection de points d'intérêt.

La figure 3.3 montre le calcul de dérivées multiéchelles, pour plusieurs valeurs de l'indice d'échelle s et de l'ordre de dérivations (i, j) .

3.3 Exemple 1 : Détection de contours

La détection de contours est l'un des premiers problèmes «historique» du traitement d'images. Il s'agit de transformer l'image en un ensemble de courbes définissant les «frontières significatives». Ces courbes représentent une simplification drastique du contenu de l'image, et on peut en extraire des descripteurs utiles dans plusieurs applications.

Il serait abusif de dire que ce problème est définitivement réglé aujourd'hui, car aucun détecteur ne fournit une réponse qui serait parfaitement adaptée à une application donnée, et encore moins conforme à une notion de frontière sémantique entre objets. Néanmoins la théorie des espaces d'échelles permet de définir très proprement les détecteurs de contours, tout en fournissant des algorithmes pratiques très utilisables.

Sur un signal 1d dérivable deux fois, la notion de contours peut être définie comme le lieu des maxima locaux de la dérivée première, ou, ce qui

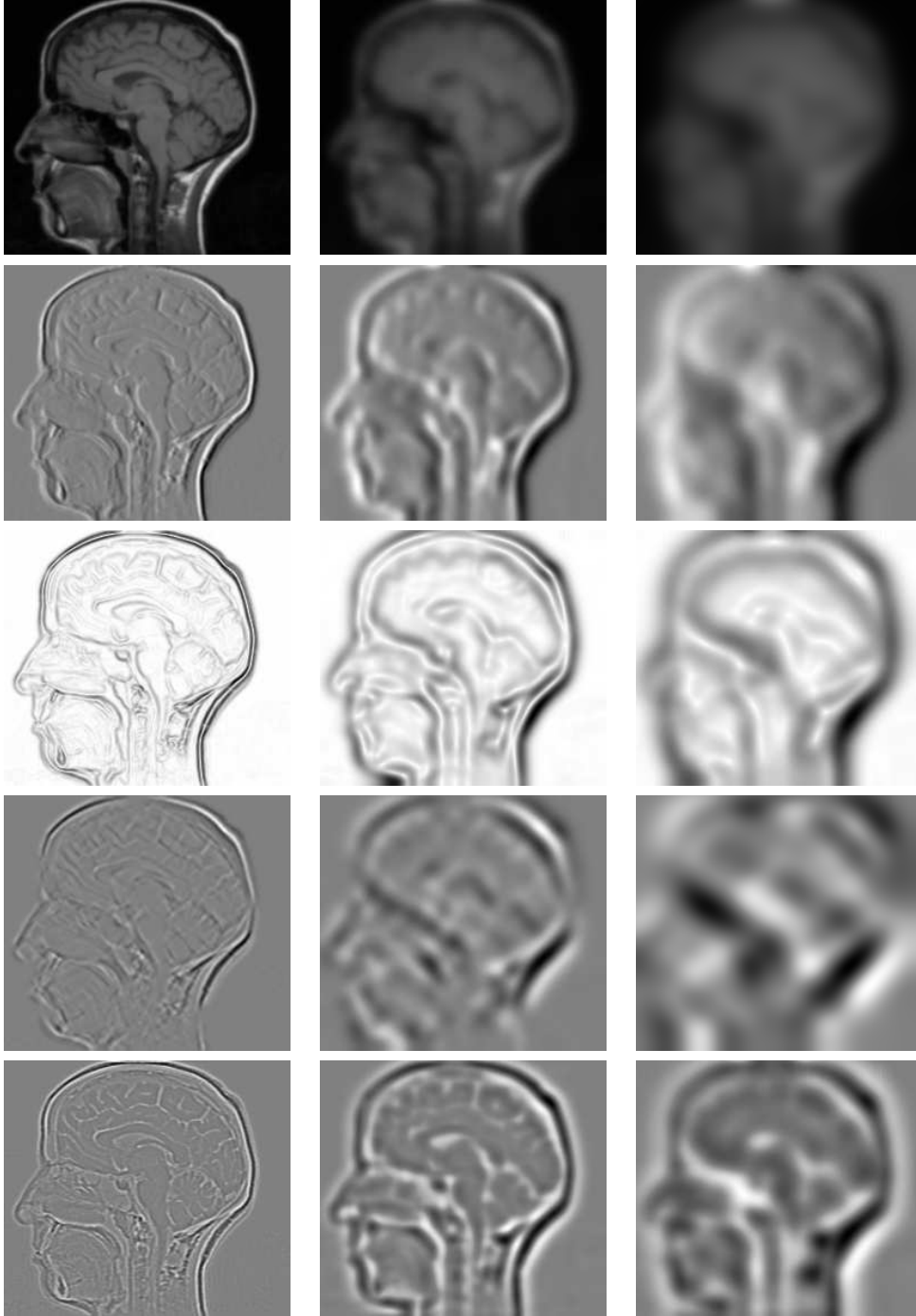


FIG. 3.3 – Espace d'échelles gaussien et dérivées. De gauche à droite : échelles $s = 0.5$, $s = 12.5$, et $s = 50.0$. De haut en bas : I , $\frac{\partial I}{\partial x}$, $\|\nabla I\|$, $\frac{\partial^2 I}{\partial x \partial y}$, et $\|\Delta I\|$. Les grandeurs différentielles sont représentées normalisées.

est équivalent, le lieu des passage par zéro de la dérivée seconde (voir Figure 3.4).

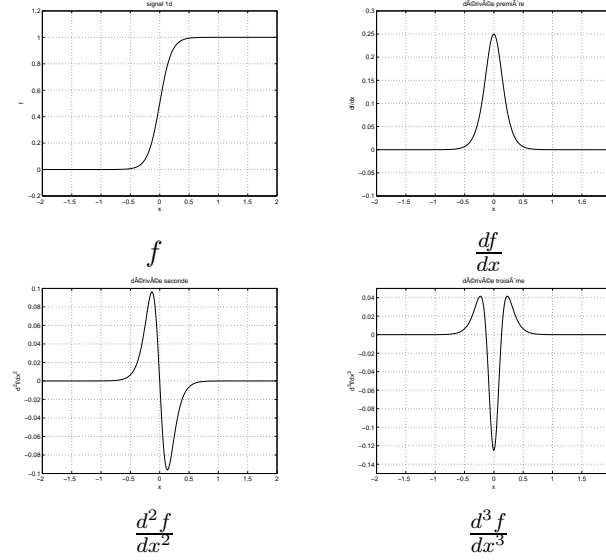


FIG. 3.4 – Un contour 1d (type échelon) en $x = 0$, et ses dérivées d'ordre 1 à 3.

Pour retrouver cette notion naturelle en 2d, il faut se placer, en chaque point, dans la direction de variation principale de l'image, c'est-à-dire la direction du gradient. Si l'on se considère l'image I au voisinage d'un pixel de contour (i, j) , en observant le signal dans la direction et le sens du gradient $\nabla I(i, j)$ (voir Figure 3.5), alors on retrouve une fonction f 1d, de type échelon, comme illustré Figure 3.4. Les contours dans les images sont donc définis comme les lieux de maxima locaux du module du gradient, dans la direction du gradient, ou encore, les lieux de passage par zéro de la dérivée seconde de l'image dans la direction du gradient. Etant, de plus, dans le sens du gradient positif, on aura toujours un échelon croissant, et donc une dérivée troisième négative (voir Figure 3.4).

Dans la suite on notera comme suit les dérivées partielles de l'image I : $I_x = \frac{\partial I}{\partial x}$, $I_y = \frac{\partial I}{\partial y}$, $I_{xx} = \frac{\partial^2 I}{\partial x^2}$, $I_{xy} = \frac{\partial^2 I}{\partial x \partial y}$, et ainsi de suite.

L'image I étant initialement définie dans un repère (x, y) , on va considérer en chaque point un repère de coordonnées locales (g, t) défini par le gradient (voir Figure 3.5), de la façon suivante :

$$\begin{cases} g = x \cos \phi + y \sin \phi \\ t = -x \sin \phi + y \cos \phi \end{cases} \quad (3.7)$$

Et la transformation inverse :

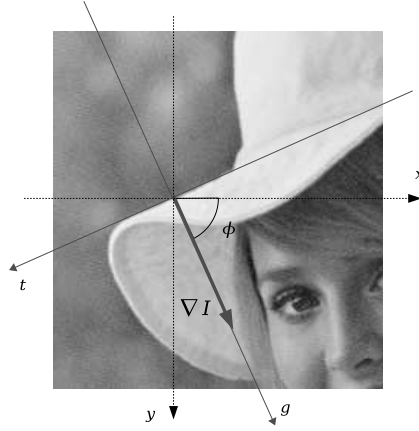


FIG. 3.5 – Le repère (g, t) du gradient local.

$$\begin{cases} x = g \cos \phi + t \sin \phi \\ y = g \sin \phi - t \cos \phi \end{cases} \quad (3.8)$$

Avec $\phi = \arg(\nabla I) = \arctan(I_y/I_x)$. La coordonnée g correspond donc à la composante de l'image dans la direction du gradient, tandis que la coordonnée t correspond à la composante tangentielle, perpendiculaire au gradient.

Si l'on dérive l'image dans ces nouvelles composantes (g, t) , on obtient :

$$\begin{cases} I_g = I_x \cos \phi + I_y \sin \phi \\ I_t = I_x \sin \phi - I_y \cos \phi \end{cases} \quad (3.9)$$

Soit, comme $\cos \phi = \frac{I_x}{\|\nabla I\|}$, et $\sin \phi = \frac{I_y}{\|\nabla I\|}$:

$$\begin{cases} I_g = \|\nabla I\| \\ I_t = 0 \end{cases} \quad (3.10)$$

La composante tangentielle t correspond donc à la direction de l'isophote, ou ligne de niveau de la fonction I .

La composante g correspond à la direction principale de variation de la fonction I , donc les maxima locaux dans cette direction coïncident avec les passage par zéros de la dérivée seconde selon g : $I_{gg} = 0$.

Calculons les dérivées secondes dans le repère (g, t) :

$$\begin{cases} I_{gg} = I_{xx} \cos^2 \phi + 2I_{xy} \cos \phi \sin \phi + I_{yy} \sin^2 \phi \\ I_{tt} = I_{xx} \sin^2 \phi - 2I_{xy} \cos \phi \sin \phi + I_{yy} \cos^2 \phi \end{cases} \quad (3.11)$$

Et les dérivées troisièmes :

$$\begin{cases} I_{ggg} = I_{xxx} \cos^3 \phi + 3I_{xxy} \cos^2 \phi \sin \phi + 3I_{xyy} \cos \phi \sin^2 \phi + I_{yyy} \sin^3 \phi \\ I_{ttt} = I_{xxx} \sin^3 \phi - 3I_{xxy} \cos \phi \sin^2 \phi - 3I_{xyy} \cos^2 \phi \sin \phi + I_{yyy} \cos^3 \phi \end{cases} \quad (3.12)$$

Formellement, les contours peuvent être définis comme le lieu des points où la dérivée seconde selon g s'annule, et la dérivée troisième selon g est strictement négative :

$$\begin{cases} I_{gg} = 0 \\ I_{ggg} < 0 \end{cases} \quad (3.13)$$

En remplaçant $\cos \phi$ et $\sin \phi$ par leur expression en fonction des coordonnées du gradient :

$$\begin{cases} I_{gg} = \frac{I_{xx}I_x^2 + 2I_{xy}I_xI_y + I_{yy}I_y^2}{I_x^2 + I_y^2} = 0 \\ I_{ggg} = \frac{I_{xxx}I_x^3 + 3I_{xxy}I_x^2I_y + 3I_{xyy}I_xI_y^2 + I_{yyy}I_y^3}{I_x^2 + I_y^2} < 0 \end{cases} \quad (3.14)$$

Comme on ne s'intéresse qu'au signe, on peut exclure le dénominateur, et la définition des contours devient finalement :

$$\begin{cases} I_{xx}I_x^2 + 2I_{xy}I_xI_y + I_{yy}I_y^2 = 0 \\ I_{xxx}I_x^3 + 3I_{xxy}I_x^2I_y + 3I_{xyy}I_xI_y^2 + I_{yyy}I_y^3 < 0 \end{cases} \quad (3.15)$$

L'implantation discrète consiste à calculer, en tout point (i, j) les deux quantités précédentes, soit :

$$\begin{cases} \hat{I}_{gg} = I_{xx}I_x^2 + 2I_{xy}I_xI_y + I_{yy}I_y^2 \\ \hat{I}_{ggg} = I_{xxx}I_x^3 + 3I_{xxy}I_x^2I_y + 3I_{xyy}I_xI_y^2 + I_{yyy}I_y^3 \end{cases} \quad (3.16)$$

Et à détecter les changements de polarité de \hat{I}_{gg} sur des voisinages 2×2 , sous condition de négativité de \hat{I}_{ggg} , soit :

$$C_1[i, j] = 1 \Leftrightarrow \begin{cases} \bigvee_{(di, dj) \in \{(0,1), (1,0), (1,1)\}} (\hat{I}_{gg}[i, j] \cdot \hat{I}_{gg}[i - di, j - dj] < 0) \\ \bigwedge_{(di, dj) \in \{0,1\}^2} (\hat{I}_{ggg}[i - di, j - dj] < 0) \end{cases} \quad (3.17)$$

Avec \vee (resp. \wedge), le OU (resp. le ET) logique. La figure 3.6 montre le résultat de cet opérateur sur une image naturelle. On voit que l'application de la seule condition $(I_{gg} = 0; I_{ggg} < 0)$ fournit beaucoup de contours, en effet, aucune condition sur la valeur du contraste n'est spécifiée, et de nombreuses structures imperceptibles sur l'image originale apparaissent. Pour limiter ce problème, on ajoute une condition en imposant, en plus, que le module du gradient soit supérieur à un certain seuil. Cependant, il est très

3.3. EXEMPLE 1 : DÉTECTION DE CONTOURS

difficile de trouver avec un seul seuil le bon compromis entre élimination des structures parasites et bonne continuité des contours intéressants (voir Figure 3.6, images (2) et (3)).

Pour cela, on utilise un seuillage par hystérésis, qui consiste à calculer 2 seuils, un seuil haut s_h et un seuil bas s_b . Chaque seuil fournit un ensemble de pixels, certains étant connectés entre eux pour former des courbes connexes. Le principe est de ne retenir que les courbes connexes de s_b qui ont une intersection non vide avec s_h . Cet opérateur correspond à la reconstruction géodésique de s_h dans s_b (voir Cours de Morphologie Mathématique).

L'opérateur \mathcal{C}_1 , bien que conceptuellement simple, est relativement complexe du point de vue calculatoire, car il implique le calcul de 9 grandeurs différentielles. Une approximation nettement moins coûteuse peut être obtenue en négligeant la condition de négativité de la dérivée troisième et en remplaçant les passages par zéro de la dérivée seconde dans la direction du gradient par les passages par zéros du laplacien. En observant l'équation 3.11, on peut remarquer que :

$$I_{gg} + I_{tt} = I_{xx} + I_{yy} \quad (3.18)$$

Le laplacien est donc un invariant vis-à-vis de la rotation. D'autre part, la dérivée seconde dans la direction tangentielle I_{tt} , correspond à la courbure de la ligne de niveau. Si l'on suppose que cette courbure est faible, le laplacien coïncide alors avec la dérivée seconde dans la direction du gradient :

$$I_{tt} \simeq 0 \Rightarrow I_{gg} \simeq I_{xx} + I_{yy} \quad (3.19)$$

La définition formelle des contours devient donc :

$$\Delta_I = 0 \quad (3.20)$$

où $\Delta_I = I_{xx} + I_{yy}$ désigne le laplacien de I . Comme dans le cas précédent, on calcule le laplacien en tout point (i, j) , puis on détecte les changements de polarité de Δ_I sur des voisinages 2×2 :

$$\mathcal{C}_2[i, j] = 1 \Leftrightarrow \bigvee_{(di, dj) \in \{(0,1), (1,0), (1,1)\}} (\Delta_I[i, j] \cdot \Delta_I[i - di, j - dj] < 0) \quad (3.21)$$

La figure 3.7 montre le résultats de cet opérateur, dans les mêmes conditions que l'opérateur précédent.

Les figures 3.8 et 3.9 montrent les résultats obtenus par les opérateurs \mathcal{C}_1 , et \mathcal{C}_2 , pour différentes échelles.

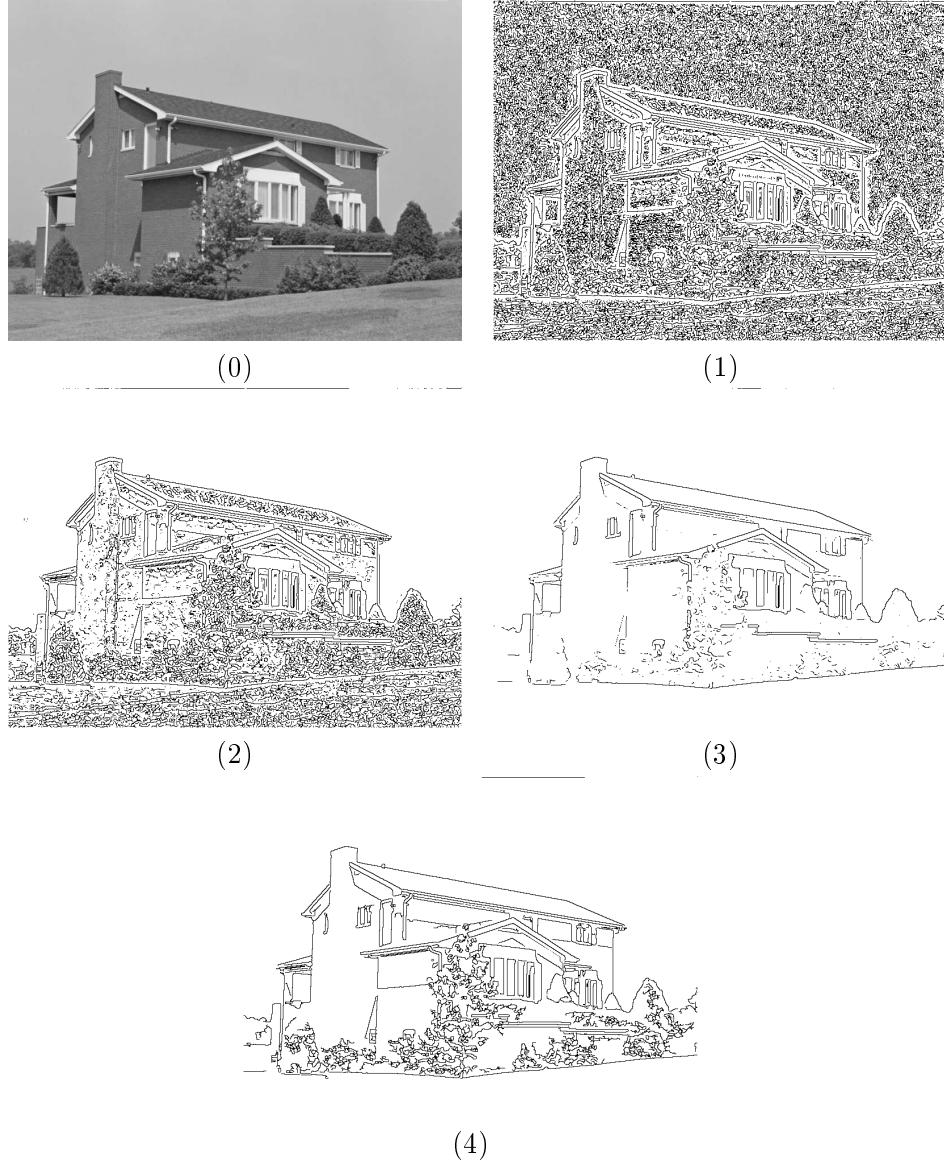


FIG. 3.6 – Détection de contours par l'opérateur $I_{gg} = 0; I_{ggg} < 0$, pour l'échelle $s = 1.125$. (0) Image originale, (1) Détection sans seuil, (2) Détection avec seuil $\|\nabla I\| > 1.0$, (3) Détection avec seuil $\|\nabla I\| > 10.0$, (4) Détection avec seuillage par hystérésis (seuil bas : 1.0, seuil haut : 10.0)

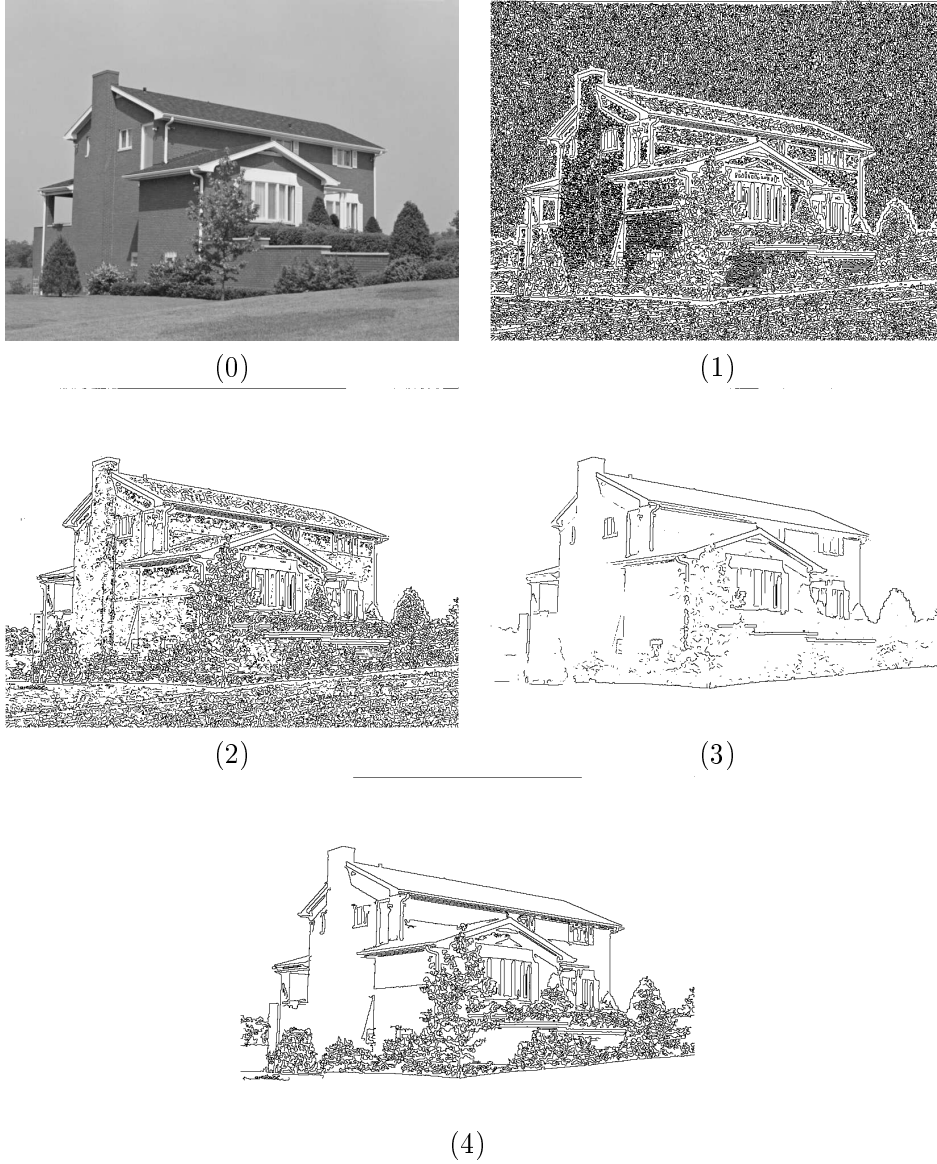


FIG. 3.7 – Détection de contours par l'opérateur $\Delta_I = 0$, pour l'échelle $s = 1.125$. (0) Image originale, (1) Détection sans seuil, (2) Détection avec seuil $\|\nabla I\| > 1.0$, (3) Détection avec seuil $\|\nabla I\| > 10.0$, (4) Détection avec seuillage par hystérésis (seuil bas : 1.0, seuil haut : 10.0)

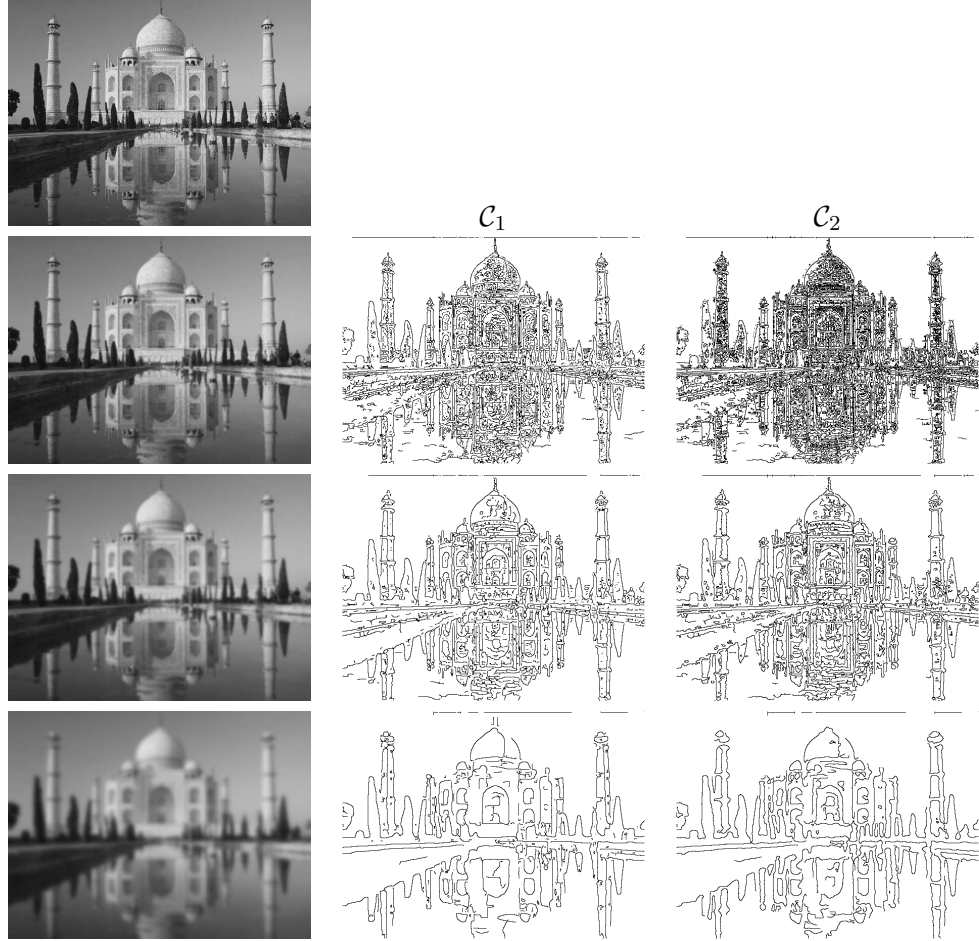


FIG. 3.8 – Detections de contours et espaces d'échelles, pour les échelles $s = 0.5, 2, 8$. La colonne de gauche montre l'image lissée par la gaussienne de variance $2s$. La colonne du milieu montre les contours définis par $(I_{gg} = 0; I_{ggg} < 0)$, la colonne de droite les contours définis par $\Delta_I = 0$. Le même seuillage par hystérésis du gradient (seuil bas : 1.0 ; seuil haut : 2.0) a été appliqué pour toutes les échelles.

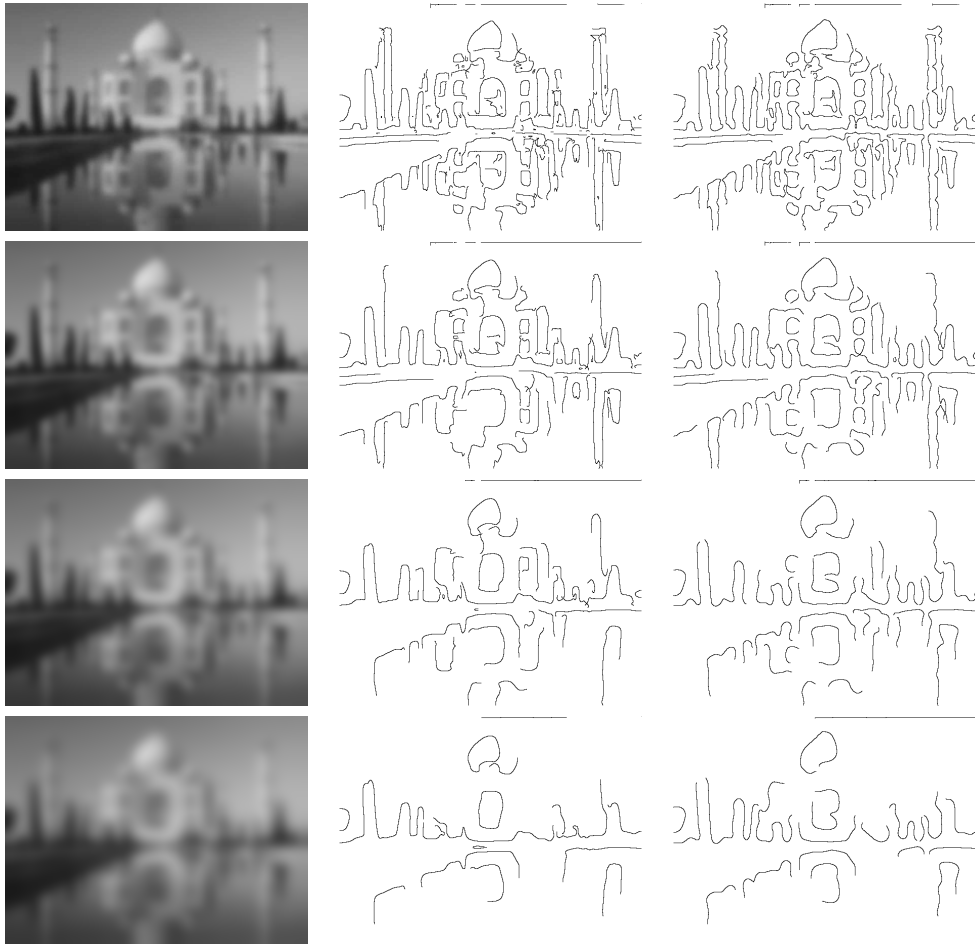


FIG. 3.9 – Détection de contours et espaces d'échelles, pour les échelles $s = 18, 32, 50, 72$. La colonne de gauche montre l'image lissée par la gaussienne de variance $2s$. La colonne du milieu montre les contours définis par $(I_{gg} = 0; I_{ggg} < 0)$, la colonne de droite les contours définis par $\Delta_I = 0$. Le même seuillage par hystérésis du gradient (seuil bas : 1.0 ; seuil haut : 2.0) a été appliqué pour toutes les échelles.

3.4 Exemple 2 : Détection de points d'intérêt

Les points d'intérêt constituent, au même titre que les contours, un opérateur de réduction d'information très important en traitement d'images. L'objectif est de trouver un ensemble de points isolés dans l'image, tel que chaque point soit représentatif de l'information locale, et que cette information soit suffisamment riche pour être discriminée du reste de l'image. Sont donc exclus les points appartenant à des régions homogènes, mais aussi les points qui se trouvent le long d'un contour rectiligne ou de faible courbure, car dans ce cas, les voisinages de tous les points qui se trouvent le long de ce contour sont semblables.

Les points d'intérêt sont donc associés à des points anguleux, aux voisinages desquels l'image varie significativement dans plusieurs directions.

Soit (x, y) un pixel. Soit $W_{(x,y)}$ une fenêtre centrée en (x, y) , représentant un voisinage du pixel. Une mesure des variations locales de l'image I au point (x, y) , et associée au déplacement $(\Delta x, \Delta y)$, est fournie par la fonction d'autocorrélation Chi , définie comme suit :

$$\chi_{(x,y)}(\Delta x, \Delta y) = \sum_{(i,j) \in W_{(x,y)}} (I(i, j) - I(i + \Delta x, j + \Delta y))^2 \quad (3.22)$$

La formule de Taylor appliquée au premier ordre donne :

$$I(i + \Delta x, j + \Delta y) \simeq I(i, j) + \begin{pmatrix} \frac{\partial I}{\partial x}(i, j) & \frac{\partial I}{\partial y}(i, j) \end{pmatrix} \cdot \begin{pmatrix} \Delta x \\ \Delta y \end{pmatrix} \quad (3.23)$$

Et donc :

$$\chi_{(x,y)}(\Delta x, \Delta y) = \sum_{(i,j) \in W_{(x,y)}} \left(\begin{pmatrix} \frac{\partial I}{\partial x}(i, j) & \frac{\partial I}{\partial y}(i, j) \end{pmatrix} \cdot \begin{pmatrix} \Delta x \\ \Delta y \end{pmatrix} \right)^2 \quad (3.24)$$

Ce qu'on peut réécrire sous la forme matricielle suivante :

$$\chi_{(x,y)}(\Delta x, \Delta y) = \begin{pmatrix} \Delta x & \Delta y \end{pmatrix} \Xi_I(x, y) \begin{pmatrix} \Delta x \\ \Delta y \end{pmatrix} \quad (3.25)$$

Où $\Xi_I(x, y)$ est dite *matrice d'autocorrélation* de l'image I en (x, y) , avec :

$$\Xi_I(x, y) = \begin{pmatrix} \sum_{(i,j) \in W_{(x,y)}} \left(\frac{\partial I}{\partial x}(i, j) \right)^2 & \sum_{(i,j) \in W_{(x,y)}} \left(\frac{\partial I}{\partial x}(i, j) \cdot \frac{\partial I}{\partial y}(i, j) \right) \\ \sum_{(i,j) \in W_{(x,y)}} \left(\frac{\partial I}{\partial x}(i, j) \cdot \frac{\partial I}{\partial y}(i, j) \right) & \sum_{(i,j) \in W_{(x,y)}} \left(\frac{\partial I}{\partial y}(i, j) \right)^2 \end{pmatrix} \quad (3.26)$$

La matrice d'autocorrélation $\Xi_I(x, y)$ représente la variation locale de l'image I au point (x, y) . Le pixel (x, y) sera considéré comme un point d'intérêt si, pour tous les déplacements $(\Delta x, \Delta y)$, la quantité $(\Delta x, \Delta y) \Xi_I(x, y) (\Delta x, \Delta y)^t$ est grande.

Les points d'intérêt sont donc les points (x, y) tels que la matrice d'autocorrélation $\Xi_I(x, y)$ a 2 valeurs propres grandes. Cela correspond aux points pour lesquels il existe localement une base de vecteurs propres de Ξ_I décrivant les principales directions de variations de l'image.

Le *détecteur de Harris* est un opérateur basé sur la matrice d'autocorrélation. Il ne calcule pas explicitement les valeurs propres de Ξ_I , mais une fonction d'intérêt Θ_I définie par une combinaison linéaire du déterminant et de la trace de Ξ_I :

$$\Theta_I(x, y) = \det \Xi_I(x, y) - \alpha \operatorname{trace} \Xi_I(x, y) \quad (3.27)$$

Le premier terme correspond au produit des valeurs propres, le second terme sert à pénaliser les points de fort contraste avec une seule forte valeur propre. α est un paramètre de pondération déterminé de façon empirique (les valeurs rencontrées dans la littérature varient peu, entre 0.04 et 0.06).

Les figures 3.10 et 3.11 montrent le calcul à différentes échelles de la fonction d'intérêt, pour une image faisant apparaître différents types de courbures. Noter l'effet du terme négatif faisant intervenir la trace, qui fait apparaître des valeurs négatives sur les contours rectilignes. Pour calculer la fonction d'intérêt de Harris Θ_I^s à l'échelle s , on calcule les dérivées premières I_x^s et I_y^s comme les convolutions de I avec les dérivées de gaussienne de variance $2s$, selon x et selon y respectivement. Puis, on calcule les composantes de la matrice Ξ_I^s en remplaçant la somme des dérivées dans la fenêtre W par un lissage gaussien de variance $2s'$. s et s' doivent être proportionnels ; $s' = s$ ou bien $s' = 2s$ sont des choix possibles.

Une fois la fonction d'intérêt calculée, les points d'intérêt sont définis comme les maxima locaux de la fonction d'intérêt (c'est-à-dire les pixels qui ont une valeur de Θ_I supérieure ou égale à celle de tous leurs voisins. On fixe aussi habituellement un seuil sur la valeur de la fonction d'intérêt (par exemple 1% de la valeur maximale sur toute l'image de la fonction d'intérêt). La figure 3.12 montre les points d'intérêt détectés sur un détail de l'image «Guernica», pour différentes échelles.

3.5 Espaces d'échelles non linéaires

Coming soon...



FIG. 3.10 – Fonction d'intérêt de Harris calculée sur l'image «Guernica», à l'échelle $s = 0.5$

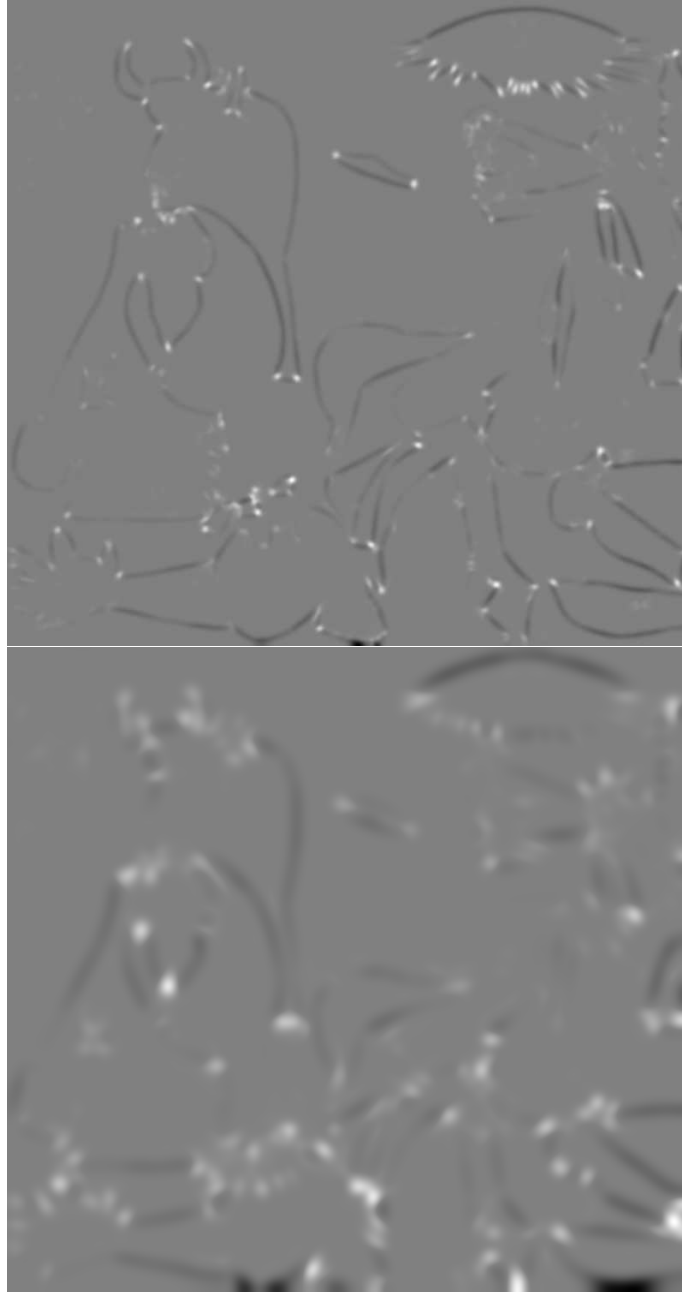


FIG. 3.11 – Fonction d'intérêt de Harris calculée sur l'image «Guernica», aux échelles $s = 4.5$, et $s = 32.0$.

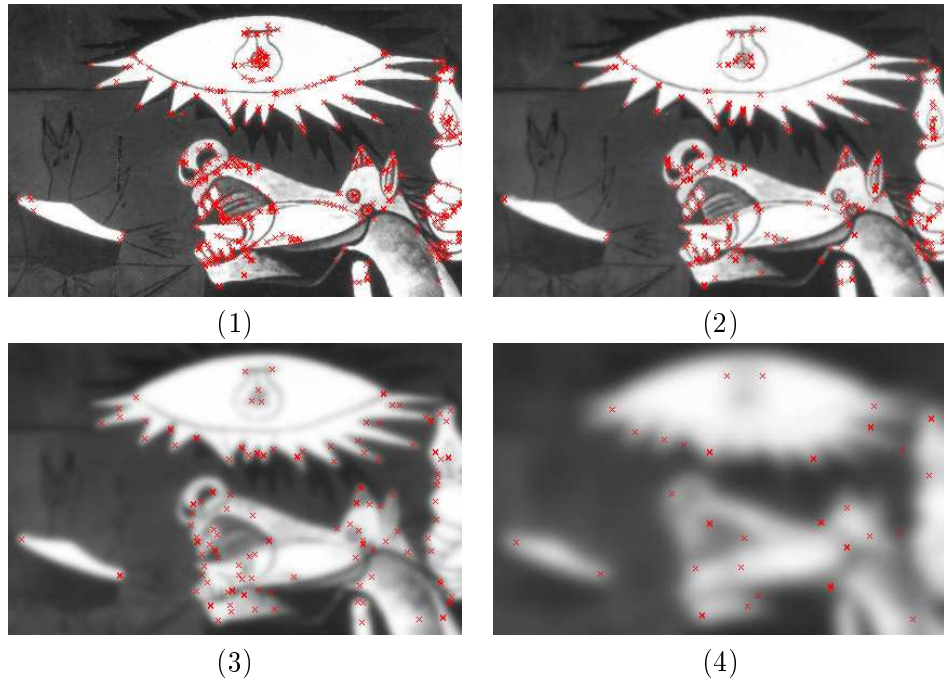


FIG. 3.12 – Détection de points d'intérêt par la méthode de Harris multi-échelle, pour les échelles : (1) $s = 0.125$, (2) $s = 0.5$, (3) $s = 4.5$, et (4) $s = 32.0$. Les points affichés (croix en rouge) correspondent aux maxima locaux de la fonction d'intérêt Θ_I , dont la valeur est supérieure à 1% du maximum (global) de Θ_I