

Estimation du niveau de bruit par arbre des formes et statistiques non paramétriques

Baptiste ESTEBAN, Guillaume TOCHON, Thierry GÉRAUD

Laboratoire de Recherche et Développement de l'EPITA (LRDE)
14-16 rue Voltaire, 94270 Le Kremlin-Bicêtre, France
prénom.nom@lrde.epita.fr

Résumé – La connaissance du niveau de bruit dans une image est précieuse pour de nombreuses applications en traitement d'images. L'estimation de la fonction de niveau de bruit requiert l'identification des zones homogènes sur lesquelles les paramètres du bruit peuvent être calculés. Sutour *et al.* [11] ont proposé une méthode d'estimation de la fonction de niveau de bruit se basant sur la recherche de zones homogènes de forme carrée, donc inadaptées au contenu local de l'image. Nous généralisons cette méthode à la recherche de zones homogènes de forme quelconque en nous basant sur la représentation par arbre des formes de l'image étudiée, permettant ainsi une estimation plus robuste de la fonction de niveau de bruit.

Abstract – The knowledge of the noise level within an image is valuable for many image processing applications. Estimating the noise level function requires the identification of homogeneous regions, upon which the noise parameters are computed. Sutour *et al.* [11] have proposed a method to estimate the noise level function based on the search for homogeneous areas with squared shape. We generalize this method to the search of homogeneous areas with arbitrary shape thanks to the tree of shape representation of the image under study, thus allowing a more robust estimation of the noise level function.

1 Introduction

Les images naturelles sont nativement corrompues par du bruit numérique résultant des imperfections de la chaîne d'acquisition (capteur, quantification, etc). La connaissance des caractéristiques de ce bruit, comme la fonction de niveau de bruit (NLF, pour *noise level function*) qui fournit la relation liant l'intensité de l'image à la variance du bruit [9], est précieuse pour de multiples applications en traitement d'images. La NLF est notamment d'intérêt pour des applications de débruitage [2] ou de segmentation. Son estimation requiert l'identification de zones homogènes dans l'image, sur lesquelles les paramètres du bruit sont évalués [1]. Sutour *et al.* [11] ont proposé une méthode non paramétrique de détection des zones homogènes, avec comme unique hypothèse la décorrélation spatiale du bruit et du contenu de l'image. Leur définition comme étant des blocs de $N \times N$ pixels n'en facilite cependant pas la détection, puisque leur forme carrée les rend inadaptée au contenu de l'image. L'utilisation de patchs s'adaptant localement à la morphologie de l'image permettrait une meilleure identification des zones homogènes, et donc une estimation plus robuste de la NLF. Cette idée de patchs adaptatifs se retrouve déjà dans la littérature pour le débruitage d'images [5, 4], mais n'a, à notre connaissance, jamais été implémentée pour l'estimation de la NLF.

Dans cet article, nous proposons d'adapter la méthode d'estimation de la NLF proposée par Sutour *et al.* [11] en définissant les zones homogènes de l'image considérée à partir de sa représentation par arbre des formes (AdF) [10]. Cette représentation hiérarchique fournit naturellement des zones dont les contours suivent les lignes de niveau de l'image, donc adaptées à son contenu morphologique. L'extraction des formes les plus perti-

nentes (au sens de la fonctionnelle de Mumford-Shah) de l'AdF permet effectivement d'obtenir une estimation de la NLF plus robuste que dans le cadre de patchs carrés.

2 Estimation de la NLF par blocs carrés

L'estimation de la NLF proposée par Sutour *et al.* [11] se décompose en deux étapes : la détection de blocs homogènes par test statistique non paramétrique, puis l'identification de la relation liant l'intensité de l'image à la variance du bruit. Une image sera défini par la suite comme une fonction $f : \Omega \subset \mathbb{Z}^2 \rightarrow \mathbb{R}$ qui associe une valeur numérique $x_i = f(i)$ à tout pixel $i \in \Omega$.

Coefficient τ de Kendall Le bloc de pixel $b_i \subset \Omega$ est dit homogène si la valeur de ses pixels ne fluctue qu'à cause du bruit. Afin de ne dépendre d'aucune hypothèse sur la nature du bruit, la détection des blocs homogènes est formulée comme un test d'hypothèse non paramétrique grâce au coefficient τ de Kendall [7]. Soit $x \in \mathbb{R}^n$ et $y \in \mathbb{R}^n$ deux séquences de n observations de deux variables aléatoires X et Y . Le coefficient $\tau(x, y)$ de Kendall est défini sur l'intervalle $[-1, 1]$ par :

$$\tau(x, y) = \frac{1}{n(n-1)} \sum_{1 \leq i, j \leq n} \text{sign}(x_i - x_j) \text{sign}(y_i - y_j) \quad (1)$$

avec $x_i \neq x_j$ et $y_i \neq y_j \quad \forall i \neq j$.

$\tau(x, y) = 0$ constitue l'hypothèse H_0 et indique une absence de corrélation entre les valeurs de x et de y . L'équation (1) peut être reformulée pour prendre en compte les éventuelles paires liées (si $x_i = x_j$ ou $y_i = y_j$) [8]. Sous H_0 , le z-score associé à

$\tau(x, y)$ suit une loi normal centrée réduite $\mathcal{N}(0, 1)$ [11, Proposition 3.6]. À contrario, l’hypothèse alternative H_1 correspond à $\tau(x, y) \neq 0$, c’est-à-dire lorsque la fluctuation des valeurs des pixels de b_i ne peut être imputée au bruit seul (présence d’un contour dans le bloc par exemple).

Détection des blocs homogènes L’image étudiée est divisée en blocs de 16×16 pixels. Chaque bloc b_i est séparé en deux séquences de valeurs x et y pour lesquelles le coefficient $\tau(x, y)$ de Kendall est calculé. Une p-valeur du test $p = \mathbb{P}(\tau(X, Y) > \tau(x, y) | H_0)$ inférieure à un seuil de détection α pré-déterminé rejette l’hypothèse H_0 et le bloc b_i est déclaré non homogène. En pratique, b_i est divisé en $K = 4$ séquences $x^{(k)}$ et $y^{(k)}$ selon les relations de voisinage horizontales, verticales, diagonales et anti-diagonales. L’hypothèse d’homogénéité de b_i est acceptée si les K p-valeurs p_k dépassent le seuil de détection :

$$\min_k \left\{ p_k = \mathbb{P} \left(\tau(X, Y) > \tau(x^{(k)}, y^{(k)}) | H_0 \right) \right\} > \alpha. \quad (2)$$

Estimation de la NLF La NLF définit la relation liant l’intensité des pixels à la variance du bruit les corrompant. Le modèle polynomial du second degré $\sigma_i^2 = \text{NLF}_{(a,b,c)}(x_i) = ax_i^2 + bx_i + c$ permet notamment d’inclure le bruit additif gaussien (NLF constante, $(a, b) = (0, 0)$), le bruit de Poisson (NLF linéaire, $(a, c) = (0, 0)$), le bruit multiplicatif (NLF parabolique, $(b, c) = (0, 0)$), et leurs différentes mixtures.

En pratique, la moyenne $\hat{\mu}_i$ et la variance $\hat{\sigma}_i^2$ de chaque bloc homogène b_i sont calculées, et les coefficients de la NLF sont estimés comme ceux permettant la meilleure correspondance entre tous les couples $(\hat{\sigma}_i^2, \text{NLF}_{(a,b,c)}(\hat{\mu}_i))$ pour la norme L^1 :

$$(\hat{a}, \hat{b}, \hat{c}) = \underset{(a,b,c) \in (\mathbb{R}^+)^3}{\operatorname{argmin}} \sum_i \|\text{NLF}_{(a,b,c)}(\hat{\mu}_i) - \hat{\sigma}_i^2\|_1 \quad (3)$$

Ce problème de minimisation est résolu grâce à l’algorithme primal-dual préconditionné de Chambolle-Pock [3].

3 Méthode proposée

La méthode proposée consiste à remplacer la sélection de blocs carrés par des patchs adaptés à la morphologie de l’image. Ces patchs sont extraits de l’AdF de l’image considérée.

3.1 L’arbre des formes

Pour $\lambda \in \mathbb{R}$, l’ensemble de niveaux inférieurs à λ est défini comme $[f < \lambda] = \{i \in \Omega \mid f(i) < \lambda\}$ et l’ensemble de niveaux supérieurs à λ comme $[f \geq \lambda] = \{i \in \Omega \mid f(i) \geq \lambda\}$. Une forme $\mathcal{C} \subset \Omega$ est une composante connexe appartenant à l’un de ces deux ensembles, avec les trous remplis [6]. L’AdF \mathcal{T} d’une image encode de manière hiérarchique la relation d’inclusion entre les différentes formes de l’image. La bordure $\partial\mathcal{C}$ de chaque forme \mathcal{C} correspondant à une ligne de niveau de l’image (vue comme un relief topographique). L’AdF représente donc l’inclusion des lignes de niveau de l’image. Un exemple d’AdF est représenté figure 1.

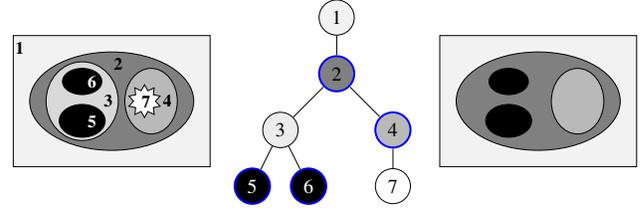


FIGURE 1 – (gauche) Image d’entrée, (milieu) arbre des formes correspondant, et (droite) simplification de Mumford-Shah pour les nœuds cerclés de bleu.

3.2 Simplification de l’AdF

Afin de ne retenir que les formes \mathcal{C}_i les plus adaptées au contenu de l’image, une simplification de l’AdF selon la fonctionnelle de Mumford-Shah est tout d’abord appliquée. Si $\pi = \bigsqcup_i R_i$ est une partition de Ω , et si \bar{f}_i est la valeur moyenne de f sur la région $R_i \subset \Omega$, la fonctionnelle de Mumford-Shah (constante par morceaux) de f sur π est définie comme

$$E_\nu(f, \pi) = \sum_{R_i \in \pi} \left(\iint_{R_i} (\bar{f}_i - f)^2 dx dy + \frac{\nu}{2} |\partial R_i| \right) \quad (4)$$

où $|\partial R_i|$ est la longueur de la frontière de R_i et ν est un paramètre de régularisation de la fonctionnelle.

Dans le cas de l’AdF, la simplification de Mumford-Shah revient à supprimer certaines formes de \mathcal{T} pour aboutir à un AdF simplifié \mathcal{T}'_ν dont la partition constituée par l’ensemble des lignes de niveau $\partial\mathcal{T}'_\nu$ minimise la fonctionnelle (4) (voir [12] pour plus de détails) :

$$\min_{\mathcal{T}'_\nu} E_\nu(f, \partial\mathcal{T}'_\nu) \quad (5)$$

Un exemple de simplification est représenté figure 1.

3.3 Détection des formes homogènes

L’AdF construit sur l’image f est composé de formes aux contours très irréguliers à cause du bruit, et avec une profondeur d’inclusion limitée. Un pré-filtrage de f par filtre gaussien avant la construction de l’AdF régularise les contours des formes, et augmente leur nombre ainsi que leur profondeur d’inclusion. L’AdF obtenu est simplifié selon la fonctionnelle de Mumford-Shah pour plusieurs valeurs du paramètre ν afin de constituer un ensemble de formes candidates $\mathcal{C}^\cup = \bigcup_\nu \{\mathcal{C} \in \mathcal{T}'_\nu\}$ de différentes tailles et complexités.

Le test d’homogénéité est ensuite appliqué à toutes les formes candidates $\mathcal{C}_i \in \mathcal{C}^\cup$, en utilisant les valeurs des pixels de l’image bruitée $f(\mathcal{C}_i)$ et non ceux de l’image filtrée. Le calcul du coefficient $\tau(x, y)$ de Kendall nécessite de répartir les valeurs des pixels de \mathcal{C}_i en deux séquences x et y . Or, la forme de \mathcal{C}_i étant quelconque (par opposition à un bloc carré de 16×16 pixels), cela ne permet pas de garantir la validité des relations de voisinage utilisées par Soutour *et al.* [11]. La répartition en deux séquences x et y s’effectue donc de manière aléatoire (en imposant que les deux séquences soient de même longueur). En pratique, K répartitions aléatoires en séquences $x^{(k)}$ et $y^{(k)}$ sont effectuées, et les K p-valeurs p_k des coefficients $\tau(x^{(k)}, y^{(k)})$

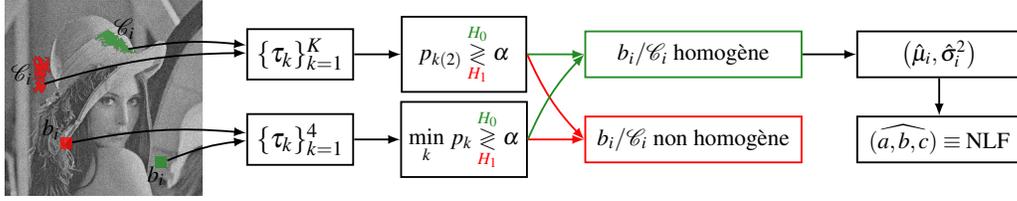


FIGURE 2 – Illustration des différentes étapes de l’estimation de la NLF à partir de blocs b_i et de formes C_i .

sont calculées. La forme C_i est déclarée homogène si la statistique d’ordre de rang 2 (la deuxième plus petite p-valeur) $p_{k(2)}$ excède le seuil de détection α (pour prévenir le cas où un test serait rejeté à cause d’une répartition aléatoire “particulière” et non de l’inhomogénéité de la forme considérée). La moyenne $\hat{\mu}_i$ et la variance $\hat{\sigma}_i^2$ de chaque forme homogène permettent par la suite l’estimation des coefficients $(\widehat{a}, \widehat{b}, \widehat{c})$ de la NLF selon l’équation (3). La figure 2 illustre ces différentes étapes.

4 Résultats

Nous évaluons notre méthode sur les 6 images de dimensions 720×540 pixels présentées sur la figure 3. Celles-ci sont extraites d’une base d’images¹ haute définition, pour lesquelles le bruit d’acquisition peut être considéré comme négligeable. Afin d’évaluer la qualité et la robustesse de notre estimation de la NLF, nous procédons à deux expériences. Dans un premier temps, nous rajoutons artificiellement un bruit mixte dont les coefficients a , b et c de la NLF sont tirés d’une loi normale de moyenne 0.01 et de variance 0.003 (l’intensité des pixels de l’image ayant été au préalable ramenée dans l’intervalle $[0;1]$). Dans un second temps, nous ajoutons un bruit non-mixte en faisant varier la valeur du coefficient de la NLF le contrôlant, et en maintenant les deux autres coefficients à 0.

Chaque image bruitée est pré-filtrée par un filtre gaussien de variance $\sigma^2 = 1$ sur laquelle est construit l’AdF. Ce dernier est simplifié selon la fonctionnelle de Mumford-Shah pour des valeurs du paramètre de régularisation ν de 200, 500, 1000, 2000 et 5000. Seulement les formes C_i de plus de 250 pixels des images simplifiées résultantes sont considérées, constituant un ensemble d’environ 700 régions sur lequel le test d’homogénéité est appliqué pour une image donnée. Cette taille minimale garantit une estimation fiable du coefficient de Kendall de C_i , ainsi que de sa moyenne $\hat{\mu}_i$ et sa variance $\hat{\sigma}_i^2$ (à noter que dans [11], la taille minimale est de $16 \times 16 = 256$ pixels pour un bloc b_i). Le nombre de répartitions aléatoires des pixels en séquences $x^{(k)}$ et $y^{(k)}$, et donc le nombre de tests de Kendall pour une forme donnée, est fixé à $K = 10$. La deuxième plus petite p-valeur $p_{k(2)}$ est utilisée pour valider l’homogénéité de la forme si elle excède le seuil de détection α fixé à $\alpha = 0.4$. Cette valeur très restrictive permet de vraiment garantir l’homogénéité des formes détectées comme telles.

Afin de mesurer la qualité de l’estimation des paramètres de la

NLF, nous calculons l’erreur relative moyenne (ERM) entre la NLF de référence $\text{NLF}_{(a,b,c)}$ et la NLF estimée $\text{NLF}_{(\widehat{a}, \widehat{b}, \widehat{c})}$ par

$$\text{ERM}(\widehat{a}, \widehat{b}, \widehat{c}) = \frac{1}{|I|} \sum_{x_i \in I} \frac{|\text{NLF}_{(a,b,c)}(x_i) - \text{NLF}_{(\widehat{a}, \widehat{b}, \widehat{c})}(x_i)|}{\text{NLF}_{(a,b,c)}(x_i)} \quad (6)$$

avec I une discrétisation de l’intervalle des intensités de l’image. La table 1 présente les résultats de qualité et de robustesse d’estimation de la NLF calculée à partir de blocs carrés ou de formes homogènes. Pour chaque image, nous simulons d’abord 20 bruits mixtes, et reportons la moyenne, l’écart-type ainsi que le pourcentage de fois (sur les 20 tirages) où l’ERM est plus faible lorsque la NLF est estimée à partir des formes de l’AdF plutôt que par blocs carrés (dernière colonne). La moyenne et l’écart-type de l’ERM de la NLF estimée à partir de l’AdF est significativement plus faible pour les 6 images étudiées. Cela confirme que l’utilisation de formes adaptées à la morphologie de l’image permet une estimation plus fiable et plus robuste de la NLF. Dans un second temps, un bruit non-mixte (multiplicatif, Poisson ou gaussien) est ajouté, et le paramètre a , b ou c de la NLF le contrôlant est réglé successivement à 0.01, 0.015, 0.02, 0.025, 0.03 et 0.5 (les deux autres paramètres étant maintenu à 0). Nous décomptons combien de fois chacune des deux méthodes estime une NLF caractéristique de la bonne nature du bruit (en estimant les deux autres coefficients à 0). L’estimation de la NLF par l’AdF obtient systématiquement de meilleurs résultats que la méthode par blocs, sauf pour le bruit multiplicatif a sur l’image *monument*. D’une manière générale, l’estimation d’un bruit gaussien semble nettement plus aisée que pour un bruit multiplicatif ou de Poisson.

La figure 4 présente les blocs et les formes détectées comme homogènes pour un bruit mixte de coefficients $a = 1.22 \cdot 10^{-2}$, $b = 0.82 \cdot 10^{-2}$ et $c = 0.82 \cdot 10^{-2}$, ainsi que les NLF estimées dans chacun des cas. Dans le premier cas, 104 blocs carrés sont détectés comme étant homogènes, conduisant à l’estimation $\hat{a} = 1.12 \cdot 10^{-2}$, $\hat{b} = 1.07 \cdot 10^{-2}$ et $\hat{c} = 0.85 \cdot 10^{-2}$ et une ERM de $7.56 \cdot 10^{-2}$. Dans le deuxième cas, 81 formes homogènes ont été détectées, induisant une estimation $\hat{a} = 0.93 \cdot 10^{-2}$, $\hat{b} = 0.98 \cdot 10^{-2}$ et $\hat{c} = 0.78 \cdot 10^{-2}$ et une ERM de $2.96 \cdot 10^{-2}$. Puisqu’il n’y a pas de seuil sur la taille maximale des formes homogènes C_i lors de leur détection, ces dernières apparaissent nettement plus grosses que les blocs carrés homogènes (de taille 16×16 pixels) puisqu’elles s’adaptent au contenu de l’image. Cela permet une meilleure estimation empirique de leur moyenne $\hat{\mu}_i$ et leur variance $\hat{\sigma}_i^2$, et donc une dispersion moindre

1. <http://www.gipsa-lab.grenoble-inp.fr/laurent.condat/imagebase.html>



FIGURE 3 – Images *coquillage*, *immeuble*, *monument*, *mur*, *pavé* et *ville*.

TABLE 1 – Moyenne (et écart type entre parenthèses) de l'ERM entre NLF de référence et NLF estimée en utilisant des blocs et en utilisant l'AdF, ainsi que nombre de bonne identification du type de bruit $a/b/c$.

	Par blocs [11]		Par AdF		
<i>coquillage</i>	0.1212 (0.0743)	0/1/0	0.0385 (0.0124)	2/3/6	100%
<i>immeuble</i>	0.0481 (0.0234)	1/0/1	0.0349 (0.0134)	6/4/4	65%
<i>monument</i>	0.2235 (0.0966)	6/0/0	0.0552 (0.0201)	0/5/4	95%
<i>mur</i>	0.1075 (0.1064)	0/0/1	0.0513 (0.0102)	0/0/6	75%
<i>pavé</i>	0.8701 (0.7031)	0/0/0	0.0346 (0.0121)	4/2/6	100%
<i>ville</i>	0.0926 (0.1007)	0/1/3	0.0462 (0.0116)	4/3/5	85%

des couples $(\hat{\mu}_i, \hat{\sigma}_i^2)$ autour de la courbe NLF de référence. L'utilisation de l'AdF permet donc une meilleure estimation de la NLF, tout en nécessitant moins de formes que dans le cadre de patches carrés.

5 Conclusion

En conclusion, nous avons présenté dans cet article une généralisation de la méthode d'estimation de la NLF de Sutour *et al.* Dans ce but, nous avons remplacé la recherche de patches carrés homogènes par celle de formes homogènes, extraites à partir d'un AdF simplifié par la fonctionnelle de Mumford-Shah. Le calcul des statistiques du bruit sur des zones adaptées à la morphologie de l'image rend l'estimation de la NLF plus robuste et plus précise à l'identification de la nature et des caractéristiques du bruit polluant l'image.

Références

- [1] L. Beaupaire, K. Chehdi, and B. Vozel. Identification of the nature of noise and estimation of its statistical parameters by analysis of local histograms. In *Proc. of ICASSP*, volume 4, pages 2805–2808. IEEE, 1997.
- [2] A. Buades, B. Coll, and J.-M. Morel. A review of image denoising algorithms, with a new one. *Multiscale Modeling & Simulation*, 4(2) :490–530, 2005.
- [3] A. Chambolle and T. Pock. A first-order primal-dual algorithm for convex problems with applications to imaging. *Journal of mathematical imaging and vision*, 40(1) :120–145, 2011.

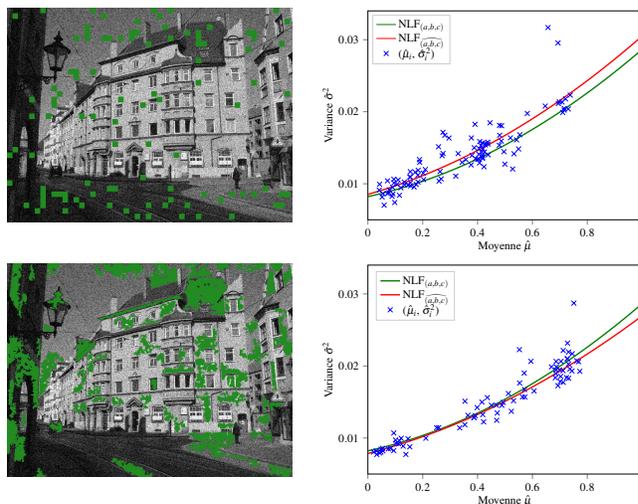


FIGURE 4 – Première ligne : blocs détectés comme homogènes pour un bruit mixte et NLF estimée par rapport à la NLF de référence ($ERM = 7.56 \cdot 10^{-2}$). Deuxième ligne : formes détectées comme homogènes pour le même bruit, et NLF estimée ($ERM = 2.96 \cdot 10^{-2}$).

- [4] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian. BM3D image denoising with shape-adaptive principal component analysis. In *Proc. of SPARS*, 2009.
- [5] C.-A. Deledalle, V. Duval, and J. Salmon. Non-local methods with shape-adaptive patches (NLM-SAP). *Journal of Mathematical Imaging and Vision*, 43(2) :103–120, 2012.
- [6] T. Géraud, E. Carlinet, S. Crozet, and L. Najman. A quasi-linear algorithm to compute the tree of shapes of nD images. In *Proc. of ISMM*, pages 98–110. Springer, 2013.
- [7] M. G. Kendall. A new measure of rank correlation. *Biometrika*, 30(1/2) :81–93, 1938.
- [8] M. G. Kendall. The treatment of ties in ranking problems. *Biometrika*, 33(3) :239–251, 1945.
- [9] C. Liu, W. T. Freeman, R. Szeliski, and S. B. Kang. Noise estimation from a single image. In *Proc. of CVPR*, volume 1, pages 901–908. IEEE, 2006.
- [10] P. Monasse and F. Guichard. Fast computation of a contrast-invariant image representation. *IEEE Transactions on Image Processing*, 9(5) :860–872, 2000.
- [11] C. Sutour, C.-A. Deledalle, and J.-F. Aujol. Estimation of the noise level function based on a nonparametric detection of homogeneous image regions. *SIAM Journal on Imaging Sciences*, 8(4) :2622–2661, 2015.
- [12] Y. Xu, T. Géraud, and L. Najman. Hierarchical image simplification and segmentation based on Mumford–Shah-salient level line selection. *Pattern Recognition Letters*, 83 :278–286, 2016.