

Estimation de la fonction de niveau de bruit pour des images couleurs en utilisant la morphologie mathématique

Baptiste ESTEBAN, Guillaume TOCHON, Edwin CARLINET, Didier VERNA

Laboratoire de Recherche et Développement de l'EPITA (LRDE)

14-16 rue Voltaire, 94270 Le Kremlin-Bicêtre, France

prenom.nom@lrde.epita.fr

Résumé – Le niveau de bruit est une information importante pour certaines applications de traitement d'image telles que la segmentation ou le débruitage. Par le passé, nous avons proposé une méthode pour estimer ce niveau de bruit en s'adaptant au contenu d'une image en niveau de gris et nous avons montré que ses performances dépassent celle de l'état de l'art. Dans cet article, nous proposons une extension de cette méthode aux images couleurs dont les valeurs multivariées, dénuées de relation d'ordre naturelle, impliquent de nouvelles problématiques. Afin de les résoudre, nous utilisons deux outils provenant de la morphologie mathématique : l'arbre des formes multivarié et l'apprentissage de treillis complet. Enfin, nous confirmons les conclusions de nos précédents travaux pour l'estimation de la fonction de niveau de bruit couleur, montrant que l'adaptation au contenu d'une image donne de meilleures performances que l'utilisation de blocs carrés.

Abstract – The noise level is an important information for several image processing tasks such as segmentation or denoising. In the past, we proposed a method to estimate the noise level in grayscale images using a content-adapted method and we demonstrated that it was more efficient than state-of-art methods. In this paper, we propose an extension to color images, whose multivariate values, lacking of a natural order relationship, prevent the usage of our previous method. To solve this issue, we introduce two tools from the field of mathematical morphology: the multivariate tree of shapes and the complete lattice learning. Finally, we compare the efficiency of our estimation with a block-based one.

1 Introduction

Les images naturelles peuvent être corrompues par un bruit dû à des phénomènes physiques se produisant durant l'acquisition. La connaissance de ce bruit, et plus particulièrement du niveau de bruit représentant sa variance en fonction de l'intensité des pixels, améliore les traitements effectués sur ces images tels que le débruitage [3]. De nombreuses méthodes existent pour estimer cette information, mais nécessitent généralement d'émettre une hypothèse sur la nature du bruit. Beaurepaire *et al.* [1] proposent d'estimer la nature du bruit avant d'en estimer son niveau, mais le bruit mixte n'est pas défini. Sutour *et al.* [12] résolvent ce problème en modélisant une fonction de niveau de bruit (FNB) [10] comme une fonction quadratique où chaque coefficient correspond à un type de bruit. Pour estimer la FNB, ils proposent une méthode non-paramétrique s'appuyant sur la détection de blocs homogènes en utilisant un test statistique basé sur le rang des valeurs des pixels d'une image. Dans [6], nous adaptons cette méthode au contenu de l'image en utilisant l'arbre des formes (AdF) pour détecter des régions homogènes, et nous montrons que cela améliore la précision de l'estimation par rapport à l'utilisation de blocs.

Néanmoins, ces deux méthodes sont restreintes aux images en niveau de gris. Dans cet article, nous proposons d'étendre notre précédente approche aux images couleurs, dont les valeurs ne possèdent pas de relation d'ordre naturelle, empêchant la construction d'un AdF et le calcul d'une corrélation basée sur le rang. Ainsi, nous contourignons ce problème par le biais de deux outils appartenant au domaine de

la morphologie mathématique : l'arbre des formes multivarié (AdFM) [4] et l'apprentissage de treillis complets [9]. Nous confirmons ensuite les résultats de [6] déclarant que l'adaptation au contenu d'une image améliore la précision de l'estimation par rapport à l'utilisation de blocs.

2 Estimation de la FNB pour des images en niveau de gris

Dans cette section, nous rappelons la méthode d'estimation de la FNB proposée dans [6] pour une image en niveau de gris $f : \Omega \rightarrow \mathcal{I} \subset \mathbb{R}$ avec Ω son domaine de définition et \mathcal{I} l'espace contenant les valeurs de cette image. La FNB est modélisée par Sutour *et al.* [12] comme une fonction quadratique $FNB(x) = ax^2 + bx + c = \sigma^2$, avec σ^2 la variance du bruit en fonction de l'intensité x et $(a, b, c) \in (\mathbb{R}^+)^3$.

2.1 Définition et simplification de l'AdF

L'ensemble des niveaux supérieurs d'une image f est défini par $[f > \lambda] = \{p \in \Omega \mid f(p) > \lambda\}$ et l'ensemble des niveaux inférieurs par $[f \leq \lambda] = \{p \in \Omega \mid f(p) \leq \lambda\}$. Une forme \mathcal{C} est une composante connexe appartenant à l'un de ces deux ensembles dont les trous sont remplis. L'arbre des formes (AdF) \mathcal{T} [11] est une représentation hiérarchique encodant la relation d'inclusion de ces formes dans une image, illustré par la Figure 1b représentant l'AdF de la Figure 1a. Le contour d'une forme \mathcal{C} , noté $\partial\mathcal{C}$, représente une ligne de niveau de l'image, l'AdF encodant ainsi la relation d'inclusion de l'ensemble des lignes de niveau d'une image.

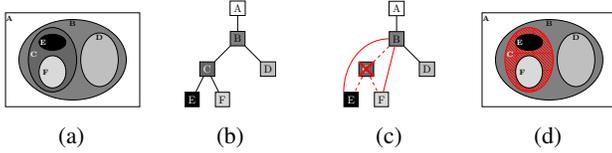


FIGURE 1 – Simplification d'image par arbre des formes

L'AdF s'est avéré performant pour plusieurs applications de traitement d'image telles que la détection d'objet ou la segmentation d'image. Certaines méthodes de segmentation se basent sur un principe de minimisation d'énergie, en utilisant notamment la fonctionnelle de Mumford-Shah. Appliquée à un AdF \mathcal{T} [13], la fonctionnelle de Mumford-Shah constante par morceaux est définie par :

$$E(\mathcal{T}) = \sum_{C_i \in \mathcal{T}} \sum_{\substack{x \in C_i \\ x \notin E(C_i)}} \|f(x) - \tilde{f}(x)\|_2^2 + \lambda |\partial \mathcal{T}| \quad (1)$$

avec $E(C)$ le domaine de définition des enfants du nœud C , \tilde{f} une fonction constante au sein d'une image (ici la moyenne d'une composante), $|\partial \mathcal{T}|$ la longueur des lignes de niveaux $\partial \mathcal{T}$ et λ un paramètre de la fonctionnelle. Minimiser cette fonctionnelle consiste à effectuer un filtrage des nœuds de l'AdF. La Figure 1c et 1d illustrent l'impact de la suppression d'un nœud de l'arbre sur une image.

2.2 Détection de régions homogènes

La présence de bruit seul dans une région d'une image équivaut à une absence de corrélation entre les pixels de cette région, qui est dite *homogène*. Sutour *et al.* [12] proposent de détecter ces régions en utilisant un test statistique basé sur un coefficient de corrélation de rang, le coefficient τ de Kendall, qui mesure la corrélation entre les valeurs de deux séquences $X \in \mathbb{R}^n$ et $Y \in \mathbb{R}^n$, et qui est défini par

$$\tau(X, Y) = \frac{1}{n-1} \sum_{1 \leq i < j \leq n} \text{signe}(x_i - x_j) \text{signe}(y_i - y_j) \quad (2)$$

avec x_k (resp. y_k) le k^{e} élément de X (resp. Y), $x_i \neq x_j$ et $y_i \neq y_j$. Cette définition ne prenant pas en compte les paires liées ($x_i = x_j$ ou $y_i = y_j$), elle est généralisée dans [8] et utilisée dans le cadre de l'estimation de la FNB.

Sous l'hypothèse nulle H_0 du test statistique, indiquant une absence de corrélation entre les deux séquences x et y ($\tau(x, y) = 0$), le z -score associé au coefficient τ de Kendall suit une distribution normale $\mathcal{N}(0, 1)$. L'hypothèse alternative H_1 est déclarée quand $\tau(x, y) \neq 0$, indiquant que la fluctuation des valeurs ne peut être expliquée par la fluctuation du bruit uniquement.

2.3 Estimation de la FNB

L'estimation de la FNB est divisée en 3 étapes. Premièrement, l'image est floutée en utilisant un filtre gaussien, ce dernier réduisant le nombre de lignes de niveau présent dans une image. Ensuite, l'AdF est construit sur

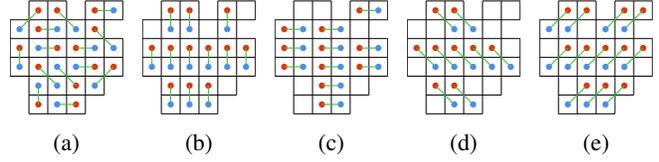


FIGURE 2 – Divisions de région en 2 séquences x et y

cette image floutée. Cet AdF est simplifié plusieurs fois en utilisant différents paramètres λ de la fonctionnelle de Mumford-Shah, résultant en plusieurs segmentations de l'image. Les composantes connexes de ces segmentations forment ainsi l'ensemble des régions sur lesquelles le test d'homogénéité est effectué, à l'exception des régions ne respectant pas un critère de taille minimum.

Le test d'homogénéité consiste à diviser chaque région en deux séquences de pixels de manière aléatoire, tel qu'illustré Figure 2a, et à calculer la valeur p sur ces séquences. Ce processus est répété $t \geq 2$ fois et la seconde plus petite valeur p est comparée à un seuil de détection α . Si cette valeur p est supérieure ou égale à α , alors la région est détectée comme homogène. La FNB représentant la variance du bruit en fonction de l'intensité des pixels d'une image, elle peut être représentée par $\hat{\sigma}_i^2 \approx \text{FNB}(\hat{\mu}_i) = a\hat{\mu}_i^2 + b\hat{\mu}_i + c$. avec $\hat{\mu}_i$ la moyenne empirique et $\hat{\sigma}_i^2$ la variance empirique d'une région homogène indexée à i . Ainsi, l'estimation de la FNB est réduite au problème suivant :

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

résolu par l'algorithme primal-dual de Chambolle-Pock [5].

3 Extension aux images couleurs

La méthode décrite dans la section 2 pour estimer la FNB pour des images en niveau de gris se base sur le coefficient τ de Kendall et l'AdF, deux outils nécessitant une relation d'ordre totale entre les valeurs d'une image. Pour une image couleur $f : \Omega \rightarrow \mathcal{I} \subset \mathbb{R}^3$, avec Ω son domaine de définition, les éléments de \mathcal{I} sont des vecteurs $(x_r, x_g, x_b) \in \mathbb{R}^3$ dénombrés par $m = \text{card}(\mathcal{I})$. La méthode décrite dans la section 2 n'a pas d'extension directe à cause de l'absence de relation d'ordre naturelle sur \mathbb{R}^3 . Dans cette section, nous contournons ce problème en utilisant deux outils provenant du domaine de la morphologie mathématique multivariée pour estimer la FNB d'une image couleur.

3.1 Apprentissage de treillis complet

Soit $\mathcal{R} : \mathcal{I} \rightarrow [1, m]$ l'application bijective retournant le rang d'un vecteur de \mathcal{I} . Cette opérateur nécessite que \mathcal{I} soit un treillis complet, noté (\mathcal{I}, \leq) , avec \leq la relation d'ordre totale équipant \mathcal{I} . Les valeurs multivariées étant dépourvues de relation d'ordre naturelle, Lézoray [9] propose d'équiper \mathcal{I} d'une relation d'ordre \leq_h basée sur une application bijective $h : \mathcal{I} \rightarrow \mathcal{L}$ apprise à partir des valeurs de \mathcal{I} . Ainsi, pour $u, v \in \mathcal{I}$, $u \leq_h v \iff h(u) \leq h(v)$.

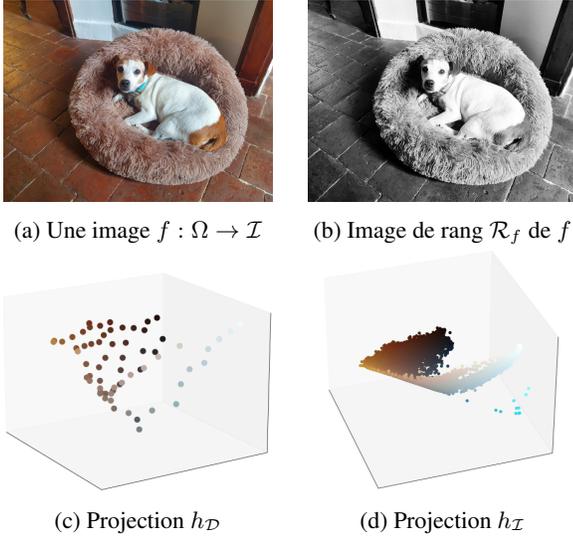


FIGURE 3 – Apprentissage de treillis complet

La création de l'application h s'effectue en utilisant un algorithme d'apprentissage de variété, le Laplacian Eigenmaps [2]. Cet algorithme ayant une complexité quadratique dépendant du nombre de valeurs contenues dans \mathcal{I} , ces valeurs sont au préalable échantillonnées dans un dictionnaire \mathcal{D} composé de $p \ll m$ valeurs en utilisant un algorithme de quantification vectorielle tel que les k -moyennes. Le Laplacian Eigenmaps est ensuite appliqué sur les valeurs de \mathcal{D} , résultant en une projection $h_{\mathcal{D}}$, illustrée sur la Figure 3c. Pour cela, la matrice de similarité $D_{\mathcal{D}}$ des valeurs de \mathcal{D} est calculée en utilisant un noyau gaussien, puis les vecteurs propres de sa matrice Laplacienne normalisée sont retournées, constituant une nouvelle représentation pour obtenir l'ensemble \mathcal{L} de la projection $h_{\mathcal{D}}$. Néanmoins, cette projection n'étant valable que pour les vecteurs du dictionnaire \mathcal{D} , elle nécessite d'être étendue aux valeurs de \mathcal{I} . Cela s'effectue en utilisant l'extension de Nyström, qui interpole les vecteurs propres de la matrice de similarité des éléments de \mathcal{I} , résultant en la projection $h_{\mathcal{I}}$, illustrée sur la Figure 3d. Enfin, pour un élément $x_j \in \mathcal{I}$, le vecteur retourné par la projection $h_{\mathcal{I}}$ est de la forme $(\phi_j^1, \phi_j^2, \dots, \phi_j^p)$ avec ϕ_j^i la j^{e} valeur du i^{e} vecteur propre.

Les vecteurs retournés par la projection $h_{\mathcal{I}}$ sont comparés lexicographiquement, permettant ainsi d'avoir le treillis complet (\mathcal{I}, \leq_h) et donc d'utiliser l'application \mathcal{R} . Appliquée à une image f , cette application retourne une image de rang \mathcal{R}_f , illustrée par la Figure 3b, qui est une représentation équivalente à une relation d'ordre totale. Ainsi, les opérations dépendant uniquement du rang d'une valeur sont appliquées sur \mathcal{R}_f .

3.2 Estimation de la FNB couleur

Imposer une relation d'ordre totale aux valeurs multivariées permet d'effectuer plusieurs opérations dépendant de l'ordre des pixels sur des images couleurs mais n'est pas la solution optimale dans tous les cas. En effet, Carlinet *et al.* [4] montrent

qu'utiliser la relation d'inclusion des formes entre les AdFs construits canal par canal est plus efficace et ils proposent un arbre des formes multivarié (AdFM) basé sur cette propriété. Ainsi, l'AdFM est utilisé pour obtenir l'ensemble des régions sur lesquelles le test statistique est effectué.

Le test statistique est calculée à partir des valeurs de la carte de rang obtenue par l'apprentissage de treillis complet. Néanmoins, contrairement à [6] où la division d'une région en séquences X et Y s'effectue aléatoirement, la relation d'adjacence utilisée par Sutour *et al.* [12] pour diviser les blocs est adaptée à la morphologie irrégulière des régions obtenues à partir de l'AdFM. Cette relation, illustrée par les Figures 2b à 2e, s'est avérée plus efficace que la division aléatoire dans le cas d'image couleur.

Enfin, les régions homogènes obtenues sont utilisées pour estimer la FNB. Néanmoins, n'étant valide que pour les images en niveau de gris sous sa forme quadratique, elle est étendue aux images couleurs sous la forme d'un triplet de FNB tel que $\text{FNBC}(x) = (\text{FNB}_r(x_r), \text{FNB}_g(x_g), \text{FNB}_b(x_b))$ où FNBC est la fonction de niveau de bruit couleur et $x = (x_r, x_g, x_b) \in \mathcal{I}$ la valeur d'un pixel couleur. Ainsi, estimer la FNBC consiste à résoudre le problème (3) sur chaque canal de l'image couleur.

4 Expérimentations et résultats

Pour rester concordant avec les expérimentations effectuées dans [12] et [6], nous avons étendu l'estimation par blocs développée par Sutour *et al.* [12] aux images couleurs, mais en utilisant un test statistique canal par canal, celui-ci résultant en de meilleures performances que l'utilisation de la carte de rang dans le contexte de la division de l'image en blocs. Pour évaluer l'efficacité de notre méthode, nous utilisons la base de 150 images de Laurent Condat¹, dont le bruit à l'acquisition est négligeable dû à la haute définition des images. Ces images sont bruitées 20 fois, avec des bruits différents dont les coefficients de la FNBC sont tirés aléatoirement à partir d'une distribution Gaussienne $\mathcal{N}(0.01, 0.003)$. Durant l'estimation, le niveau de détection α est fixé à 0.4 et les images sont simplifiées avec différents paramètres de la fonctionnelle λ fixés empiriquement à 200, 500 et 700. Chacune de ces simplifications réduit en moyenne le nombre de noeuds de l'AdFM de 57%, 63% et 64% respectivement. Enfin, les régions testées ont une taille minimum de 250 pixels. Pour comparer les résultats obtenus à partir des estimations, nous utilisons l'erreur relative moyenne (ERM) définie par :

$$\text{ERM} = \frac{1}{|\mathcal{I}|} \sum_{x \in \mathcal{I}} \frac{|\text{FNB}(x) - \widehat{\text{FNB}}(x)|}{\text{FNB}(x)} \quad (4)$$

avec $|\mathcal{I}|$ le cardinal de l'ensemble des valeur de l'image bruitée, $\widehat{\text{FNB}}$ la fonction résultant de l'estimation et FNB la fonction dont les coefficients ont été utilisés pour bruite l'image.

Les résultats de cette expérimentation, affichés en Figure 4, sont une comparaison des moyennes des ERM par images

1. <https://lcondat.github.io/imagebase.html>

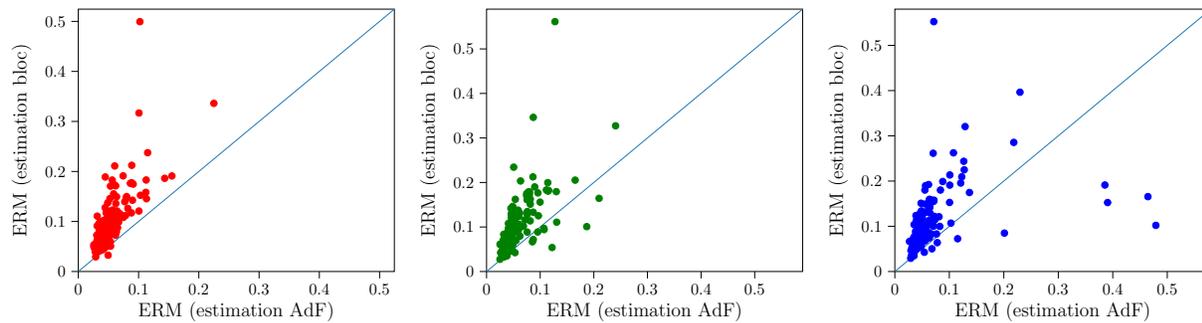


FIGURE 4 – Comparaison entre l’estimation par blocs et l’estimation utilisant les régions extraites de l’AdFM. Par soucis de lisibilité des figures, un point a été supprimé pour chaque canal, respectivement à (3.9, 0.9), (5.0, 1.0) et (2.1, 1.0)

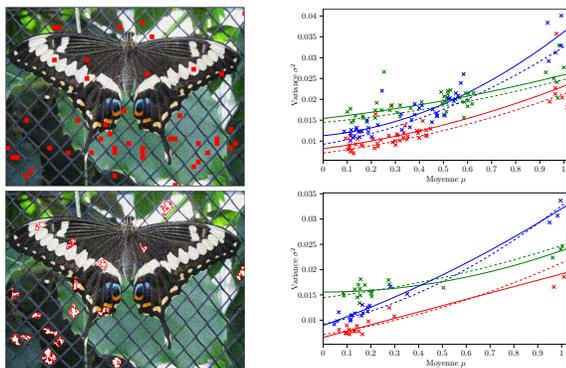


FIGURE 5 – Comparaison sur une image de l’estimation par bloc (1^{re} ligne) et l’estimation par régions (2^e ligne)

et par canal pour chaque type d’estimation. Ainsi, un point au-dessus de la ligne bleu indique que l’estimation proposée est en moyenne meilleure que l’estimation par blocs. Pour chaque canal, l’estimation proposée est plus efficace que l’estimation par blocs pour respectivement 147, 137 et 139 images, confirmant que la précision d’une estimation dépend de son adaptation au contenu de l’image.

Enfin, la Figure 5 illustre les résultats des estimations par blocs et par régions sur une image. On observe que les régions obtenues sont moins nombreuses que les blocs mais permettent d’obtenir une estimation plus fiable grâce à la précision des statistiques calculées sur les zones homogènes.

5 Conclusion

Dans cet article, nous avons étendu l’estimation de la fonction de niveau de bruit aux images couleurs en contournant l’absence de relation d’ordre totale due à la nature multivariée des valeurs d’une image couleur en utilisant l’AdFM et une carte de rang résultant d’un apprentissage de treillis complet. Nous avons confirmé les résultats obtenus dans [6] dans le cadre de l’estimation sur des images couleurs. En l’état, la méthode proposée n’utilise qu’un sous-ensemble des noeuds de la hiérarchie pour estimer le niveau de bruit. Dans le futur, nous allons utiliser une hiérarchie non-simplifiée et étudier d’autres types de hiérarchie, dont la construction ne dépend pas des valeurs de l’image mais de la dissimilarité entre ces valeurs.

De plus, il a été démontré [7] que dans le traitement du bruit sur des images couleurs, la corrélation de ce dernier sur les trois canaux (due aux méthodes d’acquisition) est un facteur important. Ainsi, nous prévoyons de prendre en compte cette corrélation dans nos futurs travaux.

Références

- [1] L BEAUREPAIRE, K CHEHDI et B VOZEL. “Identification of the nature of noise and estimation of its statistical parameters by analysis of local histograms”. In : *ICASSP*. T. 4. IEEE. 1997, p. 2805-2808.
- [2] M BELKIN et P NIYOGI. “Laplacian eigenmaps for dimensionality reduction and data representation”. In : *Neural computation* 15.6 (2003), p. 1373-1396.
- [3] A BUADES, B COLL et JM MOREL. “A review of image denoising algorithms, with a new one”. In : *Multiscale Modeling and Simulation : A SIAM Interdisciplinary Journal* 4.2 (2005), p. 490-530.
- [4] E CARLINET et T GÉRAUD. “MToS : A tree of shapes for multivariate images”. In : *TIP* 24.12 (2015), p. 5330-5342.
- [5] A CHAMBOLLE et T POCK. “A first-order primal-dual algorithm for convex problems with applications to imaging”. In : *JMIV* 40.1 (2011), p. 120-145.
- [6] B ESTEBAN, G TOCHON et T GÉRAUD. “Estimating the noise level function with the tree of shapes and non-parametric statistics”. In : *CAIP*. Springer. 2019, p. 377-388.
- [7] Q JIN, G FACCILO et JM MOREL. “A review of an old dilemma : Demosaicking first, or denoising first?”. In : *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*. 2020, p. 514-515.
- [8] M KENDALL. “The treatment of ties in ranking problems”. In : *Biometrika* 33.3 (1945), p. 239-251.
- [9] O LÉZORAY. “Complete lattice learning for multivariate mathematical morphology”. In : *JVCIR* 35 (2016), p. 220-235.
- [10] C LIU et al. “Noise estimation from a single image”. In : *CVPR 06*. T. 1. IEEE. 2006, p. 901-908.
- [11] P MONASSE et F GUICHARD. “Fast computation of a contrast-invariant image representation”. In : *TIP* 9.5 (2000), p. 860-872.
- [12] C SUTOUR, CA DELEDALLE et JF AUJOL. “Estimation of the noise level function based on a nonparametric detection of homogeneous image regions”. In : *SIAM Journal on Imaging Sciences* 8.4 (2015), p. 2622-2661.
- [13] Y XU, T GÉRAUD et L NAJMAN. “Salient level lines selection using the Mumford-Shah functional”. In : *ICIP*. IEEE. 2013, p. 1227-1231.