

Débruitage Non-Local Itératif fondé sur un Critère de Similarité Probabiliste

Charles-Alban DELEDALLE¹, Loïc DENIS², Florence TUPIN¹

¹Institut Télécom, Télécom ParisTech, CNRS LTCI
46, rue Barrault, 75634 Paris Cedex 13, France

²École Supérieure de Chimie Physique Électronique de Lyon et Laboratoire Hubert Curien, CNRS UMR 5516
Université Jean Monnet, 18 rue Pr Benoît Lauras, 42000 Saint-Etienne, France

Charles-Alban.Deledalle@telecom-paristech.fr, Loic.Denis@cpe.fr
Florence.Tupin@telecom-paristech.fr

Résumé – Une nouvelle méthodologie est proposée pour généraliser le filtre à moyennes non-locales de Buades *et al.* [2] aux bruits non-Gaussiens et/ou non-additifs. La similarité entre patches, exprimée par Buades *et al.* comme une distance quadratique, est ici définie à partir du modèle paramétrique du bruit considéré. L'image débruitée est obtenue par maximisation de la vraisemblance pondérée. Les poids sont affinés itérativement en exprimant la similarité entre patches bruités et entre les patches débruités obtenus à l'itération précédente. Le processus itératif améliore significativement la qualité du débruitage. Les résultats sont donnés sur trois images présentant des bruits intrinsèquement différents : un bruit additif Gaussien, un bruit impulsionnel et un bruit de speckle multiplicatif.

Abstract – A new methodology is proposed to generalize the non-local means algorithm introduced by Buades *et al.* [2] to non-additive and/or non-Gaussian noises. The patch similarity used by Buades *et al.* is substituted by a statistically-grounded similarity criterion that we derive from the noise distribution model. The denoised image is estimated by maximizing the weighted likelihood. The weights are iteratively refined by defining the similarity on both the noisy image and the previous estimated image. The iterative process noticeably improves the denoising performance. The results are given on three images corrupted by different noises: an additive Gaussian noise, an impulsive noise and a multiplicative speckle noise.

1 Introduction

En traitement des images, il est souvent nécessaire de débruitier les images puisque le bruit peut rendre difficile l'interprétation visuelle ou automatique de la scène imagée. Le bruit est généralement bien modélisé par des distributions paramétriques. Débruitier revient alors à estimer les valeurs de ces paramètres en tout point de l'espace. De nombreuses stratégies ont été proposées. La plupart considèrent un modèle de bruit blanc additif Gaussien. D'autres sont spécifiquement conçues pour traiter des bruits non-Gaussien et/ou non-additif. Peu d'entre elles fournissent une méthodologie générale qui s'applique à différents types de bruits. Leurs différences résident dans les hypothèses formulées sur la nature de la scène et/ou du bruit. La qualité du débruitage dépend de l'adéquation des hypothèses avec les images traitées. Récemment, Buades *et al.* ont introduit le filtre à moyennes non-locales [2] qui suppose qu'un motif quelconque de l'image se retrouve plusieurs fois dans celle-ci. Cette hypothèse semble bien appropriée aux images naturelles, puisque le filtre à moyennes non-locales offre de bonnes performances pour réduire le bruit blanc additif Gaussien. Cet article présente une approche non-locale plus générale, qui s'applique à différents types de bruits.

2 Max. de Vraisemblance Pondérée

Soit une image v bruitée. On suppose que le bruit est décorrélé en chaque pixel s et que la valeur associée v_s est la réalisation d'un phénomène aléatoire gouverné par une loi de probabilité connue de paramètre inconnu θ_s^* . Filtrer l'image v revient à trouver en chaque pixel s une estimation $\hat{\theta}_s$ de θ_s^* expliquant au mieux la valeur v_s .

Pour chaque pixel s , l'Estimateur au sens du Maximum de Vraisemblance (EMV) définit une estimation $\hat{\theta}_s$ des paramètres sous-jacents θ_s^* à partir d'un ensemble $\mathcal{S}_{\theta_s^*}$ de réalisations indépendantes et identiquement distribuées disponibles dans l'image :

$$\begin{aligned}\hat{\theta}_s^{(EMV)} &\triangleq \arg \max_{\theta} \sum_{t \in \mathcal{S}_{\theta_s^*}} \log p(v_t | \theta) \\ &= \arg \max_{\theta} \sum_t \delta_{\mathcal{S}_{\theta_s^*}}(t) \log p(v_t | \theta),\end{aligned}\quad (1)$$

avec $\delta_{\mathcal{S}_{\theta_s^*}}$ l'indicatrice de l'ensemble $\mathcal{S}_{\theta_s^*}$ (i.e., $\delta_{\mathcal{S}_{\theta_s^*}}(t) = 1$ si $t \in \mathcal{S}_{\theta_s^*}$, 0 sinon). L'EMV est non biaisé et asymptotiquement efficace. Malheureusement, on ne connaît pas l'ensemble $\mathcal{S}_{\theta_s^*}$ pour chaque site s . On est donc obligé de remplacer l'indicatrice $\delta_{\mathcal{S}_{\theta_s^*}}(t)$ par un poids $w(s, t) \geq 0$ approchant au mieux

$\delta_{\mathcal{S}_{\theta_s^*}}(t)$. On parle d'Estimation au sens du Maximum de Vraisemblance Pondérée (EMVP) définie comme suit :

$$\hat{\theta}_s^{(EMVP)} \triangleq \arg \max_{\theta} \sum_t w(s, t) \log p(v_t | \theta). \quad (2)$$

L'EMVP diminue l'erreur quadratique moyenne en réduisant la variance de l'estimation au détriment d'un biais introduit par les échantillons qui suivent une distribution de paramètre θ_t^* différent de θ_s^* [6]. L'EMVP a déjà été appliquée en traitement d'image par Polzehl et Spokoiny [9]. Etant donné ce compromis biais-variance, l'EMVP améliore les performances de l'EMV dès que les poids sont bien choisis, ce qui est l'objectif de la section suivante.

3 Poids Probabilistes et Non-Locaux

Une approche naïve, basée sur le principe de cohérence spatiale, définit les poids $w(s, t)$ localement, en fonction de la proximité des pixels s et t . Ceci produit un effet de flou au niveau des contours et des textures qui invalide l'hypothèse de cohérence spatiale. Par opposition aux approches locales et en s'inspirant de l'approche de synthèse de texture de Efros et Leung [5], Buades *et al.* proposent dans leur filtre à moyennes non-locales (ou NL means) de définir les poids $w(s, t)$ en fonction de la similarité entre les valeurs des voisinages de s et de t [2]. Cette similarité est définie par la distance Euclidienne entre deux patchs Δ_s et Δ_t (fenêtres rectangulaires centrées respectivement autour des pixels s et t). Cette définition est bien adaptée en présence d'un bruit blanc additif et Gaussien, mais peu appropriée à d'autres types de bruits.

Cet article cherche à définir une EMVP performante et adaptée à tout type de modèle statistique de bruit. Étant données les bonnes performances du filtre NL means, on s'inspire de la définition non-locale des poids $w(s, t)$ fondée sur la similarité entre les valeurs des fenêtres Δ_s et Δ_t . Rappelons que nous recherchons des poids $w(s, t)$ approchant l'indicatrice $\delta_{\mathcal{S}_{\theta_s^*}}(t)$ de l'ensemble $\mathcal{S}_{\theta_s^*} = \{t | \theta_t^* = \theta_s^*\}$. Ainsi, pour ne plus être restreint aux bruits blancs additifs Gaussien et dans une approche non-locale, nous proposons de définir les poids $w(s, t)$ directement en fonction de la probabilité de l'évènement $\theta_{\Delta_s}^* = \theta_{\Delta_t}^*$ étant donnée l'observation v :

$$w(s, t) \triangleq p(\theta_{\Delta_s}^* = \theta_{\Delta_t}^* | v)^{1/h}. \quad (3)$$

avec $\theta_{\Delta_s}^*$ et $\theta_{\Delta_t}^*$ les sous-images extraites de l'image de paramètres θ^* dans les fenêtres respectives Δ_s et Δ_t , et h une valeur réelle positive. Le paramètre h provient du filtre à moyennes non-locales et agit sur le taux de débruitage. Supposons maintenant qu'une estimation $\hat{\theta}^{i-1}$ de θ^* nous soit fournie. La probabilité de similarité des sites s et t peut alors être étendue en introduisant la connaissance de $\hat{\theta}^{i-1}$:

$$w(s, t) \triangleq p(\theta_{\Delta_s}^* = \theta_{\Delta_t}^* | v, \hat{\theta}^{i-1})^{1/h}. \quad (4)$$

Sous hypothèse d'indépendance, la probabilité de similarité $p(\theta_{\Delta_s}^* = \theta_{\Delta_t}^* | v, \hat{\theta}^{i-1})$ peut se réécrire comme le produit suivant $\prod_k p(\theta_{s,k}^* = \theta_{t,k}^* | v_{s,k}, v_{t,k}, \hat{\theta}^{i-1})$, où les indices s, k et

t, k réfèrent respectivement aux k -ièmes pixels dans les fenêtres Δ_s et Δ_t . Dans une approche Bayésienne, en l'absence de connaissances sur $p(v_{s,k}, v_{t,k} | \hat{\theta}^{i-1})$, et en supposant que l'évènement $v_{s,k}, v_{t,k} | \theta_{s,k}^* = \theta_{t,k}^*$ est indépendant de $\hat{\theta}^{i-1}$, on en déduit la décomposition suivante :

$$p(\theta_{s,k}^* = \theta_{t,k}^* | v_{s,k}, v_{t,k}, \hat{\theta}^{i-1}) \propto \underbrace{p(v_{s,k}, v_{t,k} | \theta_{s,k}^* = \theta_{t,k}^*)}_{\text{attache à l'image bruitée}} \times \underbrace{p(\theta_{s,k}^* = \theta_{t,k}^* | \hat{\theta}^{i-1})}_{\text{attache à l'estimation}}. \quad (5)$$

3.1 Attache à l'image bruitée

Le terme d'attache à l'image bruitée $p(v_{s,k}, v_{t,k} | \theta_{s,k}^* = \theta_{t,k}^*)$ mesure la vraisemblance que deux valeurs bruitées aient été générées par la même distribution de paramètre θ quelque soit cette valeur θ . En utilisant la règle de Bayes, ceci se traduit comme suit :

$$\frac{p(v_{s,k}, v_{t,k}, \theta_{s,k}^* = \theta_{t,k}^*)}{p(\theta_{s,k}^* = \theta_{t,k}^*)} = \frac{\int p(v_{s,k}, v_{t,k} | \theta_{s,k}^* = \theta, \theta_{t,k}^* = \theta) p(\theta_{s,k}^* = \theta, \theta_{t,k}^* = \theta) d\theta}{\int p(\theta_{s,k}^* = \theta, \theta_{t,k}^* = \theta) d\theta}$$

Sous hypothèse d'indépendance entre les valeurs bruitées $v_{s,k}$ et $v_{t,k}$, on obtient finalement

$$p(v_{s,k}, v_{t,k} | \theta_{s,k}^* = \theta_{t,k}^*) \propto \int_D p(v_{s,k} | \theta_{s,k}^* = \theta) p(v_{t,k} | \theta_{t,k}^* = \theta) d\theta \quad (6)$$

où D est le domaine de définition du paramètre θ . Le dénominateur $\int p(\theta_{s,k}^* = \theta, \theta_{t,k}^* = \theta) d\theta$ est considéré comme constant ainsi que la quantité $p(\theta_{s,k}^* = \theta, \theta_{t,k}^* = \theta)$. Lorsque D est un continuum comme \mathbb{R} , \mathbb{R}^+ ou \mathbb{C} , $p(\theta_{s,k}^* = \theta, \theta_{t,k}^* = \theta)$ correspond à une densité uniforme impropre.

3.2 Attache à l'estimation

Le terme d'attache à l'estimation mesure la validité de $\theta_{s,k}^* = \theta_{t,k}^*$ étant donnée l'estimation $\hat{\theta}^{i-1}$. On suppose que l'égalité $\theta_{s,k}^* = \theta_{t,k}^*$ est d'autant plus vérifiée que les distributions décrites par les paramètres $\hat{\theta}_{s,k}^{i-1}$ et $\hat{\theta}_{t,k}^{i-1}$ sont proches. Ce terme est donc fonction de la similarité entre ces deux distributions. Polzehl et Spokoiny utilisent la divergence de Kullback-Leibler entre les estimations comme un test statistique de l'hypothèse $\theta_{s,k}^* = \theta_{t,k}^*$ [9]. Nous proposons aussi d'utiliser une version symétrique de la divergence de Kullback-Leibler à travers une fonction exponentielle décroissante

$$p(\theta_{s,k}^* = \theta_{t,k}^* | \hat{\theta}^{i-1}) \propto \exp\left(-\frac{1}{T} \int_D \left(p(t | \hat{\theta}_{s,k}^{i-1}) - p(t | \hat{\theta}_{t,k}^{i-1})\right) \log \frac{p(t | \hat{\theta}_{s,k}^{i-1})}{p(t | \hat{\theta}_{t,k}^{i-1})} dt\right) \quad (7)$$

avec D le domaine des valeurs des pixels et T une valeur réelle positive. Cela correspond au noyau basé sur la divergence de Kullback-Leibler utilisé dans [8].

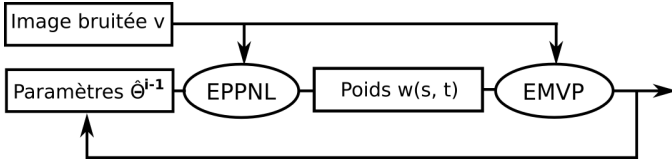


FIG. 1 – L’Estimateur des Poids Probabilistes et Non-Locaux (EPPNL) calcule les poids $w(s, t)$ à partir des images v et $\hat{\theta}^{i-1}$. L’Estimateur du Maximum de Vraisemblance Pondéré (EMVP) calcule les paramètres $\hat{\theta}^i$ en utilisant les poids $w(s, t)$ et l’image v . La procédure est répétée jusqu’à convergence.

La définition du filtre proposé est intrinsèquement itérative. A l’itération $i - 1$, l’ensemble des estimations $\hat{\theta}_s$ fournissent l’estimation $\hat{\theta}^{i-1}$ utilisée à l’itération i . En pratique, l’estimation converge rapidement vers un point fixe dépendant peu du choix de l’image de paramètre initial $\hat{\theta}^1$.

La Figure 1 illustre l’ensemble de la procédure :

- 1 les poids $w(s, t)$ sont estimés à partir de l’image bruitée v pour le terme d’attache à l’image bruitée (Équation 6) et à partir de l’estimation $\hat{\theta}^{i-1}$ pour le terme d’attache à l’estimation (Équation 7) ;
- 2 l’estimateur au sens du maximum de vraisemblance pondéré calcule les nouveaux paramètres $\hat{\theta}^i$ à partir des poids estimés $w(s, t)$ et de l’image bruitée v (Équation 2) ;
- 3 les étapes 1 et 2 sont répétées tant que des changements persistent entre deux itérations successives.

4 Algorithmes et Implémentation

Notre méthodologie a permis d’extraire trois filtres non-locaux adaptés respectivement au bruit additif Gaussien, au bruit impulsif, et au bruit de speckle multiplicatif. Les lois génératives des bruits utilisées sont respectivement :

$$p(v_s|\theta_s)^{Gauss.} \triangleq \frac{1}{\sqrt{2\pi}\sigma_G} \exp\left(-\frac{(v_s - \theta_s)^2}{2\sigma_G^2}\right) \quad (8)$$

$$p(v_s|\theta_s)^{Impuls.} \triangleq \frac{P}{N} + \frac{(1-P)}{Z_{\theta_s}} \exp\left(-\frac{(v_s - \theta_s)^2}{2\sigma_I^2}\right) \quad (9)$$

$$p(v_s|\theta_s)^{Speckle} \triangleq \frac{2L^L}{\Gamma(L)\theta_s^L} v_s^{2L-1} \exp\left(-\frac{Lv_s^2}{\theta_s}\right) \quad (10)$$

où σ_G est l’écart-type du bruit Gaussien, N le nombre de niveaux de gris de l’image dans le cas du bruit impulsif, σ_I l’écart-type de la composante Gaussienne du modèle de bruit impulsif, Z_{θ_s} une constante de normalisation et L le nombre de vues équivalent du bruit de speckle.

A partir des équations 2 à 7, nous avons dérivé les trois filtres adaptés. Dans le cas du bruit Gaussien et du bruit de speckle, toutes ces équations s’expriment analytiquement [4] et sont données dans le Tableau 1. Dans le cas du bruit impulsif, les quantités sont calculées numériquement : le maximum de vraisemblance est recherché exhaustivement parmi les N niveaux de gris possibles et les mesures de similarités sont calculées au préalable de façon exacte pour chaque couple de valeurs et stockées dans une matrice de taille $N \times N$.

	Gaussien	Speckle
Eq. 2	$\frac{\sum_t w(s,t)v_t}{\sum_t w(s,t)}$	$\frac{\sum_t w(s,t)v_t^2}{\sum_t w(s,t)} 2^{L-1}$
Eq. 6	$\exp\left(-\frac{ v_{s,k} - v_{t,k} ^2}{4\sigma^2}\right)$	$\left(\frac{v_{s,k}v_{t,k}}{v_{s,k}^2 + v_{t,k}^2}\right)^{2L-1}$
Eq. 7	$\exp\left(-\frac{ \hat{\theta}_{s,k}^{i-1} - \hat{\theta}_{t,k}^{i-1} ^2}{T\sigma^2}\right)$	$\exp\left(-\frac{L}{T} \frac{ \hat{\theta}_{s,k}^{i-1} - \hat{\theta}_{t,k}^{i-1} ^2}{\hat{\theta}_{s,k}^{i-1}\hat{\theta}_{t,k}^{i-1}}\right)$

TAB. 1 – Formes analytiques du maximum de vraisemblance (Eq. 2), de l’attache à l’image bruitée (Eq. 6) et de l’attache à l’estimation (Eq. 7) pour les bruits Gaussien et de Speckle.

5 Expériences et Résultats

Les trois filtres obtenus ont été appliqués sur trois images présentées sur la Figure 2. La première est l’image *Maison* dégradée par un bruit additif Gaussien d’écart-type $\sigma_G = 40$, la seconde est l’image *Avion* dégradée par un bruit impulsif sur $N = 256$ niveaux avec un taux $P = 0.5$ et une contribution Gaussienne $\sigma_I = 10$ et la troisième est l’image radar *CNES* de Toulouse acquise par le système TerraSAR-X et présentant un bruit de speckle multiplicatif mono-vue ($L = 1$). Pour chaque type de bruit, notre méthode a été comparée avec un filtre récent : le filtre BM3D [3] pour le bruit additif Gaussien, le filtre MAP-UWD-S [1] pour le bruit de speckle et par minimisation exacte de la variation totale sur un champs aléatoire Markovien pour le bruit impulsif [10]. Comme dans [2], notre méthode a été appliquée sur un voisinage 21×21 avec des fenêtres 7×7 . Nous utilisons 25 itérations à partir d’une image constante. Le paramètre de filtrage h est fonction du 92-ième quantile de la distribution de notre critère [7]. Le paramètre $T = 0.2$ a été déterminé empiriquement. Dans les trois cas, l’image est bien homogénéisée et les contours bien conservés. L’information restituée est bien estimée sans introduire d’artefacts. Un défaut du filtre, commun à de nombreuses approches, est la suppression des détails fins dans l’image régularisée. Les valeurs du rapport signal sur bruit (SNR), connues pour les images synthétiques *Maison* et *Avion*, sont comparables à celles obtenues avec les autres méthodes.

6 Conclusion

Une méthodologie générale a été proposée pour le débruitage d’images qui peut être adaptée à différents modèles de distribution du bruit. Cette méthode se base sur le filtre à moyennes non-locales [2]. Une procédure itérative est proposée pour améliorer la qualité du débruitage. Le bruit présent dans l’image originale est bien atténué dans les régions homogènes et les contours des objets sont bien conservés (préservation de la résolution). Le filtre offre un cadre général pour l’estimation de paramètres des images lorsqu’un modèle de bruit décorréolé est fourni. Finalement, l’élaboration du filtre proposé a permis de définir un nouveau critère de similarité entre patches d’images qui pourrait être ré-utilisé dans d’autres applications de télédétection.

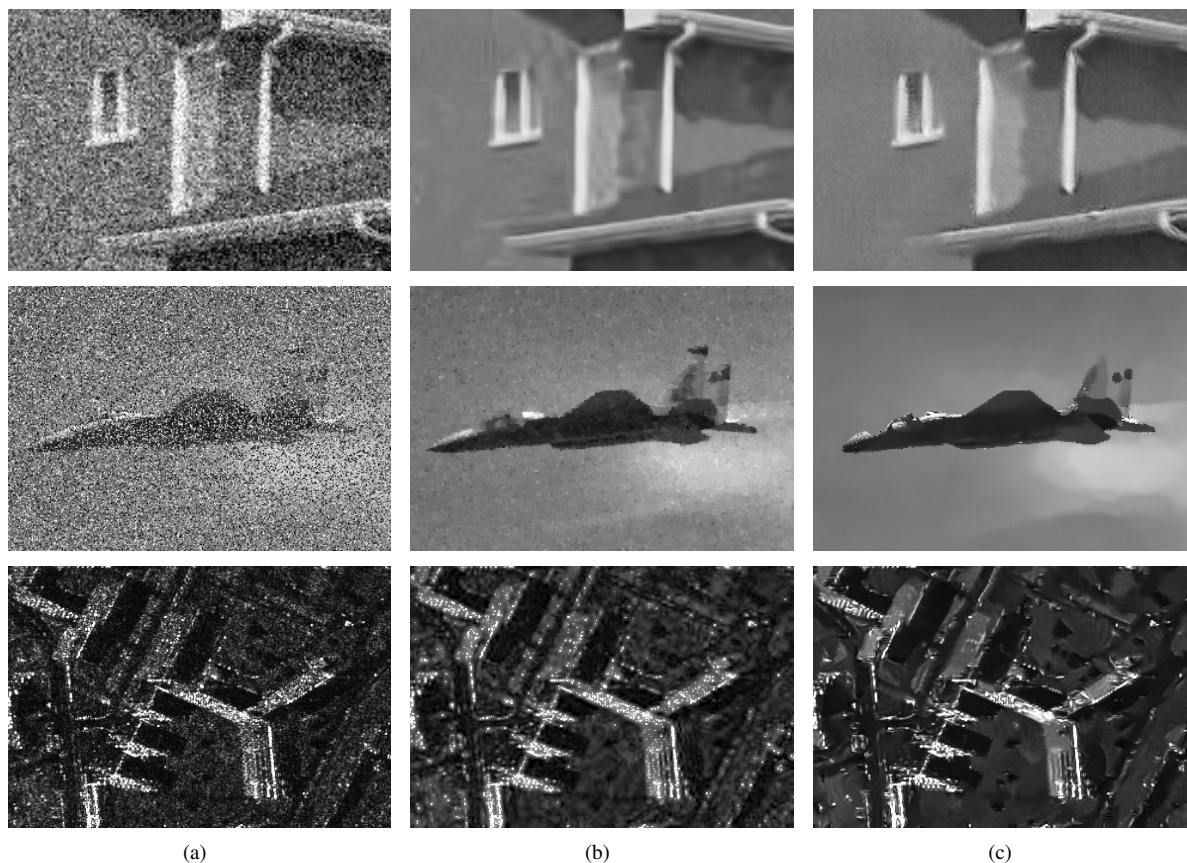


FIG. 2 – Première ligne : (a) l'image *Maison* dégradée par un bruit additif Gaussien (SNR 1.45), (b) débruitée par l'algorithme BM3D (SNR 15.78) et (c) débruitée par le filtre proposé (SNR 14.20). Deuxième ligne : (a) l'image *Avion* dégradée par un bruit impulsif (SNR -4.59), (b) débruitée par minimisation exacte de la variation totale sur un champs aléatoire Markovien (SNR 9.55) et (c) débruitée par le filtre proposé (SNR 9.77). Troisième ligne : (a) l'image radar *CNES* présentant un bruit de speckle multiplicatif mono-voie ©DLR, (b) débruitée par le filtre MAP-UWD-S et (c) débruitée par le filtre proposé.

Références

- [1] T. Bianchi, F. Argenti, and L. Alparone. Segmentation-Based MAP Despeckling of SAR Images in the Undecimated Wavelet Domain. *IEEE Transactions on Geoscience and Remote Sensing*, 46(9) :2728–2742, 2008.
- [2] A. Buades, B. Coll, and J.M. Morel. A Non-Local Algorithm for Image Denoising. *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*, 2, 2005.
- [3] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian. Image denoising by sparse 3-D transform-domain collaborative filtering. *IEEE Transactions on image processing*, 16(8) :2080, 2007.
- [4] C. Deledalle, L. Denis, and F. Tupin. Débruitage par maximum de vraisemblance pondérée par une méthode non-locale itérative et probabiliste. Technical Report 2009D004, Telecom ParisTech, 2009. <http://www.tsi.enst.fr/~deledall/files/ppb.pdf>.
- [5] AA Efros and TK Leung. Texture synthesis by non-parametric sampling. In *Computer Vision, 1999. The Proceedings of the Seventh IEEE International Conference on*, volume 2, 1999.
- [6] J. Fan, M. Farnen, and I. Gijbels. Local maximum likelihood estimation and inference. *Journal of the Royal Statistical Society. Series B, Statistical Methodology*, pages 591–608, 1998.
- [7] C. Kervrann and J. Boulanger. Local Adaptivity to Variable Smoothness for Exemplar-Based Image Regularization and Representation. *International Journal of Computer Vision*, 79(1) :45–69, 2008.
- [8] P.J. Moreno, P. Ho, and N. Vasconcelos. A Kullback-Leibler divergence based kernel for SVM classification in multimedia applications. *Advances in Neural Information Processing Systems*, 16, 2004.
- [9] J. Polzehl and V. Spokoiny. Propagation-separation approach for local likelihood estimation. *Probability Theory and Related Fields*, 135(3) :335–362, 2006.
- [10] L. Rudin, S. Osher, and E. Fatemi. Nonlinear total variation based noise removal algorithms. *Physica D*, 60(1-4) :259–268, 1992.