

Étude du centrage des patches et modélisation de leur moyenne pour le débruitage d'images

Antoine HOUDARD

CNRS, Univ. Bordeaux, IMB,
F-33400 Talence, France

antoine.houdard@math.u-bordeaux.fr

Résumé – Dans le contexte du débruitage d'images par patches, avec un modèle de bruit blanc gaussien, ce travail propose la dérivation des modèles de bruit induits sur les patches centrés et leur moyenne. Nous traitons ensuite séparément ces deux composantes. Enfin, nous proposons des expériences utilisant la méthode HDMI [3] qui illustrent l'amélioration que procure cette approche, notamment pour le bruit résiduel basse fréquence.

Abstract – This work is devoted to patch-based image denoising with additive white Gaussian noise. We derive corresponding models on centered patches and on their mean component. Then we treat the two components separately. Finally, we provide experiments with the HDMI method [3] showing denoising improvement, particularly for residual low frequency noise.

1 Introduction

Dans le contexte du débruitage d'images par patches, une image est représentée par l'ensemble de tous ses patches. Ces derniers sont extraits avec recouvrement et sont vus comme des vecteurs de taille $p = s \times s$. On considère sur ces patches un modèle de bruit blanc gaussien donné par

$$Y_i = x_i + N_i, \quad (1)$$

où $Y_i \in \mathbf{R}^p$ est le vecteur aléatoire modélisant le i^e patch, $x_i \in \mathbf{R}^p$ le patch sans bruit sous-jacent et $N_i \sim \mathcal{N}(0, \sigma^2 \mathbf{I}_p)$ un bruit blanc gaussien. Une partie des méthodes de la littérature utilise une modélisation *a priori* des patches de l'image pour résoudre ce problème [4, 6, 8, 5, 3]. L'idée est de modéliser le patch sans bruit x_i par un vecteur aléatoire X_i . Le modèle se réécrit alors

$$Y_i = X_i + N_i, \quad (2)$$

et le théorème de Bayes permet d'obtenir un estimateur du patch x_i par l'espérance conditionnelle

$$\hat{x}_i = \mathbb{E}[X_i | Y_i = y_i]. \quad (3)$$

Les modèles *a priori* gaussiens [4] ou de type mélange de gaussiennes [6, 5, 3] permettent de calculer (3) explicitement et ont été largement utilisés dans la littérature. L'étude de ces modèles montre qu'ils sont capables d'encoder, à travers la matrice de covariance, des structures locales à différents niveaux de contraste [2]. Cette invariance au contraste permet le regroupement de plus de patches représentant la même structure et permet ainsi une meilleure estimation des paramètres du modèle. En contrepartie, le paramètre de moyenne du modèle devient très peu informatif. L'estimateur de chaque patch est alors tiré vers le patch moyen, ce qui produit un bruit basse fréquence

résiduel. Ce phénomène apparaît généralement dans le résultat de méthodes de débruitage par patches comme l'illustre la figure 1 (b) dans le cas de la méthode HDMI [3] pour des patches de taille 7×7 . Une partie de ce bruit basse fréquence est due à une mauvaise estimation de la moyenne de chaque patch. En effet, dans la figure 1, l'image (d) formée des moyennes de chaque patch de l'images bruitée (a) présente les mêmes motifs de bruit que l'image (c) formées des moyennes des patches de l'image débruitée (b). Ainsi, lorsque l'on corrige la moyenne de chaque patch débruité de l'images (b) par sa valeur oracle (f), le débruitage (c) obtenu est nettement amélioré. Il est donc intéressant de chercher à débruiter l'image (e) séparément. C'est à ce problème de débruitage des moyennes des patches que nous nous intéressons ici. De plus, certaines méthodes de la littérature [5, 8] suggèrent de retirer la moyenne de chaque patch afin d'apprendre le modèle sur des données centrées, nous discutons également ce point.

Contribution Nous proposons d'étudier la décomposition du problème de débruitage (2) en deux. Pour cela, nous définissons le vecteur aléatoire $Y_i^c = Y_i - \bar{Y}_i \mathbf{1}_p$ modélisant les patches centrés observés, où

$$\bar{Y}_i = \frac{1}{p} \sum_{j=1}^p Y_i(j), \quad (4)$$

est la moyenne de Y_i et $\mathbf{1}_p = (1, \dots, 1) \in \mathbf{R}^p$. Le problème (2) se décompose en deux problèmes

$$\bar{Y}_i = \bar{X}_i + \bar{N}_i \in \mathbf{R}, \quad (5)$$

$$\text{et } Y_i^c = X_i^c + N_i^c \in \mathbf{R}^p. \quad (6)$$

Dans ce travail, nous déduisons du modèle de bruit (2) les modèles des composantes de bruit N_i^c (section 2) et \bar{N}_i (section 3)

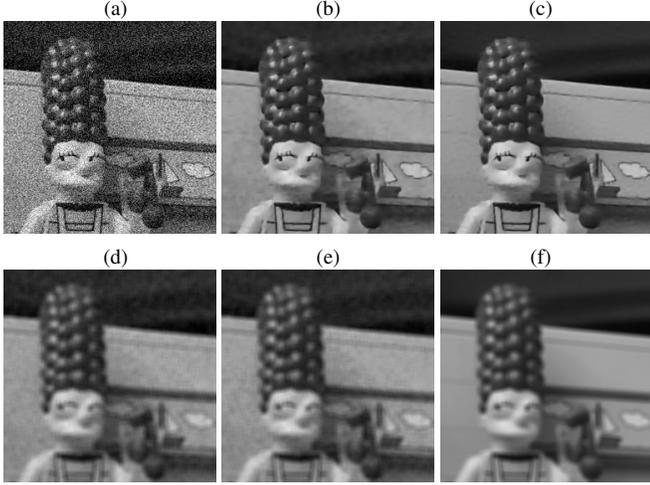


FIGURE 1 – (a) Image bruitée (écart-type 30/255). (b) Image débruitée avec HDMI [3] (patches 7×7), $PSNR$ 31.92 dB. (c) Image obtenue à partir des patches de (b) dont la moyenne a été corrigée par la valeur oracle (f), $PSNR$ 33.78 dB. (d) Image des moyennes des patches de (a). (e) Image des moyennes des patches de (b). (f) Image des moyennes de l’image oracle.

de ces deux problèmes. Puis, nous proposons en section 4 d’utiliser cette décomposition afin d’améliorer le résultat des méthodes qui débruitent patch par patch comme [4, 5, 3]. Enfin, nous présentons des expériences numériques illustrant les résultats obtenus dans le cas de la méthode HDMI [3].

2 Modélisation du bruit centré

Le bruit centré est défini par $N_i^c = N_i - \bar{N}_i \mathbf{1}_p$. Puisque que N_i est un vecteur gaussien, N_i^c est également un vecteur gaussien dont chaque composante de sa moyenne est

$$\mathbb{E}[N_i^c(j)] = \mathbb{E}[N_i(j)] - \frac{1}{p} \sum_{k=1}^p \mathbb{E}[N_j(k)] = 0, \quad (7)$$

et pour sa matrice de covariance, $\forall k, l \in \{1, \dots, p\}$

$$\mathbb{E}[N_i^c(k)N_i^c(l)] = \sigma^2 \left(\delta_{kl} - \frac{1}{p} \right). \quad (8)$$

Finalement, $N_i^c \sim \mathcal{N}(0, \Sigma_{N_i^c})$ avec $\Sigma_{N_i^c} = \frac{\sigma^2}{p} (p\mathbf{I}_p - \mathbf{J}_p)$, où \mathbf{J}_p est la matrice de taille p constituée de 1. Puisque $\Sigma_{N_i^c}$ est symétrique réelle, elle est diagonalisable en base orthonormale. Sachant que ses valeurs propres sont p (de multiplicité $p-1$) et 0, nous pouvons définir une matrice orthogonale Q telle que

$$\Sigma_{N_i^c} = Q \begin{pmatrix} \sigma^2 \mathbf{I}_{p-1} & 0 \\ 0 & 0 \end{pmatrix} Q^T. \quad (9)$$

Le changement de base Q^T appliqué au bruit centré donne $Q^T N_i^c \sim \mathcal{N}(0, \text{diag}(\sigma^2 \mathbf{I}_{p-1}, 0))$. Finalement, travailler sur le bruit centré permet de réduire le problème d’une dimension et la variance totale du bruit centré $\text{TVar}(N_i^c) := \mathbb{E}[\|N_i^c\|_2^2] = \frac{p-1}{p} \text{TVar}(N_i)$ est réduite d’un facteur $(p-1)/p$.

3 Modélisation des moyennes

Puisque N_i est un vecteur gaussien, sa moyenne est une variable aléatoire gaussienne $\bar{N}_i \sim \mathcal{N}(0, \frac{\sigma^2}{p})$. Ainsi, en réorganisant le problème (5) sous la forme d’une image, nous pouvons le récrire comme le problème de débruitage d’image

$$\bar{Y} = \bar{X} + \bar{N}, \quad (10)$$

où \bar{Y} , \bar{X} et \bar{N} sont des images de \mathbf{R}^n dont les valeurs au pixel i sont \bar{Y}_i , \bar{X}_i et \bar{N}_i . La seule différence ici est que le bruit additif gaussien est coloré. En effet, si l’on considère deux pixels de bruit \bar{N}_i et \bar{N}_j dans la même zone de $s \times s$ pixels, ils sont issus de deux patches qui se chevauchent et ne sont donc pas indépendants.

Cependant, il est toujours possible de se ramener à un problème de bruit blanc gaussien. En effet, considérons des patches de taille $p = s \times s$ de cette nouvelle image en définissant $Z_i = \pi_i(\bar{Y})$, $W_i = \pi_i(\bar{X})$ et $M_i = \pi_i(\bar{N})$, où π_i est l’opérateur d’extraction du i^e patch. Le modèle sur ces patches s’écrit

$$Z_i = W_i + M_i, \quad (11)$$

avec M_i le vecteur aléatoire modélisant le bruit. Puisque $M_i = (\bar{N}_{i_1}, \dots, \bar{N}_{i_p})$, toutes ses composantes sont des combinaisons linéaires de variables aléatoires gaussiennes, ainsi M_i est un vecteur gaussien. Afin de déterminer sa loi, calculons sa moyenne et sa matrice de covariance.

La moyenne de M_i est trivialement 0_p et les coefficients de la matrice de covariance Σ_{M_i} sont donnés par

$$(\Sigma_{M_i})_{kl} = \mathbb{E}[\bar{N}_{i_k}, \bar{N}_{i_l}] = \frac{\sigma^2}{p^2} C_{kl}, \quad (12)$$

où C_{kl} est le nombre de pixels en commun entre deux patches de l’image d’origine desquels \bar{N}_{i_k} et \bar{N}_{i_l} sont issus. Après dénombrement, on obtient

$$\Sigma_{M_i} = \frac{\sigma^2}{p^2} B \otimes B, \quad (13)$$

où

$$B = \begin{pmatrix} s & (s-1) & \dots & 1 \\ (s-1) & s & \ddots & \vdots \\ \vdots & \ddots & \ddots & (s-1) \\ 1 & \dots & (s-1) & s \end{pmatrix}, \quad (14)$$

et \otimes est le produit de Kronecker. Afin d’utiliser une méthode de débruitage qui a été conçue pour un modèle de bruit blanc gaussien, exprimons les données dans une base qui rende le bruit blanc. Pour cela, nous cherchons une décomposition $\Sigma_{M_i} = LL^T$. Une telle décomposition existe de façon unique si Σ_{M_i} est symétrique définie positive. Pour le montrer, commençons par montrer que B est symétrique définie positive. En utilisant le critère de Sylvester, il suffit de montrer que tous les mineurs principaux de B sont positifs. Ces mineurs de taille $d \in \{1, \dots, s\}$ sont donnés par

$$m_d = \begin{vmatrix} s & (s-1) & \dots & (s-d+1) \\ (s-1) & s & \ddots & \vdots \\ \vdots & \ddots & \ddots & (s-1) \\ (s-d+1) & \dots & (s-1) & s \end{vmatrix}. \quad (15)$$

En ajoutant la première colonne à la dernière, puis en soustrayant la deuxième et la dernière colonne à la première, on obtient

$$m_d = (2s - d + 1) \begin{vmatrix} 2 & (s-1) & \cdots & 1 \\ 0 & s & \ddots & \vdots \\ \vdots & \ddots & \ddots & 1 \\ 0 & \cdots & (s-1) & 1 \end{vmatrix}. \quad (16)$$

Enfin, le développement du déterminant par rapport à la première colonne donne

$$m_d = (2s - d + 1)2^{d-2} > 0. \quad (17)$$

Ainsi, B est définie positive et $B \otimes B$ aussi en tant que produit de Kronecker de deux matrices définies positives. La décomposition de Cholesky donne donc une matrice inversible L telle que $B \otimes B = LL^T$. Par conséquent, le problème

$$L^{-1}Z_i = L^{-1}W_i + L^{-1}M_i, \quad (18)$$

est un problème de débruitage avec un bruit blanc gaussien de variance σ^2/p^2I_p . Ainsi, une méthode de débruitage par patch telle que la méthode HDMI peut être utilisée pour trouver un estimateur $\widehat{L^{-1}W_i}$, et donc un estimateur de W_i donné par $\widehat{W_i} = LL^{-1}\widehat{W_i}$.

4 Expériences numériques

Dans cette partie, nous proposons d'utiliser la modélisation du problème de bruit sur les moyennes des patches afin d'améliorer la qualité de débruitage des méthodes de débruitage patch à patch. Pour les expériences, nous proposons d'utiliser la méthode de débruitage HDMI [3] dont le principe est le suivant :

- les patches non bruités sont modélisés par un GMM avec de faibles dimensions intrinsèque ;
- les paramètres de ce modèle sont inférés par un algorithme d'espérance-maximisation (EM) ;
- les patches débruités sont estimés par l'espérance conditionnelle (3).

Nous proposons des expériences pour des images couleur ainsi qu'une brève comparaison des résultats avec l'état-de-l'art.

4.1 Débruitage des moyennes

Dans la section 3, nous avons proposé de réorganiser les moyennes des patches en une nouvelle image afin d'appliquer une méthode de débruitage par patches à cette dernière. Le modèle de bruit sur les patches de cette image est blanc gaussien dans la base L , par conséquent, nous pouvons directement appliquer une méthode de débruitage par patches au problème (18) afin de résoudre le problème des moyennes (5). La figure 2 montre le résultat du débruitage de l'image des moyennes des patches en utilisant la méthode HDMI [3] avec des patches de taille 7×7 . Les résultats obtenus sont très bon, aussi bien visuellement qu'en terme de PSNR. Cela n'est pas étonnant puisque le problème (18) est plus simple que le problème d'origine (2) au sens où le rapport signal sur bruit du problème sur



FIGURE 2 – Gauche : image des moyennes bruitée (écart-type 50/255). Milieu : image des moyennes débruitée PSNR 38.58 dB. Droite : image des moyennes oracle.

les moyennes est réduit d'un facteur p par rapport au problème de départ.

Finalement, cette étape permet d'obtenir pour chaque patch $i \in \{1, \dots, n\}$ de taille $s \times s$ d'une image u à n pixels, un bon estimateur $\widehat{X_i}$ de sa moyenne comme suit :

- construction d'une image U dont chaque pixel i est la moyenne du patch \bar{Y}_i ;
- extraction des patches Z_i de M de taille $s \times s$;
- débruitage des patches transformés $L^{-1}Z_i$ par un opérateur de débruitage $f_{denoise}$;
- estimation des patches débruités $\widehat{W_i} = L f_{denoise}(L^{-1}Z_i)$, puis de l'image \widehat{U} ;
- l'estimateur de la moyenne de chaque patch i est alors donné par $\widehat{X_i} = \widehat{U}_i$.

4.2 Débruitage des patches

Afin d'obtenir un débruitage final de chaque patch, et donc de l'image, nous avons étudié deux stratégies :

1. débruiter les patches centrés du problème (6) puis ajouter l'estimateur des moyennes de la section 4.1 ;
2. débruiter les patches originaux de (2), retirer leur moyenne *a posteriori* puis ajouter l'estimateur des moyennes de la section 4.1.

Les résultats obtenus par la première stratégie étaient toujours moins bon que pour la seconde : les zones constante étaient aussi bien débruitées, mais les parties complexe comme les textures devenaient plus floues avec la première stratégie. Cela s'explique par le fait que le rapport signal sur bruit du problème sur les patches centrés (6) est plus faible que celui sur les patches originaux (2) (puisque le centrage des patches à réduit leur dynamique). Dans cette section, nous nous intéressons donc à l'amélioration du résultat de débruitage par la correction *a posteriori* des moyennes des patches que nous détaillons ci-après.

Amélioration par correction des moyennes *a posteriori* La méthode proposée ici pour améliorer le débruitage peut être utilisée avec n'importe quel opérateur de débruitage $f_{denoise}$ pour bruit blanc gaussien. Les grandes étapes en sont les suivantes :

1. obtenir un estimateur $\widehat{X_i} = f_{denoise}(Y_i)$ de chaque patch de l'image ;
2. obtenir un estimateur $\widehat{X_i}$ de la moyenne de chaque patch (section 4.1) ;

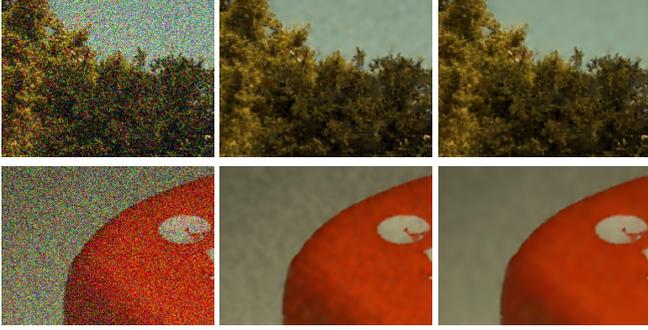


FIGURE 3 – De gauche à droite. images bruitées (écart-type 50/255). Images débruitées par HDMI (haut 26.96dB, bas 34.62dB). Images améliorées par correction des moyennes (haut 27.08dB, bas 36.06dB).

3. corriger chaque patch par la nouvelle estimation de sa moyenne :

$$S(X_i) = \widehat{X}_i - \widehat{X}_i \mathbf{1}_p + \widehat{X}_i \mathbf{1}_p. \quad (19)$$

Cette stratégie permet d’améliorer la qualité du débruitage par rapport à la méthode d’origine. La figure 3 illustre cette amélioration avec $f_{denoise} = \text{HDMI}$ pour deux images. Le bruit basse fréquence est réduit visuellement et les résultats sont également meilleurs en terme de PSNR, +0.12dB pour l’image du haut, plus texturée, et +1.4dB pour l’image du bas, plus lisse.

4.3 Discussion

La table 1 présente les résultats obtenus pour différentes tailles de patch, différents niveaux de bruit et trois images (figure 4). Notre approche permet systématiquement d’améliorer la qualité du débruitage, avec une amélioration nettement plus significative pour les images les moins texturées. Cette amélioration permet à la méthode HDMI de se rapprocher des performances de la méthode état-de-l’art FFDnet [7] qui utilise de l’apprentissage profond et surpasse la méthode par patches BM3D [1]. Une autre tendance qui apparaît est que l’amélioration est plus nette dans le cas d’une forte variance que dans le cas d’une faible variance. Le bruit basse-fréquence résiduel étant proportionnel à la variance, celui-ci est bien réduit. Enfin, on constate que l’amélioration est d’autant plus grande que les patches sont petits. Ceci s’explique par le fait que la variance du bruit de la moyenne des patches est de l’ordre de σ^2/p où p est le nombre de pixels du patch. Ainsi, pour des grands patches, le bruit dans la moyenne est déjà faible et le corriger améliore peu le résultat final.



FIGURE 4 – Images utilisées : Traffic, Dice et Flower.

TABLE 1 – Résultats en PSNR (dB). HDMI_s indique que l’algorithme a été utilisé avec des patches $s \times s$, basic correspond à HDMI [3], new correspond à l’amélioration proposée et diff. indique la différence. Les résultats pour BM3D [1] et FFDnet [7] sont indiqués à titre de comparaison.

	σ	HDMI ₅			HDMI ₇			HDMI ₁₀			BM3D	FFDnet
		basic	new	diff.	basic	new	diff.	basic	new	diff.		
Traffic	20	31.42	31.53	+0.11	31.48	31.53	+0.05	31.29	31.31	+0.02	30.81	31.74
	30	29.24	29.40	+0.16	29.39	29.46	+0.07	29.25	29.28	+0.03	28.83	29.79
	40	27.74	27.95	+0.21	27.99	28.09	+0.10	27.94	27.98	+0.04	27.45	28.48
	50	26.60	26.86	+0.26	26.97	27.07	+0.10	26.97	27.02	+0.05	26.43	27.52
Dice	20	38.51	40.00	+1.49	39.76	40.74	+0.98	40.49	41.01	+0.52	39.98	41.06
	30	36.08	37.79	+1.71	37.48	38.76	+1.28	38.55	39.30	+0.75	38.01	39.36
	40	34.33	36.14	+1.81	35.90	37.21	+1.31	37.01	37.95	+0.94	36.52	38.01
	50	32.98	34.88	+1.90	34.61	36.06	+1.45	35.89	36.85	+0.96	35.19	36.72
Flower	20	36.25	36.71	+0.46	36.76	36.95	+0.19	36.81	36.85	+0.04	35.89	37.19
	30	33.91	34.53	+0.62	34.58	34.86	+0.28	34.73	34.81	+0.08	33.74	35.18
	40	32.22	32.94	+0.72	33.00	33.36	+0.36	33.26	33.37	+0.11	32.13	33.73
	50	30.90	31.74	+0.84	31.76	32.16	+0.40	32.11	32.27	+0.16	30.94	32.51
Moyenne	20	35.39	36.08	+0.69	36.00	36.41	+0.41	36.20	36.39	+0.19	35.56	36.66
	30	33.08	33.91	+0.83	33.82	34.36	+0.54	34.18	34.46	+0.29	33.53	34.78
	40	31.43	32.34	+0.91	32.30	32.89	+0.59	32.74	33.10	+0.36	32.03	33.41
	50	30.16	31.16	+1.00	31.11	31.76	+0.65	31.66	32.05	+0.39	30.85	32.25

5 Conclusion

Dans ce travail, nous avons étudié la décomposition du problème de bruit blanc gaussien sur les patches en un problème sur les patches centrés et un problème sur les moyennes. Nous avons proposé une modélisation des patches bruités centrés et une modélisation des moyennes des patches bruités. Ces modélisations nous ont menés à une stratégie d’amélioration de la qualité du débruitage lorsque l’on dispose d’un débruiteur $f_{denoise}$ pour les patches. La stratégie proposée permet une amélioration du débruitage de la méthode HDMI utilisée ici pour les expériences, notamment pour la réduction du bruit basse fréquence.

Par la suite, nous aimerions explorer les liens entre cette approche et les *frameworks* multi-échelles.

Références

- [1] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian. Image denoising by sparse 3-d transform-domain collaborative filtering. *Image Processing, IEEE Transactions*, 2007.
- [2] J. Delon and A. Houdard. *Gaussian Priors for Image Denoising*, chapter , pages 125–149. Springer International Publishing, Cham, 2018.
- [3] A. Houdard, C. Bouveyron, and J. Delon. High-dimensional mixture models for unsupervised image denoising (hdmi). *SIAM Journal on Imaging Sciences*, 11(4) :2815–2846, 2018.
- [4] M. Lebrun, A. Buades, and J. M. Morel. A Nonlocal Bayesian Image Denoising Algorithm. *SIAM J. Imaging Sci.*, 2013.
- [5] A. M. Teodoro, M. S. Almeida, and M. A. Figueiredo. Single-frame image denoising and inpainting using gaussian mixtures. In *ICPRAM (2)*, pages 283–288, 2015.
- [6] Y.-Q. Wang and J.-M. Morel. SURE Guided Gaussian Mixture Image Denoising. *SIAM J. Imaging Sci.*, 2013.
- [7] K. Zhang, W. Zuo, and L. Zhang. Ffdnet : Toward a fast and flexible solution for cnn based image denoising. *IEEE Transactions on Image Processing*, 2018.
- [8] D. Zoran and Y. Weiss. From learning models of natural image patches to whole image restoration. In *2011 Int. Conf. Comput. Vis.*, pages 479–486. 2011.