

Débruitage de données de Poisson par BeamCurvelets 3-D

Arnaud WOISSELLE¹, Jean-Luc STARCK², Jalal FADILI³

¹Sagem Défense Sécurité, 72-74 rue de la Tour Billy, 95101 Argenteuil CEDEX, France

²CEA, IRFU, SEDI-Service d'Astrophysique, F-91191 GIF-Sur-YVETTE, France.

³GREYC CNRS UMR 6072, Image Processing Group, ENSICAEN 14050 Caen Cedex, France.
arnaud.woiselle@sagem.com, jean-luc.starck@cea.fr, jalal.fadili@greyc.ensicaen.fr

Résumé – Dans cet article, nous introduisons une méthode pour stabiliser la variance des transformées décimées en utilisant deux transformées stabilisatrices de variance (VST). Ces VST sont appliquées aux ondelettes de Meyer 3-D qui sont au coeur des curvelets 3-D première génération. Cela nous permet de mettre à profit ces curvelets 3-D pour le débruitage du bruit de Poisson, que nous illustrons ensuite sur une simulation cosmologique.

Abstract – In this paper, we introduce a method to stabilize the variance of decimated transforms using two variance stabilizing transforms (VST). These VSTs are applied to the 3-D Meyer wavelet transform which is the core of the first generation 3D curvelets. This allows us to extend these 3-D curvelets to handle Poisson noise, that we apply to the denoising of a simulated cosmological volume.

1 Introduction

Lors de l'acquisition d'images à faible flux, les données sont dégradées par un bruit de Poisson. Ce phénomène apparait souvent en imagerie médicale avec la microscopie confocale [3], ainsi qu'en astrophysique [4] où l'on cherche à observer des objets toujours plus lointains et plus faibles. Le signal observé $X = (X_i)_i$ est constitué d'un ensemble d'échantillons X_i , réalisations supposées mutuellement indépendantes d'un processus de Poisson d'intensités λ_i : $X_i \sim \mathcal{P}(\lambda_i)$. La suppression du bruit de poisson dans des données correspond à la recherche de cette intensité sous-jacente $(\lambda_i)_i$.

2 La stabilisation de variance multi-échelle

Une transformation stabilisatrice de variance, VST, telle que la transformée d'Anscombe [1, 2], transforme le bruit de Poisson en bruit approximativement additif blanc Gaussien centré standard. Zhang *et al.* [6] ont étendu ce type de fonctions à la stabilisation de processus filtrés (par un filtre discret h), plus efficaces lorsque le nombre de coups (λ_i) devient très bas (de

l'ordre du photon par pixel, voire moins).

Etant donné un filtre h de moments $\tau_k = \sum_i h[i]k^k \in [1, 4]$, et un processus de Poisson $X = (X_i)_i$, $X_i \sim \mathcal{P}(\lambda_i)$, nous souhaitons étudier le processus filtré $(Y_j)_j$, $Y_j = \sum_i h[i]X_{j-i}$ obtenu par convolution de X par le filtre h .

La VST proposée est

$$T(Y) = b \cdot \text{sign}(Y + c)|T + c|^{1/2},$$

où b et c sont les paramètres de la transformée, à ajuster au filtre h pour obtenir une convergence optimale (au sens de la variance) vers un bruit Gaussien de variance unitaire.

En posant $b = \text{sign}(\tau_1)/\sqrt{\tau_1}$ et sous l'hypothèse que le processus de Poisson est homogène, i.e. λ constant, sur le support du filtre h , on peut montrer que

$$Y_k \xrightarrow[\lambda \rightarrow \infty]{\mathcal{D}} \mathcal{N}(0, \sigma_h^2),$$

où σ_h ne dépend que du filtre h , et donc peut être précalculé une fois pour toutes. Pour certains types de filtres h , sa valeur peut être calculée théoriquement, mais dans le cas général, on peut l'estimer facilement à partir de cartes de bruit simulées.

Cette VST a été appliquée aux ondelettes non décimées (UWT), aux ondelettes isotropes non décimées (IUWT), et aux curvelets 2-D première

génération.

Ces représentations transforment un signal a_0 en un ensemble de $J + 1$ bandes de coefficients $\{d_1, \dots, d_J, a_J\} = \{\mathcal{W}a_0, a_J\}$, les (d_j) étant les coefficients de détail aux échelles j , et a_J l'approximation basse fréquence. Dans le cas des IUWT, en notant h_j le filtre passe-bas à l'échelle j et T_j la VST associée au filtre global équivalent¹, la transformée en IUWT d'un signal a_0 avec $J + 1$ échelles est

$$\begin{aligned} & \text{IUWT} \begin{cases} a_j = h_{j-1} \star a_{j-1} \\ d_j = a_{j-1} - a_j \end{cases} \\ \Rightarrow & \text{MSVST} \\ & + \\ & \text{IUWT} \begin{cases} a_j = h_{j-1} \star a_{j-1} \\ d_j = T_{j-1}(a_{j-1}) - T_j(a_j). \end{cases} \end{aligned}$$

Dans le cas des ondelettes non décimées classiques (UWT), en utilisant les filtres passe-bas et passe-haut h_j et g_j , l'algorithme est le suivant :

$$\begin{aligned} & \text{UWT} \begin{cases} a_j = h_{j-1} \star a_{j-1} \\ d_j = g_{j-1} \star a_{j-1} \end{cases} \\ \Rightarrow & \text{MSVST} \\ & + \\ & \text{UWT} \begin{cases} a_j = h_{j-1} \star a_{j-1} \\ d_j = g_{j-1} \star T_{j-1}(a_{j-1}). \end{cases} \end{aligned}$$

La reconstruction des IUWT est la somme des bandes, et il est facile de voir que dans le cas MSVST+IUWT, la reconstruction est faite d'une manière très similaire, menant à

$\hat{x} = T_0^{-1} \left(T_J(a_J) + \sum_{j=1}^J d_j \right)$. Dans le cas de la MSVST+UWT, on ne peut définir de reconstruction directe, et il faut utiliser un algorithme itératif, projetant alternativement le résultat sur l'orthant positif dans l'espace direct, et sur l'ensemble d'attache aux données défini par \mathcal{M} comme suit. En notant \mathcal{W} la transformée directe et \mathcal{R} son pseudo-inverse, soit \mathcal{M} l'ensemble des coefficients significatifs du signal a_0 ,

$$\mathcal{M} = \{(i, j) | (\mathcal{W}a_0)_j[i] \text{ is significant}\}.$$

On définit l'ensemble d'attache aux données comme

$$\mathcal{S} = \{d | \forall (i, j) \in \mathcal{M}, d_j[i] = (\mathcal{W}a_0)_j[i]\}.$$

Ces deux projections sont notées \mathcal{P}_S et $\mathcal{Q}_+ = \mathcal{W}\mathcal{P}_+\mathcal{R}$, \mathcal{P}_+ étant le projecteur sur l'orthant positif. Zhang *et al.* [6] ont proposé l'Algorithme 1 pour obtenir une reconstruction avec ces contraintes et attache aux données. Afin d'obtenir de meilleures performances de reconstruction, notamment en termes de photométrie, cet algorithme est utilisé aussi dans le cas des IUWT.

¹ T_j est associé à $h_j^{\text{global}} = h_j \star h_{j-1} \star \dots \star h_1$, d'où $a_j = h_j^{\text{global}} \star a_0$.

Algorithm 1: Algorithme de Reconstruction

Data: a_0 , un nombre d'itérations N .
initialiser $\mathbf{d}^{(0)} = \mathcal{W}\mathcal{P}_+a_0$
for $n = 1$ **to** N **do**
 $\hat{\mathbf{d}} = \mathcal{P}_S \circ \mathcal{Q}_+ \mathbf{d}^{(n-1)}$.
 $\mathbf{d}^{(n)} = \mathcal{S}\mathcal{T}_{\beta_n}(\hat{\mathbf{d}})$, $\mathcal{S}\mathcal{T}_{\beta_n}$ est le seuillage
 doux de seuil β_n .
 Diminuer le seuil β_n .

Result: the estimated density $\mathcal{P}_+\mathcal{R}\mathbf{d}^N$.

3 MSVST et BeamCurvelets 3-D

La transformée en BeamCurvelets (BC), introduite par Woiselle *et al.* [5] en tant qu'extension des curvelets 2-D première génération est une représentation multi-résolution adaptée aux données filamenteuses. Cette décomposition consiste à appliquer des Beamlet 3-D par blocs sur une transformée en ondelettes isotropes des données. Voir [5] pour de plus amples détails.

A l'image du travail de [6], nous appliquons la VST sur les ondelettes 3-D et laissons le reste de la transformée inchangée à l'exception des coefficients de normalisation. La principale différence avec ses travaux est qu'à notre connaissance, les MSVST (VST multi-échelles) ne sont définies que sur des transformations non décimées suivant les schémas présentés en section 2, et l'utilisation d'une telle transformée pour les BeamCurvelets n'est pas possible en termes de charge de calcul et d'utilisation mémoire. Il faut utiliser une transformée décimée donc pyramidale. La transformée en ondelettes de Meyer utilisée est un pavage de Fourier en J couronnes dyadiques et une approximation basse fréquence, notées d_1, \dots, d_J, a_J . Soit h_j , $j \in [0, J]$, les filtres passe-bas (symétriques pairs) de cette transformée dont les transformées de Fourier \hat{h}_j sont de supports inclus dans des cubes imbriqués. Les couronnes sont définies comme les complémentaires de ces cubes, $\hat{g}_j = \sqrt{1 - \hat{h}_j^2}$. Chaque échelle est décimée avec l'opérateur \mathbf{D}_j qui consiste en l'extraction de la partie centrale non nulle en Fourier. Soit \mathbf{U}_j l'opérateur inverse, qui effectue un *zero-padding* de la transformée de Fourier. Cette transformée en ondelettes appelée par la suite BC_C (c pour cohérente avec la transformée gaussienne) suit le schéma suivant:

$$a_0 = x, \quad \forall j \in [1, J] : \begin{cases} d_j = g_{j-1} \star a_{j-1} \\ a_j = \mathbf{D}_j(h_{j-1} \star a_{j-1}), \end{cases}$$

et sa version stabilisée en variance

$$a_0 = x, \quad \forall j \in [1, J] : \begin{cases} d_j = g_{j-1} \star T_j(a_{j-1}) \\ a_j = \mathbf{D}_j(h_{j-1} \star a_{j-1}). \end{cases}$$

Ce processus n'étant pas inversible, il faut utiliser l'Algorithme 1 pour obtenir une estimée de l'intensité du processus de Poisson avec cette représentation.

Pour pallier l'absence d'opérateur de reconstruction, on peut changer la transformée pyramidale, en utilisant une représentation proche des IUWT en 2-D, mais décimée. On définit ainsi le banc des filtres $\hat{\mathbf{h}}_i = \hat{h}_i^2$ et $\hat{\mathbf{g}}_i = \hat{g}_i^2$, qui satisfait la condition d'admissibilité $\hat{\mathbf{h}}_i \hat{\mathbf{h}}_i + \hat{\mathbf{g}}_i \hat{\mathbf{g}}_i = 1$ avec les filtres de synthèse $\tilde{\mathbf{h}}_i$ et $\tilde{\mathbf{h}}_i$ égaux à δ , engendrant ainsi une reconstruction par simple sommation. La transformée en ondelettes devient alors

$$a_0 = x, \quad \forall j \in [1, J] : \begin{cases} \tilde{a}_j = \mathbf{h}_{j-1} \star a_{j-1} \\ d_j = a_{j-1} - \tilde{a}_j \\ a_{j+1} = \mathbf{D}_j \tilde{a}_j, \end{cases}$$

la reconstruction étant

$$\forall j \in [1, J] : \begin{cases} \tilde{a}_j = \mathbf{U}_j a_j \\ a_{j-1} = d_j + \tilde{a}_j. \end{cases}, \hat{x} = a_0.$$

Cette transformée, nommée BC_S (s pour somme) permet d'intégrer la VST comme suit:

$$a_0 = x, \quad \forall j \in [1, J] : \begin{cases} \tilde{a}_j = \mathbf{h}_{j-1} \star a_{j-1} \\ d_j = T_{j-1}^D(a_{j-1}) - T_j^U(\tilde{a}_j) \\ a_{j+1} = \mathbf{D}_j \tilde{a}_j, \end{cases}$$

où l'on peut noter l'utilisation de deux familles de transformées stabilisatrices T^D and T^U , qui ont des paramètres différents et sont respectivement appliquées aux familles décimées et non décimées des approximations $(a_j)_j$ et $(\tilde{a}_j)_j$. La reconstruction est obtenue par un algorithme légèrement plus complexe que la somme, la décimation et les deux VST devant être prises en compte :

$$\forall j \in [1, J] : \begin{cases} \tilde{a}_j = \mathbf{U}_j a_j \\ a_{j-1} = (T_{j-1}^D)^{-1}(d_j + T_j^U(\tilde{a}_j)). \end{cases}$$

Ce changement dans la transformée en ondelettes mène à des atomes d'analyse très similaires, et donne une reconstruction directe. Cette représentation peut aussi être utilisée en analyse uniquement, et la reconstruction obtenue par l'algorithme itératif présenté plus haut.

4 Application aux données astrophysique

Les données 3-D issues des simulations de matière noire à très grande échelle contiennent princi-

palement des amas très denses et des filaments beaucoup plus faibles. Ces données étant obtenues par des codes physiques d'interaction à N corps, le bruit de Poisson y est prépondérant, et un débruitage de ces données permet d'accélérer les temps de calcul. Par ailleurs, les données réelles (relevés de galaxies) suivent naturellement une distribution de Poisson, la matière visible étant interprétable comme une réalisation de la densité de matière noire recherchée.

Nous avons utilisé la transformée présentée – avec ses deux constructions – pour débruiter une simulation de matière noire, et comparons d'une part à un débruitage utilisant des ondelettes 3-D isotropes non décimées stabilisées en variance [6], et d'autre part aux BC gaussiennes utilisées en stabilisant les données par la transformée d'Anscombe. Dans cette dernière méthode, le masque Mc est estimé sur la transformée BC gaussienne de la transformée d'Anscombe des données, et le même algorithme itératif est utilisé pour la reconstruction.

Les résultats sur deux cubes de simulations astrophysique ayant un nombre moyen $\bar{\lambda}$ de 0.17 et 0.05 coups par voxel sont présentés dans les Figures 1 et 2.

Nous présentons aussi les résultats utilisant conjointement les ondelettes isotropes et les BC, afin de prendre en compte la diversité morphologique – amas isotropes et filaments – des données. On observe que le résultat utilisant les ondelettes est bon sur les amas mais perd les filaments, et les résultats BC génèrent un fort bruit d'approximation au voisinage des amas, alors que la combinaison des deux donne un résultat meilleur sur toutes les structures.

Finalement, la Table 1 propose une analyse quantitative des résultats, en termes d'erreur absolue moyenne (MAE), mesure classique en présence de bruit de Poisson, de formule $MAE = -20 \log(\sum_i |\lambda_i - \hat{x}_i|)$, λ_i et \hat{x}_i étant les densités théoriques et estimées du voxel i . Nous observons que les meilleurs résultats – avec les métriques proposées – sont obtenus en utilisant conjointement les ondelettes et BC stabilisées, et que la construction BC_S donne de meilleurs résultats sur ces données que la construction cohérente.

5 Conclusion

Nous avons présenté une nouvelle construction pour les ondelettes pyramidales utilisées dans les curvelets 3-D première génération (BC), et avons montré comment les stabiliser en variance

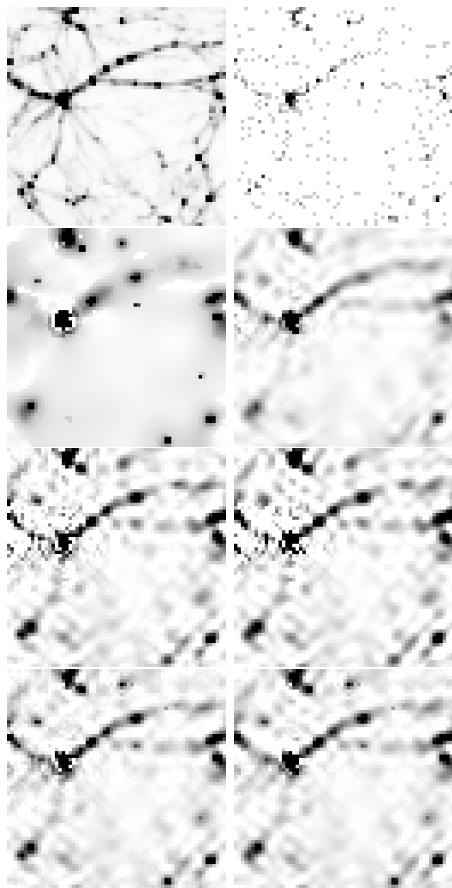


FIG. 1: Dans l'ordre de lecture : original ($\bar{\lambda} = 0.17$), réalisation bruitée, résultats avec : IUWT3D+VST; Anscombe+BC; BC_C ; BC_S

à l'aide de deux VST. Cette représentation filamenteuse adaptée au bruit de Poisson donne de bons résultats en débruitage de simulations astrophysiques. Ces outils seront appliqués par la suite à des relevés réels de galaxies, et les mesures de performances telles que le PSNR ou le MAE n'étant pas forcément des indicateurs pertinents pour la mesure de la restauration des structures, nous envisageons une analyse du genre ou des fonctions de Minkowski dans le futur.

References

- [1] F.J. Anscombe. The transformation of poisson, binomial and negative-binomial data. *Biometrika*, 35:246–254, 1948.
- [2] D.L. Donoho. Nonlinear wavelet methods for recovery of signals, densities and spectra from indirect and noisy data. *Proc. Symp. Applied Mathematics: Different Perspectives on Wavelets*, 47:173–205, 1993.

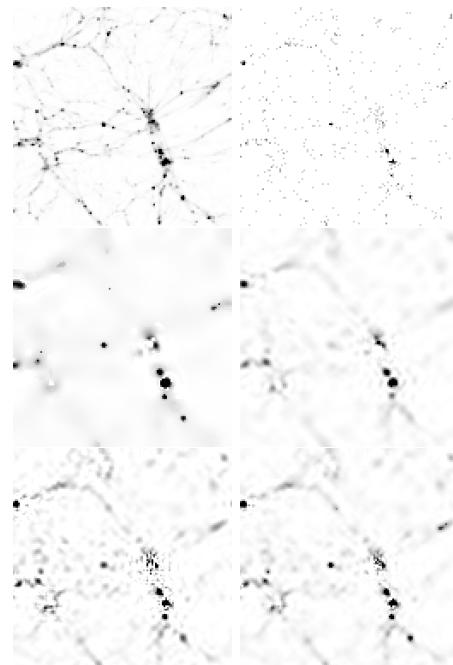


FIG. 2: Dans l'ordre de lecture : original ($\bar{\lambda} = 0.05$), réalisation bruitée, résultats avec : IUWT+VST; Anscombe+BC; BC_S ; IUWT+VST et BC_S

TAB. 1: MAE (en dB) entre la densité initiale et les résultats de débruitage. Résultats en différence par rapport à la donnée bruitée.

Méthode \ MAE(dB)	$\lambda = 0.17$	$\lambda = 0.05$
Données	15.17	24.85
IUWT+VST	+3.91	+4.44
Anscombe+BC	+2.36	+2.47
BC_C	+2.23	+2.70
BC_S	+3.14	+3.27
IUWT+VST+ BC_S	+4.96	+5.01

- [3] J. Pawley. *Handbook of Biological Confocal Microscopy*, 3rd ed. Springer, 2006.
- [4] J.-L. Starck, F. Murtagh, and A. Bijaoui. *Image Processing and Data Analysis: The Multiscale Approach*. Cambridge University Press, 1998.
- [5] A. Woiselle, J.L. Starck, and M.J. Fadili. 3D curvelet transforms and astronomical data restoration. *ACHA*, 28(2):171–188, March 2010.
- [6] B. Zhang, M.J. Fadili, and J.L. Starck. Wavelets, ridgelets, and curvelets for poisson noise removal. *IEEE Transactions on Image Processing*, 17(7):1093–1108, Jul. 2008.