

Débruitage de séquences d'images dynamiques par ondelettes espace-temps hyperboliques

Younes FAROUJ¹, Laurent NAVARRO², Marianne CLAUSEL³, Philippe DELACHARTRE¹

¹Université de Lyon, CREATIS, CNRS UMR5220, Inserm U1044, Villeurbanne, France

²École Nationale Supérieure des Mines, CIS-EMSE, CNRS UMR 5307, LGF, F-42023 Saint-Etienne, France

³ University of Grenoble-Alpes, Laboratoire Jean Kuntzmann, CNRS UMR 5224, Grenoble, France

younes.farouj@creatis.insa-lyon.fr, navarro@emse.fr
marianne.clausel@imag.fr, philippe.delachartre@creatis.insa-lyon.fr

Résumé – Nous introduisons un nouvel algorithme pour la restauration de séquences d'images bruitées. La démarche proposée est basée sur un seuillage de coefficients d'ondelettes tridimensionnelles isotropes en espace et hyperboliques en espace/temps. Cette construction permet de prendre en compte la forte anisotropie entre les régularités spatiale et temporelle. De plus, la régularité en temps est en general élevée, ce qui fournit des estimateurs avec de bonnes vitesses de convergence. Nous mettons ceci en évidence en rappelant des résultats récents liés à l'étude des performances d'estimateurs en ondelettes par des approches maxiset. Enfin, nous présentons des résultats numériques qui montrent le gain dû à la prise en compte de la dimension temporelle et l'apport de l'approche hyperbolique.

Abstract – We introduce a novel algorithm for image sequences noise removal. The proposed approach is based on a thresholding on the coefficients set of a tridimensional wavelets basis which is isotropic in space and hyperbolic in space/time. This construction allows one to take into account the high anisotropy between the spatial and temporal regularities. Moreover, The temporal pattern is, in general, very regular which provides estimators with good convergence rates. We demonstrate this by highlighting some recent results related to the study of wavelet estimators performances via maxiset approaches. Finally, we present some numerical results that show the gain due to the inclusion of the time dimension and the contribution of the hyperbolic setting.

1 Introduction

Le débruitage est une tâche classique en traitement d'images qui constitue depuis longtemps un axe de recherche très actif. Dans le cas d'un bruit blanc *gaussien*, ce problème consiste à trouver une fonction $f(\bar{x})$, $\bar{x} = (x_1, \dots, x_d) \in [0, 1]^d$ à partir d'une observation

$$Y(\bar{x}) = f(\bar{x}) + \varepsilon W(\bar{x}), \quad (1)$$

où W est un bruit blanc *gaussien* et $\varepsilon > 0$ sa variance. Afin d'estimer f , une approche simple mais particulièrement puissante a été proposée par Donoho et Johnstone dans [6] et consiste à construire un estimateur par troncature dans une certaine base d'ondelettes $\mathcal{B} = \{e_i\}_i$

$$\hat{f}_\varepsilon = \sum_{e_i \in \mathcal{B}_Y^T} \langle Y, e_i \rangle e_i \quad ; \quad \mathcal{B}_Y^T = \{e_i \mid \langle Y, e_i \rangle > T\} \quad (2)$$

Quand l'observation est une image, une base d'ondelettes isotropes est typiquement utilisée, ce qui garantit des vitesses de convergence optimales ou presque sur le risque quadratique

$\mathbb{E}[\|\hat{f}_\varepsilon - f\|_2^2]$. De plus un seuil universel permettant d'atteindre ces taux de convergence est donné par

$$T = \gamma \sqrt{\log_2(N)} \varepsilon, \quad (3)$$

où N est le nombre de pixels dans l'image et γ un réel positif. Le succès de cette technique s'explique par le fait que les ondelettes isotropes fournissent des représentations compressibles pour des fonctions appartenant à des classes de Besov isotropes et donc une grande partie des images manipulées en pratique [4]. L'extension de cette démarche à des données spatio-temporelles consisterait alors à traiter la séquence comme un volume $3D$. Intuitivement, les régularités spatiale et temporelle n'ont aucune raison d'être similaires et donc l'hypothèse d'isotropie n'est plus valable. En effet, une séquence d'image $I(t; \mathbf{x}) : \mathbb{R}^3 \rightarrow \mathbb{R}$ vérifie l'équation de transport suivante¹ [10].

$$\partial_t I + \mathbf{v} \cdot \nabla_{\mathbf{x}} I = 0, \quad (4)$$

où $\partial_t I$ et $\nabla_{\mathbf{x}}$ sont, respectivement, les dérivées spatiale et temporelle, alors que \mathbf{v} représente la vitesse dans l'image. Les ondelettes isotropes ne sont pas adéquates pour représenter des phénomènes physiques de ce genre [5]. L'idée du présent travail est de réaliser l'opération du seuillage dans une base qui

¹Y.F bénéficie d'un financement de la "Région Rhône-Alpes" dans le cadre de l'ARC 6.

¹Quand l'inconnue est la vitesse \mathbf{v} , cette équation est appelée : équation du flot optique.

respecte à la fois l'image et l'évolution temporelle en tenant compte des deux régularités. Pour cela, nous utilisons la base formée par le produit tensoriel d'une ondelette spatiale isotrope $2D$ et une ondelette $1D$ qui évolue dans la dimension temporelle. Cette construction en temps, dite "hyperbolique", permet de représenter d'une façon optimale des espaces de Besov anisotropes. Notons qu'une base similaire a été utilisée dans le cadre de l'échantillonnage compressif de données hyperspectrales [7].

2 Ondelettes et régularité

2.1 Quelques définitions

Considérons une ondelette mère unidimensionnelle ψ à laquelle on associe ses versions dilatées et translatées $\psi_{j,k}(\cdot) = (2^j \cdot - k)$ pour $j \geq 0$ et $k \geq 0$. De la même façon, on définit une fonction d'échelle ϕ et ses translatées $\phi_k(\cdot) = \phi(\cdot - k)$. La décomposition en ondelettes d'un signal $f \in \mathbb{L}^2([0, 1])$ s'écrit alors

$$f = \sum_{k \in \mathbb{Z}} \langle f, \phi_k \rangle \phi_k + \sum_{j=0}^J \sum_{k \in 2^{-j}\mathbb{Z}} \langle f, \psi_{j,k} \rangle \psi_{j,k} \quad (5)$$

Dans le cadre multidimensionnel $\mathbb{L}^2([0, 1]^d)$, $d > 1$, la construction de la fonction d'échelle est immédiate, alors que la fonction ondelette est construite à l'aide d'un produit tensoriel

$$\begin{aligned} \psi_{\vec{j}, \vec{k}} &= \psi_{j_1, \dots, j_d, k_1, \dots, k_d}(x) \\ &= \psi_{j_1, k_1}(x_1) \otimes \dots \otimes \psi_{j_d, k_d}(x_d). \end{aligned} \quad (6)$$

L'extension de la décomposition (5) n'est évidemment pas unique. En effet, suivant la règle de sommation double imposée, on obtient différentes définitions du produit tensoriel (6). Si le produit est réalié à partir d'échelles fixes, on obtient les ondelettes isotropes classiques qui forment une analyse multi-résolution de $\mathbb{L}^2([0, 1]^d)$ [12]. Les ondelettes anisotropes sont obtenues dans le cas où le rapport entre les échelles constituant la somme est fixe². Les ondelettes hyperboliques correspondent au cas où aucune restriction n'est faite sur l'échelle. Dans la suite on utilisera la notation $\beta_{\vec{j}, \vec{k}} = \langle f, \psi_{\vec{j}, \vec{k}} \rangle$.

Notons $\vec{p} = \{p_1, \dots, p_d\}$ et $\vec{s} = \{s_1, \dots, s_d\}$. La classe de fonctions dont la s_i -ème dérivée est L^{p_i} est appelée classe de Besov $B_{p,q}^{\vec{s}}$. La classe est dite isotrope si les $\{p_i\}_{i=1 \dots d}$ sont égaux et anisotrope sinon. On peut caractériser les boules dans cette classe par le comportement des coefficients d'ondelettes de la façon suivante [9]

$$B_{p,q}^{\vec{s}}(A) = \left\{ f \in \mathbb{L}^2([0, 1]^d) \text{ telles que } \left(\sum_{\vec{j}} 2^q \sum_{i=1}^d j_i s_i^* \left(\sum_{\vec{k}} |\theta_{\vec{j}, \vec{k}}|^p \right)^{\frac{q}{p}} \right)^{\frac{1}{q}} < A \right\},$$

où $s_i^* = s_i + 1/2 - 1/p$. Nous nous intéressons à présent, à l'estimation des fonctions appartenant à des classes de Besov et les choix appropriés des bases d'ondelettes utilisées.

2.2 Estimateurs en ondelettes

D'un point de vue régularité, résoudre (1) par l'estimateur (2) revient à trouver \hat{f} en imposant son appartenance à une certaine classe de Besov. Ces classes ont été identifiées comme un bon modèle pour les images car elles ont l'avantage de tolérer les discontinuités, ce qui n'est pas le cas pour des classes plus classiques comme les classes de Sobolev. Comme les bases d'ondelettes sont inconditionnelles dans \mathbb{L}^2 , l'estimateur (2) procède de façon adaptative. Cela signifie que l'on peut atteindre des estimations optimales, à un terme *logarithmique* près, sans une connaissance exacte des paramètres de régularités (\vec{p}, q, s) . Cependant, dans le cas où f dépend de plusieurs variables, on peut aussi ajouter une information sur le rapport entre les régularités suivant les différentes directions où bien prendre en compte le fait que des variables sont plus importantes que d'autres pour améliorer la vitesse de convergence. En d'autres termes on cherche des propriétés d'anisotropie et les ondelettes hyperboliques se sont imposées comme un outil puissant et naturel du traitement anisotrope des données [1]. On formalise cette idée d'un point de vue maxiset.

Pour tout estimateur \hat{f}_ε on définit un maxiset $MS(\hat{f}_\varepsilon, \|\cdot\|_p^p, r_\varepsilon)$ par le plus grand ensemble sur lequel le risque $\mathbb{E}[\|\hat{f}_\varepsilon - f\|_p^p]$ converge à une vitesse r_ε .

$$\sup_{\varepsilon} r_\varepsilon^{-1} \mathbb{E}[\|\hat{f}_\varepsilon - f\|_p^p] < \infty$$

Nous nous intéressons naturellement à des vitesses de convergences presque optimales $r_\varepsilon = C_\varepsilon^{2\mu(\vec{s}, p)}$ avec

$$C_\varepsilon = \varepsilon (\log \varepsilon^{-1})^\alpha, \text{ pour un } \alpha > 0$$

et

$$\mu(\vec{s}, p) = 2\tilde{s}/(2\tilde{s} + 1), \quad \tilde{s} = 1/\left(\sum_{i=1}^d s_i^{-1}\right).$$

Pour une telle vitesse de convergence, il a été montré récemment [2] que les $B_{p,q}^{\vec{s}}$ anisotropes ne sont pas inclus dans le maxiset de l'estimateur (2) si \mathcal{B} est une base d'ondelettes isotrope et ils le sont si \mathcal{B} est hyperbolique. Ces résultats orientent vers l'utilisation des ondelettes hyperboliques quand on traite des données anisotropes. Cependant, la qualité de l'estimation ne dépend pas que de l'anisotropie. En effet, l'expression de μ laisse penser qu'il faut qu'au moins l'un des s_i soit suffisamment grand pour avoir une bonne estimation, ce qui n'est

²Si ce rapport est fixé à un, on retrouve les ondelettes isotropes.

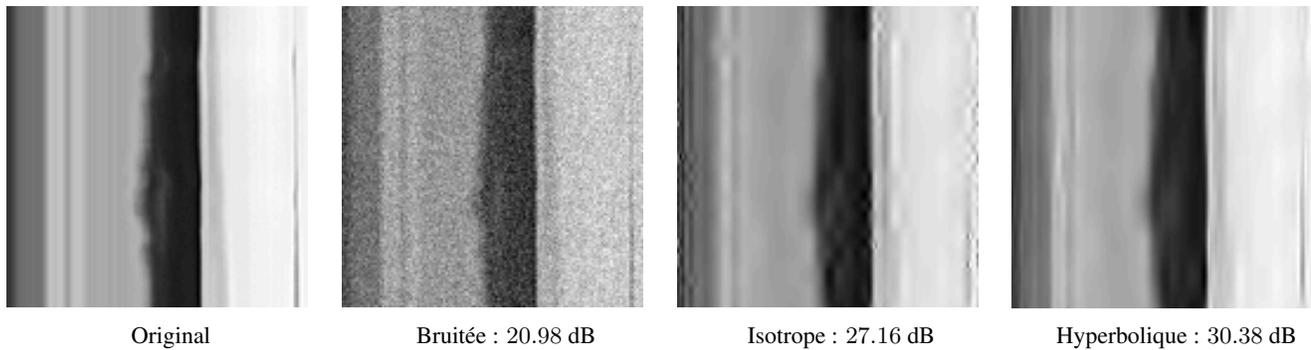


FIG. 1: Exemple de débruitage sur coupe en temps de la séquence "Aniko".

pas le cas par exemple pour les images $2D$. Dans notre application, la régularité sur la dimension temporelle peut être élevée comme le suggère la Figure ???. Cette figure montre un exemple de débruitage, par un seuillage sur des ondelettes de Daubechies à 6 moments nuls, d'une coupe bi-dimensionnelle d'une séquence $2D + t$. Ceci revient à fixer l'une des dimensions spatiales. La comparaison visuelle et du PSNR mettent en évidence l'apport de l'approche hyperbolique qui fournit une bonne reconstruction des lignes droites dans le sens du temps. Ces résultats ont motivé le choix d'une base particulière pour le débruitage de séquences dynamiques que nous présentons dans la section suivante.

3 Débruitage de séquences

La régularité isotrope est bien évidemment un cas spécial du cas anisotrope, mais comme les expressions des vitesses de convergence le montrent, l'amélioration qui peut être apportée par la troncature dans des bases hyperboliques est plus importante quand on est confronté à des données avec un minimum d'anisotropie et une régularité suffisante sur au moins l'une des dimensions. Nous proposons donc, d'utiliser comme base celle qui résulte d'un produit tensoriel sans restriction d'une ondelette $2D$ isotrope dans l'espace (x_1, x_2) et une ondelette $1D$ pour x_3 . La base construite est ainsi hyperbolique en temps.

La Figure 2 met en évidence deux points remarquables. Le premier est l'avantage qu'il y a à faire un procédé spatio-temporel lors du débruitage au lieu de traiter la séquence image par image. Le deuxième point est l'amélioration due à la prise en compte de la régularité temporelle à travers les coefficients d'ondelettes hyperboliques. Cela se manifeste en terme de PSNR, d'après l'aperçu visuel et la consistance par rapport à la variance du bruit (Figure 4). Les situations où il y a apparition d'objets dans la scène créent des discontinuités dans le plan orthogonal à l'axe du temps. La Figure 3 en présente un exemple. Notre approche arrive à surmonter ces singularités. Notons que l'explication de la performance de notre algorithme peut avoir une inclination vers un phénomène appelé fléau de la dimensionalité. En effet, les vitesses de convergences de n'importe quel estimateur dépendent de la dimension du problème. Dans

notre problématique, les variables spatiales et la variable temporelle n'interagissent pas forcément entre elles, ce qui fait que la dimension effective du problème est égale à 2. Ceci permet d'augmenter l'optimalité de l'estimateur [13]

4 Discussions et perspectives

Nous avons présenté un algorithme simple et efficace pour le débruitage de séquences d'images dynamiques. Ceci peut être une étape préliminaire à plusieurs tâches en vision et traitement d'images comme l'estimation de mouvement où bien la segmentation spatio-temporelles. Cependant, Nous tenons à préciser que nous attachons seulement de mettre en évidence le gain dû au traitement anisotropique par ondelettes hyperboliques dans le cas d'estimateurs en ondelettes, ce qui a fait la base de notre comparaison. Nous notons, donc, l'existence de plusieurs méthodes sophistiquées dans la littérature. On cite, à titre d'exemple, le très performant algorithme dans [11] qui se base sur des raffinements par compensation de mouvement et filtrage de Wiener le rendant, par ailleurs, couteux en terme de temps de calcul et difficile à paramétrer. Des pistes diverses peuvent être envisagées pour améliorer nos résultats et prétendre à concurrencer ce type de méthodes. On pourrait par exemple utiliser des ondelettes non décimées ou bien complexes pour profiter de leur invariance par translation. On pourrait aussi améliorer le procédé de troncature. En effet, dans [2], il est montré que le seuillage par blocs permet d'atteindre des vitesses de convergence optimales (sans le terme *logarithmique*). D'une autre part, dans plusieurs applications, on est confronté à des bruits non-*gaussiens*. On peut penser alors à une présentation variationnelle [8], dans laquelle le terme d'attache aux données prend en compte la nature du bruit [3].

References

- [1] P. Abry, M. Clausel, S. Jaffard, S. Roux, B. Vedel. *Hyperbolic wavelet transform: an efficient tool for multifractal analysis of anisotropic textures*. Revista Matemática Iberoamericana, European Mathematical Society, 31 (1), pp.10.4171/rmi /836, 2015.
- [2] F. Autin, G. Claeskens et J.M. Freyermuth. *Hyperbolic wavelet thresholding methods and the curse of dimensionality through*



FIG. 2: Comparaison visuelle sur une frame de la séquence "Aniko". PSNR calculé sur toute la vidéo.

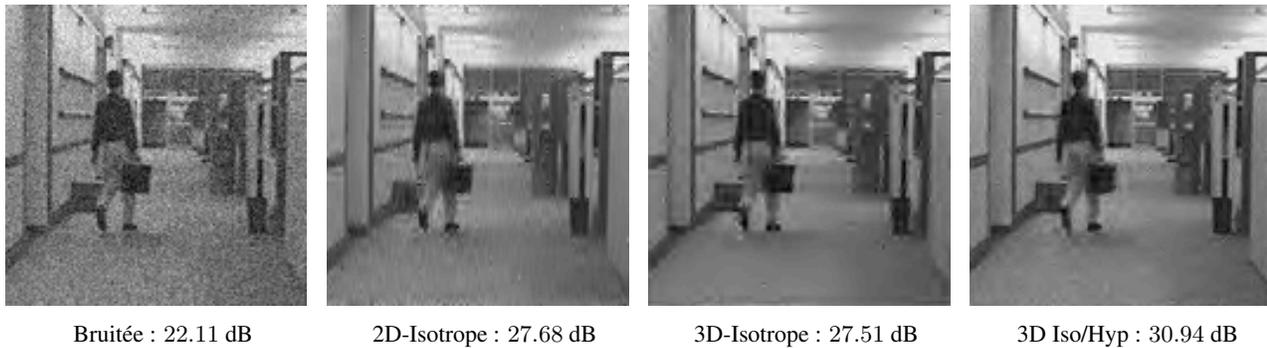


FIG. 3: Comparaison visuelle sur une frame de la séquence "Hall". PSNR calculé sur toute la vidéo.

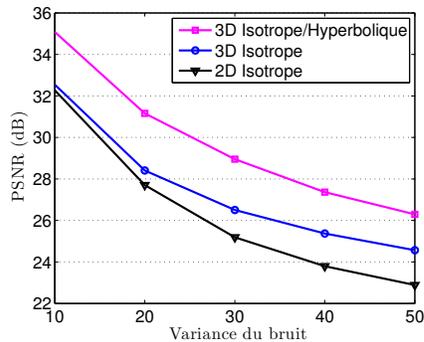


FIG. 4: Performances des trois techniques sur la séquence "Aniko", PSNR calculé en moyennant sur 20 tests.

the maxiset approach. Appl. Comput. Harmon. Anal., volume 36, 239-255. 2014

- [3] C.-A. Deledalle, L. Denis et F. Tupin *How to compare noisy patches? Patch similarity beyond Gaussian noise*. International Journal of Computer Vision, vol. 99, no. 1, pp. 86-102, 2012
- [4] R.A. DeVore, B. Jawerth et B.J. Lucier. *Image compression through wavelet transform coding*. Information Theory, IEEE Transactions on 38 (2), 719-746. 1992
- [5] R.A. DeVore, G. Petrova et P. Wojtaszczyk. *Anisotropic smoothness spaces via level sets*. Communications on Pure and Applied Mathematics, Volume 61, Issue 9, pages 1264-1297, September 2008

- [6] D. Donoho et I. Johnstone. *Ideal spatial adaptation via wavelet shrinkage*. Biometrika, 81:425-455, Dec 1994.
- [7] M. F. Duarte, et R. G. Baraniuk. *Kronecker compressive sensing*. Image Processing, IEEE Transactions on, vol. 21, no. 2, pp. 494-504, Feb., 2012
- [8] S. Durand, J. Fadili et M. Nikolova. *Multiplicative noise removal using L1 fidelity on frame coefficients*. Journal of Mathematical Imaging and Vision 36 (3), 201-226, 2010.
- [9] W. Heping, *Representation and approximation of multivariate function with mixed smoothness by hyperbolic wavelets*. J. Math. Anal. Appl., 291, 698-715. 2004
- [10] B.K.P. Horn et B.G. Schunck. *Determining Optical Flow*. AI(17), No. 1-3, pp. 185-203, August 1981
- [11] M. Maggioni, G. Boracchi, A. Foi, et K. Egiazarian. *Video Denoising, Deblocking and Enhancement Through Separable 4-D Nonlocal Spatiotemporal Transforms*. Image Processing, IEEE Transactions on, vol. 21, no. 9, pp. 3952-3966, September 2012.
- [12] S. G. Mallat. *A theory for multiresolution signal decomposition: the wavelet representation*. Pattern Analysis and Machine Intelligence, IEEE Transactions on, Volume.11, Issue. 7, pp. 674-693, Jul 1989.
- [13] M. H. Neumann *Multivariate wavelet thresholding in anisotropic function spaces*. Stat. Sin. 10, pp. 399-431. 2000.