

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 hyperboliques en espace/temps correspondant à la tensorisation d’ondelettes isotropes en espace et d’ondelettes unidimensionnelles en temps. Cette construction permet de prendre en compte la forte différence de régularité entre espace et temps. La régularité en temps est en général plus élevée que la régularité en espace. Cela est néanmoins suffisant pour avoir des estimateurs avec de bonnes vitesses de convergence. Nous mettons ceci en évidence en rappelant des résultats théoriques liés à l’étude des performances d’estimateurs en ondelettes. 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 hyperbolic wavelet basis which defined as the tensor product of an isotropic wavelet basis in space and an unidimensional one in time. This construction allows one to take into account the difference of regularity between space and time. Moreover, the temporal pattern is, in general, very regular which provides estimators with good convergence rates. We demonstrate this by highlighting some theoretical results related to the study of wavelet estimators performances. 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 et 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 $\varepsilon > 0$ le niveau de bruit. Afin d’estimer f , une approche simple mais particulièrement puissante a été proposée par Donoho et Johnstone dans [5] et consiste à construire un estimateur par troncature dans une certaine base d’ondelettes $\mathcal{B} = \{\psi_{j,k}\}_{j,k}$

$$\hat{f}_\varepsilon = \sum_{(j,k) \in I_\varepsilon} \langle Y, \psi_{j,k} \rangle \psi_{j,k}, \quad (2)$$

où t_ε désigne un seuil positif dépendant du niveau de bruit ε et $I_\varepsilon = \{(j,k) \mid |\langle Y, \psi_{j,k} \rangle| \geq t_\varepsilon\}$ est l’ensemble des indices des coefficients d’ondelettes au dessus de ce seuil.

Y.F bénéficie d’un financement de la “Région Rhône-Alpes” dans le cadre de l’ARC 6. Les travaux de M.C et Ph.D sont soutenus, respectivement, par les projets ANR ASTRES et Labex CELYA.

Quand l’observation est une image, on utilise très souvent une base d’ondelettes isotropes, ce qui garantit des performances optimales ou quasi-optimales de cet estimateur mesuré à partir du risque quadratique $\mathbb{E}[\|\hat{f}_\varepsilon - f\|_2^2]$. De plus un seuil dit universel permettant d’atteindre ces performances est donné par

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

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 la plupart des images, usuellement considérées comme des fonctions appartenant à des classes d’espaces fonctionnels classiques : les espaces de Besov [3].

L’extension de cette démarche à des données spatio-temporelles va consister à traiter la séquence comme un volume 3D. Intuitivement, les régularités spatiale et temporelle n’ont aucune raison d’être similaires. Les ondelettes isotropes ne sont donc pas adéquates pour représenter des phénomènes physiques de ce type [4]. Il s’agit donc d’être capable de traiter différemment espace et temps.

L’idée du présent travail est de réaliser l’opération de seuillage dans une base prenant en compte à la fois chaque image composant la séquence considérée et son évolution tem-

porelle. Notre approche permet de tenir compte de la différence de caractéristiques en temps et en espace. Pour cela, nous utilisons une base construite comme le produit tensoriel d’une base d’ondelettes spatiales isotropes bidimensionnelles et d’une base d’ondelettes unidimensionnelles en temps. Cette construction en temps, dite “hyperbolique”, permet de représenter d’une façon optimale des séquences d’images. Les ondelettes hyperboliques sont un outil efficace, déjà utilisé dans le cadre de l’analyse d’image. Citons par exemple [1] où une procédure d’estimation du degré d’anisotropie d’une image basée sur les ondelettes hyperboliques est proposée. Une approche analogue a été aussi proposée pour l’échantillonnage compressif de données hyperspectrales [6]. L’étude des performances statistiques de notre algorithme nécessitera de modéliser nos séquences d’images comme des objets appartenant à une classe d’espaces fonctionnels adaptée : les espaces de Besov anisotropes.

2 Ondelettes et régularité

2.1 Ondelettes hyperboliques et régularité spatio-temporelle

Nous introduisons tout d’abord notre outil d’analyse multi-résolution espace/temps que sont les bases d’ondelettes hyperboliques. Commençons par considérer une ondelette mère unidimensionnelle ψ à laquelle on associe ses versions dilatées et translatées $\psi_{j,k}(\cdot) = 2^{j/2}\psi(2^j \cdot - k)$ 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

$$f = \sum_{k \in \mathbb{Z}} \langle f, \phi_k \rangle \phi_k + \sum_{j=0}^{+\infty} \sum_{k \in 2^{-j}\mathbb{Z}} \langle f, \psi_{j,k} \rangle \psi_{j,k} \quad (4)$$

Dans le cadre unidimensionnel [5], il est bien connu que les performances de l’estimateur défini par (2) sont liées à l’indice de régularité du signal f mesuré dans la classe des espaces de Besov $B_{p,\infty}^s([0, 1])$, définis pour $p > 1$ et $s > 0$ comme l’ensemble des fonctions de $L^p([0, 1])$ telles que

$$\sup_{j \geq 0} 2^{pj(s-1/p+1/2)} \sum_{k \in 2^{-j}\mathbb{Z}} |\theta_{j,k}|^p < \infty.$$

où on note $\theta_{j,k} = \langle f, \psi_{j,k} \rangle$. Dans le cadre multidimensionnel $\mathbb{L}^2([0, 1]^d)$, $d > 1$, la fonction d’échelle est le produit tensoriel des fonctions d’échelles unidimensionnelles alors que l’ondelette mère est définie pour tout, $\vec{j} = (j_1, \dots, j_d)$ et $\vec{k} = (k_1, \dots, k_d)$ comme

$$\begin{aligned} \psi_{\vec{j},\vec{k}} &= \psi_{j_1,\dots,j_d,k_1,\dots,k_d}(x), \\ &= \psi_{j_1,k_1}(x_1) \otimes \dots \otimes \psi_{j_d,k_d}(x_d). \end{aligned} \quad (5)$$

Notons que contrairement aux ondelettes isotropiques qui font intervenir un seul facteur de dilatation 2^j , les ondelettes hyperboliques font intervenir autant de facteurs de dilatation que la dimension, ce qui les rend beaucoup plus adaptées lorsqu’il

s’agit de traiter des données anisotropes. Dans la suite on utilisera la notation $\theta_{\vec{j},\vec{k}} = \langle f, \psi_{\vec{j},\vec{k}} \rangle$.

Nous définissons maintenant une notion de régularité qui nous permettra de mesurer directement les performances de notre méthode en tenant compte de la dissymétrie entre temps et espace. Nous utilisons une généralisation multidimensionnelle de la classe des espaces de Besov dont nous avons rappelé la définition plus haut. Notons $\vec{p} = (p_1, \dots, p_d)$ des réels appartenant à $[1, +\infty]$ et $\vec{s} = (s_1, \dots, s_d)$ des réels positifs. L’espace de Besov anisotrope $B_{\vec{p},\infty}^{\vec{s}}([0, 1]^d)$ est ainsi la classe de fonctions $f \in \mathbb{L}^2([0, 1]^d)$ telle que [13, 1]

$$\sup_{j_1, \dots, j_d} 2^{p \sum_{\ell=1}^d j_\ell (s_\ell + 1/2 - 1/p)} \left(\sum_{\vec{k}} |\theta_{\vec{j},\vec{k}}|^p \right) < \infty.$$

Nous modélisons maintenant nos séquences d’images comme des fonctions tridimensionnelles dont la régularité est mesurée naturellement dans des espaces du type $B_{\vec{p},q}^{\vec{s}}([0, 1]^d)$ avec $\vec{s} = (s_{space}, s_{space}, s_{time})$, le paramètre s_{space} mesurant la régularité en espace de chaque image tandis que le paramètre s_{time} mesure la régularité en temps de la séquence. Dans la section suivante, nous donnons les résultats théoriques sur lesquels est fondée notre procédure d’estimation.

2.2 Performances statistiques du débruitage par ondelettes hyperboliques

Comme nous allons l’expliquer, les performances de l’estimateur (2) sont directement liées à la régularité de la séquence d’images mesurée dans les espaces de Besov anisotropes. Soulignons néanmoins que cet estimateur possède une propriété très importante à savoir qu’il est *adaptatif*. Cela signifie que l’on peut atteindre des performances optimales, à un terme *logarithmique* près, sans une connaissance préalable de la régularité (\vec{p}, \vec{s}) . Ceci est bien sûr essentiel dans les applications pratiques puisqu’on peut difficilement se permettre une estimation préalable de la régularité des données.

L’utilisation d’ondelettes hyperboliques pour débruiter des données ayant des caractéristiques différentes suivant les directions permet d’améliorer notablement la performance de la procédure.

Dans [10], la vitesse de reconstruction d’une fonction f appartenant à un espace de Besov $B_{\vec{p},\infty}^{\vec{s}}$ pour un niveau de bruit $\varepsilon > 0$ donné est reliée à la fois à la régularité du signal et à son degré d’anisotropie. Plus précisément, on a

$$\mathbb{E} \|f - f_\varepsilon\|_p^p \leq r_\varepsilon,$$

où $r_\varepsilon = \varepsilon (\log \varepsilon^{-1})^{2\mu(\vec{s},p)}$, avec

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

Par ailleurs, il est aussi montré que l’estimateur défini par (2) est quasi-optimal, c’est à dire qu’il n’existe pas d’estimateur permettant d’obtenir de meilleures vitesses de convergence si

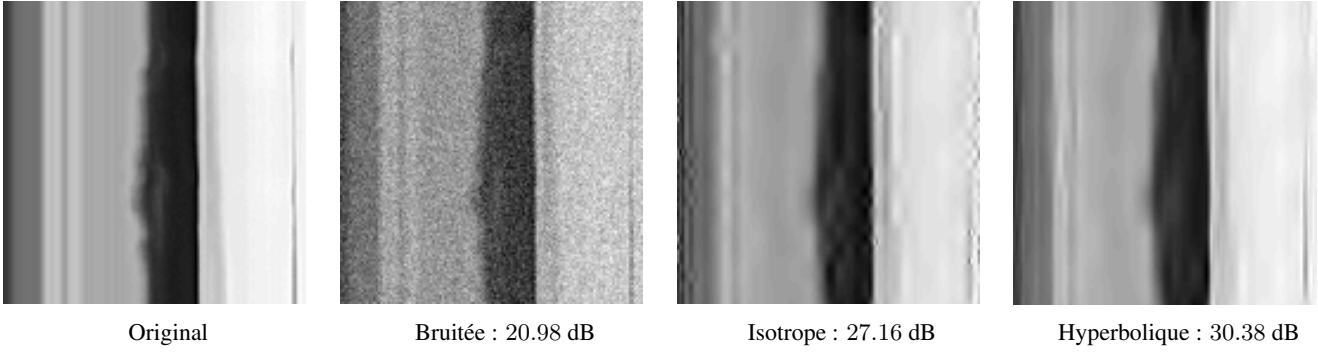


FIG. 1: Exemple de débruitage sur coupe en temps de la séquence “Aniko”.

on s’autorise une perte en $\log(\varepsilon)$. 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 les données sont fortement anisotropes et suffisamment régulières. Des résultats complémentaires sur ce type d’estimateurs peuvent être aussi trouvés dans [2], détaillant en particulier la classe maximale des fonctions pouvant être reconstruites à une vitesse donnée (approche maxiset).

3 Application à des données réelles

Nous nous sommes intéressés au débruitage de deux séquences d’images : les séquences “Aniko” et “Hall”¹. Dans chacun des exemples considérés, les séquences vidéos sont beaucoup plus régulières en temps qu’en espace. Dans un premier temps, nous avons considéré une coupe bi-dimensionnelle d’une séquence $2D + t$ en fixant l’une des dimensions spatiales. Nous avons testé les performances de l’estimateur (2) en utilisant des ondelettes de Daubechies à 6 moments nuls et une décomposition à 4 échelles. Nous avons comparé les deux cas : isotrope et hyperbolique avec le même seuil (3) pour $\gamma = 1$. La comparaison visuelle et le rapport signal/bruit de crête (PSNR) mettent en évidence l’apport de l’approche hyperbolique qui fournit une bonne reconstruction temporelle. Les résultats encourageants qui ont été ainsi obtenus ont motivé la construction d’une base particulière dans le cas tridimensionnel des séquences d’images vidéo. En effet, nous proposons d’utiliser la base résultant d’un produit tensoriel d’une ondelette $2D$ isotrope en espace et une ondelette unidimensionnelle en temps. La base construite correspond ainsi à une variante des bases hyperboliques précédemment introduites. Nous comparons les résultats obtenus par la troncature dans la nouvelle base au cas isotrope en $2D$ et en $3D$ en fixant le niveau de décomposition à 4. Nous avons optimisé le paramètre γ du seuil pour chaque cas.

Nos résultats sur les séquences vidéo sont résumés dans la Figure 2. On peut ainsi constater qu’il est plus intéressant de prendre en compte l’aspect spatio-temporel des données lors du débruitage au lieu de traiter la séquence image par image.

L’amélioration due à la prise en compte de la régularité temporelle à travers les coefficients d’ondelettes hyperboliques se manifeste en terme de PSNR (voir figure 4). Le cas où il y a une apparition d’objets dans la scène, et donc création de discontinuités dans le plan orthogonal à l’axe du temps, est illustré dans la figure 3. Notons que notre approche est robuste vis à vis des discontinuités pouvant apparaître dans une séquence d’image.

4 Discussions et perspectives

Nous avons présenté un algorithme de seuillage, dans une base d’ondelettes adaptée, pour le débruitage de séquences d’images dynamiques. Ceci peut être une étape préliminaire à plusieurs tâches en vision et en traitement d’images comme l’estimation de mouvement où bien la segmentation spatio-temporelle. Les résultats montrent le gain dû à l’utilisation de bases d’ondelettes hyperboliques. D’autres techniques comme la méthode BM3D [8] basée sur des raffinements par compensation de mouvement et filtrage de Wiener, peuvent être plus efficaces mais aussi plus coûteuses en terme de temps de calcul et plus difficiles à paramétrer. L’algorithme que nous avons proposé est lui à la fois simple d’utilisation et très performant. Des pistes d’amélioration de notre approche peuvent être envisagées. 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 des méthodes héréditaires permettent d’atteindre des vitesses de convergence optimales (sans le terme *logarithmique*). D’une autre part, dans plusieurs applications, le bruit est non-*gaussien*. Des méthodes alternatives intéressantes sont les approches variationnelles [7], dans laquelle le terme d’attache aux données prendrait en compte la nature du bruit.

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.

¹Disponibles sur http://see.xidian.edu.cn/vipsl/database_Video.html



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

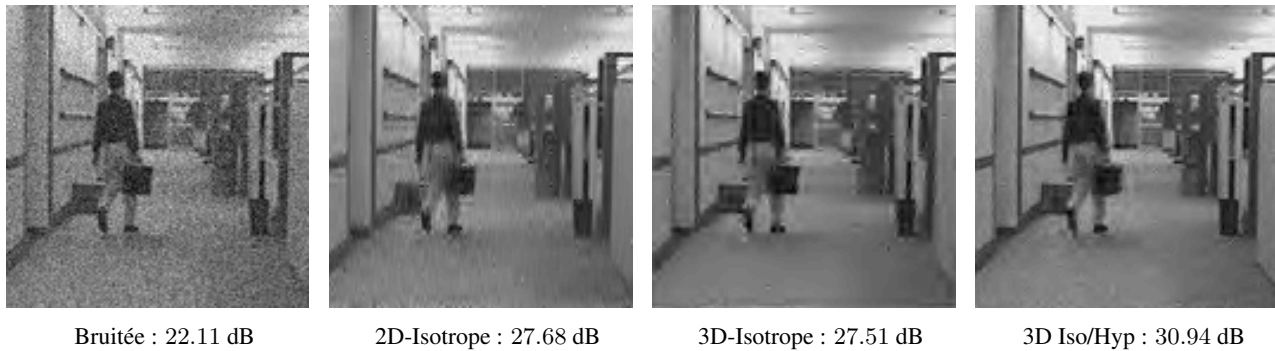


FIG. 3: Comparaison visuelle sur une frame de la séquence “Hall”. PSNR calculé sur toute la vidéo.

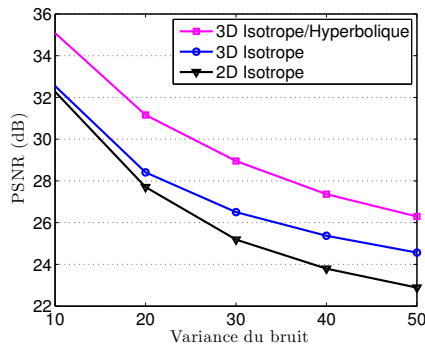


FIG. 4: Performances des trois techniques sur la séquence “Aniko”, PSNR calculé en moyennant sur 20 tests.

- [2] F. Autin, G. Claeskens et J.M. Freyermuth. *Hyperbolic wavelet thresholding methods and the curse of dimensionality through the maxiset approach*. Appl. Comput. Harmon. Anal., volume 36, 239-255. 2014
- [3] 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
- [4] 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
- [5] D. Donoho et I. Johnstone. *Ideal spatial adaptation via wavelet shrinkage*. Biometrika, 81:425-455, Dec 1994.
- [6] M. F. Duarte, et R. G. Baraniuk. *Kronecker compressive sensing*. Image Processing, IEEE Transactions on, vol. 21, no. 2, pp. 494-504, Feb., 2012
- [7] 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.
- [8] 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.
- [9] 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.
- [10] M. H. Neumann *Multivariate wavelet thresholding in anisotropic function spaces*. Stat. Sin. 10, pp. 399-431. 2000.
- [11] C. P. Rosiène and T. Q. Nguyen, *Tensor-product wavelet vs. Mallat decomposition: a comparative analysis*, in ISCAS 99, vol. 3, Jul. 1999, pp. 431-434.
- [12] S.G. Roux, M. Clausel, B. Vedel, S. Jaffard, P. Abry. *The Hyperbolic Wavelet Transform for self similar anisotropic texture analysis*. IEEE Trans. Imag. Proc., vol 22(11), pp 4353-4363 (2013).
- [13] H. Triebel, *Interpolation theory, function spaces, differential operators.*, Amsterdam, North-Holland, 1978.