Ultrasound spatiotemporal despeckling via Kronecker wavelet-Fisz thresholding

Younes Farouj, Laurent Navarro, Jean-Marc Freyermuth, Marianne Clausel, Philippe Delachartre

Noname manuscript No.
(will be inserted by the editor)

Ultrasound Spatio-temporal Despeckling via Kronecker Wavelet-Fisz Thresholding

Younes Farouj[1] · **Laurent Navarro** · **Jean-Marc Freyermuth** · **Marianne Clausel** · **Philippe Delachartre**

Received: date / Accepted: date

Abstract We propose a novel framework for despeckling ultrasound image sequences while respecting the structural details. More precisely, we use thresholding in an adapted wavelet domain that jointly takes into account for the non-Gaussian statistics of the noise and the differences in spatial and temporal regularities. The spatio-temporal wavelet is obtained via the *Kronecker* product of two sparsifying wavelet bases acting respectively on the spatial and temporal domain. Besides enabling a structured sparse representation of the time-space plan, it also makes it possible to perform a variance stabilization routine on the spatial domain through a Fisz transformation. The proposed method enjoys adaptability, easy tuning and theoretical guaranties. We propose the corresponding algorithm together with results that demonstrate the benefits of the proposed spatio-temporal approach over the successive spatial treatment. Finally, we describe a data-driven extension of the proposed method that is based on temporal pre-filtering.

Keywords Dynamic ultrasound imaging · despeckling · Fisz transformation · variance stabilization · structured sparsity

Younes Farouj
Medical Image Processing Lab , Institute of Bioengineering, EPFL, CH-1015 Lausanne, Switzerland
E-mail: younes.farouj@epfl.ch

Laurent Navarro
École Nationale Supérieure des Mines;
CNRS UMR 5307; LGF, F-42023 Saint-Etienne, France

Jean-Marc Freyermuth
Institute of Statistics, Université de Neuchâtel,
Av. de Bellevaux 51 CH - 2000 Neuchâtel, Switzerland

Marianne Clausel
Université de Grenoble; Laboratoire Jean Kuntzmann;
CNRS UMR 5224, Grenoble, France

Philippe Delachartre
Université de Lyon; CREATIS; CNRS UMR5220;
Inserm U1044; INSA-Lyon; Lyon 69622, France

1 Introduction

Ultrasound (US) Imaging allows safe and cheap exploration of biological tissues and organs in real-time. This medical modality continues to be an attractive tool for clinical diagnostic and investigation. Unfortunately, US images are affected by "speckle" noise resulting from the accumulation of individual scattered beams from tissue inhomogeneities. Although speckle might give information about the properties of underlying tissues, it also affects the quality of these images and makes them challenging for human perception and for post-processing tasks such as segmentation and registration. Hence, pre-filtering is an important first step in the US image analysis pipeline. Some structural details are, however, important for diagnosis. Therefore, it is important to preserve them while performing the denoising task.

Throughout the years, various US despeckling approaches have been proposed in the literature. Early contributions were simply based on classical spatial approaches such as median filtering [15]. More sophisticated techniques were developed to deal with the multiplicative nature of the speckle noise following methodologies introduced in the context of Synthetic Aperture Radar (SAR) imaging. These include, for example, the work of Lee [14]. More recently, adaptations of more advanced filtering techniques were proposed to deal with

noise characteristics, e.g; complex diffusion [13]. The main drawback of these methods is that many of the structural details are not preserved. This is a common feature of linear filtering methods. This suggests the use of non-linear methods obtained naturally when using of local or even multiscale transforms [9]. Nevertheless, transform domain approaches [12] are more adapted to additive Gaussian white noise removal. Researchers also understood early how this limitation is not fully surmounted in ultrasound even when logarithmic functions is applied to transform the multiplicative noise model into an additive one. Achim *et al.* [2] were first to highlight the non-Gaussian statistics of wavelet coefficients of US images after logarithmic transformations, which makes the use of classical wavelet thresholding schemes not straightforward. Thus, they proposed to model the statistics of these coefficients using a family of heavy-tailed distributions, called alpha-stable distribution, and proposed an adapted thresholding scheme in a Bayesian framework. In the present work, we pursue the adjustment of wavelet-thresholding methods to deal with the statistics and the regularity of US data. It is surprising that research in speckle reduction has been focused on images without taking into account temporal information, while US imaging itself is often related to dynamic phenomena. One of the few contributions taking into account the temporal aspect is the recent work in [11]. Here, however, the spatial aspect is completely overlooked. We propose, therefore, an important extension of a despleckling method recently introduced in [8] for US images, in order to deal with spatio-temporal sequences. More precisely, we construct a *Kronecker* product based wavelet that Gaussianize the noise statistics in the transform domain. Using this novel *Kronecker* construction for the spatio-temporal wavelet atoms has two advantages: (i) It is well suited for sparsifying multi-dimensional signals that are measured progressively on subsets of the coordinates (US image sequences) and (ii) it makes it also possible to apply variance stabilization on subsets of the coordinates (e.g. spatial variables), this is of particular benefit for computational purposes. Such constructions were introduced by [7] in the context of compressive sensing. In the present paper, the proposed construction acts separately on the spatial and temporal variables. This yields to a simple scheme based on the following steps: (1) apply a spatial 2-dimensional wavelet transform to the spatio-temporal sequence, (2) use the low-frequency outputs of the wavelet transform as a local means estimation, (3) Fisz transform: use these local means to stabilize the variance of the wavelet coefficients, (3)*Kronecker* product: apply a temporal 1-dimensional wavelet transform to the stabilized coefficients, (4) perform thresholding on the coefficients obtained in step (3). The rest of the paper is organized as follows. We start by motivating the noise model we are considering by some statistical experiments on a real US sequence, in Section 2. In Section 3, we describe the proposed approach; in particular, we show how the combination of spatial Fisz transformation and the *Kronecker* product construction stabilizes the variance of the wavelet coefficients allowing the use of simple thresholding procedures. Finally, we provide experimental results on synthetic and real data in Section 4 to support our claims.

2 US image analysis & problem formulation

Before performing despeckling, a first step consists in choosing an adequate mathematical description of the noise component. A relevent noise model was introduced by Loupas [15]. It stipulates that the standard deviation of the speckle noise is proportional to a power of the intensity at each time point:

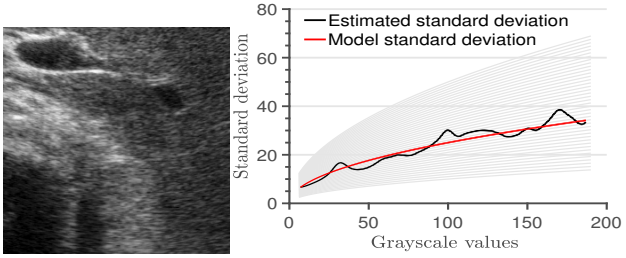
$$f_\varepsilon(\mathbf{x},t) = f(\mathbf{x},t)+\varepsilon(\mathbf{x},t),\ \varepsilon(\mathbf{x},.) \sim \mathcal{N}(0,\sigma f^\gamma(\mathbf{x},.)), \quad (1)$$

with $\gamma > 0$. $\mathbf{x} = (x_1, x_2)$ and t are respectively the spatial and temporal variable while f, ε and f_ε, are, respectively, the unknown sequence, the noise component and the observed corrupted sequence. To assess this model, we made a simple experiment; we estimated the shape of the variance function of the noise component using a pre-estimation based on temporal averaging and a Nadaraya-Watson estimator [17]. Let us describe this routine. We choose a *"Liver"* image sequence of 256×256 pixels and 128 time-points. We start by a applying a straightforward temporal averaging [1] to obtain a smooth image $f_\varepsilon^{\text{mean}}(\mathbf{x})$. By fixing a time point $\bar{t}$, we can obtain an estimation of the noise component on the current image $f(\mathbf{x},\bar{t})$ by computing, first, the residual:

$$\hat{\varepsilon}(\mathbf{x},\bar{t}) = f(\mathbf{x},\bar{t}) - f_\varepsilon^{\text{mean}}(\mathbf{x}), \quad (2)$$

and then evaluate the variance of $\hat{\varepsilon}$ using a Nadaraya-Watson estimator. We are interested in establishing a possible relation between gray-level intensity values and the variance. Thus, we start by defining a discrete uniform interval of intensities $\mathcal{G}$ with values ranging from $\min_{\mathbf{x}}(f_\varepsilon^{\text{mean}}(\mathbf{x}))$ to $\max_{\mathbf{x}}(f_\varepsilon^{\text{mean}}(\mathbf{x}))$. Then, for each point $g \in \mathcal{G}$, an estimation of the variance function h is given as:

$$h(g) = \frac{\langle K_b(g), \hat{\varepsilon}^2\rangle}{\hat{K}_b(g)}, \quad K_b(g) = \frac{1}{N^2 b} K\Big(\frac{f_\varepsilon^{\text{mean}} - g}{b}\Big), \quad (3)$$



(a) Single image (b) Standard deviation estimation

Fig. 1: The *"Liver"* sequence: The 40^{th} image of the sequence along with the estimated standard deviation and the theoretical one from the model introduced in [15].

where K_b is a Gaussian smoothing Kernel K of bandwidth b and N denotes the length of $\mathcal{G}$. We applied this routine to real US data; The *"Liver"* sequence with $b = 3$ was used. The results are shown in Fig. 1. We can observe that the estimated standard deviation (black line), $\sqrt{h}$, is not constant. This means that the noise component is not a white Gaussian, but it is rather depending on the underlying intensity. with $\gamma > 0$. The red curve in Fig. 1 is obtained with $\gamma = 1/2$ and $\sigma = 2.5$. The gray curves correspond to standard deviation plots with different values of σ ranging from 1 to 5. Model (1) was, for instance, exploited by Tenbrinck *et al.* [19] to perform robust histogram-based optical flow estimation in US images sequences. Coupé *et al.* [6] adapted the non-local means paradigm to this model. In [8], the authors of the present work showed how the wavelet-Fisz methodology introduced in [10] can be used in this context. In the next section, we recall this framework and we show how a structured sparse representation based on *Kronecker* product of wavelet bases can be used to perform wavelet-Fisz despeckling in dynamic US imaging.

3 Methodology

3.1 Notations

Let ψ and φ be two one-dimensional functions, commonly known as the mother wavelet and the scaling function, respectively. Note by $\psi_{j,k}(.) = 2^{j/2}\psi(2^j. - k)$ and $\varphi_{j,k}(.) = 2^{j/2}\varphi(2^j. - k)$ with $j, k \in \mathbf{N} \times \mathbf{Z}$ their dilated and translated versions. Then, one-dimensional wavelet bases are defined by the family

$$\{\varphi_{0,k}\}_{k\in\mathbf{Z}} \cup \{\psi_{j,k}\}_{(j,k)\in\mathbf{N}\times\mathbf{Z}}, \tag{4}$$

Any function $f \in L^2([0,1])$ can be decomposed on this basis leading to a set of approximation coefficients $a_{0,k}(f) = \langle f, \varphi_{0,k}\rangle$ and wavelet coefficients $w_{j,k}(f) = \langle f, \psi_{j,k}\rangle$. In two dimensional settings, the most common analogue construction to the one given in (4) relies on wavelet product given for translations $\mathbf{k} = (k_1, k_2)$ by:

$$\{\Phi_{0,\mathbf{k}}\}_{\mathbf{k}\in\mathbb{Z}^2} \cup \{\Psi^H_{j,\mathbf{k}}, \Psi^V_{j,\mathbf{k}}, \Psi^D_{j,\mathbf{k}}\}_{(j,\mathbf{k})\in\mathbb{N}\times\mathbb{Z}^2}, \tag{5}$$

where the functions $\Phi_{0,k_1,k_2}(\mathbf{x}) = \varphi_{0,k_1}(x_1)\varphi_{0,k_2}(x_2)$ lead to the approximation coefficients, while the horizontal, vertical and diagonal wavelet coefficients are obtained, respectively, by decomposing on:

$$\begin{aligned}
\Psi^H_{j,\mathbf{k}}(x_1, x_2) &= \varphi_{j,k_1}(x_1)\psi_{j,k_2}(x_2),\\
\Psi^V_{j,\mathbf{k}}(x_1, x_2) &= \psi_{j,k_1}(x_1)\varphi_{j,k_2}(x_2),\\
\Psi^D_{j,\mathbf{k}}(x_1, x_2) &= \psi_{j,k_1}(x_1)\psi_{j,k_2}(x_2).
\end{aligned} \tag{6}$$

The coefficients are computed in the same way as in the one-dimensional case. In the sequel, DWT and 2-D DWT will denote the discrete wavelet transforms leading to the wavelet decomposition in one and two dimensions. If not specified, $a_{j,\mathbf{k}}$ and $w_{j,\mathbf{k}}$ can refer to approximation and wavelet coefficients in one or two dimensions. In the sequel we denote $\{\Psi_{j,\mathbf{k}}\}_{(j,\mathbf{k})\in\mathbb{N}\times\mathbb{Z}^2} = \{\Psi^H_{j,\mathbf{k}}, \Psi^V_{j,\mathbf{k}}, \Psi^D_{j,\mathbf{k}}\}_{(j,\mathbf{k})\in\mathbb{N}\times\mathbb{Z}^2}$

3.2 Proposed approach

Consider we are observing an corrupted image sequence f_ε obeying model (1). By applying a 2- DWT to this model, with respect to spatial variables and fixing the time variable, we obtain:

$$w_{j,\mathbf{k}}(f_\varepsilon(.,t)) = w_{j,\mathbf{k}}(f(.,t)) + w_{j,\mathbf{k}}(\varepsilon(.,t)), \tag{7}$$

at each time point t. with $(j,\mathbf{k}) \in I \subset \mathbb{N} \times \mathbb{Z}^d$, $d = 1, 2$. The vectors $\{w_{j,\mathbf{k}}(\varepsilon(.,t))\}_{j,\mathbf{k}}$ are not Gaussian. Hence, the hard-thresholding procedure is not applicable directly. In the subsequent subsections we introduce the proposed approach for spatio-temporal despeckling in order to adapt the thresholding to the non-Gaussianity assumption.

3.2.1 Spatial wavelet-Fisz transform

We can notice that ε is asymptotically Gaussian in the sense that at fine scales, under local piece-wise constancy assumptions on f, f^γ is locally constant and thus ε and $\{w_{j,\mathbf{k}}(\varepsilon(.,t))\}_{j,\mathbf{k}}$ are locally Gaussian. To have a globally normal vector $\{w_{j,\mathbf{k}}(\varepsilon(.,t))\}_{j,\mathbf{k}}$, its variance should be stabilized for each couple $(j,\mathbf{k})$. A way to do so is the wavelet-Fisz procedure introduced by Fryzlewicz and his collaborators in a series of papers [10,16] for the one dimensional case. It consists in dividing the wavelet coefficients by local means computed on the

same supports as the wavelet bases. A fancy computation of these local means involves the use of approximation coefficients at each scale. If we denote two-dimensional function scaling:

$$\Phi_{j,\mathbf{k}}(x_1,x_2)=\varphi_{j,k_1}(x_1)\varphi_{j,k_2}(x_2), \tag{8}$$

Then, the scaling coefficients $\{a_{j,\mathbf{k}}\}_{j,\mathbf{k}}$ are simply given as:

$$a_{j,\mathbf{k}}(t)=\langle f_\varepsilon(.,t),\Phi_{j,\mathbf{k}}\rangle. \tag{9}$$

These coefficients can be seen as local approximations of f in the supports of the scaling and wavelet functions:

$$a_{j,\mathbf{k}}(t)\approx f_{j,k}(\mathbf{x},t)=f(2^jx_1-k_1,2^jx_2-k_2,t).$$

The Fisz-transformation consists in normalizing equation (7) by dividing it by $a^\gamma_{j,\mathbf{k}}$ for each couple $(j,\mathbf{k})$. Let us denote by $d_{j,\mathbf{k}}=\dfrac{w_{j,\mathbf{k}}}{a^\gamma_{j,\mathbf{k}}}$. Then:

$$d_{j,\mathbf{k}}(f_\varepsilon(.,t))=d_{j,\mathbf{k}}(f(.,t))+d_{j,\mathbf{k}}(\varepsilon(.,t)). \tag{10}$$

Note that in equation (10) the variable t is fixed. In the next paragraph, we address the temporal aspect by constructing a well-adapted spatio-temporal wavelet taking into account the variance stabilization described above and the differences in spatial and temporal regularities.

3.2.2 Kronecker Wavelet-Fisz despeckling

To incorporate the temporal aspect we apply a DWT to (10) with respect to the temporal variable. Integrating (10) against a temporal one-dimensional wavelet function $\psi_{l,m}$ gives:

$$d^{l,m}_{j,\mathbf{k}}(f_\varepsilon)=d^{l,m}_{j,\mathbf{k}}(f)+d^{l,m}_{j,\mathbf{k}}(\varepsilon), \tag{11}$$

with

$$\begin{aligned} d^{l,m}_{j,\mathbf{k}}(f_\varepsilon)&=\langle d_{j,\mathbf{k}}(f_\varepsilon(.,t)),\psi_{l,m}(t)\rangle,\\ d^{l,m}_{j,\mathbf{k}}(f)&=\langle d_{j,\mathbf{k}}(f(.,t)),\psi_{l,m}(t)\rangle,\\ d^{l,m}_{j,\mathbf{k}}(\varepsilon)&=\langle d_{j,\mathbf{k}}(\varepsilon(.,t)),\psi_{l,m}(t)\rangle. \end{aligned} \tag{12}$$

with $(j,\mathbf{k},l,m)\in\bar{I}\subset\mathbb{N}\times\mathbb{Z}^2\times\mathbb{N}\times\mathbb{Z}$. Equation (10) has the appropriate properties for wavelet thresholding. In particular the wavelet coefficients of the noise component have Gaussian distribution thanks to the following lemma.

Lemma 1 *Let $\{\Psi_{j,\mathbf{k}}\}_{(j,\mathbf{k})\in\mathbb{N}\times\mathbb{Z}^2}$ and $\{\psi_{l,m}\}_{(l,m)\in\mathbb{N}\times\mathbb{Z}}$ be two normalized wavelet bases, such that $\|\Psi\|_2^2=1$ and $\|\psi\|_2^2=1$. Let $u_{\underline{j},\underline{k}}$ denote the restriction of u to the support of the function $\psi_{\underline{j},\underline{k}}$. Assume that local approximations $a_{j,\mathbf{k}}$ converge to $f_{j,\mathbf{k}}$ as $j\to\infty$. Then we have:*

$$\Big\{d^{l,m}_{j,\mathbf{k}}(\varepsilon)\Big\}_{j,\mathbf{k},l,m}\xrightarrow{d}\mathcal{N}(0,\sigma),\quad as\ j\to\infty\ . \tag{13}$$

Proof We recall that $\varepsilon=f^\gamma(\mathbf{x},.)\epsilon$, with $\epsilon\sim\mathcal{N}(0,\sigma)$.

$$\begin{aligned} \mathrm{Var}\Big\{d^{l,m}_{j,\mathbf{k}}(\varepsilon)\Big\}&=\mathrm{Var}\Big\{\sum_t\frac{\psi_{l,m}(t)}{a^\gamma_{j,\mathbf{k}}(t)}\sum_{\mathbf{x}}\Psi_{j,\mathbf{k}}(\mathbf{x})f^\gamma(\mathbf{x},t)\epsilon\Big\},\\ &=\sigma^2\sum_t\frac{\psi^2_{l,m}(t)}{a^{2\gamma}_{j,\mathbf{k}}(t)}\sum_{\mathbf{x}}\Psi^2_{j,\mathbf{k}}(\mathbf{x})f^{2\gamma}(\mathbf{x},t),\\ &=\sigma^2\sum_t\frac{\psi^2_{l,m}(t)}{a^{2\gamma}_{j,\mathbf{k}}(t)}\sum_{\mathbf{x}}\Psi^2_{j,\mathbf{k}}(\mathbf{x})f^{2\gamma}_{j,\mathbf{k}}(\mathbf{x},t),\\ &=\sigma^2\sum_t\psi^2_{l,m}(t)\sum_{\mathbf{x}}\frac{f^{2\gamma}_{j,\mathbf{k}}(\mathbf{x},t)}{a^{2\gamma}_{j,\mathbf{k}}(t)}\Psi^2_{j,\mathbf{k}}(\mathbf{x}). \end{aligned}$$

Finally, since when $j\to\infty$, $a_{j,\mathbf{k}}$ converges to $f_{j,\mathbf{k}}$, then for each $\mathbf{x}$ in the support of $\Psi_{\underline{j},\underline{k}}$, we have:

$$\lim_{j\to\infty}\frac{f^{2\gamma}_{j,\mathbf{k}}(\mathbf{x},t)}{a^{2\gamma}_{j,\mathbf{k}}(t)}=1.$$

Thus: $\mathrm{Var}\Big\{d^{l,m}_{j,\mathbf{k}}(\varepsilon)\Big\}=\sigma^2||\psi||_2^2||\Psi||_2^2=\sigma^2$, which ends the proof.

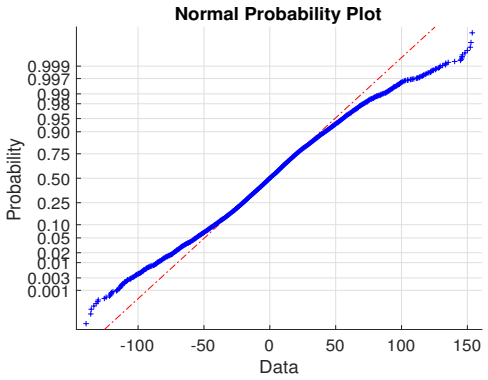
Now, using equation (11), we are able to perform wavelet thresholding on the entire $2D{+}t$ data. Precisely, we start by keeping only coefficients $d^{l,m}_{j,\mathbf{k}}(f_\varepsilon)$ larger than the threshold given $\sigma\sqrt{\log(\mathrm{Card}(\bar{I}))}$. Here, We tried to present a line of reasoning to help the reader to understand the succession of ideas. Nevertheless, we can link the final equation (11) to the initial model (1) by the use of an appropriate wavelet basis. Such a basis is the product of a 1-D temporal basis as in (4) and a rescaled version of the 2-D spatial basis in (5):

$$\{\Phi_{0,\mathbf{k}}\}\cup\{\frac{\Psi^H_{j,\mathbf{k}}}{\sigma a^\gamma_{j,\mathbf{k}}},\frac{\Psi^V_{j,\mathbf{k}}}{\sigma a^\gamma_{j,\mathbf{k}}},\frac{\Psi^D_{j,\mathbf{k}}}{\sigma a^\gamma_{j,\mathbf{k}}}\}\otimes\{\varphi_{0,m}\}\cup\{\psi_{l,m}\}, \tag{14}$$

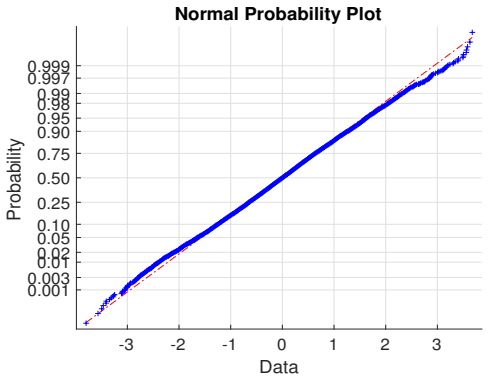
where $\otimes$ refers to the *Kronecker* product of vectors. This means that the resulting basis has all combinations of products between the elements of each basis. In particular, as the atoms are not obliged to be at the same scale to be crossed, this results in isotropic filters in the spatial plan and anisotropic filters in the space-time plan. As we showed, besides its ability to produce a structured sparsifying basis. This construction allows to apply variance stabilization techniques such as the Fisz-transformation on subsets of coordinates. It also enables the use of different families of wavelets on different variables to fit the directional regularity of the analyzed function. We finish this section by some remarks on discrete settings and the implementation of our approach that we coined *Kronecker* Wavelet-Fisz (KWF) Despeckling.

3.3 Implementation

Many of the choices we made for the construction of the KWF Despeckling approach were driven by practical reasons. The choice for hard-thresholding over the soft-thresholding paradigm is motivated by our use of the non-decimated wavelet transform (NDWT) [5] which has proven to outperform classical (decimated) wavelet transforms [18]. In the sequel, we denote by SNDWT the 2-D dimensional spatial NDWT and TNDWT the 1-D dimensional temporal NDWT. We denote their inverses, respectively, ISNDWT and ITNDWT. We should mention that the local means estimation $\{a^{\gamma}_{j,\mathbf{k}}\}_{j,\mathbf{k}}$ is not accurate at all scales. In particular, at very fine scales the observation is not large enough for an accurate estimation. Accordingly, at coarse scales, the rough piece-wise constancy constraint is not verified anymore. This issue is naturally handled by the hard-thresholding approach; small coefficients corresponding to fine scales are not kept as they are considered as noise, while coarse scale coefficients are kept as they are larger than the threshold. Finally, the *Kronecker* product procedure can be implemented as a sequence of two independent wavelet transforms [7]. Algorithm 1 describes the overall routine of the KWF Despeckling procedures. Fig. 2 shows the distribution of a portion of the coefficients $d^{l,m}_{j,\mathbf{k}}(f_\varepsilon)$, before and after stabilization, at the finest thresholding scale of the diagonal sub-band. The figure highlights the ability of the spatial Fisz-transform at normalizing the coefficients. This obviates the need to apply the Fisz transformation after the temporal analysis which motivates our approach.



(a) Without stabilization



(b) With stabilization

Fig. 2: Normal plot showing the distribution of the 3-dimensional wavelet coefficients $d^{l,m}_{j,\mathbf{k}}(f_\varepsilon)$, at the finest scale of the diagonal sub-band, with and without the stabilization step (*"Liver"* sequence).

Algorithm 1 KWF Algorithm

Input: $f_\varepsilon, \sigma, \gamma$
Output: Estimate $\hat{f}$
  $[w_{j,k}, a_{j,k}] \longleftarrow$ SNDWT(f_ε)
  for each couple (j,k) **do**
    $c_{j,k} = \sigma \times (a_{j,k})^{\gamma}$
    $d_{j,k} = |w_{j,k}|/c_{j,k}$
    $[d^{l,m}_{j,k}] \longleftarrow$ TNDWT$(d_{j,k})$
    for each couple (l,m) **do**
      if $d^{l,m}_{j,k} < T$ **then** $d^{l,m}_{j,k} = 0$
      end if
      $\hat{d}_{j,k} \longleftarrow$ ITNDWT$(d^{l,m}_{j,k})$
    end for
  end for
  $\hat{w}_{j,k} = \hat{d}_{j,k} \times c_{j,k}$
  $\hat{f} =$ ISNDWT$(\hat{w}_{j,k})$

3.4 A data-driven extension

Similarly to the majority of wavelet based approaches, the KWF Despeckling is adaptive, in the sense that no prior knowledge on the regularity of the unknown is needed. Still, the method is not data-driven as one needs to fix parameters σ and γ to run Algorithm 1. This is not desirable in practice. To overcome this limitation, it is possible to use variance estimation techniques. For example, by applying the Nadaraya-Watson estimation, for each point of a given discretization of the grayscale values range, a value of the variance is obtained (*cf.* 1). Then, the data-driven extension (D-DKWF) consists in replacing step 3 in Algorithm 1, where σ and γ are needed, by $c_{j,k} = \hat{h}(a_{j,k})$ where $\hat{h}$ is the estimated variance. This variance estimation technique was also used in the original work on the wavelet-Fisz methodology [10]. Since we have, now, a model to rely on, it would also be valuable to consider the fact that the variance function is expected to be non-decreasing. As in [10], this can be done by applying, an isotonic regression seeking for the non-increasing function that is the closest to the estimated variance in the mean square error sense by pooling on which the assumption is violated with their neighbors.

4 Experiments & Discussion

In this section we present some experiments to assess the performance of KWF Despeckling for both synthetic and real data. Note that the WF approach [8] has proven to perform as well as the state-of-the-art adaptation of non-local means paradigm [6] coined Optimized Bayesian non-local means (ONBLM). As this method does not have spatio-temporal extension in the literature, we restrict ourselves to a comparison with spatial WF and spatial ONBLM filtering[1]. Through these comparison, we aim essentially at showing the benefits of the spatio-temporal processing provided by the KWF. For all experiments, we used Daubechies wavelets with 6 vanishing moments in the spatial domain while Haar wavelets are used in the temporal coordinate. The use of Haar wavelets in the temporal domain favors piecewise constancy that can be observed over the sequence because of the preservation of pixel intensity values. For the OBNLM algorithm, the parameters α and M controlling the number of blocks and the size of the search window were fixed at 3 and 6, as in the original paper, and the filtering parameter was optimized for each noise level scenario.

4.1 Synthetic data

Our synthetic data is a clean sequence, *"Heart"*, of a beating heart provided by the simulation software in [4]. Even though the motion does not necessary fit the heart motion that can be observed in US imaging, this is not disturbing as we only consider the noise removal task. Here we assume that $\gamma = 1/2$. This is suggested by our experiments in Section 2. The same value was used, for instance, in [6]. Our experiments were conducted for $\sigma = \{1, 2, 3, 4\}$. Accordingly, the corresponding filtering parameter for OBNLM was fixed at $\{1, 1, 1.5, 2\}$. We evaluated the peak-signal-to-noise-ratio (PSNR) over the entire sequence. Moreover, to assess the details preservation we measured the structural similarity index measure (SSIM) [20]. The results are reported in Table I: The results show the constancy

σ	1	2	3	4
Noisy	29.76/0.78	23.74/0.56	20.21/0.42	17.72/0.32
WF	32.93/0.93	30.44/0.88	28.48/0.84	26.90/0.80
OBNLM	33.30/0.89	31.83/0.87	29.43/0.83	27.72/0.77
KWF	**34.31/0.96**	**33.28/0.95**	**32.27/0.93**	**31.34/0.91**

Table 1: Quantitative comparison (PSNR (dB)/SSIM) for different methods applied to *"Heart"* at different noise levels.

[1] Available on `https://sites.google.com/site/pierrickcoupe/softwares/denoising-for-medical-imaging/speckle-reduction/obnlm-package`

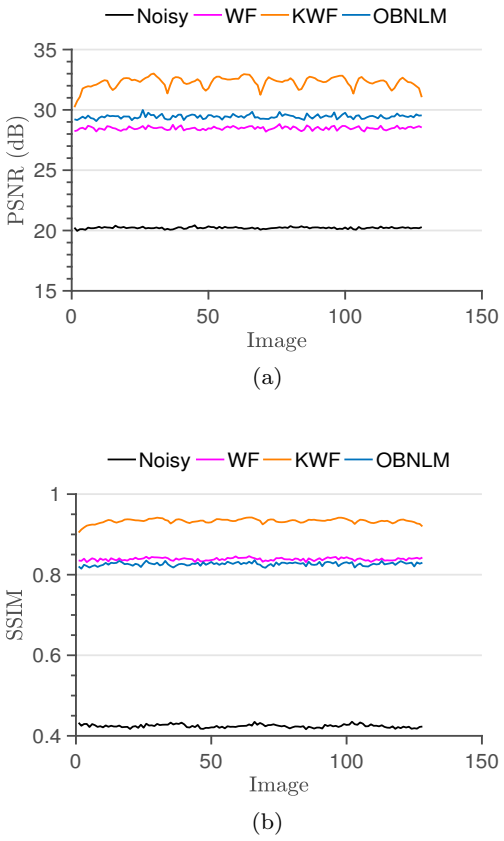


Fig. 3: Quantitative evolution ((a) PSNR & (b) SSIM) of different methods applied to the sequence *"Heart"* ($\sigma = 3$): frame by frame evolution.

of KWF despeckling at outperforming OBNLM and WF filtering, at a global level, for different values of σ. OBNLM gives good PSNR results as it is constructed to minimize the noise variance. For the SSIM measure, the two wavelet based techniques are better. Fig.3 helps to understand these performances, and the local behavior of the three approaches. This figure illustrates the PSNR and SSIM on the spatial coordinates at each time-point; for each image of the sequence. Note that the KWF approach also outperforms locally the WF approach. It is also interesting to not that the PSNR in the WF and OBNLM approach is stationary as each image is treated individually. On the other hand, the PSNR in the KWF shows variations in time. Specifically, we can observe a periodic behavior and the presence of two modes in each period. This is related to the cardiac motion. In fact, the discontinuities in motion affect the temporal regularity and hence the despeckling performance. This explains the drops in PSNR when the cardiac cycle phase switches from relaxation (diastole) to contraction (systole). Finally, a visual comparison is shown in Fig. 4. It highlights the visual improvement

(a) Original (b) Noisy ($\sigma = 2$) (c) WF (d) OBNLM (e) KWF

Fig. 4: Example of results obtained by the two methods for the 20^{th} image of the sequence *"Heart"*.

provided by the KWF. We can observe that the artifacts related to wavelet atoms are less visible and the contours are nicely preserved.

4.2 Real data

For our experiment on real data, we used the *"Liver"* sequence presented earlier and the *"Median"* sequence that shows the median nerve. We ran the algorithms with $\alpha = 2.5$ and $\gamma = 1/2$ as suggested by Fig. 1. We fixed $h = 1$ for the OBNLM filter. Fig. 5 and Fig. 6 show images in the original sequences along with the recovered images from the two approaches. Here again, we can observe that the output of the WF approach has more artifacts related to wavelet atoms, while the KWF and OBNLM results are more visually pleasant and less blurred. A video is provided, as an accompanying media to this paper, for a comparison over the sequences. Though the results are convincing, the choice of parameters σ and γ was visually estimated from the curve in 1. The D-DKWT makes it possible to perform despeckling, without fixing these parameters, using the variance estimation procedure. The results are shown in Fig. 7. Isotonic regression forces the estimated variance to be non-decreasing. This variance is plugged directly in the stabilization step. The despeckled results are even less blurred with sharp edges. A video of the D-DKWF is also available in the accompanying media. We are convinced that data-driven (blind) denoising approaches are very appreciated in medical imaging in general and in US imaging in particular, as they are easier to use in the image analysis pipeline effectuated by practitioners.

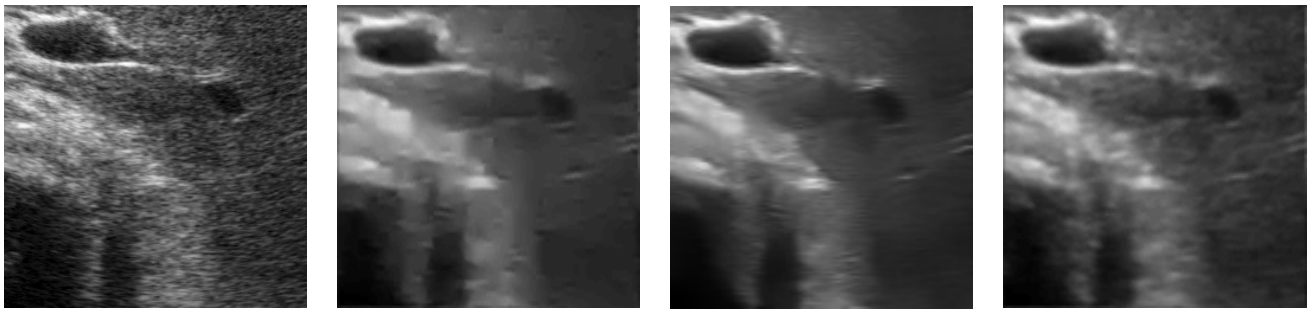
5 Conclusion

We presented the KWF and D-DKWF Despeckling algorithms for dynamic US imaging. These approaches are based on a combination of multi-scale variance stabilization and a structured sparse representation provided by the *Kronecker* product of spatial and temporal wavelet bases. The results confirmed our claims on the relevance of this construction. The data-driven extension, which we believe is of great interest in practice, give convincing results with incorporating priors on the model parameters. Future work will focus on performing simultaneously deconvolution and despeckling via the presented approach. A way to address this task is through a variational framework with a regularization term controlling the KWF coefficients and data-driven term taking into account the blur operator. Another possible improvement is the use of motion compensation techniques [3] to improve the spatio-temporal filtering in the presence of highly non-uniform motion.

Acknowledgements This work was supported by "Région Rhône-Alpes" under the ARC 6. L. Navarro's research is supported by the (ANR) under reference ANR-15-CE19-0002 (LBSMI). P. Delachartre is within the framework of the Labex CELYA (ANR-10-LABX-0060) and Labex PRIMES (ANR-11-LABX-0063) of the Université de Lyon.

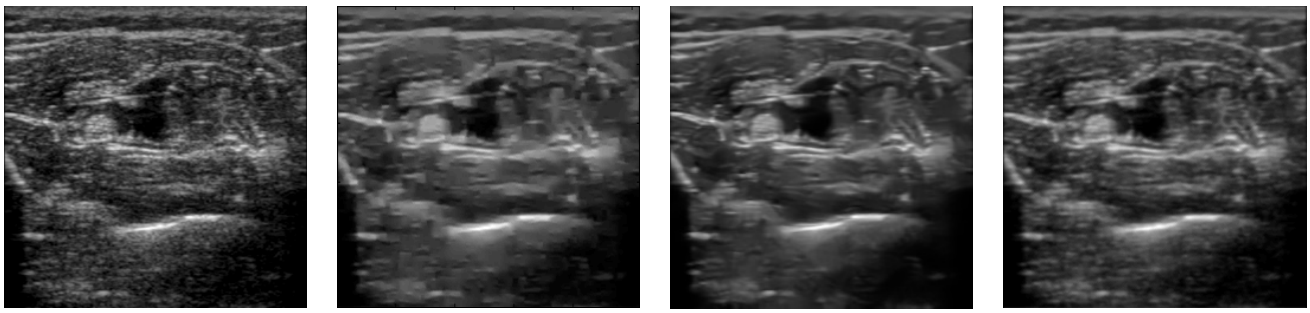
References

1. Abbott, J.G., Thurstone, F.: Acoustic speckle: Theory and experimental analysis. Ultrasonic imaging **1**(4), 303–324 (1979)
2. Achim, A., Bezerianos, A., Tsakalides, P.: Novel Bayesian multiscale method for speckle removal in medical ultrasound images. IEEE Trans. Med. Imag. **20**(8), 772–783 (2001)
3. Amiot, C., Girard, C., Chanussot, J., Pescatore, J., Desvignes, M.: Spatio-temporal multiscale denoising of fluoroscopic sequence. IEEE Trans. Med. Imag. **35**(6), 1565–1574 (2016)
4. Clarysse, P., Tafazzoli, J., Delachartre, P., Croisille, P.: Simulation based evaluation of cardiac motion estimation methods in tagged-mr image sequences. Journal of Cardiovascular Magnetic Resonance **13**(Suppl 1), P360 (2011)
5. Coifman, R.R., Donoho, D.L.: Translation-invariant denoising. in Antoniadis & Oppenheim, Wavelets and Statistics. Lecture Notes in Statistics pp. 125–150 (1995)
6. Coupé, P., Hellier, P., Kervrann, C., Barillot, C.: Nonlocal means-based speckle filtering for ultrasound images. IEEE Trans. Imag. Proc. **18**(7), 2221–2229 (2009)



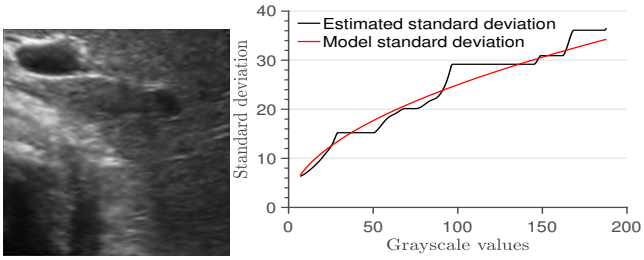
(a) Original (b) WF (c) OBNLM (d) KWF

Fig. 5: *"Liver"* sequence: the 30^{th} image of the sequence, original and despeckled using the two methods.



(a) Original (b) WF (c) OBNLM (d) KWF

Fig. 6: *"Median"* sequence: the 80^{th} image of the sequence, original and despeckled using the two methods.



(a) D-DKWF (b) Estimated standard deviation

Fig. 7: *"Liver"* sequence: the 40^{th} image of the sequence. (a): Estimated standard deviation after isotonic regression together with the theoretical standard deviation ($\alpha = 2.5$ and $\gamma = 1/2$). (b): The result of the data-driven (blind) approach.

7. Duarte, M.F., Baraniuk, R.G.: Kronecker compressive sensing. IEEE Trans. Imag. Proc. **21**(2), 494–504 (2012)
8. Farouj, Y., Freyermuth, J.M., Navarro, L., Clausel, M., Delachartre, P.: Hyperbolic wavelet-fisz denoising for a model arising in ultrasound imaging. IEEE Trans. Computational Imaging **3**(1), 1–10 (2017)
9. Forouzanfar, M., Moghaddam, H.A., Gity, M.: A new multiscale bayesian algorithm for speckle reduction in medical ultrasound images. Signal, Image and Video Processing **4**(3), 359–375 (2010)
10. Fryzlewicz, P.: Data-driven wavelet-Fisz methodology for nonparametric function estimation. Electronic Journal of Statistics **2**, 863–896 (2008)
11. Gifani, P., Behnam, H., Sani, Z.A.: Noise reduction of echocardiographic images based on temporal information. IEEE Trans. Ultrasonics, Ferroelectrics, and Frequency Control **61**(4), 620–630 (2014)
12. Gupta, D., Anand, R.S., Tyagi, B.: Despeckling of ultrasound medical images using ripplet domain nonlinear filtering. Signal, Image and Video Processing **9**(5), 1093–1111 (2015)
13. Jidesh, P., Bini, A.A.: Image despeckling and deblurring via regularized complex diffusion. Signal, Image and Video Processing **11**(6), 977–984 (2017)
14. Lee, J.S.: Digital image enhancement and noise filtering by use of local statistics. IEEE Transactins on Pattern Anal. Machine Intel. **2**(2), 165–168 (1980)
15. Loupas, T., McDicken, W., Allan, P.: An adaptive weighted median filter for speckle suppression in medical ultrasonic images. IEEE Trans. Circuits and Systems **36**, 129–135 (1989)
16. Motakis, E., Nason, G.P., Fryzlewicz, P., Rutter, G.: Variance stabilization and normalization for one-color microarray data using a data-driven multiscale approach. Bioinformatics **22**(20), 2547–2553 (2006)
17. Nadaraya, E.A.: On estimating regression. Theory of Probability & Its Applications **9**(1), 141–142 (1964)
18. Nason, G.: Wavelet Methods in Statistics with R. Springer Science & Business Media (2010)
19. Tenbrinck, D., Schmid, S., Jiang, X., Schäfers, K.P., Stypmann, J.: Histogram-based optical flow for motion estimation in ultrasound imaging. Journal of Mathematical Imaging and Vision **47**(1-2), 138–150 (2013)
20. Wang, Z., Bovik, A.C., Sheikh, H.R., Simoncelli, E.P.: Image quality assessment: From error visibility to structural similarity. IEEE Trans. Imag. Proc. **13**(4), 600–612 (2004)