Generalizing Wiener estimator to frame operators

Daniela De Canditiis

Istituto per le Applicazioni del Calcolo-CNR,
Via dei Taurini 19, 00185 Roma, Italy
d.decanditiis@iac.cnr.it

Abstract

In this paper the Wiener estimator for signal-denoising is generalized to finite frame operators. In particular, a two-stage procedure which results in a non-linear and non-diagonal estimator is proposed. Advantages and disadvantages with respect to the classical Wiener estimator used with orthonormal basis operator are discussed showing results on standard and real test signals.

Keywords: Frame-operators, Signal-denoising, Wiener-estimator.

1. Introduction

Among literature the best results for signal denoising are obtained using spatially adaptive statistical models, as [1], [2], [3] and [4]. However, in this paper a different point of view is considered avoiding any *a priori* statistical hypothesis on the unknown signal. The problem is that of recovering a deterministic signal $f \in R^n$, from the set of its contaminated samples $x_i = f_i + \delta_i$, $i = 1, ..., n$, with $\delta \sim N(0, \sigma^2 I_n)$.

When the noise level σ^2 is known or well estimated, the denoising problem divides into two sub problems:

- choice of the transformation which better separates the noise from the signal

- choice of the signal recovering method.

The first problem is usually pursued by looking for the sparsest representation of the underlying signal and it is not explored here since no a priori hypothesis are given for f. In this paper, we concentrate on the second problem given a specific representation. In particular, we deal with the case of a given frame representation. The advantage of frames with respect to classical orthonormal bases (e.g. dyadic wavelet, Fourier, polynomial) is that they can furnish an

efficient representation of a more broad class of signals as well as more adaptivity for their parsimonious representation, see [5]. For example, signals which have fast oscillating behavior as audio, speech, sonar, radar, EEG and stock market are much more well represented by a frame (with similar oscillating characteristic) than by a classical orthonormal basis, although the frame representation for such kind of signals can be not highly sparse. The disadvantage of working by frames with respect to orthonormal bases is that the synthesis and the analysis approaches are no more equivalent see [6] and the computational cost increases significantly. Here, we consider synthesis approaches, so that the data are transformed by the frame operator Wx, the signal coefficients are recovered/estimated in some way and then the signal is reconstructed by the pseudo-inverse matrix frame operator W^+ applied to the recovered coefficients. The pseudo-inverse operator, W^+, assures that the recovered signal is the minimum (L_2-norm) energy signal corresponding to the obtained coefficients and does not pursuit sparseness by definition. This observation is crucial since it guarantees that also small coefficients are considered as informative for the signal reconstruction, which can make the difference when the signal frame decomposition is not highly sparse. Sparsity is pursuit by the data frame coefficients processing through a Bayesian approach or a variational approach or a classical shrinkage/thresholding approach. A Bayesian approach promotes sparsity through specific prior hypothesis on frame coefficients, as for example in [7]. A variational approach looks for the best coefficients' vector which minimizes an objective function sum of a data fidelity term plus a penalization term which promotes sparsity, as for example the Basis Pursuit [5] and the Matching Pursuit [8] denoising algorithms. A shrinkage/thresholding approach works on data frame coefficients by procedures inspired to classical orthogonal bases based denoising schemes. For example, in [9] a frame based hard thresholding procedure is applied for seismic data denoising, or more elaborately, in [10] a Block thresholding procedure is proposed for audio signal denoising.

However, both in [9] and [10] the thresholding procedures are applied to the frame coefficients as they were obtained by an orthogonal transformation of the data, which is not the case neither in [9] where a non-dyadic WT (Wavelet Transform) is considered as well as in [10] where a STFT (Short Time Fourier Transform) is considered. Both cases are special case of frame transformations which need to be considered adequately when shrinkage/thresholding procedure are applied. From a practical point of view when a frame instead of an orthonormal bases is considered, the matrix operator W applied to the data vector x is a non square matrix instead of an orthonormal one, then the procedures need to be properly accommodated. In this respect, there are already some papers in literature which fill this gap, for example, [11] adapts to the case of a Gabor frame with a Blackman window the universal hard thresholding procedure, [12] adapts the SURE approach to the case of undecimated wavelet transform and image denoising, while [13] and [14] makes a survey of these procedures and in particular of the Wiener estimator for a general finite frame operator. In fact, in [13] the formal derivation of the classical Wiener estimator as the *best linear diagonal* estimator in the Mean Square Error (MSE) sense is obtained for a gen-

eral finite frame operator which include non-dyadic wavelet, Gabor, redundant basis as well as any matrix W with specific rank characteristic. Since the Wiener estimator results in an oracle procedure, i.e. it involves the unknown true signal, from a practical point of view it is not so appealing, this is the motivation of this paper which proposes a new two-stage procedure for signal-denoising based on the generalization of the Wiener filter to frame operator. In the proposed procedure, the unknown signal is first estimated by a fast thresholding scheme and then this first estimate is used to design the Wiener estimator. The proposed procedure has two principal advantages, first it brings to a closed form expression of the error, second, it generalizes the procedure successfully adopted in [15] for the classical Wiener estimator, employing one or two different frame operators.

The paper is organized as follows. In Section 1 the definition of the generalization of the Wiener estimator to frame operators is recalled. In Section 2 the new two-stage empirical Wiener estimator is presented. In Section 3 some numerical examples performed on both standard test and real signals are showed.

2. The generalized Wiener estimator

In this section we present the Wiener estimator generalized to the frame operators. First let us recall the definition of frames.

In the space of discrete signals of length n, a collection of N vectors $w_i \in R^n$, $i = 1, \cdots, N$, which together form matrix $W \in R^{N \times n}$, forms a *frame* if there exist two positive constants A and B such that, for any $f \in L_2(R^n)$ one has

$$A\|f\|^2 \leq f^* W^* W f \leq B\|f\|^2 \tag{1}$$

where W^* is the transpose of W, see [16]. The latter guarantees that eigenvalues of matrix $W^* W$ are bounded above and below and, therefore, $W^* W$ is invertible. N/n is the redundancy factor of the frame operator. If frame bounds are equal to each other and normalized, $A = B = 1$, then the frame is called *normalized tight*. In this case the generalized Parseval's identity holds and $W^* W = I_n$. Let us observe that in this definition are included any finite frame operator like non-dyadic wavelet, STFT, Gabor, redundant basis, union of basis as well as any matrix W which satisfies (1).

Consider, now, the classical denoising problem of recovering a vector $f \in R^n$ from its noisy observations

$$x = f + \delta, \quad \delta \sim N(0, \sigma^2 I_n), \tag{2}$$

and consider the following algorithm:

step 1 apply frame operator W to data x represented in (2), obtain y

$$y = \theta + \epsilon, \quad \epsilon \sim N(0, \sigma^2 U) \tag{3}$$

where $y = Wx$, $\theta = Wf$, $\epsilon = W\delta$ and $U = WW^* \in R^{N \times N}$.

3

step 2 process data frame coefficients y to get an estimate $\hat{\theta}$ of the unknown signal coefficients

step 3 estimate f by

$$\hat{f} = W^+\hat{\theta}, \tag{4}$$

where $W^+ = (W^*W)^{-1}W^*$ is the pseudo-inverse (Mooreenrose) of matrix W. Note that in the case of tight frame $W^+ = W^*$.

In classical denoising scheme, estimation of $\hat{\theta}$ in the *step 2* is achieved by shrinking or thresholding the empirical frame coefficients y. Among these procedures the Wiener estimator is the best linear diagonal estimator and it is obtained minimizing the MSE, i.e. minimizing $\mathbf{E}\|\hat{f} - f\|^2$. The next statement provides the analytic expression of the Wiener estimator for frame based operator. We assume that the vector of frame coefficients θ is estimated by $\hat{\theta} = \Gamma y$ where $\Gamma = diag(\gamma_1, \cdots, \gamma_N)$ is a data-independent diagonal matrix.

Theorem 1: Let data y follows model in (3) and let $\hat{f} = W^+\Gamma y$ be an estimator of f, then if vector $\gamma = diag(\Gamma)$ is data-independent, the minimum MSE is reached for

$$\gamma = \left(\theta\theta^* \circ U^- + \sigma^2 U \circ U^-\right)^{-1} \left(\theta\theta^* \circ U^-\right) e_N \tag{5}$$

where $U^- = (W^+)^*W^+$, e_N is the vertical vector with all entries equal 1 and $\circ$ denotes the Hadamard (i.e. element-wise) matrix product. If the frame is tight, then substitute U^- with U.

Proof of Theorem 1 is given in [13]. It is worth to observe that automatically the best *linear diagonal* estimator (5) is no more diagonal being each coefficient θ_i shrinked by γ_i which depends from other coefficients. The Wiener estimator becomes then an overlapping block shrinkage procedure where the length of the block is automatically decided by the correlation inherits from the frame operator. Let us consider the expression of the MSE:

$$\mathbf{E}\|\hat{f} - f\|^2 = \mathbf{E}Tr\left[W^+(\Gamma y - \theta)(\Gamma y - \theta)^*(W^+)^*\right] \tag{6}$$

and denote E_{opt} expression (6) when evaluated for matrix $\Gamma = diag(\gamma)$ given in (5). E_{opt} results to be the optimal minimum error reached when recovering the unknown function f by the estimator $\hat{f} = W^+\Gamma y$, with Γ a data-independent diagonal matrix.

Note that, in the case of orthonormal basis, W is an orthonormal matrix ($U^- = U = WW^* = I$), there is no redundancy ($N = n$) and expression (5) reduces to the well known Wiener estimator

$$\gamma = \left(\theta\theta^* \circ I_n + \sigma^2 I_n \circ I_n\right)^{-1} \left(\theta\theta^* \circ I_n\right) e_N, \tag{7}$$

which element-wise is $\gamma_i = \theta_i^2/(\theta_i^2 + \sigma^2)$. Moreover in this case

$$E_{opt} = Tr[(I_N - \Gamma)\theta\theta^*(I_N - \Gamma) + \sigma^2\Gamma^2] = \sum_{i=1}^{N} \sigma^2\theta_i^2/(\theta_i^2 + \sigma^2)$$

as showed in [17].

3. The empirical Wiener estimator

The Wiener estimator is an oracle estimator since the optimal diagonal matrix (5) requires the knowledge of the true signal f, i.e. of the true coefficients $\theta = Wf$. Hence to make it practical in *step 2* of the algorithm we propose a new two-stage empirical procedure which exploits recursive estimation: in a first stage, a first guess $\hat{\theta}$ of the unknown coefficients is obtained in some way then, in a second stage, $\hat{\theta}$ is plugged into eq. (5) obtaining $\hat{\Gamma}$ and hence a new estimate of the unknown coefficients comes out, $\hat{\hat{\theta}} = \hat{\Gamma}y$. The final estimator is of course given in *step 3* by $\hat{f} = W^+\hat{\Gamma}y$. This new procedures has two principal advantage, first it brings to a closed form expression of the error, second, it generalizes the WienerChop algorithm presented in [15]. Let us first look at the error by the following theorem which gives the expression for the MSE of the proposed estimator.

Theorem 2: Let y follows model in (3) and let $\hat{f} = W^+\hat{\Gamma}y$ be the empirical Wiener estimator, then

$$MSE = E_{opt} + E_{mis} \tag{8}$$

where E_{opt}, defined in Section 1, is the MSE of the oracle Wiener estimator and
$$E_{mis} = e_N^*(\Gamma - \hat{\Gamma})\left(\theta\theta^* \circ U^- + \sigma^2 U \circ U^-\right)(\Gamma - \hat{\Gamma})e_N.$$
proof:

We can write the MSE for the empirical Wiener estimator as

$$
\begin{aligned}
\mathbf{E}\|\hat{f} - f\|^2 &= \mathbf{E}Tr\left[W^+(\hat{\Gamma}y - \theta)(\hat{\Gamma}y - \theta)^*(W^+)^*\right] \\
&= Tr\left[W^+(I_N - \hat{\Gamma})\theta\theta^*(I_N - \hat{\Gamma})(W^+)^*\right] + Tr\left[W^+\hat{\Gamma}\mathbf{E}(\epsilon\epsilon^*)\hat{\Gamma}^*(W^+)^*\right] \\
&= Tr\left[U^-(I_N - \hat{\Gamma})\theta\theta^*(I_N - \hat{\Gamma})\right] + \sigma^2 Tr\left[\hat{\Gamma}U\hat{\Gamma}U^-\right],
\end{aligned}
$$

where we have used $y = \theta + \epsilon$, $\mathbf{E}(\epsilon) = 0$, $\mathbf{E}(\epsilon\epsilon^*) = \sigma^2 U$, $U^- = (W^+)^*W^+$ and the the classical identity $Tr[AB] = Tr[BA]$. By adding and subtracting Γ near the term $\hat{\Gamma}$ and using the properties of the trace operator we have the following expression $\mathbf{E}\|\hat{f} - f\|^2 =$

$$E_{opt} + Tr\left[U^-(\Gamma - \hat{\Gamma})\theta\theta^*(\Gamma - \hat{\Gamma}) + \sigma^2(\Gamma - \hat{\Gamma})U(\Gamma - \hat{\Gamma})U^-\right].$$

Finally, since for all matrices A and B and for all diagonal matrix Δ of appropriate size, it holds $Tr\left[A\Delta B\Delta\right] = \sum_{ij} a_{ij}\delta_j b_{ji}\delta_i = diag(\Delta)^*(A \circ B^*)diag(\Delta) = e_N^*\Delta A \circ B^*\Delta e_N$, the theorem remains proved.

Let us observe that, the error term E_{mis} is the error resulting from the mismatch between Γ and $\hat{\Gamma}$. In the case of ortnonormal basis E_{mis} simplifies to $e_N(\Gamma - \hat{\Gamma})\left(\theta\theta^* \circ I + \sigma^2 I\right)(\Gamma - \hat{\Gamma})e_N$, which element-wise reduces to $(\theta_i^2 + \sigma^2)(\gamma_i - \hat{\gamma}_i)^2$ as appeared in [17]. Since the matrix $(\theta\theta^* \circ U^- + \sigma^2 U \circ U^-)$ is non-negative definite (being the hadamard product of two non-negative matrices) then there exist eigen-vectors v_i, $i = 1, ..., N$ and eigen-values $\lambda_1 \geq \lambda_2 \geq ... \geq \lambda_N \geq 0$ such

that $E_{mis} = \sum_{i=1}^{N} \lambda_i < \gamma - \hat{\gamma}, v_i >^2$; then it is much more penalizing that the difference $\gamma - \hat{\gamma}$ is concorde to the eigen-vectors which correspond to the highest eigen-values. Similar conclusions for the mismatch error were obtained in [17] for the case of orthonormal basis, since in that case v_i are canonical vectors and $\lambda_i = \theta_i + \sigma^2$.

As observed in practice and in some more detailed discussion on error analysis in [18], it is more appropriate to use as first guess an overestimate of the unknown coefficients instead of an underestimate. For that reason and for computational reason, here we propose to give a first guess of the frame coefficients θ by a fast thresholding procedure, then to use it for the construction of the proposed two-stage algorithm.

Specifically, let us call $W1$ the frame matrix used in *step 1* to obtain the data frame coefficients y, then in *step 2* first obtain $\hat{\theta}_1$ by a fast thresholding estimator, then plug $\hat{\theta}_1$ into formula (5) obtaining a more accurate estimate $\hat{\hat{\theta}}_1 = \hat{\Gamma} y$, finally in *step 3* use $\hat{\hat{\theta}}_1$ to obtain $\hat{f} = W^+ \hat{\hat{\theta}}_1 = W^+ \hat{\Gamma} y$. The proposed procedure generalizes the one proposed in [15] and than it naturally leads to the generalization of the WienerChop algorithm of [15] to the case of frame operator. Indeed, as explained by the authors of [15], it can be advantageous to use a different frame operator $W2 \neq W1$ for constructing the first guess of the signal coefficients for improving the scheme's performance. Specifically, if $\hat{\theta}_2$ is the estimator obtained by a fast thresholding procedure applied to the data coefficients in frame $W2$, then expression $W1W2^+\hat{\theta}_2$ can be substituted into formula (5) in place of the true signal coefficients, obtaining an hybrid empirical Wiener estimator. Of course, this hybrid version is computationally more expensive then the first since it requires a synthesis plus an analysis operation more to transform estimator $\hat{\theta}_2$ into the $W1$ frame. However, as it will be seen in the numerical experiments, this strategy always gives performance improvement which sometime can justify its heavier computational cost.

We conclude, observing that the Empirical Wiener estimator is non-linear since it depends on an first guess estimation of the signal itself which is non-linear.

4. Experimental results

In this section we consider the rational-dilatation wavelet frame presented in [19] (here after denoted RDWT). It is a normalized tight frame (then $W^*W = I$ and $U = U^-$) and it is particularly suitable for fast oscillating signals. The RDWT frame is defined as $\left\{ (\frac{q}{p})^{k/2} \psi \left((\frac{q}{p})^k t + \frac{sp}{q} l \right) \right\}_{k,l \in Z}$ where ψ is a wavelet function and (p, q, s) is a triplet of parameters which gives the time-scale characteristic of the frame. In particular $q/p > 1$ is the rational dilatation factor and $\frac{p}{s(q-p)}$ is the redundant factor with $s > 1$. In particular, when $q = 2$, $p = 1$ and $s = 1$ the frame reduces to the classical (dyadic) wavelet basis, which is not the case of our experiments. Given a finite energy signal x of length n and $J \in N$ levels of decomposition, the RDWT transform is obtained by a sequence of proper

down-sampling operations and fast Fourier transforms; it ends up with $\lceil \frac{np^J}{q^J} \rceil$ scaling coefficients (low-pass filtering) and $\lceil \frac{np^j}{q^j s} \rceil$ wavelet coefficients (high-pass filtering) at each level $j=1,..J$.

The numerical simulations in this section are meant to be illustrative rather then exhaustive for the problem of signal denoising by frame, for that reason we concentrate on the different between classical and generalized Wiener estimator. We consider two standard test signals HISINE and LICHIRPS which have a fast oscillating characteristic and two real ones. The last being respectively a piece of a SPEECH signal and a piece of a GLOCK sound. The four test signals are reported in Fig.1. All the results are presented in terms of MSE obtained over 100 independent simulated error runs in the case of signal length $n = 1500$ and severe level of noise, i.e. with SNR (the ratio of the standard deviation of the signal and the standard deviation of the noise) equals to 1. In Table1 the difference between considering and ignoring the frame structure when evaluating the performance of the Wiener oracle estimator is showed. Specifically, we compare E_{opt} in the case the diagonal of Γ is given by (5) and in the case it is given by (7), denoting these estimators GW and CW respectively. In Table1 results are reported for the same family of frames with $p = 7$, $q = 8$ and with different level of redundancy, $s = 2, 3, 5, 7$. It is very clear to the reader how the difference between the two estimators reduces when the redundancy of the frame reduces, the difference would be exactly zero if an orthonormal basis were used with no redundancy. On the other hand, Table1 shows how increasing frame redundancy can improve MSE for those signals whose small coefficients are still important for reconstruction, as it happens for GLOCK signal. On the other hand, for compactly represented signals, as HISINE, a low-redundant frame (as well as it could be a simple orthonormal basis) is enough to get good results.

Next, we consider the performance of the empirical generalized Wiener estimator. In particular, the following procedure is applied to establish appropriate level-wise thresholds:

for $i = 1, .., I$

- generate a normalized Gaussian white noise vector b_i of length n

- apply the RDWT to b_i, to get frame coefficients $(r_i^{(1)}, ..., r_i^{(J)} r_i^{(J+1)})^* = W b_i$

- at each level j, evaluate $v_i^{(j)} = max(|r_i^{(j)}|)$

estimate thresholds $\lambda_j = \frac{1}{I} \sum_i v_i^{(j)}$. A first guess of the frame coefficients, $\hat{\theta}$, is estimated trough a level-wise Soft thresholding algorithm with thresholds λ_j; then $\hat{\theta}$ is plugged into the generalized estimator design (5) and into the classical one (7), we denote these estimators $eGW11$ and $eCW11$ respectively. The same procedures is also applied using a different frame $W2$ to get a first guess estimation of θ, then a frame changing matrix is applied $W1W2^*$ before plugging the first guess into the estimator designs, we denote these estimators $eGW21$ and $eCW21$ respectively (21 stands for the employment of two different
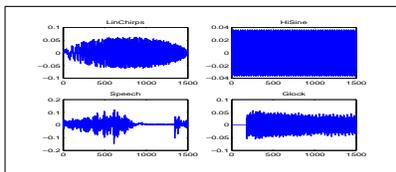
Figure 1: test signals.

Table 1: Results for the generalized oracle Wiener estimator and for the classical one. $p = 7$ $q = 8$

frame type	linchirps		hisine		speech		glock	
	GW	CW	GW	CW	GW	CW	GW	CW
$s = 2$	0.1052	0.1959	0.1713	0.2691	0.1613	0.2252	0.0659	0.1191
$s = 3$	0.1093	0.1422	0.0749	0.1299	0.1651	0.2083	0.0583	0.0738
$s = 5$	0.1207	0.1289	0.0677	0.0718	0.1689	0.1869	0.1410	0.1540
$s = 7$	0.1371	0.1404	0.0706	0.0731	0.1775	0.1817	0.1586	0.1604

frame matrices). Finally, estimator BP is considered: it is defined as $\hat{f} = W^*\hat{\theta}$, where $\hat{\theta} = argmin_\theta ||\theta||_1$ subject to $||W^*\theta - x||_2 \leq \sigma\sqrt{n}\sqrt{(1 + 2\sqrt{2}/\sqrt{n})}$ is obtained by the log-barrier method implemented in *l1qc-logbarrier* available at *http://users.ece.gatech.edu/ justin/l1magic/*.

In Table 2 results for BP, $eGW11$, $eCW11$, $eGW21$ and $eCW21$ are reported in the case of $SNR = 1$, as well as the average computing time in *sec* employed for estimating $\hat{f}$ on a 2.2 GHz processor. Table 2 shows how the generalized Wiener estimator, at the price of a significant increment of computational time, outperforms the classical Wiener estimator when such kind of first guess is considered. Moreover, results confirms the advantage of hybrid procedures for both classical (as already noted in [17]) and generalized Wiener estimators. Finally, it is instructive to note that the searching strategy of BP estimator is computationally more expensive then the generalized empirical Wiener estimator which has an analytic closed form expression.

Table 2: $p = 7, q = 8, s = 3, J = 10$ for $W1$ and $p = 7, q = 8, s = 4, J = 10$ for W2.

SNR	INPUT	BP	$eCW11$	$eGWn11$	$eCW21$	$eGW21$
linchirps						
MSE	0.9989	0.6276	0.6225	0.5299	0.6236	0.4952
time		246.5121	0.0190	44.1518	4.0449	47.0566
hisine						
MSE	0.9989	0.5714	0.7577	0.6096	0.6936	0.4109
time		378.4901	0.0177	47.7936	4.0518	59.2092
speech						
MSE	0.9623	0.6303	0.6099	0.5592	0.5922	0.5169
time		249.5071	0.0171	37.4308	4.0369	45.7847
glock						
MSE	1.0021	0.6910	0.2234	0.1512	0.2098	0.1130
time		244.6096	0.0195	42.6305	4.0516	44.0899

References

[1] J. Portilla, V. Strela, M.J. Wainwright and E.P. Simoncelli, "Image denoising using scale mixtures of Gaussians in the wavelet domain", *IEEE Transactions on Image Process.*, vol.12 n.11, pp. 1338-1351, 2003.

[2] S.A.G. Chang, B. Yu and M. Vetterli, "Adaptive wavelet thresholding for image denoising and compression", *IEEE Trans. Image Processing*, vol. 9, pp. 1532-1546, 2000.

[3] A. Pizurica and W. Philips, "Estimating the probability of the presence of a signal of interest in mutiresolution single- and multiband image denoising" *IEEE Trans Image Process.*, vol. 15 n.3, pp. 654-665, 2006.

[4] L. Sendur and I.W. Selesnick, "Bivariate shrinkage with local variance estimation", *IEEE Signal Process. Lett.*, vol. 9 n.12, pp.438-441, 2002.

[5] S.S.Chen, D. Donoho and M.A. Saunders, "Atomic decomposition by basis pursuit", *SIAM J. Sci. Comput.*, vol.20,no.1,pp.33-61,1999.

[6] I.W. Selesnick and M.A.T. Figueiredo, "Signal restoration with overcomplete wavelet transforms: comparison of Analysis and Synthesis priors", *Proc. SPIE* 7446, 74460D; http://dx.doi.org/10.1117/12.826663, 2009.

[7] P.J. Wolfe, S.J. Godsill and W-J. Ng, "Bayesian variable selection and regularization for time-frequency surface estimation", *J. R. Statistic Soc. B*, vol. 66, pp. 575-589, 2004.

[8] S. Mallat and Z. Zhang, "Matching pursuits with time-frequency dictionary", *IEEE Trans. Signal Process*, vol.41, 3397-3415, 1993.

[9] R. Zhang and T.J. Ulrych, "Physical wavelet frame denoising", *Geophysiscs*, vol. 68, n.1, 2251, 2003.

[10] G. Yu, S. Mallat and E. Bacry, "Audio denoising by time-frequency block thresholding", *IEEE Trans. on Sig. Proc.* vol 56, no.5, 2008.

[11] J.S. Walker and Y.-J. Chen, "Denoising Gabor transforms", *available at http://www.uwec.edu/ walkerjs/primer/pages/BooksPapers.htm*, 2006.

[12] T. Blu and F. Luisier, "The SURE-LET approach to image denoising", *IEEE Trans. Image Process.*, vol. 16 n. 11, 2007.

[13] D. De Canditiis, M. Pensky and P. J. Wolfe, "Denoising by frames", *manuscript available at http://www.iac.rm.cnr.it/ danielad/*

[14] D. De Canditiis, "A frame based shrinkage procedure for fast oscillating functions", *Comp. Stat. Data Analysis*, vol. 75, 142-150, 2014.

[15] S.P. Ghael, A.M. Sayeed and R.G. Baraniuk, "Improved wavelet denoising via empirical wiener filtering", *Proceedings of SPIE*, 1997.

[16] S. Mallat, "A wavelet tour of signal processing", Academic Publishing, 1999

[17] H. Choi and R.G. Baraniuk, "Analysis of wavelet-domain wiener filter", Proceeding of IEEE-SP 1998.

[18] Y.Ephraim and D. Malah "Speech enhancement using a minimum mean-square error short-time spectral amplitude estimator", *IEEE. Trans. Acoust., Speech, Signal process.*, vol.32, no.6, pp.1109-1121, Dec.1984

[19] I.Bayram and I.W. Selesnick, "Frequency-Domain design of overcomplete rational-dilation wavelet transform", *.IEEE trans. Signal Processing*, vol. 57, n.8, pp. 2957-2972, 2009.