Wavelet based image decomposition by variational functionals

I. Daubechies[a] and G. Teschke[b]

[a]Princeton University, PACM, Washington Road, Princeton, NJ 08544-1000, USA
[b]University of Bremen, ZeTeM, P.O.Box 33 04 40, 28334 Bremen, Germany

ABSTRACT

We discuss a wavelet based treatment of variational problems arising in the context of image processing, inspired by papers of Vese–Osher and Osher–Solé–Vese, in particular, we introduce a special class of variational functionals, that induce a decomposition of images in oscillating and cartoon components. Cartoons are often modeled by BV functions. In the setting of Vese et.el. and Osher et.al. the incorporation of BV penalty terms leads to PDE schemes that are numerically intensive. We propose to embed the problem in a wavelet framework. This provides us with elegant and numerically efficient schemes even though a basic requirement, the involvement of the space BV, has to be softened slightly. We show results on test images of our wavelet algorithm with a $B^1_1(L_1)$ penalty term, and we compare them with the BV restorations of Osher–Solé–Vese.

Keywords: Contour and texture analysis, near BV restoration, non-linear wavelet filtering

1. INTRODUCTION

One important problem in image processing is the reconstruction of the 'true' image from an observation. In almost all applications the observation is a noisy and blurred version of the true image. In principle, this problem can be understood as an inverse problem. Consequently, the reconstruction can be done be means of regularization techniques and minimization of related variational functionals.

In this paper we consider a special class of variational problems which induce a decomposition of images in oscillating and cartoon components. Vese–Osher and Osher–Solé–Vese suggested incorporating BV penalty terms in the variational treatment of this problem in order to allow edges and contours in the reconstructed images. However, the resulting PDE based scheme is numerically intensive. Solving the variational problem in a wavelet regime would provide efficient techniques, even though the current state of the art of wavelet theory does not allow the involvement of the space BV, but of the smaller space $B^1_1(L_1)$. Nevertheless, despite this drawback, it seems to be worthwhile to apply wavelet based methods. Here we show how this wavelet method works, and we investigate the discrepancy between BV restoration and the wavelet based method for reconstructing 'near' BV image components.

In order to give a description of the underlying variational problems, we recall the methods proposed by Vese–Osher and Osher–Solé–Vese.[1, 2] They follow the idea of Y. Meyer in a total variation framework of L. Rudin, S. Osher and E. Fetami. In principle, the models can be understood as a decomposition of an image of the form $f = u + v$, where u represents the cartoon part and v the texture part. In the Vese–Osher model[1] the decomposition is given by $\inf_{u,g_1,g_2} G_p(u, g_1, g_2)$, where

$$G_p(u, g_1, g_2) = \int_\Omega |\nabla u| + \lambda \|f - (u + \mathrm{div} g)\|^2_{L_2(\Omega)} + \mu \| |g| \|_{L_p(\Omega)} \ , \tag{1.1}$$

where $f \in L_2(\Omega)$, $\Omega \subset \mathbf{R}^2$, and $v = \mathrm{div} g$ with $g = (g_1, g_2) \in L_\infty(\Omega) \times L_\infty(\Omega)$ (i.e. $p \to \infty$). The first term is the total variation of u. If $u \in L_1$ and $\int_\Omega |\nabla u| < \infty$, then $u \in BV(\Omega)$. This space allows discontinuities, therefore edges and contours generally appear in u. The minimization of (1.1) is performed with respect to the unknown components u and g. To model v they use the space of oscillating functions introduced by Y. Meyer which is in some sense dual to $BV(\Omega)$.[3] This model was extensively studied and many convincing test examples were presented. Under the assumption that $g = \nabla P + Q$, where P is a single–valued function and Q

is a divergence–free vector field, Osher–Solé–Vese show how the v–penalty term can be expressed by means of Sobolev norms.[2] By setting $p = 2$, one has

$$\||g|\|_{L_2(\Omega)} = \left(\int_\Omega |\nabla(\Delta)^{-1} v|^2 \right)^{1/2} = \|v\|_{H^{-1}(\Omega)} \ .$$

This H^{-1} calculus is allowed as long as we deal with zero mean oscillatory texture/noise components. With this simplification, the functional (1.1) reduces to

$$\tilde{G}_p(u, v) = \int_\Omega |\nabla u| + \lambda \|f - (u + v)\|^2_{L_2(\Omega)} + \mu \|v\|_{H^{-1}(\Omega)} \ , \tag{1.2}$$

In the H^{-1} framework Osher–Solé–Vese have also proposed the following model

$$\inf_u \mathcal{F}_f(u) = \int_\Omega |\nabla u| + \lambda \|f - u\|^2_{H^{-1}(\Omega)} \ , \tag{1.3}$$

which leads to an equivalent fourth order Euler-Lagrange PDE. In general, one drawback is that the minimization of 1.1 as well as 1.3 lead to numerically intensive schemes.

Instead of solving PDE systems, we propose a wavelet based treatment of problem (1.1). We are encouraged by the fact that many function spaces of interest can be characterized by means of wavelet coefficients. Moreover, it is well-known that elementary methods based on wavelet shrinkage solve extremal problems, e.g. problem (1.3) where $BV(\Omega)$ is replaced by the Besov space $B^1_1(L_1(\Omega))$. Since $BV(\Omega)$ can not be simply described in terms of wavelet coefficients, it is not clear that $BV(\Omega)$ minimizers can be obtained in this way. When one restricts oneself to Haar wavelets, it is shown, exploiting $B^1_1(L_1(\Omega)) \subset BV(\Omega) \subset B^1_1(L_1(\Omega)) - weak$, that methods based on the involvement of Haar systems provide near $BV(\Omega)$ minimizers.[4] However, so far there exists no closed theory incorporating general wavelet systems. We naturally propose to replace $BV(\Omega)$ by $B^1_1(L_1(\Omega))$. Moreover, we limit ourselves to the case $p = 2$, i.e. we aim at incorporating the H^{-1} norm. Altogether this leads to a somewhat different variational problem

$$\inf_{u,v} \mathcal{F}_f(v, u) = \|f - (u + v)\|^2_{L_2(\Omega)} + \lambda \|v\|^2_{H^{-1}(\Omega)} + 2\alpha |u|_{B^1_1(L_1(\Omega))} \ , \tag{1.4}$$

where the minimization is performed with respect to u and v.

This paper is organized as follows. In Section 2 we summarize some results on wavelets and function spaces. Section 3 is mainly concerned with minimizing the variational functional and with creating a somewhat advanced wavelet scheme. In Section 5 we present results obtained with our wavelet based model, and, moreover, we discuss and compare the results with other proposed models.

2. PRELIMINARIES

In this section, we briefly recall some facts on wavelets which are needed later on. For a comprehensive introduction and in order to solve the variational problem (1.4) we would like to incorporate smoothness characterization properties of wavelets, which means that one can determine the membership of a function in many different function spaces by examining its wavelets coefficients.[5–12]

Suppose H is a Hilbert space. Let $\{V_j\}$ be a sequence of closed nested subspaces of H whose union is dense in H while their intersection is zero. By the requirements $f \in V_j \leftrightarrow f(M^j \cdot) \in V_0$ and V_0 is shift–invariant the sequence $\{V_j\}$ forms a multi-resolution analysis. In many cases of practical relevance the spaces V_j are spanned by single scale bases $\Phi_j = \{\phi_{j,k} : k \in I_j\}$ which are uniformly stable. Successively updating a current approximation in V_j to a better one in V_{j+1} can be facilitated if stable bases $\Psi_j = \{\psi_{j,k} : k \in J_j\}$ for some complement W_j of V_j in V_{j+1} are available. Any $f_n = \sum_{k \in I_n} c_{n,k} \phi_{n,k} \in V_n$ has then an alternative multi-scale representation $f_n = \sum_{k \in I_0} c_{0,k} \phi_{0,k} + \sum_{j=0}^{n} \sum_{k \in J_j} d_{j,k} \psi_{j,k}$. Of course, there is a continuum of possible choices of such complements. The essential constraint on the choice of W_j is that $\Psi = \bigcup_j \Psi_j$ forms a Riesz-basis of H, i.e. every $f \in H$ has a unique expansion

$$f = \sum_j \sum_{k \in J_j} \langle f, \tilde{\psi}_{j,k} \rangle \psi_{j,k} \;\; \text{such that} \;\; \|f\|_H \sim \left(\sum_j \sum_{k \in J_j} |\langle f, \tilde{\psi}_{j,k} \rangle|^2 \right)^{\frac{1}{2}} , \tag{2.1}$$

where $\tilde{\Psi}$ forms a bi-orthogonal system and is in fact also a Riesz-basis for H.[5]

In order to adjust this setting to the needs of image processing, we assume that any function $f \in L_2(I)$ can be extended periodically to all of $\mathbf{R}^2$. Here I is assumed to be the unit square $(0,1]^2$. The periodization is performed by mirroring f down- and sidewards (now supported in $\Omega = [-1,1]^2$) followed by k–shifts, where $k \in 2\mathbf{Z}^2$. The resulting function f represents our image. Throughout this paper we only consider compactly support wavelet systems, e.g., such as Daubechies' orthogonal wavelets[6] or symmetric bi-orthogonal wavelets by Cohen, Daubechies, and Feauveau.[7] We assume that we have scaling functions ϕ, $\tilde{\phi}$ and wavelets ψ_i, $\tilde{\psi}_i$, where $i = 1,2,3$, constructed by tensor products of one-dimensional bi-orthogonal wavelet systems. Moreover, it will be convenient to introduce the following notation. Let

$$J := \{\lambda = (i,j,k) : k \in J_j, j \in \mathbf{Z}, i = 1,2,3\},$$

$$J_{j_0} := \{\lambda = (i,j,k) : k \in J_j, j \geq j_0, i = 1,2,3\} ,$$

and define $|\lambda| := j$ if $\lambda \in J_j$; then any function $f \in L_2(\Omega)$ can be represented by

$$f = \sum_{\lambda \in J} \langle f, \tilde{\psi}_\lambda \rangle \psi_\lambda = \sum_{k \in I_{j_0}} \langle f, \tilde{\phi}_{j_0,k} \rangle \phi_{j_0,k} + \sum_{\lambda \in J_{j_0}} \langle f, \tilde{\psi}_\lambda \rangle \psi_\lambda . \tag{2.2}$$

We are interested in characterizations of Besov spaces.[12] For $\beta > 0$ and $0 < p, q \leq \infty$ the Besov space $B_q^\beta(L_p(\Omega))$ of order β is the set of functions

$$B_q^\beta(L_p(\Omega)) = \{f \in L_p(\Omega) : \; |f|_{B_q^\beta(L_p(\Omega))} < \infty\} ,$$

where $|f|_{B_q^\beta(L_p(\Omega))} = \left(\int_0^\infty (t^{-\beta}\omega_l(f;t)_p)^q dt/t\right)^{1/q}$ and ω_l denotes the l-th modulus of smoothness, $l > \beta$. This space is equipped with the norm $\|f\|_{B_q^\beta(L_p(\Omega))} = \|f\|_{L_p(\Omega)} + |f|_{B_q^\beta(L_p(\Omega))}$. However, what is important to us is that one can determine whether a function is in $B_q^\beta(L_p(\Omega))$ simply by examining its wavelet coefficients. Here we are only interested in the especially simple case $p = q$. To this end, suppose that ϕ has R continuous derivatives and ψ has vanishing moments of order M. Then, as long as $\beta < \min(R, M)$, one can prove that for all $f \in B_p^\beta(L_p(\Omega))$ one has

$$|f|_{B_p^\beta(L_p(\Omega))} \sim \left(\sum_{\lambda \in J_{j_0}} 2^{|\lambda| sp} |\langle f, \tilde{\psi}_\lambda \rangle|^p \right)^{1/p} \tag{2.3}$$

and

$$\|f\|_{B_p^\beta(L_p(\Omega))} \sim |f|_{B_p^\beta(L_p(\Omega))} + \left(\sum_{k \in I_{j_0}} |\langle f, \tilde{\phi}_{j_0,k} \rangle|^p \right)^{1/p} , \tag{2.4}$$

where $s = \beta + 1 - 2/p$. In case of $p = q = 2$ the space $B_2^\beta(L_2(\Omega))$ is the Bessel potential space $H^\beta(\Omega)$. In analogy with the special case of Bessel potential spaces $H^\beta(\Omega)$, the Besov space $B_p^\beta(L_p(\Omega))$ with $\beta < 0$ is understood as the dual space of $B_{p'}^{\beta'}(L_{p'}(\Omega))$, where $\beta' = -\beta$ and $1/p + 1/p' = 1$. Here we are interested in $B_1^1(L_1(\Omega))$, $B_2^{-1}(L_2(\Omega)) = H^{-1}(\Omega)$, and $B_2^0(L_2(\Omega)) = L_2(\Omega)$.

3. VARIATIONAL PROBLEM AND MINIMIZATION

In this section, we consider the family of variational problems (1.4) that naturally give rise to a parametrized class of solutions: Given positive parameters (γ, α), the Besov space $B_1^1(L_1(\Omega))$, and the dual Bessel potential space $H^{-1}(\Omega)$; find functions $\tilde{v}_{\gamma,\alpha} \in H^{-1}(\Omega)$ and $\tilde{u}_{\gamma,\alpha} \in B_1^1(L_1(\Omega))$ minimizing the functionals

$$\mathcal{F}_f(v,u) = \|f-(u+v)\|^2_{L_2(\Omega)} + \lambda\|v\|^2_{H^{-1}(\Omega)} + 2\alpha|u|_{B_1^1(L_1(\Omega))} \ .$$

Applying the stability property (2.1) and the smoothness characterization of wavelets, see 2.3, 2.4, we have

$$\begin{aligned} \|f-(u+v)\|^2_{L_2(\Omega)} &\sim \sum_{\lambda\in J} |f_\lambda - (u_\lambda + v_\lambda)|^2, \\ \|v\|^2_{H^{-1}(\Omega)} &\sim \sum_{\lambda\in J} 2^{-2|\lambda|}|v_\lambda|^2, \\ |u|_{B_1^1(L_1(\Omega))} &\sim \sum_{\lambda\in J_{j_0}} |u_\lambda|, \end{aligned}$$

where f_λ, v_λ, u_λ denote the λ–th wavelet coefficients. Combining these sequence sums, we have the following equivalent convex sequence based functional

$$\mathcal{S}_f(v,u) = \sum_{\lambda\in J}\left(|f_\lambda - (u_\lambda+v_\lambda)|^2 + \gamma 2^{-2|\lambda|}|v_\lambda|^2 + 2\alpha|u_\lambda|\cdot 1_{\{\lambda\in J_{j_0}\}}\right) \ . \tag{3.1}$$

The sequence space variational functional (3.1) can be minimized by minimizing separately for each term. Let $[\cdot]_\lambda$ denote the λ–th addend in (3.1), then the derivative with respect to v_λ is given by

$$D_{v_\lambda}\left[\mathcal{S}_f(v,u)\right]_\lambda = -2(f_\lambda - u_\lambda) + 2(1+\gamma 2^{-2|\lambda|})v_\lambda \ .$$

Consequently, the λ–th wavelet coefficient of the minimizer $\tilde{v}_{\gamma,\alpha}$ has to satisfy

$$(\tilde{v}_{\gamma,\alpha})_\lambda = (f_\lambda - u_\lambda)(1+\gamma 2^{-2|\lambda|})^{-1} \ . \tag{3.2}$$

Replacing v_λ by (3.2) in $\left[\mathcal{S}_f(v,u)\right]_\lambda$ yields

$$\left[\mathcal{S}_f(\tilde{v}_{\gamma,\alpha},u)\right]_\lambda = \mu_{1,\lambda}(f_\lambda - u_\lambda)^2 + 2\mu_{2,\lambda}|u_\lambda| \ , \tag{3.3}$$

where $\mu_{1,\lambda} = \gamma 2^{-2|\lambda|}(1+\gamma 2^{-2|\lambda|})^{-1}$ and $\mu_{2,\lambda} = \alpha\cdot 1_{\{\lambda\in J_{j_0}\}}$. Hence, the derivative of (3.3) with respect to u_λ is given by

$$D_{u_\lambda}\left[\mathcal{S}_f(\tilde{v}_{\gamma,\alpha},u)\right]_\lambda = -2\mu_{1,\lambda}(f_\lambda - u_\lambda) + 2\mu_{2,\lambda}\mathrm{sgn}(u_\lambda) \ .$$

Consequently, the wavelet coefficients of the minimizer $\tilde{u}_{\gamma,\alpha}$ must fulfill

$$(\tilde{u}_{\gamma,\alpha})_\lambda = f_\lambda - \frac{\mu_{2,\lambda}}{\mu_{1,\lambda}}\mathrm{sgn}((\tilde{u}_{\gamma,\alpha})_\lambda) \ \text{, i.e. } \ (\tilde{u}_{\gamma,\alpha})_\lambda = S_{\mu_{2,\lambda}/\mu_{1,\lambda}}(f_\lambda) \ , \tag{3.4}$$

where S_t denotes the soft threshold operator with threshold t. Finally, by replacing u_λ by $(\tilde{u}_{\gamma,\alpha})_\lambda$ in (3.2), we obtain

$$(\tilde{v}_{\gamma,\alpha})_\lambda = \left(f_\lambda - S_{\mu_{2,\lambda}/\mu_{1,\lambda}}(f_\lambda)\right)(1+\gamma 2^{-2|\lambda|})^{-1} \ . \tag{3.5}$$

Altogether, we summarize the results (3.5) and (3.4) in the following

Proposition 3.1. *Let f be a given function. The functional* (1.4) *is minimized by the parametrized class of functions $\tilde{v}_{\gamma,\alpha}$ and $\tilde{u}_{\gamma,\alpha}$ given by the following non-linear filtered wavelet series of f*

$$\tilde{v}_{\gamma,\alpha} = \sum_{\lambda\in J_{j_0}} (1+\gamma 2^{-2|\lambda|})^{-1}\left[f_\lambda - S_{\alpha(2^{2|\lambda|}+\gamma)/\gamma}(f_\lambda)\right]\psi_\lambda$$

and

$$\tilde{u}_{\gamma,\alpha} = \sum_{k\in I_{j_0}} \langle f, \tilde{\phi}_{j_0,k}\rangle \phi_{j_0,k} + \sum_{\lambda\in J_{j_0}} S_{\alpha(2^{2|\lambda|}+\gamma)/\gamma}(f_\lambda)\psi_\lambda \ .$$

4. SEVERAL REFINEMENTS

The advantages of the non-linear filtering rule (Proposition 3.1) are given by explicit descriptions of $\tilde{v}$ and $\tilde{u}$ and by related fast discrete wavelet schemes. However, non-redundant filtering very often creates artifacts in terms of unmeant oscillations. This mainly results in edge blurring. Moreover, we are probably faced with poor directional selectivity. These problems might be circumvented by high redundancy, by designs improving the directional selectivity[13–15] , and by additional edge dependent penalty weights. Sometimes one is also confronted with restoration problems in the presence of blur. In this case one might apply the idea of surrogate functionals which induce iterative schemes.[16]

4.1. Redundancy by translation invariance

One way to achieve redundancy is given by translation invariant representations. This can be described as follows. For $b \in \Omega$, we introduce the translation operator $T_b f(x) = f(x-b)$. At first, we derive the wavelet representation of $T_{-b}f$. Secondly, we apply T_b to this wavelet representation. Finally, we have to average over all $b \in \Omega$.

The wavelet expression of $T_{-b}f$ is given by

$$T_{-b}f = \sum_{k\in I_{j_0}} \langle f, \tilde{\phi}_{j_0,k}(\cdot - b)\rangle \phi_{j_0,k} + \sum_{j\geq j_0, k\in J_j, i} \langle f, \tilde{\psi}_{i,j,k}(\cdot - b)\rangle \psi_{i,j,k}\ . \tag{4.1}$$

Applying T_b to (4.1) and averaging yield

$$\begin{aligned} f &= \int_\Omega T_b(T_{-b}f)db \\ &= \int_\Omega \left(\sum_{k\in I_{j_0}} \langle f, \tilde{\phi}_{j_0,k}(\cdot - b)\rangle \phi_{j_0,k}(\cdot - b) + \right. \\ &\qquad \left. \sum_{j\geq j_0, k\in J_j, i} \langle f, \tilde{\psi}_{i,j,k}(\cdot - b)\rangle \psi_{i,j,k}(\cdot - b) \right) db \\ &= \sum_{k\in I_{j_0}} \int_\Omega \langle f, \tilde{\phi}_{j_0,2^{j_0}b}\rangle \phi_{j_0,2^{j_0}b} db + \sum_{j\geq j_0, k\in J_j, i} \int_\Omega \langle f, \tilde{\psi}_{i,j,2^j b}\rangle \psi_{i,j,2^j b} db \qquad (4.2) \\ &= 2^{2j_0} \int_\Omega \langle f, \tilde{\phi}_{j_0,2^{j_0}b}\rangle \phi_{j_0,2^{j_0}b} db + \sum_{j\geq j_0, i} 2^{2j} \int_\Omega \langle f, \tilde{\psi}_{i,j,2^j b}\rangle \psi_{i,j,2^j b} db\ . \qquad (4.3) \end{aligned}$$

Formula (4.2) is no longer dependent on translations k. Consequently, (4.3) represents a shift invariant wavelet series expression of f. In practical applications, there is only a finite amount of data. Hence, we cannot compute (4.3) for all $j \geq j_0$ and all translations $b \in I$. To this end, we assume that we are given 2^M rows of 2^M pixels, where each pixel is the average of f on a square $2^{-M} \times 2^{-M}$. Thus we calculate (4.3) for all $j_0 \leq j < M$ and average over 2^{2M} different translations $b = k2^{-M}$, where $k = (k_1, k_2)$ and $0 \leq k_1, k_2 < 2^M$. Under these assumptions, we obtain the following discretized version of (4.3)

$$\begin{aligned} f^M &= 2^{2(j_0-M)} \sum_k \langle f, \phi_{j_0,k2^{j_0-M}}\rangle \phi_{j_0,k2^{j_0-M}} + \\ &\qquad \sum_{M>j\geq j_0, i, k} 2^{2(j-M)} \langle f, \psi_{i,j,k2^{j-M}}\rangle \psi_{i,j,k2^{j-M}}\ . \end{aligned} \tag{4.4}$$

The translation invariant representation (4.4) has redundancy 2^{2M}. We remark that the entire calculation takes $O(M2^{2M})$ operations, instead of $O(2^{2M})$ operations for the classical discrete wavelet transform.

The resulting sequence space representation of the variational functional (3.1) has to be adapted to the redundant representation of f. To this end, we note that the Besov penalty term takes the form

$$|f|_{B_p^\beta(L_p)} \sim \left(\sum_{j\geq j_0,i,k} 2^{(js+2(j-M))} |\langle f, \tilde{\psi}_{i,j,k2^{j-M}}\rangle|^p \right)^{1/p} .$$

The norms $\|\cdot\|_{L_2}^2$ and $\|\cdot\|_{H^{-1}}^2$ change similarly. Consequently, we obtain the same minimization rule but with respect to a richer class of wavelet coefficients.

4.2. Directional sensitivity by frequency projections

In order to improve directional selectivity one should treat positive and negative frequencies separately, which might be obtained be applying the following orthogonal projections

$$\begin{aligned} \mathcal{P}^+ &: L_2 \to L_{2,+} = \{f \in L_2 : \text{ supp } \hat{f} \subseteq [0,\infty)\} \\ \mathcal{P}^- &: L_2 \to L_{2,-} = \{f \in L_2 : \text{ supp } \hat{f} \subseteq (-\infty,0]\} \end{aligned}$$

The projectors $\mathcal{P}^+$ and $\mathcal{P}^-$ may be either applied to f or to $\{\phi,\tilde{\phi}\}$ and $\{\psi,\tilde{\psi}\}$. In a discrete framework one has to approximate these projections in some suitable way. The literature suggests distinct approaches. One approach aims at creating Hilbert transform pairs of wavelets.[13,14] This approach uses the identities $\mathcal{P}^+ = (Id + iH)/2$ and $\mathcal{P}^- = (Id - iH)/2$. Another approach deals with projecting f by multiplying with certain shifted generator symbols in frequency domain.[15]

For our purposes it is convenient to focus on frequency projections. Let H denote the symbol of an orthogonal (or bi-orthogonal) generator function ϕ. This generator function must not necessarily coincide with the generator used in the decomposition scheme (2.2). We introduce projections P^+ and P^- in one dimension by defining

$$(P^+f)^\wedge(\omega) := \hat{f}(\omega)H(\omega - \pi/2) \;\text{ and }\; (P^-f)^\wedge(\omega) := \hat{f}(\omega)H(\omega + \pi/2) \;,$$

where f denotes the function to be analyzed.[15] The idea is that the shifts by $\pi/2$ attenuate negative and positive frequencies of f respectively. Obviously, P^+ and P^- only approximate $\mathcal{P}^+$ and $\mathcal{P}^-$. However, this decomposition ensures perfect reconstruction

$$\hat{f}(\omega) = (B^+P^+f)^\wedge(\omega) + (B^-P^-f)^\wedge(\omega) \;, \tag{4.5}$$

where the back-projections are given by

$$(B^+f)^\wedge = \hat{f}\overline{H(\cdot - \pi/2)} \quad \text{and} \quad (B^-f)^\wedge = \hat{f}\overline{H(\cdot + \pi/2)}$$

respectively. Formula (4.5) becomes clear by the orthogonality condition of H

$$\left(\, |H(\omega - \pi/2)|^2 + |H(\omega + \pi/2)| \, \right) \hat{f}(\omega) = \hat{f}(\omega).$$

This technique provides us with a simple multiplication scheme in Fourier, or equivalently, a convolution scheme in time domain. In a separable two dimensional framework the projections take the form

$$\begin{aligned} (P^{++}f)^\wedge(\omega_1,\omega_2) &= \hat{f}(\omega_1,\omega_2)H(\omega_1 - \pi/2)H(\omega_2 - \pi/2) \;, \\ (P^{+-}f)^\wedge(\omega_1,\omega_2) &= \hat{f}(\omega_1,\omega_2)H(\omega_1 - \pi/2)H(\omega_2 + \pi/2) \\ (P^{-+}f)^\wedge(\omega_1,\omega_2) &= \hat{f}(\omega_1,\omega_2)H(\omega_1 + \pi/2)H(\omega_2 - \pi/2) \;, \\ (P^{--}f)^\wedge(\omega_1,\omega_2) &= \hat{f}(\omega_1,\omega_2)H(\omega_1 + \pi/2)H(\omega_2 + \pi/2) \;. \end{aligned}$$

The perfect reconstruction follows immediately. We observe that

$$|f|^2 = |P^{++}f|^2 + |P^{+-}f|^2 + |P^{-+}f|^2 + |P^{--}f|^2.$$

Moreover, we note that

$$(P^{++}f)^\wedge(-\omega) = \overline{(P^{--}f)^\wedge(\omega)} \quad \text{and} \quad (P^{+-}f)^\wedge(-\omega) = \overline{(P^{-+}f)^\wedge(\omega)}.$$

Hence, the computation of $P^{-+}f$ and $P^{--}f$ can be omitted. Consequently, the modified variational functional takes the form

$$\begin{aligned}
\mathcal{F}_f(u,v) &= 2\left(\|P^{++}(f-(u+v))\|_{L_2}^2 + \|P^{+-}(f-(u+v))\|_{L_2}^2\right) + \\
&\qquad 2\lambda\left(\|P^{++}v\|_{H^{-1}}^2 + \|P^{+-}v\|_{H^{-1}}^2\right) + 2\alpha|u|_{B^1_{1(L_1)}} \\
&\leq 2\left(\|P^{++}(f-(u+v))\|_{L_2}^2 + \|P^{+-}(f-(u+v))\|_{L_2}^2\right) + \\
&\qquad 2\lambda\left(\|P^{++}v\|_{H^{-1}}^2 + \|P^{+-}v\|_{H^{-1}}^2\right) + \\
&\qquad 4\alpha\left(|P^{++}u|_{B^1_{1(L_1)}} + |P^{+-}u|_{B^1_{1(L_1)}}\right) \ ,
\end{aligned}$$

which can be minimized with respect to $\{P^{++}v, P^{++}u\}$ and $\{P^{+-}v, P^{+-}u\}$ separately. We note that the projections might be complex-valued. Parameterizing the wavelet coefficients by modulus an angle and minimizing yields the following filtering rules for the projections of $\tilde{v}_{\gamma,\alpha}$ and $\tilde{u}_{\gamma,\alpha}$

$$P^{\cdot\cdot}\tilde{v}_{\gamma,\alpha} = \sum_{\lambda\in J_{j_0}} (1+\gamma 2^{-2|\lambda|})^{-1}\left[P^{\cdot\cdot}f_\lambda - S_{\alpha(2^{2|\lambda|}+\gamma)/\gamma}(|P^{\cdot\cdot}f_\lambda|)e^{i\omega(P^{\cdot\cdot}f)}\right]\psi_\lambda$$

and

$$P^{\cdot\cdot}\tilde{u}_{\gamma,\alpha} = \sum_{k\in I_{j_0}} \langle P^{\cdot\cdot}f, \tilde{\phi}_{j_0,k}\rangle\phi_{j_0,k} + \sum_{\lambda\in J_{j_0}} (1+\gamma 2^{-2|\lambda|})^{-1} S_{\alpha(2^{2|\lambda|}+\gamma)/\gamma}(|P^{\cdot\cdot}f_\lambda|)e^{i\omega(P^{\cdot\cdot}f)}\psi_\lambda \ ,$$

where $P^{\cdot\cdot}$ denotes P^{++} and P^{+-} respectively. Finally, we have to apply the back-projections to obtain the minimizing functions

$$\tilde{v}^{BP}_{\gamma,\alpha} = B^{++}P^{++}\tilde{v}_{\gamma,\alpha} + B^{--}\overline{P^{++}\tilde{v}_{\gamma,\alpha}} + B^{+-}P^{+-}\tilde{v}_{\gamma,\alpha} + B^{-+}\overline{P^{+-}\tilde{v}_{\gamma,\alpha}}$$

and

$$\tilde{u}^{BP}_{\gamma,\alpha} = B^{++}P^{++}\tilde{u}_{\gamma,\alpha} + B^{--}\overline{P^{++}\tilde{u}_{\gamma,\alpha}} + B^{+-}P^{+-}\tilde{u}_{\gamma,\alpha} + B^{-+}\overline{P^{+-}\tilde{u}_{\gamma,\alpha}} \ .$$

4.3. Weighted penalty functions

In order to improve the capability of preserving edges we additionally introduce a positive weight sequence w_λ in the H^{-1} penalty term. Consequently, we aim at minimizing a slightly modified sequence space functional

$$\mathcal{S}^w_f(v,u) = \sum_{\lambda\in J}\left(|f_\lambda - (u_\lambda+v_\lambda)|^2 + \gamma 2^{-2|\lambda|}w_\lambda|v_\lambda|^2 + 2\alpha|u_\lambda|\cdot 1_{\{\lambda\in J_{j_0}\}}\right) \ . \tag{4.6}$$

The resulting texture and cartoon components take the form

$$\tilde{v}^w_{\gamma,\alpha} = \sum_{\lambda\in J_{j_0}} (1+\gamma w_\lambda 2^{-2|\lambda|})^{-1}\left[f_\lambda - S_{\alpha(2^{2|\lambda|}+\gamma w_\lambda)/\gamma w_\lambda}(f_\lambda)\right]\psi_\lambda$$

and

$$\tilde{u}^w_{\gamma,\alpha} = \sum_{k\in I_{j_0}} \langle f, \tilde{\phi}_{j_0,k}\rangle\phi_{j_0,k} + \sum_{\lambda\in J_{j_0}} S_{\alpha(2^{2|\lambda|}+\gamma w_\lambda)/\gamma w_\lambda}(f_\lambda)\psi_\lambda \ .$$

This refinement seems to be somewhat ad hoc. However, the main goal consists of introducing a control parameter which depends on the local structure of f. The local penalty weight w_λ should be large in the presence of an edge and small otherwise. The main problem is to localize edges. The basic idea of the procedure proposed here is similar to an edge detection algorithm proposed by Mallat and Zhong.[17]

Our scheme rests on the analysis of the wavelet coefficients f_λ at or near the same location but at different scales. We expect that the f_λ belonging to fine decomposition scales contain informations of edges (well localized) as well as oscillating components. Consequently, in order to avoid an additional penalization of the texture oscillating components it seems natural to search in coarser scales for edges, and to keep the information in order to adjust the weights in finer scales. To keep the computational effort at a reasonable level we do not apply a very sophisticated edge detector. We simply assume that wavelet coefficients representing edges must appear throughout a certain number of scales. Suppose that $f \in V_M$ and j_e denotes some 'critical' scale, then for a certain range of scales $|\lambda| = |(i,j,k)| = j \in \{j_0, \ldots, j_1 - j_e - 2, j_1 - j_e - 1\}$ we mark all positions k where $|f_\lambda|$ is larger than a level dependent threshold parameter t_j. Here the value t_j is chosen proportional to the mean value of all wavelet coefficients of level j. We say that $|f_\lambda|$ represents an edge if k was marked for all $j \in \{j_0, \ldots, j_1 - j_e - 2, j_1 - j_e - 1\}$. Finally, we adaptively choose the penalty sequence by setting

$$w_\lambda = \begin{cases} \Theta_\lambda & \text{if } j \in \{M-1, \ldots, j_1 - j_e\} \text{ and } k \text{ was marked as an edge ,} \\ \vartheta_\lambda & \text{otherwise ,} \end{cases}$$

where ϑ_λ has to be close to one and Θ_λ $(> \vartheta_\lambda)$ can be tuned manually in order to penalize the corresponding f_λ's.

4.4. Restoration in the presence of blur

Very often only a blurred version of an image is available. In this case, if K is a linear operator modeling the blur, the variational functional (1.4) reads as

$$\mathcal{F}_f(v,u) = \|f - K(u+v)\|^2_{L_2(\Omega)} + \lambda \|v\|^2_{H^{-1}(\Omega)} + 2\alpha |u|_{B^1_1(L_1(\Omega))}. \tag{4.7}$$

Minimizing (4.7) results in a coupled system of nonlinear equations for u_λ and v_λ. However, this can be circumvented by constructing a surrogate functional that removes the influence of $K^*K(u+v)$. This technique was elaborated for similar variational problems involving one non-quadratic regularizing penalty.[16] Moreover, it is shown that the resulting iterative scheme converges in norm. For our purpose we have to apply the idea of surrogate functionals to (4.7). Since K can be renormalized, we restrict ourselves without loss generality to the case $\|K^*K\| < 1$. For some $h \in L_2(\Omega)$ the surrogate functional for (4.7) is of the form

$$\mathcal{F}^{sur}_f(v,u,h) = \mathcal{F}_f(v,u) + \|u + v - h\|^2_{L_2(\Omega)} - \|K(u+v-h)\|^2_{L_2(\Omega)}. \tag{4.8}$$

One can try to approach the minimizers of $\mathcal{F}_f(v,u)$ by an iterative process. One suitable iterative scheme can be formulated as follows

$$\begin{gathered} h_0 = u_0 + v_0 \text{ arbitrary ;} \\ (u_n, v_n) = \arg\min_{u,v} \left(\mathcal{F}^{sur}_f(v,u,h_{n-1})\right) \ , \quad h_n = u_n + v_n \ ; \ n = 1,2,\ldots \end{gathered} \tag{4.9}$$

Proposition 4.1. *Let K be some bounded linear operator modeling the blur, with $\|K^*K\| < 1$, and suppose the given function f belongs to $L_2(\Omega)$. Then the n-th iteration $(u_{n,\gamma,\alpha}, v_{n,\gamma,\alpha})$ of the scheme (4.9) is given by*

$$\begin{aligned} v_{n,\gamma,\alpha} \quad = \quad & \sum_{\lambda \in J_{j_0}} (1 + \gamma 2^{-2|\lambda|})^{-1} \left[((K^*f)_\lambda + (h_{n-1})_\lambda - (K^*Kh_{n-1})_\lambda - \right. \\ & \left. S_{\alpha(2^{2|\lambda|}+\gamma)/\gamma}((K^*f)_\lambda + (h_{n-1})_\lambda - (K^*Kh_{n-1})_\lambda)\right] \psi_\lambda \end{aligned}$$

and

$$u_{n,\gamma,\alpha} = \sum_{k\in I_{j_0}} \langle K^*f + h_{n-1} - K^*Kh_{n-1}, \tilde{\phi}_{j_0,k}\rangle \phi_{j_0,k} + \sum_{\lambda\in J_{j_0}} S_{\alpha(2^{2|\lambda|}+\gamma)/\gamma}((K^*f)_\lambda + (h_{n-1})_\lambda - (K^*Kh_{n-1})_\lambda)\psi_\lambda \ ,$$

where $h_{n-1} = u_{n-1,\gamma,\alpha} + v_{n-1,\gamma,\alpha}$.

We remark that a final proof of convergence of the proposed iterative scheme has not come in yet. However, the experiments are quite convincing.

5. NUMERICAL EXAMPLES

Finally, we present some numerical results obtained with our wavelet based approach. All results were computed on the basis of redundant decompositions (either complex valued or translation invariant). In order to compare the results we have attached a resulting decomposition presented in papers of Osher–Solé–Vese and Osher–Vese, see Figure 5. In all presented examples the tuning parameters $\alpha, \gamma, \Theta_\lambda, \vartheta_\lambda$ are chosen manually.

In order to show how the proposed non-linear wavelet scheme acts on piecewise constant functions we start with a geometric image f representing cartoon components only but with sharp contours, see Figure 1. We observe that $\tilde{u}$ represents the cartoon very well (this result was computed without edge enhancement, i.e., without a special adjustment of w_λ). Note that a slight change of γ produces a cartoon with less sharp edges. The texture component $\tilde{v}$ contains some very weak contour information.

In Figure 2 we demonstrate the performance of our algorithm involving the additional local penalty weight sequence w_λ. The upper row of images shows the decomposition result where the algorithm automatically incorporates the contour structure of the image. The lower row shows the result without edge enhancements. We observe that the upper texture image $\tilde{v}$ contains less cartoon structure than the lower one.

In certain applications a segmentation of an image is required. To this end, it might be useful to decompose the image beforehand in order to separate oriented patterns. Figure 3 shows a fabric texture image. Here the cartoon mainly contains non-oriented structures whereas the texture very clearly reveals oriented structures (without any disturbing components as in the original image).

In Figure 4 non-linear wavelet filtering was applied to a woman image. The initial image contains edges, contours and several types of textures. Comparing the reconstruction $\tilde{u}$ and $\tilde{v}$ with u and v obtained with the TV model we observe that $\tilde{v}$ contains less cartoon information than v whereas the edges are somehow better preserved in u than in $\tilde{u}$. We remark that our gray scale was chosen somehow different than the gray scale used in the PDE schemes.

We finish with presenting some results in the presence of blur, see Figure 6. The original image (a section of the woman image) was multiplied in Fourier space with the Gaussian in order to create a blurred image f, i.e. K is a convolution operator. Here a reconstruction after 30 iterations is presented. As expected, a comparison of f with $\tilde{u}_{30} + \tilde{v}_{30}$ shows that blurred contours become sharper during the iteration process. A basic observation is that the iterative scheme converges very fast.

ACKNOWLEDGMENTS

The authors would like to thank S. Osher and L. Vese (UCLA) for stimulating discussions. Gerd Teschke gratefully acknowledges support by DAAD grants as well as by DFG grant Te 354/1-2, and, moreover, he thanks the PACM (Princeton University) for the hospitality during his stay in Princeton.

Figure 1. From left to right: initial geometric image f, $\tilde{u}$, $\tilde{v}+150$, Db3, $\alpha = 0.5$, $\gamma = 0.01$.

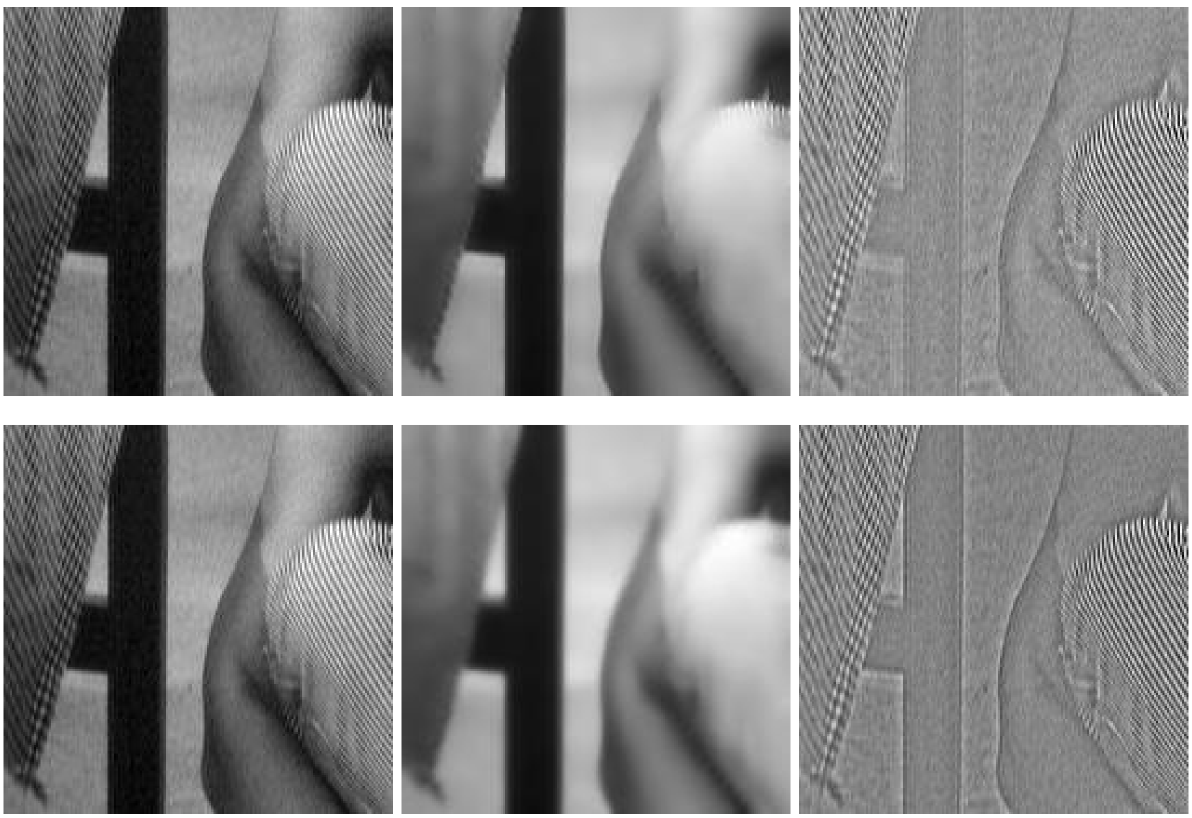

Figure 2. Up: with edge enhancement, down: without edge enhancement; From left to right: the initial image f, $\tilde{u}$, $\tilde{v}+150$, $\alpha = 0.5$, $\gamma = 0.0001$, computed with Db1 and critical scale $j_e = -3$.

REFERENCES

1. S. Osher and L. Vese, "Modeling textures with total variation minimization and oscillating patterns in image processing," Tech. Rep. 02-19, University of California Los Angeles C.A.M., 2002.

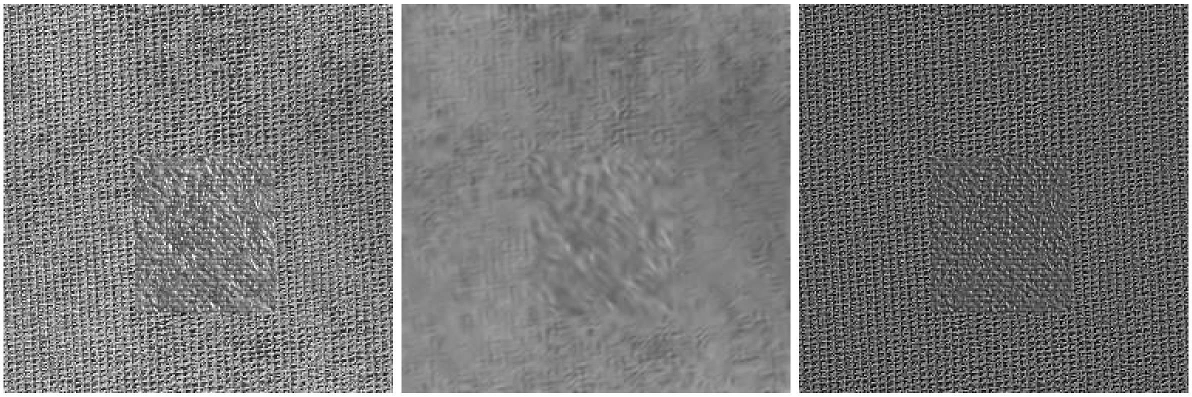

Figure 3. From left to right: initial fabric image f, $\tilde{u}$, $\tilde{v}+150$, Db4, $\alpha = 0.8$, $\gamma = 0.002$.

Figure 4. From left to right: initial woman image f, $\tilde{u}$ and $\tilde{v}+150$, Db10, $\alpha = 0.5$, $\gamma = 0.002$.

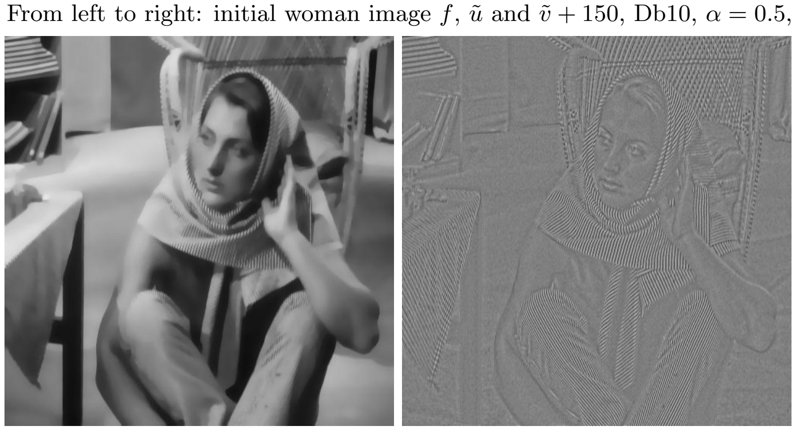

Figure 5. Form left to right: Decomposition into u and v using the TV model.

2. S. Osher, A. Sole, and L. Vese, "Image decomposition and restoration using total variation minimization and the H^{-1} norm," Tech. Rep. 02-57, University of California Los Angeles C.A.M., 2002.
3. Y. Meyer, "Oscillating Patterns in Image Processing and Nonlinear Evolution Equations," *University Lecture Series Volume 22, AMS* , 2002.
4. A. Cohen, R. DeVore, P. Petrushev, and H. Xu, "Nonlinear Approximation and the Space $BV(\mathbf{R}^2)$,"

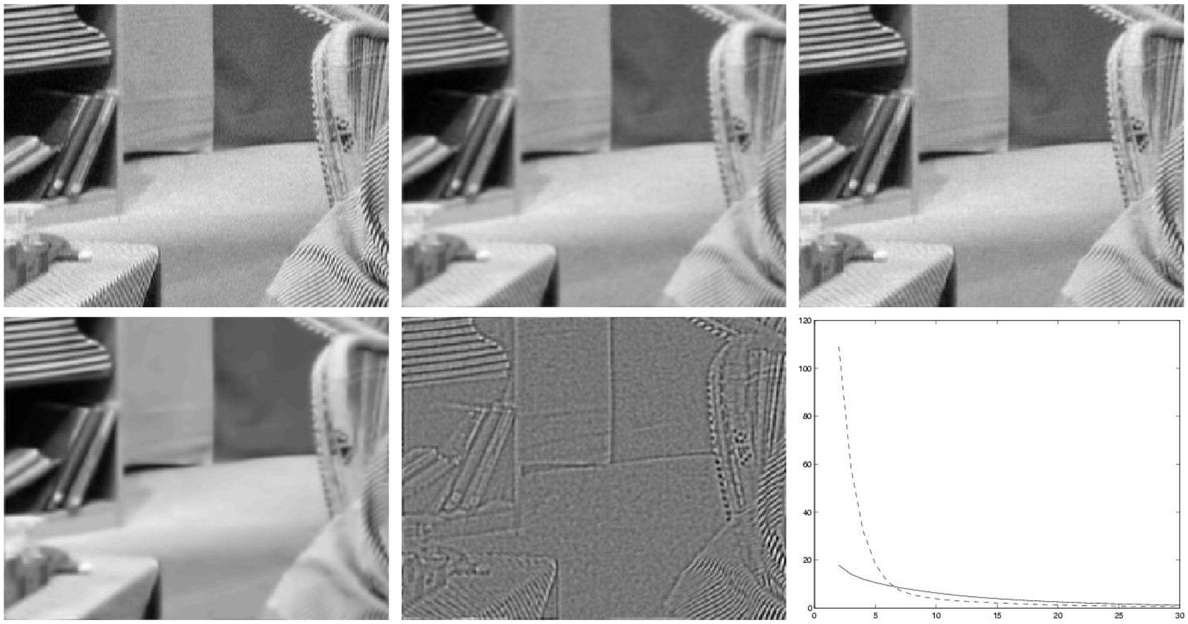

Figure 6. Up from left to right: original image, blurred image f, reconstruction $u_{30} + v_{30}$; down from left to right: u_{30}, v_{30}, and rates of convergence (solid line: $\|\tilde{v}_n - \tilde{v}_{n-1}\|_{L_2(\Omega)}$, dashed line: $\|\tilde{u}_n - \tilde{u}_{n-1}\|_{L_2(\Omega)}$).

American Journal of Mathematics (121), pp. 587–628, 1999.

5. I. Daubechies, *Ten Lectures on Wavelets*, SIAM, Philadelphia, 1992.
6. I. Daubechies, "Wavelet transforms and orthonormal wavelet bases," *Proceedings of Symposia in Applied Mathematics* (47), 1993.
7. A. Cohen, I. Daubechies, and J.-C. Feauveau, "Biorthogonal bases of compactly supported wavelets," *Comm. Pure Appl. Math.* **45**, pp. 485–560, 1992.
8. W. Dahmen, "Stability of multiscale transformations," *The Journal of Fourier Analysis and Applications* **2**, pp. 341–361, 1996.
9. R. DeVore, B. Jawerth, and V. Popov, "Compression of wavelet decompositions," *Amer. J. Math.* **114**, pp. 737–785, 1992.
10. R. DeVore, B. Jawerth, and V. Popov, "Interpolation of besov spaces," *Trans. Math. Soc.* **305**, pp. 397–414, 1988.
11. M. Frazier and B. Jawerth, "A discrete transform and decompositions of distribution spaces," *J. of Functional Anal.* **93**, pp. 34–170, 1990.
12. H. Triebel, *Interpolation Theory, Function Spaces, Differential Operators*, Verlag der Wissenschaften, Berlin, 1978.
13. N. Kinsbury, "Image processing with complex wavelets," *Phil. Trans. R. Soc. Lond.* , Sept. 1999.
14. I. W. Selesnick, "Hilbert transform pairs of wavelet bases," *IEEE Signal Processing Letters* **8**, pp. 170–173, June 2001.
15. F. C. A. Fernandes, R. v. Spaendonck, M. J. Coates, and S. Burrus, "Directional Complex-Wavelet Processing," *Proceedings of Society of Photo-Optical Instrumental Engineers—SPIE2000, Wavelet Applications in Signal Processing VIII,* San Diego. , 2000.
16. I. Daubechies, M. Defrise, and C. D. Mol, "An iterative thresholding algorithm for linear inverse problems with a sparsity constraint," *Preprint* , 2003.
17. S. Mallat and S. Zhong, "Characterization of Signals from Multiscale Edges," *IEEE Transactions on Pattern Analysis and Machine Intelligence* **14**, pp. 710–732, July 1992.