Available online at www.sciencedirect.com
ScienceDirect
Appl. Comput. Harmon. Anal. 23 (2007) 25–56
Applied and Computational Harmonic Analysis
www.elsevier.com/locate/acha

Image decompositions using bounded variation and generalized homogeneous Besov spaces ☆

John B. Garnett [a], Triet M. Le [a], Yves Meyer [b], Luminita A. Vese [a,*]

[a] *Department of Mathematics, University of California, Los Angeles, 405 Hilgard Ave., Los Angeles, CA 90095-1555, USA*
[b] *CMLA ENS Cachan, 61 Ave Président Wilson, 94235 Cachan Cedex, France*

Received 17 February 2006; revised 16 October 2006; accepted 16 January 2007

Available online 21 March 2007

Communicated by Raymond Chan

Abstract

This paper is devoted to the decomposition of an image f into $u + v$, with u a piecewise-smooth or "cartoon" component, and v an oscillatory component (texture or noise), in a variational approach. Y. Meyer [Y. Meyer, Oscillating Patterns in Image Processing and Nonlinear Evolution Equations, University Lecture Series, vol. 22, Amer. Math. Soc., Providence, RI, 2001] proposed refinements of the total variation model [L. Rudin, S. Osher, E. Fatemi, Nonlinear total variation based noise removal algorithms, Phys. D 60 (1992) 259–268] that better represent the oscillatory part v: the weaker spaces of generalized functions $G = \operatorname{div}(L^\infty)$, $F = \operatorname{div}(BMO)$, and $E = \dot{B}^{-1}_{\infty,\infty}$ have been proposed to model v, instead of the standard L^2 space, while keeping $u \in BV$, a function of bounded variation. Such new models separate better geometric structures from oscillatory structures, but it is difficult to realize them in practice. D. Mumford and B. Gidas [D. Mumford, B. Gidas, Stochastic models for generic images, Quart. Appl. Math. 59 (1) (2001) 85–111] also show that natural images can be seen as samples of scale invariant probability distributions that are supported on distributions only, and not on sets of functions. In this paper, we consider and generalize Meyer's (BV, E) model, using the homogeneous Besov spaces $\dot{B}^\alpha_{p,q}$, $-2 < \alpha < 0$, $1 \leqslant p, q \leqslant \infty$, to represent the oscillatory part v. Theoretical, experimental results and comparisons to validate the proposed methods are presented.

1. Introduction and motivations

In what follows, we assume that a given grayscale image can be represented by a function (or sometimes distribution) f, defined on a rectangle Ω of $\mathbb{R}^2$. Sometimes, we may assume that the image f is periodic (obtained by reflection), or defined on the whole plane $\mathbb{R}^2$. We limit our presentation to the two-dimensional case, but our results can be generalized to $\mathbb{R}^n$.

☆ This work was supported in part by an Alfred P. Sloan fellowship, by the National Institute of Health through the NIH Roadmap for Medical Research: Grant U54 RR021813 entitled Center for Computational Biology CCB, by the National Science Foundation (Grants NSF ITR 0113439, NSF DMS 0312222, NSF DMS 04302720, and NSF DMS 0401720), by the Institute for Pure and Applied Mathematics, and by the Dissertation Year Fellowship, University of California, Los Angeles.

* Corresponding author.
E-mail addresses: jbg@math.ucla.edu (J.B. Garnett), tle@math.ucla.edu (T.M. Le), Yves.Meyer@cmla.ens-cachan.fr (Y. Meyer), lvese@math.ucla.edu (L.A. Vese).

1063-5203/$ – see front matter © 2007 Published by Elsevier Inc.
doi:10.1016/j.acha.2007.01.005

We are interested in decomposing f into $u+v$ via an energy minimization problem

$$\inf_{(u,v)\in X_1\times X_2}\big\{\mathcal{K}(u,v)=F_1(u)+\lambda F_2(v):\ f=u+v\big\},$$

where $F_1, F_2 \geqslant 0$ are functionals, and X_1, X_2 are spaces of functions or distributions such that $F_1(u)<\infty$, $F_2(v)<\infty$, if and only if $(u,v)\in X_1\times X_2$. It is assumed that $f\in X_1+X_2$. The constant $\lambda>0$ is a tuning parameter.

An important problem in image analysis is to separate different features in images. For instance, in image denoising, f is the observed noisy version of the true unknown image u, while v represents additive Gaussian noise of zero mean. Another related problem is the separation of the geometric (cartoon) component u of f from the oscillatory component v, representing texture. In other cases, u can be seen as a geometric or structure component of f, while v is clutter (see S.C. Zhu and D. Mumford [63]). A good model for $\mathcal{K}$ is given by a choice of X_1 and X_2 such that with the above given properties of u and v, $F_1(u)\ll F_1(v)$ and $F_2(v)\ll F_2(u)$.

We give here two classical examples of image decomposition models that use the space $L^2(\Omega)$ to model the oscillatory components by variational methods. However, many other previous work (variational or non-variational) can be seen as decompositions of f into $u+v$.

In the Mumford–Shah model for image segmentation [45], $f\in L^\infty(\Omega)\subset L^2(\Omega)$ is decomposed into $u\in SBV(\Omega)$ [5,43], (a piecewise-smooth function with its discontinuity set J_u composed of a union of curves of total finite length), and $v=f-u\in L^2(\Omega)$ representing noise or texture. The problem in the weak formulation is [43,45]

$$\inf_{(u,v)\in SBV(\Omega)\times L^2(\Omega)}\left\{\int_{\Omega\setminus J_u}|\nabla u|^2+\alpha\mathcal{H}^1(J_u)+\beta\|v\|^2_{L^2(\Omega)},\ f=u+v\right\},\tag{1}$$

where $\mathcal{H}^1$ denotes the 1-dimensional Hausdorff measure, and $\alpha,\beta>0$ are tuning parameters. With the above notations, the first two terms in the energy from (1) compose $F_1(u)$, while the third term makes $F_2(v)$.

A related decomposition is obtained by the total variation minimization model of Rudin, Osher, and Fatemi [50] for image denoising, where $SBV(\Omega)$ is replaced by the slightly larger space $BV(\Omega)$ of functions of bounded variation [29]. The Rudin–Osher–Fatemi restoration model [50] can be written as a decomposition model [42] by

$$\inf_{(u,v)\in BV(\Omega)\times L^2(\Omega)}\big\{\mathcal{J}(u,v)=|u|_{BV(\Omega)}+\lambda\|v\|^2_{L^2(\Omega)},\ f=u+v\big\},\tag{2}$$

where $\lambda>0$ is a tuning parameter, and $|u|_{BV(\Omega)}=\int_\Omega|\nabla u|$ denotes the total variation of u or the semi-norm of $BV(\Omega)$. In the original TV model, v represents additive Gaussian noise of zero mean. This model provides a unique $(BV(\Omega),L^2(\Omega))$ decomposition of $f\in L^2(\Omega)$, for each $\lambda>0$ (see [1,18], or [60] for a more general case). The model is strictly convex, easy to solve in practice, and denoises well piecewise-constant images while preserving sharp edges. However it has some limitations. For instance, if f is the characteristic function of a convex, smooth domain D of finite perimeter, the model should produce $u=f$, $v=0$. But this is not true for any finite value of λ, [7,42,55]. Cartoon or BV pieces of f are sent to v, and the model does not always represent well texture or oscillatory details, as it has been analyzed in [42].

In [56], the authors E. Tadmor et al., have proposed a hierarchical multiscale $(BV(\Omega),L^2(\Omega))$ decomposition method, to better locate the scale λ and a splitting $f=u_\lambda+v_\lambda$. The main idea in [56] is to first compute an ROF optimal pair (u_λ,v_λ) as a splitting at scale λ with given data f. Then this process is iterated, by computing another ROF optimal pair $(u_{2\lambda},v_{2\lambda})$ with data v_λ, and so on. Thus, $f\approx\sum_k u_k$ and residual v_k gives a better splitting of f than the ROF model. Theoretical and experimental results are also presented. Also, in [2,32], it has been shown that natural images are not well represented only by functions of bounded variation.

Earlier work for image analysis and denoising by variational methods using Besov spaces in a wavelet framework are by [17,19,25–27], among others. We also mention a recent work of A.S. Carasso [16] for using $B^\alpha_{2,\infty}(\Omega)$ with $0<\alpha<1$ to model u instead of $BV(\Omega)$ in the recovery of details and piecewise-smooth regions in image deblurring, based on the Poisson kernel formulation of Besov spaces.

Substituting $L^2(\Omega)$ in (2) by $L^1(\Omega)$ has been earlier proposed in [3,4] in the discrete case, and later analyzed in [20,21,46], among others. However, oscillatory components do not have small norms in $L^2(\Omega)$ or $L^1(\Omega)$ ([37,42]). Here we are interested in a better choice for the oscillatory component v or for the space X_2, which has to give small norms for oscillatory functions, while keeping $X_1=BV(\Omega)$. Our discussion follows Y. Meyer [42], with motivations from D. Mumford and B. Gidas [44]. The idea is to use weaker norms for the oscillatory component v, instead of the

$L^2(\Omega)$ norm, and this can be done using generalized functions or distributions. For more motivations on the choice of such spaces over the standard $L^2(\Omega)$ approach, we refer the reader to the inspirational notes [42] and subsequent work [8,11–13,37,39,47–49,61,62].

Y. Meyer [42] theoretically proposed decompositions of $f \in X_2$, with $X_1 = BV(\mathbb{R}^2) \subset X_2$, via

$$\inf_{(u,v)\in BV(\mathbb{R}^2)\times X_2} \big\{\mathcal{K}(u,v) = |u|_{BV(\mathbb{R}^2)} + \lambda\|v\|_{X_2},\, f = u + v\big\},$$

as a refinement of the ROF model [50], where $(X_2, \|\cdot\|_{X_2})$ is one of the spaces denoted below by $(G(\mathbb{R}^2), \|.\|_{G(\mathbb{R}^2)})$, $(F(\mathbb{R}^2), \|.\|_{F(\mathbb{R}^2)})$, or $(E(\mathbb{R}^2), \|.\|_{E(\mathbb{R}^2)})$, that are introduced in Definitions 2, 3, and 4.

Definition 1. Let $\mathcal{S} = \mathcal{S}(\mathbb{R}^2)$ be the space of all functions ϕ on $\mathbb{R}^2$ that are infinitely differentiable and, together with all their derivatives, are rapidly decreasing (i.e., remain bounded when multiplied by arbitrary polynomials, or more precisely $|x|^k|D^\alpha\phi(x)| \to 0$ as $|x| \to \infty$, $\forall k \in \mathbb{N}$, $\forall \alpha \in \mathbb{N}^2$). The space of all tempered distributions on $\mathbb{R}^2$, i.e., the set of continuous linear forms on $\mathcal{S}$, is denoted by $\mathcal{S}' = \mathcal{L}(\mathcal{S}(\mathbb{R}^2), \mathbb{R})$. $\mathcal{S}'(\mathbb{R}^2)$ is given the weak-star topology as dual of $\mathcal{S}(\mathbb{R}^2)$.

Definition 2. (See [42].) Let $G(\mathbb{R}^2)$ consist of distributions v in $\mathcal{S}'$ which can be written as

$$v = \operatorname{div}(\vec{g}) \text{ in } \mathcal{S}', \quad \vec{g} = (g_1, g_2) \in \big(L^\infty\big(\mathbb{R}^2\big)\big)^2,$$

with

$$\|v\|_{G(\mathbb{R}^2)} = \inf\Big\{\Big\|\sqrt{(g_1)^2 + (g_2)^2}\Big\|_{L^\infty(\mathbb{R}^2)} : v = \operatorname{div}(\vec{g}) \text{ in } \mathcal{S}',\ \vec{g} \in L^\infty\big(\mathbb{R}^2, \mathbb{R}^2\big)\Big\}.$$

The space $F(\mathbb{R}^2)$ is obtained by replacing $L^\infty(\mathbb{R}^2)$ above with $BMO(\mathbb{R}^2)$, the John–Nirenberg space of bounded mean oscillation [53].

Definition 3. Let $F(\mathbb{R}^2)$ consist of generalized functions v which can be written as

$$v = \operatorname{div}(\vec{g}), \quad \vec{g} = (g_1, g_2) \in BMO\big(\mathbb{R}^2, \mathbb{R}^2\big),$$

with

$$\|v\|_{F(\mathbb{R}^2)} = \inf\big\{\big(\|g_1\|_{BMO(\mathbb{R}^2)} + \|g_2\|_{BMO(\mathbb{R}^2)}\big):\ v = \operatorname{div}(\vec{g}),\ \vec{g} = (g_1, g_2) \in BMO\big(\mathbb{R}^2, \mathbb{R}^2\big)\big\}.$$

Definition 4. Let $E(\mathbb{R}^2) = \dot{B}^{-1}_{\infty,\infty}(\mathbb{R}^2)$ consists of generalized functions $v = \Delta g$, with g satisfying the Zygmund condition

$$\sup_{|y|>0} \frac{\|g(\cdot + y) - 2g(\cdot) + g(\cdot - y)\|_{L^\infty(\cdot)}}{|y|} < \infty, \tag{3}$$

endowed with

$$\|v\|_{E(\mathbb{R}^2)} = \inf\Big\{\sup_{|y|>0} \frac{\|g(\cdot + y) - 2g(\cdot) + g(\cdot - y)\|_{L^\infty(\cdot)}}{|y|} :\ v = \Delta g,\ g \text{ satisfies (3)}\Big\}.$$

Remark 1. The spaces $G(\mathbb{R}^2)$ and $F(\mathbb{R}^2)$ (as defined above) consist of first order derivatives of vector fields in $L^\infty(\mathbb{R}^2)$ and $BMO(\mathbb{R}^2)$, respectively, while the homogeneous Besov space $E(\mathbb{R}^2) = \dot{B}^{-1}_{\infty,\infty}(\mathbb{R}^2) = \Delta(\dot{B}^1_{\infty,\infty})$ consists of second order derivatives of functions that satisfy the Zygmund condition (3). The following embeddings hold in $\mathbb{R}^2$,

$$\dot{B}^1_{1,1} \subset BV \subset L^2 \subset G \subset F \subset E, \tag{4}$$

where $\dot{B}^1_{1,1}(\mathbb{R}^2)$ is the pre-dual of $\dot{B}^{-1}_{\infty,\infty}(\mathbb{R}^2)$ and G can be seen as a good approximation of the dual of BV. Thus, the spaces G, F and E are larger spaces with weaker norms than the L^2 space, therefore less demanding on a minimizing pair $(u, v = f - u)$. However, the difficulty is how to compute (BV, G), (BV, F) or (BV, E) decompositions in practice. Analyzing such problems opens new and interesting computational challenges.

In [61,62], the authors proposed to model oscillatory components as first order derivatives of vector fields in L^p, for $1 \leqslant p < \infty$ as an approximation to Meyer's (BV, G) model. For a bounded domain Ω, this has been done by minimizing

$$\inf_{u,\vec{g}=(g_1,g_2)} \left\{ \mathcal{G}(u, g_1, g_2) = |u|_{BV(\Omega)} + \mu \left\| f - u - \operatorname{div}(\vec{g}) \right\|^2_{L^2(\Omega)} + \lambda \left\| \sqrt{g_1^2 + g_2^2} \right\|_{L^p(\Omega)} \right\}. \tag{5}$$

By this model, $f \in L^2(\Omega)$ is decomposed into $u + v + w$. As $\mu \to \infty$ and $p \to \infty$, this model approaches Meyer's (BV, G) model. The space $G(\Omega) = W^{-1,\infty}(\Omega)$ is approximated by $W^{-1,p}(\Omega)$, with $p < \infty$ (when $p = 2$, v belongs to the dual space $\dot{H}^{-1}(\Omega)$ of $\dot{H}^1(\Omega)$).

In [49], the last author together with S. Osher and A. Solé, proposed a simplified approximated method corresponding to the case $p = 2$. Let $\vec{g} = \nabla g$, with g a scalar function. Using $f - u = v = \operatorname{div}(\vec{g}) = \Delta g$, i.e., $g = \Delta^{-1}(f - u)$, (5) becomes an exact decomposition model,

$$\inf_u \left\{ \mathcal{G}(u) = \int_\Omega |\nabla u| + \lambda \int_\Omega \left| \nabla\left(\Delta^{-1}(f - u)\right) \right|^2 dx \right\}. \tag{6}$$

Note that the oscillatory component v is modeled as second order derivatives (as in E) of functions in the homogeneous Sobolev space $\dot{H}^1(\Omega)$. This model gives an exact decomposition $f = u + v$, with $u \in BV(\Omega)$ and $v \in \dot{H}^{-1}(\Omega)$, and the minimization problem has been solved using a fourth-order non-linear PDE. In [22,23], I. Daubechies and G. Teschke proposed a wavelets approach with residual term where $u \in \dot{B}^1_{1,1}(\Omega)$ and $v \in H^{-1}(\Omega)$. L. Lieu and L. Vese [38] recently proposed another method to solve (6) without having to solve a fourth-order PDE, and generalized the model [49] to H^s, $s < 0$ for modeling the oscillatory component v. We mention J.-F. Aujol and A. Chambolle [13] for modeling the v component using generalized Hilbert spaces and in particular using the space E for the texture norm, in a wavelets framework, and for an alternative approach for solving (6).

In [37], the authors proposed approximations to Meyer's (BV, F) model. These approximations have been obtained by the energy minimizations

$$\inf_{u,\vec{g}=(g_1,g_2)} \left\{ \mathcal{G}(u, g_1, g_2) = |u|_{BV(\Omega)} + \mu \left\| f - u - \operatorname{div}(\vec{g}) \right\|^2_{L^2(\Omega)} + \lambda \left[\|g_1\|_{BMO(\Omega)} + \|g_2\|_{BMO(\Omega)} \right] \right\}, \tag{7}$$

and

$$\inf_{u,g} \left\{ \mathcal{G}(u, g) = |u|_{BV(\Omega)} + \mu \|f - u - \Delta g\|^2_{L^2(\Omega)} + \lambda \left[\|g_x\|_{BMO(\Omega)} + \|g_y\|_{BMO(\Omega)} \right] \right\}. \tag{8}$$

In the continuous case, the models (7) and (8) are very similar. However, in the discrete case, (8) gives a more isotropic decomposition. Note that in (8), the component v is modeled as second order derivatives of g, where $\nabla g = (g_x, g_y) \in BMO(\Omega)^2$. Figure 1 shows a comparison between the models (7) and (8) with the Rudin–Osher–Fatemi model (2).

The present paper is a continuation of the previous work [37,49,61,62], and we model the oscillatory components v by second order derivatives of functions in the homogeneous Besov spaces $\dot{B}^\alpha_{p,q}(\Omega)$, $0 < \alpha < 2$ (or equivalently $v \in \dot{B}^{\alpha-2}_{p,q}(\Omega)$), while keeping $u \in BV(\Omega)$ as a function of bounded variation, in a variational approach. In other words, $v = \Delta g$, where g satisfies the following Lipschitz condition

$$\left(\int_\Omega \left(\frac{\|g(x+y) - 2g(x) + g(x-y)\|_{L^p(\Omega)}}{|y|^\alpha} \right)^q \frac{dy}{|y|^n} \right)^{1/q} < \infty, \quad \text{when } 1 \leqslant q < \infty, \quad \text{or} \tag{9}$$

$$\sup_{|y|>0} \left\{ \frac{\|g(x+y) - 2g(x) + g(x-y)\|_{L^p(\Omega)}}{|y|^\alpha} \right\} < \infty, \quad \text{when } q = \infty. \tag{10}$$

Here, $n = 2$ is the dimension. In practice we only consider the case $q = \infty$, therefore the rougher or largest Besov spaces, with weaker norms. Y. Meyer's (BV, E) model corresponds to the case $p = q = \infty$, and $\alpha = 1$.

Other related models for image decomposition into cartoon and texture have been proposed recently. We mention Daubechies and Teschke [22,23], Starck, Elad, and Donoho [52], Candes and Guo [15], among others for variational and wavelets approaches. In particular, we refer to Aujol et al. [11,12], and Aubert and Aujol [8] for more properties of the space G both in theory and practice, and another approximation of the Meyer's (BV, G) model on bounded domains. Aujol and Chambolle [13] discuss the properties of norms that are dual to negative Sobolev and Besov norms,

Fig. 1. Here f is a given image to be decomposed. The first decomposition $f = u_1 + v_1$ is obtained by the ROF model (2). The second decomposition $f = u_2 + v_2$ is obtained by (7) with first order derivatives [37]. The third decomposition $f = u_3 + v_3$ is obtained by (8) with second order derivatives [37]. We can notice the improvement from left to right in terms of cartoon and texture decomposition, because in u_1 and u_2 there are still more oscillations kept in the BV component.

in particular, the space E (in wavelet formulation). Malgouyres [41], Lintner and Malgouyres [40] also consider modeling the oscillatory component v with the constraint $\sup_{\psi\in\mathcal{D}} |\langle v, \psi\rangle| < \infty$, where $\mathcal{D}$ is a dictionary of features (usually union of wavelet bases), while keeping $u \in BV$. This is related with imposing $v \in \dot{B}^{\alpha}_{\infty,\infty}$, for some α (as will be seen from the definition of $\dot{B}^{\alpha}_{p,q}$ in the later section).

As we have mentioned earlier, Besov spaces often arise in image analysis in the wavelets framework. Our analysis and computations are purely in a variational and PDE approach, using the Poisson or the Gaussian kernel formulations. Finally, we refer the interested reader to the recent PhD thesis manuscripts of J.-F. Aujol [10], A. Haddad [34], T. Le [35], L. Lieu [36] and J. Gilles [31] for more complete expositions and results on the subject.

2. Function spaces

For definitions and properties of the space of functions of bounded variation $BV(\Omega)$ we refer the reader to [29], and for convex functions of measures (with linear growth at infinity, as generalization of the total variation) we refer to [24].

We recall below the definitions of the Besov spaces using the Poisson and Gaussian kernel formulations in the periodic case, since this is the framework that we use to model in practice the oscillatory component v. Equivalent definitions using wavelets, modulus of continuity, or finite differences can be given (for instance, we refer the reader to E. Stein [54], Chapter V, [54] on Lipschitz spaces, H. Triebel [59], Section 1.8, [58] for the homogeneous versions of Besov spaces, [51] for periodic spaces, and M.H. Taibleson [57] for periodic Lipschitz spaces using the Poisson or Gaussian kernel formulations, the most relevant for our work).

We can consider images as functions defined on the entire domain, to which we can associate their periodization (with a fundamental domain Ω), as explained below. Or, we can consider images defined initially on a rectangle Ω, that we can extend by reflection or mirror to the entire plane (thus obtaining again a periodic function with a fundamental domain Ω' made up of four rectangles of the same size as Ω). The second method is the standard one in image analysis.

Let $x \in \mathbb{R}^n$ and $t > 0$ (for us, $n = 2$). Denote the Gauss–Weierstrass kernel

$$W_t(x) = a_n t^{-n/2} e^{-\frac{\pi|x|^2}{2t}} \tag{11}$$

and the Cauchy–Poisson kernel

$$P_t(x) = c_n \frac{t}{(t^2 + |x|^2)^{(n+1)/2}} \tag{12}$$

with $W_0 = P_0 = Id$. Here the constants a_n and c_n are chosen such that $\int W_t(x)\,dx = \int P_t(x)\,dx = 1$. Let $g \in L^p$, $1 \leqslant p \leqslant \infty$. We denote, for $x \in \mathbb{R}^n$ and $t \geqslant 0$,

$$u(x,t) = W_t * g(x) = a_n t^{-n/2} \int e^{-\frac{\pi|x-y|^2}{2t}} g(y)\,dy = \left(e^{-2\pi t|\xi|^2} \hat{g}(\xi)\right)^{\vee}(x),$$

$$v(x,t) = P_t * g(x) = c_n \int \frac{t}{(t^2 + |x-y|^2)^{(n+1)/2}} g(y)\,dy = \left(e^{-2\pi t|\xi|} \hat{g}(\xi)\right)^{\vee}(x),$$

where $\hat{g}$ denotes the Fourier transform of g, and $^{\vee}$ the inverse Fourier transform.

Furthermore, we have

$$\frac{\partial u(x,t)}{\partial t} = \Delta u(x,t) \quad \text{in } \mathbb{R}^{n+1}_+ \quad \text{and} \quad u(x,0) = g(x),$$

$$\frac{\partial^2 v(x,t)}{\partial t^2} = -\Delta v(x,t) \quad \text{in } \mathbb{R}^{n+1}_+ \quad \text{and} \quad u(x,0) = g(x).$$

Hence $u(x,t)$ is a solution of the heat equation in $\mathbb{R}^{n+1}_+$ and $v(x,t)$ is a harmonic function in $\mathbb{R}^{n+1}_+$. In other words, $u(x,t)$ is a thermic extension of g from $\mathbb{R}^n$ to $\mathbb{R}^{n+1}_+$ and $v(x,t)$ is a harmonic extension of g from $\mathbb{R}^n$ to $\mathbb{R}^{n+1}_+$. We will define below the periodic Besov spaces, in terms of harmonic and thermic extensions.

In our numerical calculations we use $q = \infty$ and we work with homogeneous versions (modulo polynomials), in the periodic formulation using kernels. These can be seen as periodized versions of functions on $\mathbb{R}^n$ (obtained by the periodization process using the Poisson summation formula), or simply as periodic extensions by reflection from Ω to $\mathbb{R}^n$, as described before. The periodic Besov spaces are defined below, following [57] (see also [51]). Our case applies to $n = 2$.

Let $\mathbb{Z}^n$ denote the additive group of points in $\mathbb{R}^n$ having integer coordinates. We consider the coset space $\mathbb{R}^n/\mathbb{Z}^n$, and identify functions on $\mathbb{R}^n/\mathbb{Z}^n$ with periodic functions on $\mathbb{R}^n$. $\mathbb{R}^n/\mathbb{Z}^n$ is naturally identified with the n-torus $T^n = \{(e^{2\pi i x_1}, \ldots, e^{2\pi i x_n})\colon (x_1, \ldots, x_n) \in \mathbb{R}^n/\mathbb{Z}^n\}$. This identification is given by the mapping $(x_1, \ldots, x_n) \to (e^{2\pi i x_1}, \ldots, e^{2\pi i x_n})$. From this, periodic functions on $\mathbb{R}^n$ are identified with functions on $\Omega = \{x \in \mathbb{R}^n\colon -1/2 < x_j \leqslant$

$1/2,\ j = 1, \ldots, n\}$, the fundamental domain, and such that $f(x+m) = f(m)$ for all lattice points $m \in \mathbb{Z}^n$. Integration on T^n in space is defined as integration on Ω in space,

$$\int_{T^n} f\,dx = \int_{\Omega} f\,dx.$$

Denote the functional spaces $X(T^n)$ by $X(\Omega)$.

There are several approaches to periodize the kernels P_t and W_t. The first is by the following periodization process, that will be applied to the kernels P_t and W_t. Given a function f defined on $\mathbb{R}^n$, a natural process to obtain the periodic function from f is through the construction

$$\sum_{m\in\mathbb{Z}^n} f(x+m). \tag{13}$$

We refer to (13) as the periodization of f.

Thus, for example the Poisson kernel on T^n will be $P_t^*(x) = \sum_{m\in\mathbb{Z}^n} P_t(x+m)$, with P_t defined in (12), and for any function $f \in L^1(T_n)$, define its Poisson integral by $f(x) * P_t^*(x)$, which is harmonic in $\Omega \times (0,\infty)$.

A second approach amounts to define $P_t^*(x) = \sum_{m\in\mathbb{Z}^n} e^{-2\pi t|m|} e^{2\pi i m\cdot x}$, and we let, for $x \in \Omega$,

$$P_t^* * g(x) = \sum_{m\in\mathbb{Z}^n} e^{-2\pi t|m|}\hat{g}(m)e^{2\pi i m\cdot x},$$

and

$$W_t^* * g(x) = \sum_{m\in\mathbb{Z}^n} e^{-2\pi t|m|^2}\hat{g}(m)e^{2\pi i m\cdot x}.$$

The spaces $\dot{B}^{\alpha}_{p,q}(\Omega) = \dot{B}^{\alpha}_{p,q}(T^n)$ can be defined as follows (based on [57] and [51], these definitions of the periodic Besov spaces are equivalent to all other definitions; note that we use the notations P_t and W_t for the above periodic versions P_t^* and W_t^* respectively, for simplicity of notation).

Definition 5. Let $\alpha \in \mathbb{R},\ p \geqslant 1,\ 1 \leqslant q < \infty,\ m \in \mathbb{N}_0$ and $k \in \mathbb{N}_0$ with $m > \frac{\alpha}{2}$ and $k > \alpha$. Then the distribution $g \in \dot{B}^{\alpha}_{p,q}(\Omega)$ if

$$\|g\|_{\dot{B}^{\alpha}_{p,q}(\Omega)} = \left(\int_0^{\infty}\left(t^{m-(\alpha/2)}\left\|\frac{\partial^m W_t}{\partial t^m} * g\right\|_{L^p(\Omega)}\right)^q \frac{dt}{t}\right)^{1/q} < \infty \tag{14}$$

$$\approx \left(\int_0^{\infty}\left(t^{k-\alpha}\left\|\frac{\partial^k P_t}{\partial t^k} * g\right\|_{L^p(\Omega)}\right)^q \frac{dt}{t}\right)^{1/q} < \infty, \tag{15}$$

where "$\approx$" denotes equivalent norms. Similarly, for $q = \infty$, the distribution $g \in \dot{B}^{\alpha}_{p,\infty}(\Omega)$ if

$$\|g\|_{\dot{B}^{\alpha}_{p,\infty}(\Omega)} = \sup_{t\geqslant 0}\left\{t^{m-(\alpha/2)}\left\|\frac{\partial^m W_t}{\partial t^m} * g\right\|_{L^p(\Omega)}\right\} < \infty \tag{16}$$

$$\approx \sup_{t\geqslant 0}\left\{t^{k-\alpha}\left\|\frac{\partial^k P_t}{\partial t^k} * g\right\|_{L^p(\Omega)}\right\} < \infty. \tag{17}$$

Equipped with $\|\cdot\|_{\dot{B}^{\alpha}_{p,q}(\Omega)}$, $\dot{B}^{\alpha}_{p,q}(\Omega)$ becomes a Banach space (where g is defined modulo polynomials only). Equivalent definitions and equivalent norms are obtained for any integers m, k satisfying $m > \frac{\alpha}{2}$ and $k > \alpha$. Thus, m, k are not additional parameters in the definition above.

Proposition 1. *We have the following continuous embeddings*

$$\dot{B}^{\alpha_1}_{p,q_1}(T^n) \subset \dot{B}^{\alpha_2}_{p,q_2}(T^n), \tag{18}$$

if either $\alpha_2 < \alpha_1$, or $\alpha_1 = \alpha_2$ and $1 \leqslant q_1 \leqslant q_2 \leqslant \infty$.

Remark 2. Let $\alpha \in \mathbb{R}$. Then the space of second derivatives of functions in $\dot{B}^{\alpha}_{p,q}(T^n)$ coincides isometrically with the space $\dot{B}^{\alpha-2}_{p,q}(T^n)$. In other words, $\Delta \dot{B}^{\alpha}_{p,q}(T^n) = \dot{B}^{\alpha-2}_{p,q}(T^n)$ (see [57]).

Remark 3. We have the following duality results for $\alpha \in \mathbb{R}$, $1 \leqslant p, q < \infty$: $(B^{\alpha}_{p,q}(T^n))' = B^{-\alpha}_{p',q'}(T^n)$, with $\frac{1}{p} + \frac{1}{p'} = 1$ if $1 < p < \infty$ and $p' = \infty$ if $p = 1$. Therefore, these spaces are dual spaces for $p, q > 1$. We also have that these spaces are separable if $1 \leqslant p \leqslant \infty$ and $1 \leqslant q < \infty$.

3. $u + v$ decomposition models

Here we consider only the case $n = 2$ for simplicity, however, the proposed models can also be applied in any dimension n. Given a periodic image f defined on the fundamental domain Ω (a square), we are interested in decomposing f into $u + v$, where u is cartoon or piecewise smooth, and v is noise or texture. Here we are interested in modeling v as

$$v = \Delta g, \quad \text{where } g \in \dot{B}^{\alpha}_{p,q}(\Omega),\ 0 < \alpha < 2,\ 1 \leqslant p, q \leqslant \infty, \tag{19}$$

while keeping $u \in BV(\Omega)$. Note that $v \in \dot{B}^{\alpha-2}_{p,q}(\Omega)$. For simplicity, we write L^2 for $L^2(\Omega)$, BV for $BV(\Omega)$, etc, from now on.

Let $\varphi\colon \mathbb{R}^2 \to \mathbb{R}$ be positive everywhere except at 0, continuous, convex and even function, such that $\varphi(0) = 0$ and there exist $a > 0$, $b \geqslant 0$ satisfying

$$-b + a|x| \leqslant \varphi(x) \leqslant b + a|x|, \quad \forall x \in \mathbb{R}^2_+. \tag{20}$$

The above condition imposes φ to have linear growth at infinity, and examples include $\varphi(x) = |x|$, $\varphi(x) = \sqrt{1 + |x|^2} - 1$, $\varphi(x) = |x_1| + |x_2|$, $\varphi(x) = \log\cosh|x|$, among others. We refer to Demengel and Temam [24] for a study of convex functions of measures and to Bouchitté and Buttazzo [6,14] on the lower semi-continuity results on BV and other types of functionals. Denote by ∇u the distributional derivative of u, as a Radon measure.

We study both the exact and the approximated (with L^2 square residual) minimization problems,

$$\inf_u \left\{ \int_\Omega \varphi(\nabla u) + \lambda \|f - u\|_{\dot{B}^{\alpha-2}_{p,q}},\ 1 \leqslant p, q \leqslant \infty,\ 0 < \alpha < 2 \right\}, \tag{21}$$

$$\inf_{u,g} \left\{ \int_\Omega \varphi(\nabla u) + \mu \int_\Omega |f - u - \Delta g|^2 + \lambda \|g\|_{\dot{B}^{\alpha}_{p,q}},\ 1 \leqslant p, q \leqslant \infty,\ 0 < \alpha < 2 \right\}. \tag{22}$$

When $\varphi(x_1, x_2) = \sqrt{x_1^2 + x_2^2}$, we obtain $\varphi(\nabla u) = \int_\Omega |\nabla u|$ the total variation of u in Ω.

Theoretically, the model (21) is an exact decomposition $f = u + v$, therefore it is simpler, whereas (22) decomposes $f = u + v + r$, where $r \in L^2$ is a residual. Moreover, (21) only has one parameter λ, and (22) has two parameters λ and μ. However, based on numerical experiments, (22) converges faster to a satisfactory decomposition.

Remark 4. When $\alpha = 1$, $p = \infty$, and $q = \infty$, v belongs to the Besov space $\dot{B}^{-1}_{\infty,\infty}$, that Y. Meyer proposes in [42] to model oscillatory patterns.

Remark 5. (i) It is reasonable to assume that $f \in L^2(\Omega)$ (assume here that $1 < p, q \leqslant \infty$). Since $u \in BV(\Omega) \subset L^2(\Omega)$ in two dimensions, we deduce that $v = f - u \in L^2(\Omega) \cap \dot{B}^{\alpha-2}_{p,q}(\Omega)$ in (21). Similarly in (22), we have

$$f \in L^2(\Omega), \qquad u \in BV(\Omega) \subset L^2(\Omega), \qquad f - u - v = f - u - \Delta g \in L^2(\Omega),$$

thus $v = \Delta g \in L^2(\Omega) \cap \dot{B}^{\alpha-2}_{p,q}(\Omega)$. The texture component v still belongs to $L^2(\Omega)$, as in the ROF model [50], but now v is modeled as a function in $\dot{B}^{\alpha-2}_{p,q}(\Omega)$. Thus, for any $v \in \dot{B}^{\alpha-2}_{p,q}(\Omega) = (\dot{B}^{2-\alpha}_{p',q'}(\Omega))'$, we can define

$$\|v\|_{(\alpha-2)} = \|v\|_{\dot{B}^{\alpha-2}_{p,q}(\Omega)} = \sup_{w \in \dot{B}^{2-\alpha}_{p',q'}(\Omega),\ \|w\|_{\dot{B}^{2-\alpha}_{p',q'}(\Omega)} \neq 0} \frac{\int_\Omega v w\, dx}{\|w\|_{\dot{B}^{2-\alpha}_{p',q'}(\Omega)}}.$$

Then, we can show that for $v \in L^2(\Omega) \cap \dot{B}^{\alpha-2}_{p,q}(\Omega)$, we have $\int_\Omega v(x)\,dx = 0$. Indeed, assume by contradiction that $\int_\Omega v(x)\,dx \neq 0$. Then for arbitrary constant $c \neq 0$, we have

$$\frac{\int_\Omega v(w+c)\,dx}{\|w+c\|_{(2-\alpha)}} = \frac{\int_\Omega v(x)w(x)\,dx + c\int_\Omega w(x)\,dx}{\|w\|_{(2-\alpha)}}$$

which cannot stay finite if $|c| \to \infty$, unless $\int_\Omega v(x)\,dx = 0$. In practice, we may have to reinforce this property, by imposing $\frac{\partial g}{\partial \vec{n}} = 0$ on $\partial\Omega$.

(ii) The above remark will imply that, if u is such that the energy in (21) is finite, then $\int_\Omega f(x)\,dx = \int_\Omega u(x)\,dx$. We can show that a similar property holds for (22). First, if (u, g) are such that the energy in (22) is finite, then $v = \Delta g$ can be identified with an element of $\dot{B}^{\alpha-2}_{p,q}(\Omega)$, with $0 < \alpha < 2$, thus using (i) we must have $\int_\Omega v(x)\,dx = 0$. We also have that $\int_\Omega (f - u - \Delta g)\,dx = 0$ if in addition (u, g) are minimizers of (22). Indeed, in this case, we have

$$\min_{c\in\mathbb{R}} \int_\Omega |f - (u+c) - \Delta g|^2\,dx = \int_\Omega |f - u - \Delta g|^2\,dx.$$

But $c = 0$ is the unique minimizer of this quadratic function, and in addition $c = \int_\Omega (f - u - \Delta g)\,dx$. Thus, we must have $\int_\Omega (f - u - \Delta g)\,dx = 0$, therefore $\int_\Omega f\,dx = \int_\Omega u\,dx$, if (u, g) are minimizers of (22).

(iii) Both energies in (21) and (22) are bounded from below by 0 and take finite values for $f \in L^2(\Omega)$, thus their infimum values are finite. Indeed, take for both $u = \frac{\int_\Omega f\,dx}{|\Omega|}$ and $g = 0$ in (22).

3.1. Existence and uniqueness of minimizers

Next, we would like to show existence of minimizers for the proposed models in the cases $1 < p, q \leqslant \infty$. The spaces $\dot{B}^{\alpha}_{p,q}$ are reflexive for $1 < p, q < \infty$, therefore any bounded sequence g_k weekly converges to a limit g. For the cases $1 < p \leqslant \infty$, $q = \infty$ or $p = \infty$, $1 < q \leqslant \infty$, these are dual spaces of separable spaces, therefore a bounded sequence g_k converges weakly-star. In both cases, we will obtain the lower semi-continuity of the norm $\|g\|_{\dot{B}^{\alpha}_{p,q}} \leqslant \liminf_k \|g_k\|_{\dot{B}^{\alpha}_{p,q}}$. We also recall that $\|u\|_{BV(\Omega)} = \|u\|_{L^1(\Omega)} + |u|_{BV(\Omega)}$, with $|u|_{BV(\Omega)} = \int_\Omega |\nabla u|$.

Theorem 1. *Let $f \in L^2(\Omega)$, let φ be as above, and $0 < \alpha < 2$, $1 < p, q \leqslant \infty$. Then there exists a minimizer $(u, g) \in (BV, \dot{B}^{\alpha}_{p,q})$ of the energy*

$$\inf_{(u,g)} \left\{ \mathcal{J}(u, g) = \int_\Omega \varphi(\nabla u) + \mu \int_\Omega |f - u - \Delta g|^2 + \lambda \|g\|_{\dot{B}^{\alpha}_{p,q}} \right\}. \tag{23}$$

Moreover the minimizer $(u, v = \Delta g) \in (BV, \dot{B}^{\alpha-2}_{p,q})$ is unique if φ is strictly convex (if φ is only convex, then $u + v$ is unique).

Proof. Recall that $\Omega = (0, 1)^2$, or $|\Omega| = 1$. Let $(u_k, g_k) \in (BV, \dot{B}^{\alpha}_{p,q})$ be a minimizing sequence of (23). We know that, for $v_k = \Delta g_k$, we have $\int_\Omega v_k(x)\,dx = 0$, since $v_k \in L^2(\Omega) \cap \dot{B}^{\alpha-2}_{p,q}(\Omega)$. For each u_k, we can obtain another minimizing sequence $u_k + c_k$ such that c_k minimizes $\int_\Omega |f - (u_k + c_k) - \Delta g_k|^2\,dx$. The unique minimizer of this quadratic function in c_k is $c_k = \int_\Omega (f - u_k - \Delta g_k)\,dx$. Thus, for the new minimizing sequence $u_k + c_k$, we have that $\int_\Omega (f - (u_k + c_k) - \Delta g_k)\,dx = \int_\Omega (f - u_k - \Delta g_k)\,dx - c_k = 0$. Therefore, we could have assumed from the beginning that the minimizing sequence (u_k, g_k) satisfies $\int_\Omega (f - u_k - \Delta g_k)\,dx = 0$, from which we deduce $\int_\Omega f(x)\,dx = \int_\Omega u_k(x)\,dx$.

The following uniform bounds hold, for all k,

$$\int_\Omega \varphi(\nabla u_k) \leqslant C, \tag{24}$$

$$\int_\Omega |f - u_k - \Delta g_k|^2 \leqslant C, \tag{25}$$

$$\|g_k\|_{\dot{B}^{\alpha}_{p,q}} \leqslant C. \tag{26}$$

The first condition (24) together with the properties on φ imply that $|u_k|_{BV} \leqslant C$, for all k.

By Poincaré–Wirtinger inequality,

$$\left\| u_k - \frac{1}{|\Omega|} \int_\Omega u_k \right\|_{L^2} \leqslant C|u_k|_{BV}, \tag{27}$$

and we therefore have $\|u_k\|_{L^1} \leqslant C'\|u_k\|_{L^2} \leqslant C$, for all k. This implies that

$$\|u_k\|_{BV} \leqslant C, \quad \forall k. \tag{28}$$

Therefore, there exists a function $u \in BV(\Omega)$ such that, up to a subsequence, u_k converges to u in the BV-weak* topology (strongly in L^1 and $\nabla u_k \to \nabla u$ weakly in the sense of measures). We also have

$$\int_\Omega \varphi(\nabla u) \leqslant \liminf_{k\to\infty} \int_\Omega \varphi(\nabla u_k). \tag{29}$$

The condition (25) and the fact that $f - u_k \in L^2$ is uniformly bounded, imply that g_k is uniformly bounded in the homogeneous Sobolev space $\dot{H}^2$. Therefore, there exists an h in $\dot{H}^2$ such that g_k converges to h weakly in $\dot{H}^2$.

The condition (26) implies that there exists a $g \in \dot{B}^\alpha_{p,q}$ such that, up to a subsequence, g_k converges to a function g weakly (or weakly-star) in $\dot{B}^\alpha_{p,q}$. We must have $g = h$ a.e., and

$$\|g\|_{\dot{B}^\alpha_{p,q}} \leqslant \liminf_{k\to\infty} \|g_k\|_{\dot{B}^\alpha_{p,q}}. \tag{30}$$

Similarly,

$$\|f - u - \Delta g\|_{L^2} \leqslant \liminf_{k\to\infty} \|f - u_k - \Delta g_k\|_{L^2}. \tag{31}$$

Finally, conditions (29), (30), and (31) imply that $\mathcal{J}(u, g) \leqslant \mathcal{J}(u_k, g_k)$, for all k. Therefore (u, g) minimizes $\mathcal{J}$.

To discuss uniqueness, we first rewrite the minimization problem (23) as

$$\inf_{(u,v)} \left\{ \mathcal{J}(u, v) = \int_\Omega \varphi(\nabla u) + \mu \int_\Omega |f - u - v|^2 + \lambda \|v\|_{\dot{B}^{\alpha-2}_{p,q}} \right\}, \tag{32}$$

for simplicity.

The functional $\mathcal{J}(u, v)$ in (32) is a sum of three convex functions. The second term $F(u, v) = \int_\Omega |f - u - v|^2\,dx$ is not strictly convex in the direction $(w, -w)$. If φ is strictly convex, then the functional is strictly convex and we obtain uniqueness of minimizers. If φ is only convex, but not strictly convex (like the total variation), then we obtain that if (u_1, v_1) and (u_2, v_2) are minimizers, then $u_1 + v_1 = u_2 + v_2$; thus we can say that if (u, v) is a minimizer, then $u + v$ is unique. □

A similar existence proof can be used to show the following theorem (see also [8]).

Theorem 2. *Let $f \in L^2(\Omega)$, let φ be as above, $0 < \alpha < 2$, and $1 < p, q \leqslant \infty$. Then there exists $u \in BV(\Omega)$ (and $v = f - u \in L^2(\Omega) \cap \dot{B}^{\alpha-2}_{p,q}(\Omega)$) which minimizes the energy*

$$\inf_u \left\{ \mathcal{J}(u) = \int_\Omega \varphi(\nabla u) + \lambda \|f - u\|_{\dot{B}^{\alpha-2}_{p,q}} \right\}. \tag{33}$$

3.2. Interpretation of parameters μ and λ

Recall the dilating operator $\tau_\delta f(x) = f(\delta x)$, $\delta > 0$. Suppose that $\int_\Omega \varphi(\nabla(\tau_\delta u)) = \delta^s \int_\Omega \varphi(\nabla u)$. Note that if in $\mathbb{R}^n$, $\varphi(\nabla u) = |\nabla u|$, then $s = 1 - n$. Next we would like to characterize the parameter λ in the model (21) with $F_1(u) = \int \varphi(\nabla u)\,dx$, $F_2(v) = \|v\|_{\dot{B}^{\alpha-2}_{p,q}}$, when the image is being dilated (zoom in when $0 < \delta < 1$ and zoom out when $\delta > 1$).

Lemma 1. *Let F_i be functionals with the dilating properties $F_i(\tau_\delta g) = \delta^{s_i} F_i(g)$, $i = 1, 2$.*

For a fixed f and $\lambda > 0$, let $(u_\lambda, v_\lambda = f - u_\lambda)$ be a minimizer for the following energy,

$$\mathcal{J}_{(f,\lambda)} = F_1(u) + \lambda F_2(f - u). \tag{34}$$

Then for $\lambda' = \lambda\delta^{s_1 - s_2}$, $(\tau_\delta u_\lambda, \tau_\delta v_\lambda = \tau_\delta f - \tau_\delta u_\lambda)$ is a minimizer for the energy $\mathcal{J}_{(\tau_\delta f,\lambda')}$.

Proof. Applying $\tau_\delta u_\lambda$ and $\tau_\delta v_\lambda$ to $\mathcal{J}_{(\tau_\delta f,\lambda')}$, we have

$$\mathcal{J}_{(\tau_\delta f,\lambda')}(\tau_\delta u_\lambda) = F_1(\tau_\delta u_\lambda) + \lambda' F_2(\tau_\delta v_\lambda) = \delta^{s_1} F_1(u_\lambda) + \lambda'\delta^{s_2} F_2(v_\lambda).$$

Since (u_λ, v_λ) is a minimizer for $\mathcal{J}_{(f,\lambda)}$, we have $\delta^{-s_1}\mathcal{J}_{(\tau_\delta f,\lambda')}(\tau_\delta u_\lambda)$ is minimized when $\lambda' = \lambda\delta^{s_1 - s_2}$. Therefore, $(\tau_\delta u_\lambda, \tau_\delta v_\lambda)$ minimizes the energy $\mathcal{J}_{(\tau_\delta f,\lambda')}$. □

Similarly, we have the following characterization of the parameters μ and λ for the model (22), with $F_1(u) = \int \varphi(\nabla u)\,dx$, $R(f - u - v) = \|f - u - v\|_{L^2}^2$, and $F_2(v) = \|v\|_{\dot{B}^{\alpha-2}_{p,q}}$.

Lemma 2. *Suppose we are in $\mathbb{R}^n$. Let F_i and R be functionals with the dilating properties $F_i(\tau_\delta g) = \delta^{s_i} F_i(g)$, $i = 1, 2$, and $R(\tau_\delta g) = \delta^r R(g)$.*

For a fixed f, μ and $\lambda > 0$, let $(u_{\mu,\lambda}, v_{\mu,\lambda})$ be a minimizer for the following energy,

$$\mathcal{J}_{(f,\mu,\lambda)} = F_1(u) + \mu R(f - u - v) + \lambda F_2(v). \tag{35}$$

Then for $\mu' = \mu\delta^{s_1 - r}$ and $\lambda' = \lambda\delta^{s_1 - s_2}$, $(\tau_\delta u_{\mu,\lambda}, \tau_\delta v_{\mu,\lambda})$ is a minimizer for the energy $\mathcal{J}_{(\tau_\delta f,\mu'\lambda')}$.

From Lemma 1, we see that the decomposition (34) is invariant under dilation when $s_1 = s_2$, i.e., $\lambda' = \lambda$. It is understood that noise is scale invariant. However, this scale invariance property is different for texture. When the image is being dilated, we expect the scale of the texture component to change as well.

The same analysis can be applied to variational models of the form (35). Suppose

$$F_1(u) = \int_\Omega \varphi(\nabla u)\,dx = \int_\Omega |\nabla u|\,dx = |u|_{BV}.$$

In $\mathbb{R}^n$, $n = 2$, $s_1 = 1 - n = -1$. Ideally, we would like the parameter μ in (35) to be scale invariant, i.e., $r = s_1$. Therefore, in $\mathbb{R}^n$, with $n = 2$, $R(w) = \|w\|_{L^2}$ would be more appropriate than $R(w) = \|w\|_{L^2}^2$ if w denotes a small noisy residual. However, in practice, the residual term $R(f - u_{\mu,\lambda} - v_{\mu,\lambda})$ for the minimizer $(u_{\mu,\lambda}, v_{\mu,\lambda})$ is negligible for large μ. We choose to keep $R(w) = \|w\|_{L^2}^2$, since in this way we ensure uniqueness for the model. Moreover, we can also obtain a characterization of minimizers, as explained next.

3.3. Characterization of minimizers

Recall the variational problem of decomposing f via (23) which can be written as

$$\inf_{(u,v)} \left\{ \mathcal{J}(u, v) = \varphi(\nabla u)(\Omega) + \mu\|f - u - v\|_{L^2}^2 + \lambda\|v\|_{\dot{B}^{\alpha-2}_{p,q}},\ \int_\Omega u = \int_\Omega f \right\}, \tag{36}$$

for $0 < \alpha < 2$. Here we denote $\varphi(\nabla u)(\Omega) = \int_\Omega \varphi(\nabla u)$.

Definition 6. Given a function $w \in L^2$ and $\lambda > 0$, define

$$\|w\|_{*,\lambda} = \sup_{(g,h)\in(BV,\dot{B}^{\alpha-2}_{p,q})} \frac{\langle w, g + h\rangle}{\varphi(\nabla g)(\Omega) + \lambda\|h\|_{\dot{B}^{\alpha-2}_{p,q}}}, \qquad \varphi(\nabla g)(\Omega) + \lambda\|h\|_{\dot{B}^{\alpha-2}_{p,q}} \neq 0, \tag{37}$$

where $\langle\cdot,\cdot\rangle$ is the L^2 inner product.

The following result is a general characterization of the minimizer, and we refer to [37] for the proof, inspired by [42]. We assume in addition that φ is homogeneous of degree 1.

Lemma 3. *Let* (u, v) *be an optimal decomposition of* $f \in L^2$ *via* (36)*, and denote* $w = f - u - v$*. Then we have the following*:

(1) $\|f\|_{*,\lambda} \leqslant \frac{1}{2\mu} \Leftrightarrow u = 0,\ v = 0,$ *and* $w = f$.
(2) *Suppose* $\|f\|_{*,\lambda} > \frac{1}{2\mu}$*; then* (u, v) *is characterized by the two conditions*

$$\|w\|_{*,\lambda} = \frac{1}{2\mu} \quad \text{and} \quad \langle w, u + v\rangle = \frac{1}{2\mu}\big(\varphi(\nabla u)(\Omega) + \lambda\|v\|_{\dot{B}^{\alpha-2}_{p,q}}\big). \tag{38}$$

We note that the polar norm computed in (37) is the maximum of the polar norms of each term (this will be seen again below). Also, obviously, the inf-convolution of the first and last term in (36) is an interpolated norm, thus the problem becomes "similar" with the ROF model, but now the BV semi-norm is substituted by the interpolated norm.

Another interpretation of the above characterization of minimizers and additional results are included in the following analysis, aimed to discuss a problem raised by J. Gilles in his Ph.D. [31]. In the model proposed in this present paper, an image f is split into a sum $u + v + w$ between three components. Let us assume for now that we work on the entire domain $\mathbb{R}^2$. The first component u is modeling the objects which are contained in f. In other words, u yields the geometric component of the image f. J. Gilles speculated that the second piece v would be the textured component of f, and w would be the noise. In our paper, we say that v is noise or texture and w is a residual (this residual is not given a meaning). Recall that the decomposition $f = u + v + w$ is obtained in a variational approach, and the energy which is minimized over all decompositions $f = u + v + w$ of f can be rewritten as

$$E(u, v, w) = |u|_{BV} + \mu\|w\|_2^2 + \lambda\|v\|_{(-\alpha)}. \tag{39}$$

This energy contains three terms: $|u|_{BV}$ and $\|w\|_2^2$ are reminiscent of the Rudin–Osher–Fatemi model [50]. The third term is the Besov norm of v. The smoothness exponent $-\alpha$ is negative and this means that v can be a tempered distribution ($\alpha > 0$). The norm in the homogeneous Besov space $\dot{B}^{-\alpha}_{p,q}$ will be denoted by $\|\cdot\|_{(-\alpha)}$ and the norm in the pre-dual space $\dot{B}^{\alpha}_{p',q'}$ by $\|\cdot\|_{\alpha}$. Let us consider the simplest case of an image corrupted by an additive white noise (random noise function whose Fourier spectrum is a constant). Then we immediately observe that J. Gilles' conjecture is inconsistent, since white noise is not a function in $L^2(\mathbb{R}^2)$ (see, for instance, [42], p. 31, regarding "deterministic white noise", and [44] where it is recalled that white noise probability measure is supported in $\bigcap_{\epsilon>0} H^{-n/2-\epsilon}_{loc}$). Thus the role played by the texture and the noise could be intertwined. These remarks imply that there is a need to better understand the meanings of v and w, and this is the goal of the following two theorems.

For computing u, v, and w, we first write $m = u + v$ and define the norm of m in the Banach space $Y = BV + \dot{B}^{-\alpha}_{p,q}$ as

$$|||m||| = \inf\big\{|u|_{BV} + \lambda\|v\|_{(-\alpha)}\big\}, \tag{40}$$

where the lower bound is computed over all decompositions of m into a sum $m = u + v$. Then one is led to find the optimal $m \in Y$ which minimizes

$$\mu\|f - g\|_2^2 + |||m|||.$$

Once m is computed, one will optimally split m into $u + v$. The general theory in [42] tells us that one should compute the dual norm in Y^* which is $\|h\|_{**} = \max\{\frac{\|h\|_\alpha}{\lambda},\ \|h\|_*\}$, where the norm $\|h\|_* = \|\cdot\|_G$ is computed in the "dual space" of BV as explained in [42]. Our first result says what happens if μ is small.

Theorem 3. *We have* $u = v = 0$ *and* $f = w$ *if and only if* $\|f\|_* \leqslant \frac{1}{2\mu}$ *and* $\|f\|_{(\alpha)} \leqslant \frac{\lambda}{2\mu}$.

Theorem 3 tells us that we cannot expect that w be the noise contained in f. One could also say that μ should not be small. If μ is large while λ is much larger, then Theorem 3 still yields $u = v = 0$. In this case, the price paid to use the v-component is too high. The proof of Theorem 3 is straightforward if the general theory of [42] is used.

If λ and μ are fixed and if f is highly oscillating, one would expect $\|f\|_{(\alpha)} > \frac{\lambda}{2\mu}$. The following theorem will tell us about what is happening to v.

Theorem 4. *Let us write* $\eta = \frac{(4-2\sqrt{3})\pi}{\lambda\mu}$. *If* $\|f\|_* \leqslant \eta$, $\|f\|_{(-\alpha)} = \epsilon \leqslant \eta$, *and* $\|f\|_{(\alpha)} > \frac{\lambda}{2\mu}$, *then* $u = 0$, $\|w\|_{(\alpha)} = \frac{\lambda}{2\mu}$, *and* $\|w\|_2 \leqslant \sqrt{\frac{\lambda\epsilon}{\mu}}$.

This theorem says that the energy of w is small. Moreover w is smooth which is a surprise. Viewed as a function in the Besov space $\dot{B}^{\alpha}_{p',q'}$, w is not small. Most of the energy of f is carried by v and this settles a problem raised by J. Gilles. Let us assume that the given image is a sum $f + \sigma z$ where f is as above, σ is a positive parameter and z is a stationary Gaussian process with a power spectrum $|\xi|^{-\gamma}$ and γ is a small positive number. Everything is assumed to take place on the two-dimensional torus. Then the noise belongs to the "dual space" of BV. This would not be true for the standard white noise. If $\alpha > 1 - \frac{\gamma}{2}$, the Besov norm of our Gaussian process is finite and does not exceed $C\sigma$. If σ is small enough, Theorem 4 applies and says that the noise cannot appear in w since w is smooth.

Let us prove Theorem 4. We first compare the optimal decomposition of f to the trivial one given by $u = 0$, $v = f$, and $w = 0$. We then have

$$|u|_{BV} + \mu\|w\|_2^2 + \lambda\|v\|_{(-\alpha)} \leqslant \lambda\|f\|_{(-\alpha)} \leqslant \lambda\epsilon.$$

This provides us with bounds on the first two terms of the left-hand side. We obtain $|u|_{BV} \leqslant \lambda\epsilon$, $\|w\|_2 \leqslant \sqrt{\frac{\lambda\epsilon}{\mu}}$. Using the isoperimetric inequality, we have

$$\|u\|_* \leqslant \frac{1}{2\sqrt{\pi}}\|u\|_2 \leqslant \frac{1}{4\pi}|u|_{BV}, \tag{41}$$

which yields $\|u\|_* \leqslant \frac{\lambda\eta}{4\pi}$ and $\|w\|_* \leqslant \frac{1}{2\sqrt{\pi}}\sqrt{\frac{\lambda\eta}{\mu}}$. The value of η is tailored to yield $\|f - v\|_* \leqslant \|u\|_* + \|w\|_* \leqslant \frac{1}{2\mu}$. We freeze v and minimize $E(u, v, w)$ as a functional in u. The general theory of [42] gives $u = 0$, as announced. Then one is left with $\mu\|w\|_2^2 + \lambda\|v\|_{(-\alpha)}$ to which the general theory applies again.

3.4. Numerical computations of the models

From the continuous embedding (18), we restrict the decompositions (22) and (21) to cases where $q = \infty$, thus providing weaker norms for the texture component v. The computations of $\|.\|_{\dot{B}^{\alpha}_{p,\infty}}$ require taking the supremum over all $t > 0$. In practice, for discrete images defined on a bounded domain, it is sufficient to take the supremum only over a discrete set of values

$$\{t_i = 2.5\tau^i\colon\ \tau = 0.9,\ i = 1, \ldots, N = 150\}.$$

These t_i's are chosen so that discretely $P_{t_1}(x)$ is a constant and $P_{t_N}(x)$ approximates the Dirac delta function.

Our calculations involve associated Euler–Lagrange equations. These are not well defined for BV functions (this could be done in a rigorous way using the techniques from [60]). However, we formally assume in this section that we work with functions $u \in W^{1,1}(\Omega)$, a dense subset of $BV(\Omega)$ and for which these Euler–Lagrange equations are well defined. We also assume that the function φ is differentiable everywhere (therefore, we work for instance with $\varphi(x) = \sqrt{\epsilon^2 + |x|^2}$ instead of $\varphi(x) = |x|$, for a small parameter $\epsilon > 0$).

Recall the minimization problem (22):

$$\inf_{u,g}\left\{\int_{\Omega} \varphi(\nabla u)\,dx + \mu\int_{\Omega} |f - u - \Delta g|^2\,dx + \lambda\|g\|_{\dot{B}^{\alpha}_{p,\infty}}\right\}. \tag{42}$$

Let K_t^{α} denotes either $(t^{2-\alpha}\frac{\partial^2 P_t}{\partial t^2})$ or $(t^{1-\alpha/2}\frac{\partial W_t}{\partial t})$ (in the periodic formulations).

3.4.1. The case $p < \infty$

For $1 \leqslant p < \infty$, and $0 < \alpha < 2$, (42) can be written as

$$\inf_{u,g}\left\{\int_{\Omega} \varphi(\nabla u)\,dx + \mu\int_{\Omega} |f - u - \Delta g|^2\,dx + \lambda\sup_{t>0}\left(\int_{\Omega} \left|K_t^{\alpha} * g\right|^p dx\right)^{1/p}\right\}. \tag{43}$$

This can also be written as a minimax problem

$$\inf_{u,g}\Big\{\sup_{t>0} L(u,g,t)\Big\},$$

with

$$L(u,g,t)=\int_\Omega \varphi(\nabla u)\,dx+\mu\int_\Omega |f-u-\Delta g|^2\,dx+\lambda\bigg(\int_\Omega \big|K_t^\alpha * g\big|^p\,dx\bigg)^{1/p},$$

whose optimal solution $(\bar u,\bar g,\bar t)$ is a saddle point.

We apply numerically the general Uzawa's algorithm, as presented in Ekeland and Temam [28] for minimax problems. However, we do not show here the convergence of the numerical approximation to the optimal solution (we cannot directly apply the general convergence result from [28], Chapter VII, Section 1). Uzawa's method constructs the sequences u_m, g_m, t_m, defined in the following way: start with any $t_0>0$, then we calculate u_0, g_0, t_1, u_1, g_1, etc. Thus,

- t_m being known, find (u_m, g_m) solution of $\inf_{u,g} L(u,g,t_m)$;
- then define t_{m+1} as solution of $\sup_{t>0} L(u_m,g_m,t)$, and repeat.

The numerical details are as follows:

(1) Start with a given $t_0=2.5\tau^i>0$, for some $i=1,\dots,N$, and let $m=0$.
(2) For $m\geqslant 0$, t_m is known and then define (u_m,g_m) as solution of $\inf_{u,g} L(u,g,t_m)$. Thus, we compute the Euler–Lagrange equations in u_m, g_m of $L(u_m,g_m,t_m)$, which are (44) and (45) below with $(u,g)=(u_m,g_m)$

$$0=-\nabla\cdot\big(\nabla\varphi(\nabla u)\big)-2\mu(f-u-\Delta g), \tag{44}$$

$$0=-2\mu\Delta(f-u-\Delta g)+\lambda\big\|K_{t_m}^\alpha * g\big\|_{L^p}^{1-p}K_{t_m}^\alpha * \big(\big|K_{t_m}^\alpha * g\big|^{p-2}K_{t_m}^\alpha * g\big) \tag{45}$$

in Ω. Equations (44) and (45) are solved using standard finite differences (as in [9,60] for the divergence term, with a semi-implicit scheme; these Euler–Lagrange equations are well defined, as in Ekeland–Temam [28], Chapter X, Section 4.3).
(3) t_{m+1} is obtained as solution of $\sup_{t>0} L(u_m,g_m,t)$. In practice, to avoid longer nonlocal computations, we only look near t_m, so

$$t_{m+1}=\operatorname*{arg\,max}_{t\in\{2.5\tau^{i-1},t_m=2.5\tau^i,2.5\tau^{i+1}\}}\big\|K_t^\alpha * g_{m+1}\big\|_{L^p}.$$

(4) Continue iterating steps 2 and 3 until the discrete energy $L(u_m,g_m,t_m)$ that we minimize becomes stationary.

The algorithm will guarantee that $L(u_m,g_m,t_m)\leqslant L(u_{m-1},g_{m-1},t_m)$ and $L(u_{m-1},g_{m-1},t_m)>L(u_{m-1},g_{m-1},t_{m-1})$.

3.4.2. The case $p=\infty$

For $p=\infty$, we define the sup norm in space by duality. Thus we minimize

$$\inf_{u,g}\bigg\{\int_\Omega \varphi(\nabla u)\,dx+\mu\int_\Omega |f-u-\Delta g|^2\,dx+\lambda\sup_{h\in L^1(\Omega),\,\|h\|_{L^1(\Omega)}\neq 0,\,t>0}\frac{\int_\Omega (K_t^\alpha * g)h\,dx}{\|h\|_{L^1(\Omega)}}\bigg\}, \tag{46}$$

by introducing a new unknown variable h. The problem can be written as a minimax problem, as

$$\inf_{u,g}\Big\{\sup_{\{t>0,\,h,\,\|h\|_{L^1}\neq 0\}} L(u,g,t,h)\Big\},$$

where now

$$L(u,g,t,h)=\int_\Omega \varphi(\nabla u)\,dx+\mu\int_\Omega |f-u-\Delta g|^2\,dx+\lambda\frac{\int_\Omega (K_t^\alpha * g)h\,dx}{\|h\|_{L^1(\Omega)}}.$$

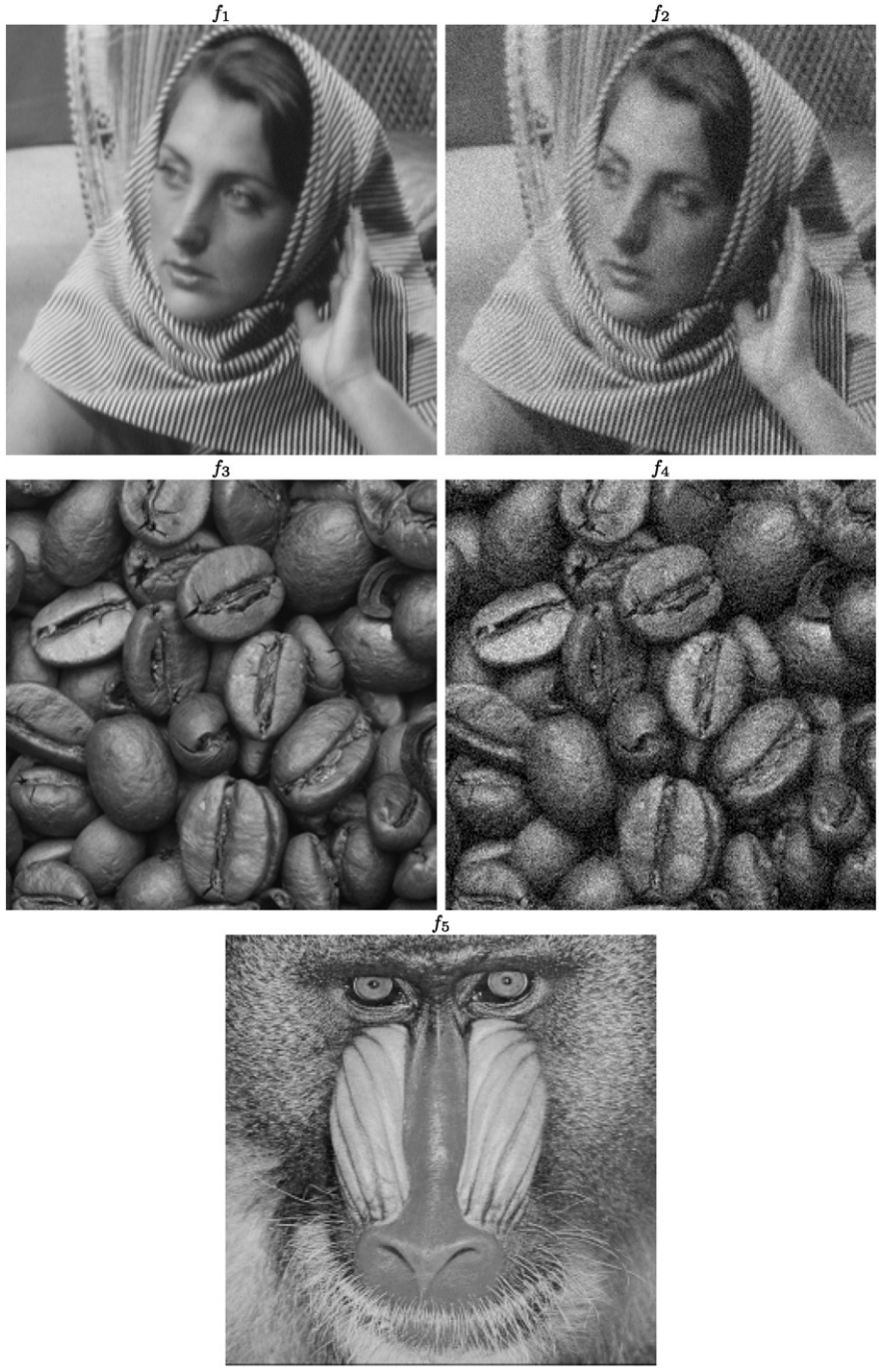

Fig. 2. Test images to be decomposed. Top four images: (left) real (noise free) images, (right) noisy (additive Gaussian) versions of images on the left. The bottom one is a noise free image.

Again the general Uzawa's algorithm [28] becomes:

- Start with $t_0 > 0$, $h_0 \in L^1(\Omega)$, $\|h\|_{L^1} \neq 0$, $m = 0$.
- Let (u_m, g_m) be solution of $\inf_{(u,g)} L(u, g, t_m, h_m)$.
- Let (t_{m+1}, h_{m+1}) be solution of $\sup_{\{t>0,\, h,\, \|h\|_{L^1}\neq 0\}} L(u_m, g_m, t, h)$, and repeat.

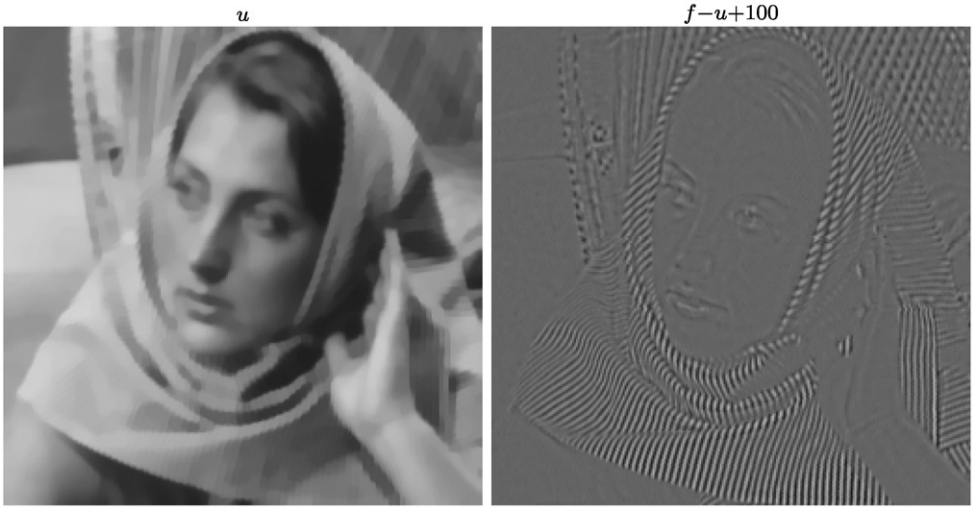

Fig. 3. A decomposition of f_1 using (42) with $\varphi(\nabla u) = |\nabla u|$, $\alpha = 1.9$, $p = 1$, $\mu = 1$, and $\lambda = 3e^{-06}$.

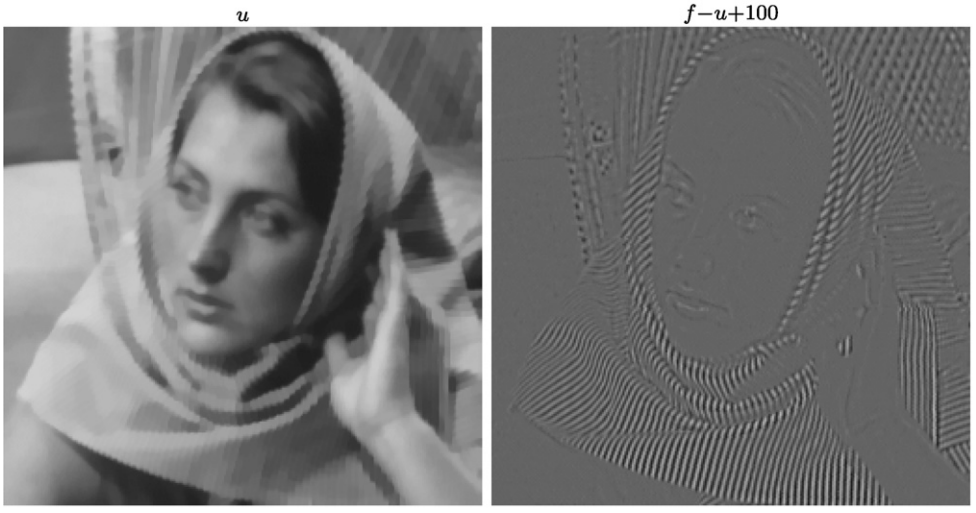

Fig. 4. A decomposition of f_1 using P_t in (42) with $\varphi(\nabla u) = |\nabla u|$, $\alpha = 1.5$, $p = 1$, $\mu = 1$, and $\lambda = 1e^{-04}$.

The details are as follows:

(1) Let $t_0 = 2.5\tau^i > 0$ given for some $i = 1, \ldots, N$, and let h_0 given, with $\|h_0\|_{L^1} \neq 0$, $m = 0$.
(2) For $m \geqslant 0$, suppose that t_m and h_m are known. Compute (u_m, g_m) as solution of the problem $\inf_{u,g} L(u, g, t_m, h_m)$ by the Euler–Lagrange equations and gradient descent (47) and (48)

$$u_T = \nabla \cdot \big(\nabla\varphi(\nabla u)\big) + 2\mu(f - u - \Delta g), \tag{47}$$

$$g_T = 2\mu\Delta(f - u - \Delta g) - \lambda \frac{K^{\alpha}_{t_m} * h_m}{\|h_m\|_{L^1}}, \tag{48}$$

and letting $(u, g) = (u_m, g_m)$.
(3) $t_{m+1} = \arg\max_{t>0} L(u_m, g_m, t, h_m)$. Again, we only look locally, near t_m, thus

$$t_{m+1} = \underset{t \in \{2.5\tau^{i-1}, t_m = 2.5\tau^i, 2.5\tau^{i+1}\}}{\arg\max} L(u_m, g_m, t, h_m).$$

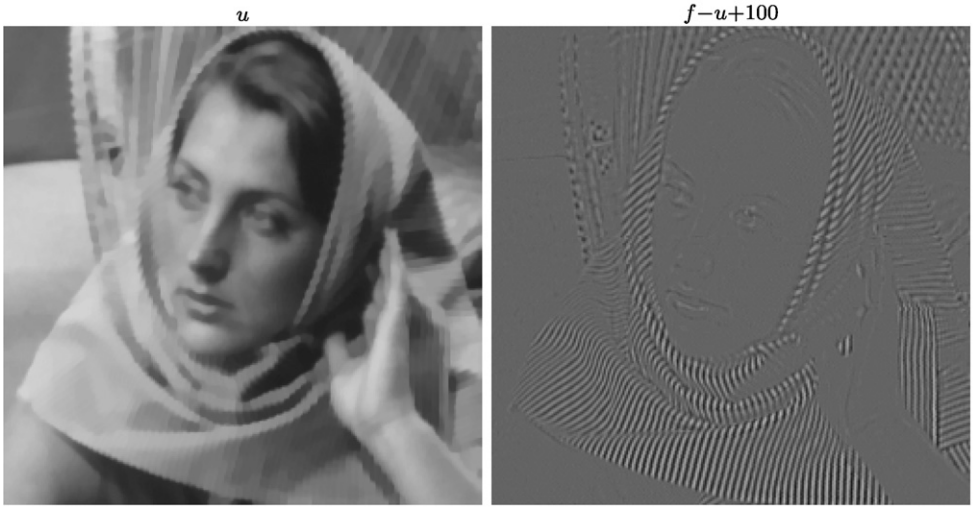

Fig. 5. A decomposition of f_1 using P_t in (42) with $\varphi(\nabla u) = |\nabla u|$, $\alpha = 1.0$, $p = 1$, $\mu = 1$, and $\lambda = 3e^{-03}$.

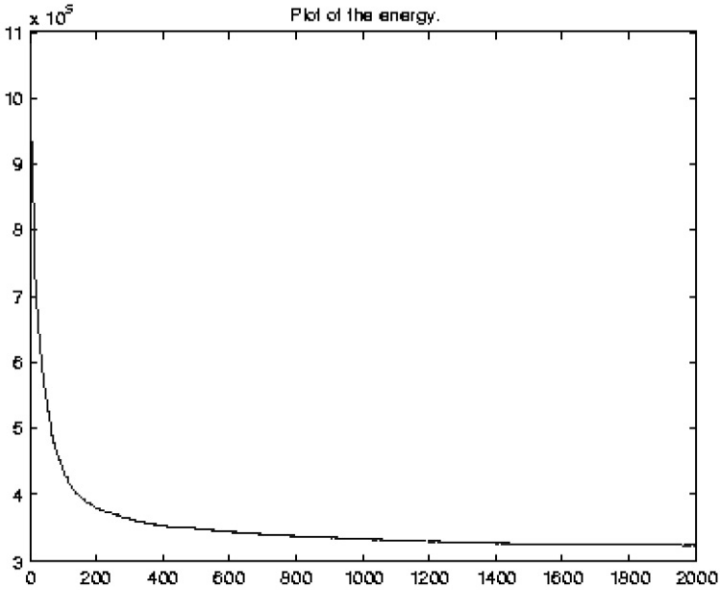

Fig. 6. Energy versus iterations.

(4) h_{m+1} is obtained as solution of $\sup_{h,\,\|h\|_{L^1}\neq 0} L(u_m, g_m, t_{m+1}, h)$, or using (49) as gradient ascent:

$$h_T = \frac{K^{\alpha}_{t_m} * g_m}{\|h\|_{L^1}} - \frac{\int_{\Omega}(K^{\alpha}_{t_m} * g_m)h\,dx}{\|h\|^2_{L^1}}\frac{h}{|h|}, \tag{49}$$

and letting $h = h_{m+1}$.

(5) Continue iterating until the discrete energy $L(u_m, g_m, t_m, h_m)$ becomes stationary.

The numerical algorithm guarantees that $L(u_m, g_m, t_m, h_m) \leqslant L(u_{m-1}, g_{m-1}, t_m, h_m)$, and that $L(u_{m-1}, g_{m-1}, t_m, h_m) \geqslant L(u_{m-1}, g_{m-1}, t_{m-1}, h_{m-1})$, etc.

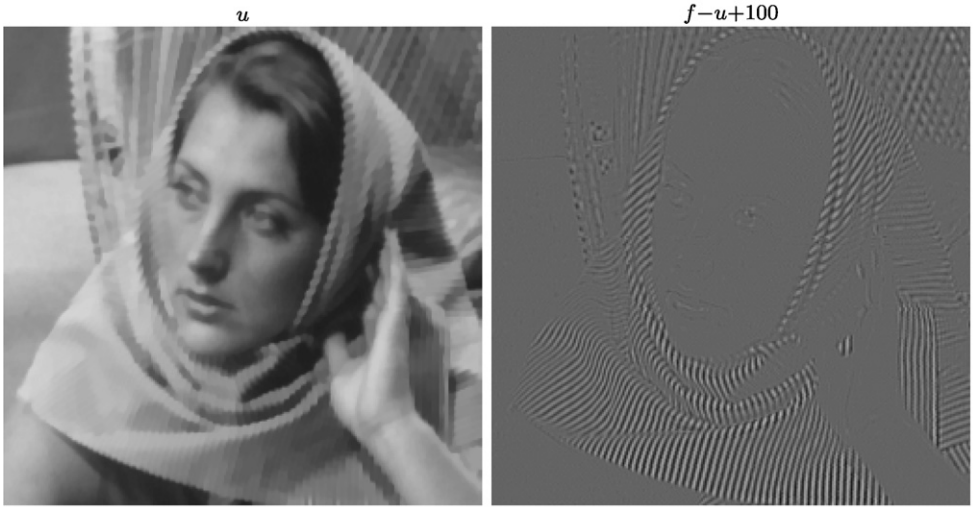

Fig. 7. A decomposition of f_1 using P_t in (42) with $\varphi(\nabla u) = |\nabla u|$, $\alpha = 0.5$, $p = 1$, $\mu = 1$, and $\lambda = 0.5$.

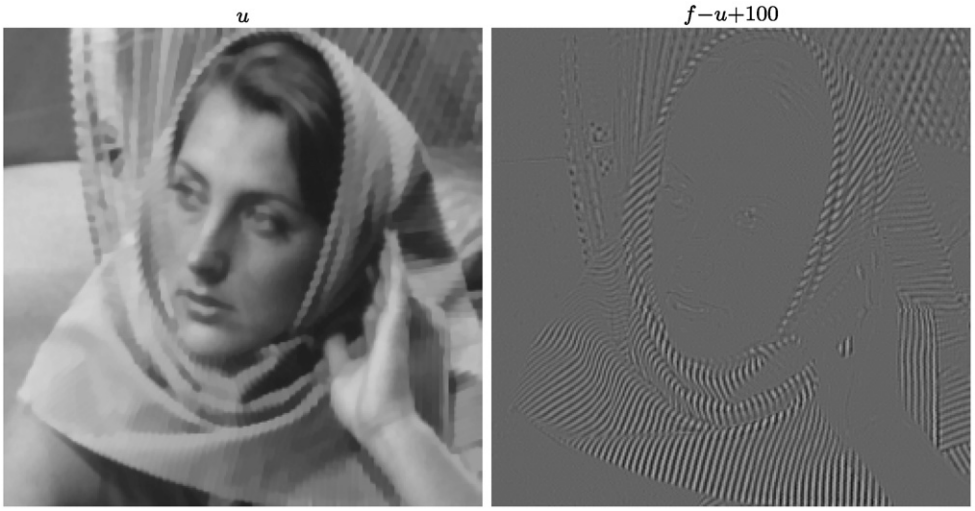

Fig. 8. A decomposition of f_1 using P_t in (42) with $\varphi(\nabla u) = |\nabla u|$, $\alpha = 0.1$, $p = 1$, $\mu = 1$, and $\lambda = 0.5$.

3.4.3. Exact decompositions

Recall the minimization problem (21):

$$\inf_u \left\{ \int_\Omega \varphi(\nabla u)\,dx + \lambda \|f-u\|_{\dot{B}^{\alpha-2}_{p,\infty}},\ 1 \leqslant p \leqslant \infty,\ 0 < \alpha < 2 \right\}. \tag{50}$$

Let H_t^α denotes either $(t^{2-\alpha} P_t)$ or $(t^{1-\alpha/2} W_t)$. For $1 \leqslant p < \infty$, we are led to consider

$$\inf_u \left\{ \int_\Omega \varphi(\nabla u)\,dx + \lambda \sup_{t>0} \left(\int_\Omega \left| H_t^\alpha * (f-u) \right|^p \right)^{1/p} \right\}, \tag{51}$$

or as minimax problem $\inf_u \{\sup_{t>0} L(u,t)\}$, with

$$L(u,t) = \int_\Omega \varphi(\nabla u)\,dx + \lambda \left(\int_\Omega \left| H_t^\alpha * (f-u) \right|^p \right)^{1/p}.$$

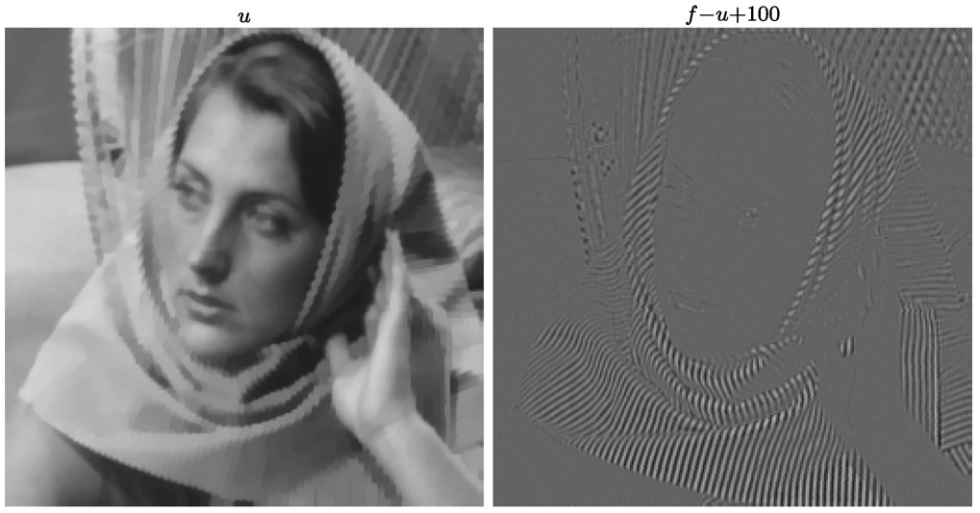

Fig. 9. A decomposition of f_1 using P_t in (50) with $\varphi(\nabla u) = |\nabla u|$, $\alpha = 0.8$, $p = 1$, $\lambda = 5000$.

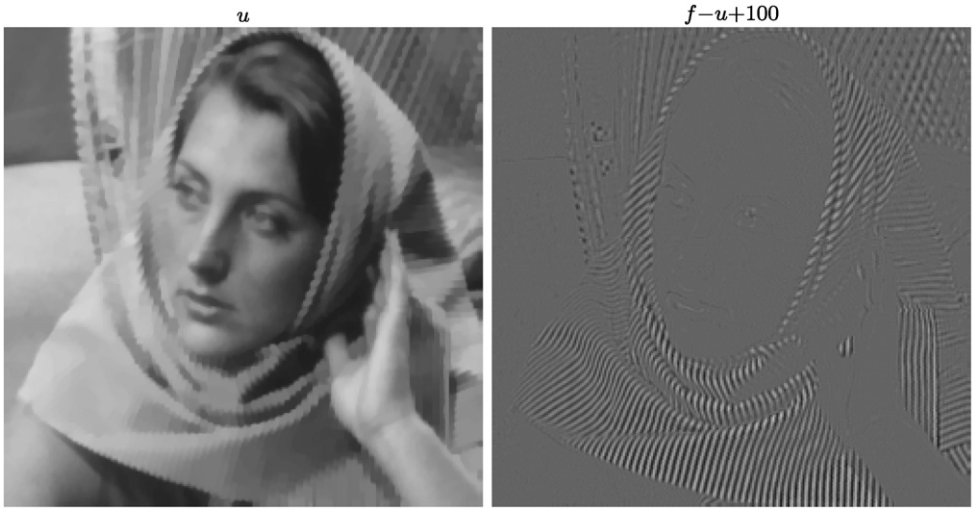

Fig. 10. A decomposition of f_1 using P_t in (50) with $\varphi(\nabla u) = |\nabla u|$, $\alpha = 1$, $p = 1$, $\lambda = 1500$.

Thus the Uzawa's steps are as before:

- Start with $t_0 > 0$, $m = 0$.
- Let u_m be solution of $\inf_u L(u, t_m)$.
- Let t_{m+1} be solution of $\sup_{t>0} L(u_m, t)$, and repeat.

Briefly the details are:

(1) Let $t_0 = 2.5\tau^i > 0$ given for some $i = 1, \ldots, N$, $m = 0$.
(2) For $m \geqslant 0$, assume that t_m is known, and let u_m be solution of $\inf_u L(u, t_m)$ by solving the Euler–Lagrange equation with gradient descent

$$u_T = \nabla \cdot \big(\nabla\varphi(\nabla u)\big) + \lambda \big\| H^\alpha_{t_m} * (f-u) \big\|^{1-p}_{L^p} H^\alpha_{t_m} * \big(\big| H^\alpha_{t_m} * (f-u) \big|^{p-2} H^\alpha_{t_m} * (f-u) \big), \tag{52}$$

and we let $u_m = u$.

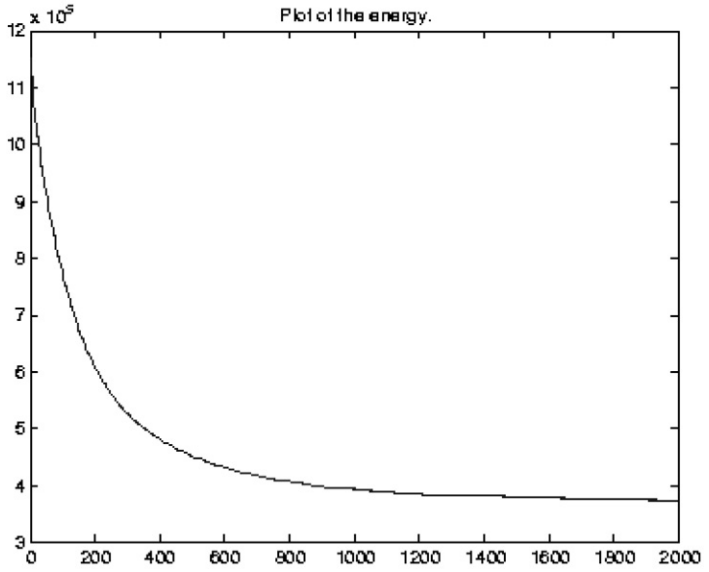

Fig. 11. Energy versus iterations.

Fig. 12. A decomposition of f_1 using P_t in (50) with $\varphi(\nabla u) = |\nabla u|$, $\alpha = 1.5$, $p = 1$, $\lambda = 50$.

(3) Let $t_{m+1} = \arg\max_{t>0}(\int_\Omega |H_t^\alpha * (f - u_m)|^p)^{1/p}$, again in practice looking only locally, near t_m.
(4) Repeat until energy becomes stationary.

For $p = \infty$, (50) can be seen as

$$\inf_u \left\{ \int_\Omega \varphi(\nabla u)\,dx + \lambda \sup_{h\in L^1,\, \|h\|_{L^1}>0,\, t>0} \frac{\int_\Omega (H_t^\alpha * (f-u))h}{\|h\|_{L^1}} \right\}. \tag{53}$$

The corresponding minimax problem is $\inf_u\{\sup_{t>0,h,\,\|h\|_{L^1}\neq 0} L(u,t,h)\}$, with

$$L(u,t,h) = \int_\Omega \varphi(\nabla u)\,dx + \lambda \frac{\int_\Omega (H_t^\alpha * (f-u))h}{\|h\|_{L^1}}.$$

The Uzawa's numerical steps are

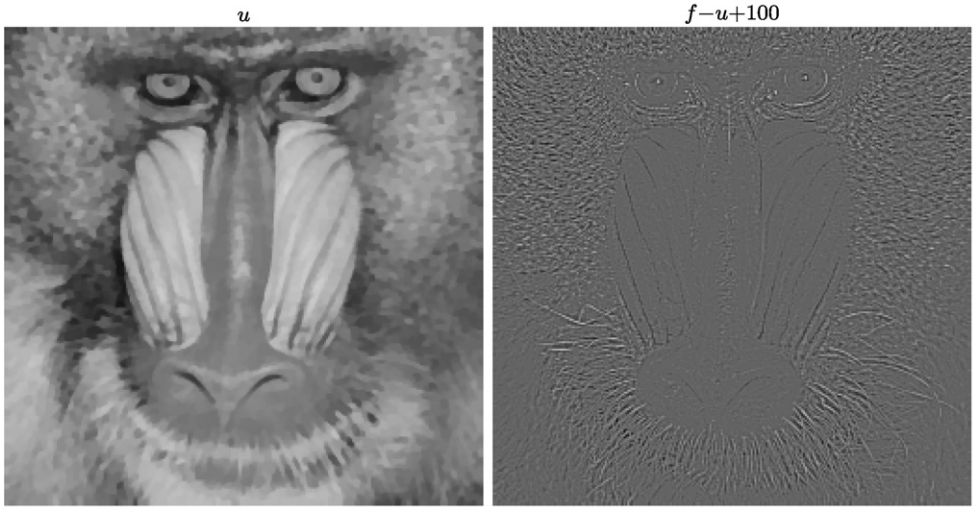

Fig. 13. A decomposition of f_5 using P_t in (50) with $\varphi(\nabla u) = |\nabla u|$, $\alpha = 1$, $p = 1$, $\lambda = 2000$.

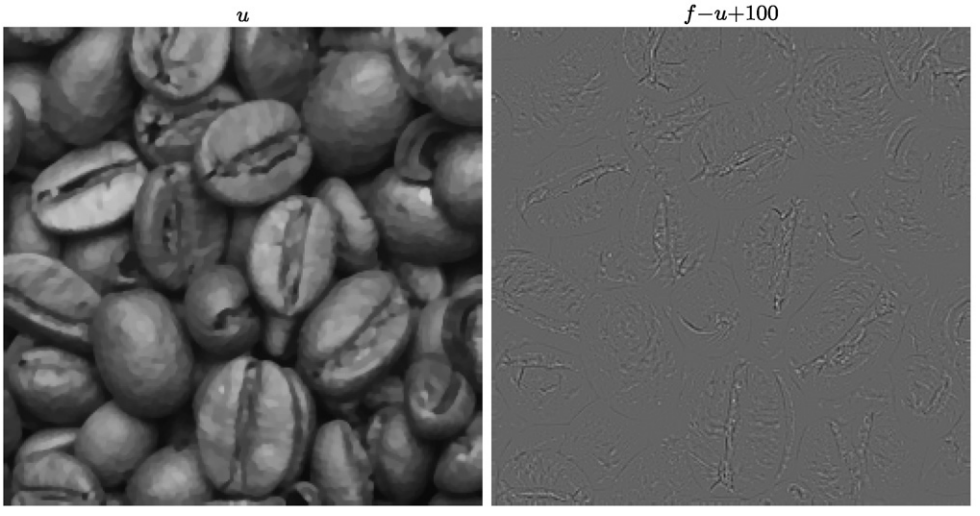

Fig. 14. A decomposition of f_3 using P_t in (50) with $\varphi(\nabla u) = |\nabla u|$, $\alpha = 1$, $p = 1$, $\lambda = 1500$.

- Start with $t_0 > 0$, h_0, $\|h_0\|_{L^1} \neq 0$, $m = 0$.
- u_m is solution of $\inf_u L(u, t_m, h_m)$.
- (t_{m+1}, h_{m+1}) is solution of $\sup_{t>0,\, h,\, \|h\|_{L^1}} L(u_m, t, h)$, and repeat.

Briefly the main details are:

(1) Let $t_0 = 2.5\tau^i > 0$ given for some $i = 1, \ldots, N$, and let h_0 given, with $\|h_0\|_{L^1} \neq 0$, $m = 0$.
(2) For $m \geqslant 0$, assume that (t_m, h_m) is known, and let u_m be solution of $\inf_u L(u, t_m, h_m)$ by solving the Euler–Lagrange equation with gradient descent

$$u_T = \nabla \cdot \big(\nabla\varphi(\nabla u)\big) + \lambda \frac{H^{\alpha}_{t_m} * h_m}{\|h_m\|_{L^1}}, \tag{54}$$

and let $u_m = u$.
(3) $t_{m+1} = \arg\max_{t>0} L(u_m, t, h_m)$, again looking locally, near t_m.

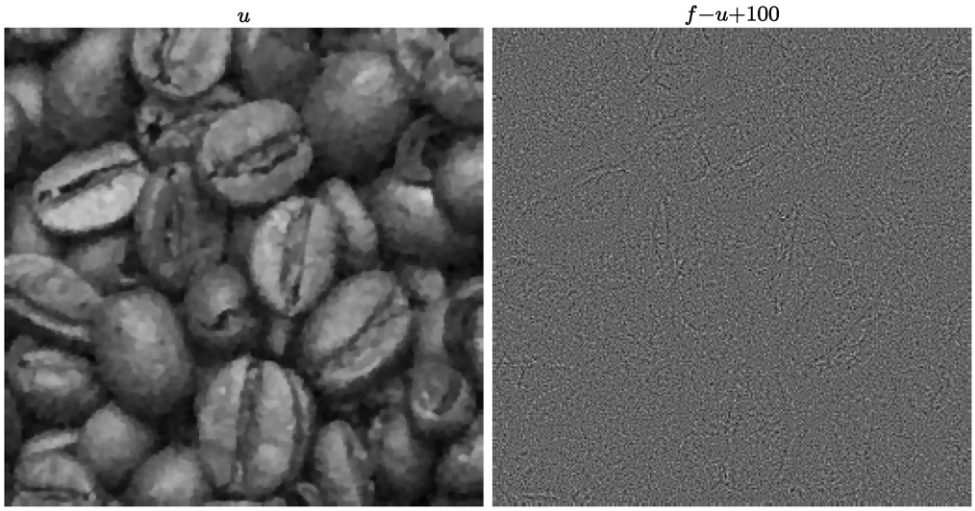

Fig. 15. A decomposition of the noisy image f_4 using P_t in (50) with $\varphi(\nabla u) = -\beta + \sqrt{|\nabla u|^2 + \beta^2}$, $\beta = \sqrt{10}$, $\alpha = 1$, $p = 1$, $\lambda = 1500$.

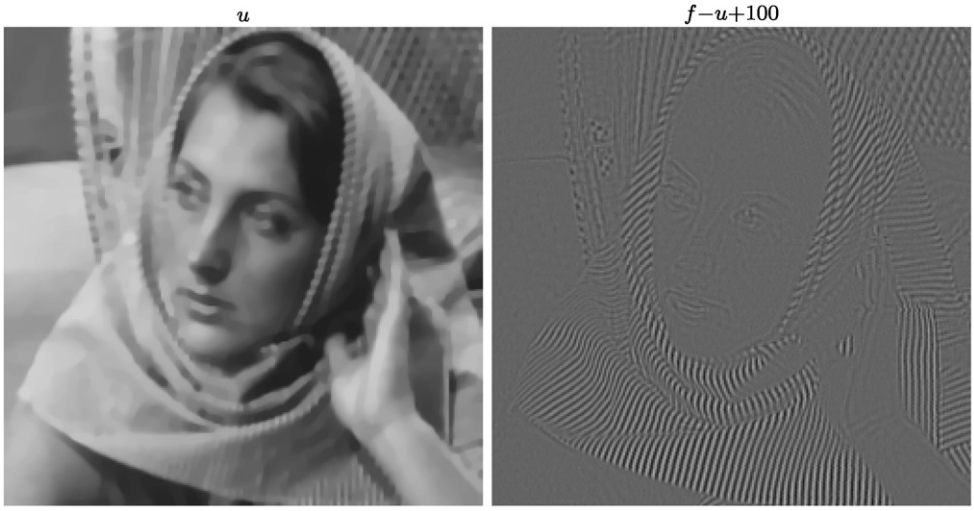

Fig. 16. A decomposition of f_1 using P_t in (42) with $\varphi(\nabla u) = |\nabla u|$, $\alpha = 1$, $p = \infty$, $\mu = 10$, and $\lambda = 1$.

(4) $$h_{m+1} = \underset{h,\, \|h\|_{L^1}>0}{\arg\max}\ L(u_m, t_{m+1}, h), \quad \text{or} \quad h_{m+1} = \underset{h,\, \|h\|_{L^1}>0}{\arg\max}\ \frac{\int_\Omega (H^\alpha_{t_{m+1}} * (f - u_m))h\,dx}{\|h\|_{L^1}},$$

will be solved as before by Euler–Lagrange equation and gradient ascent (similar with (49)).

(5) Repeat the above steps.

Remark 6. In practice, we solve all these models on the original image domain with Neumann boundary conditions on u (we only extend periodically by reflection the functions to compute the convolutions using the Fast Fourier Transform).

Again, we do not show here convergence of our numerical algorithms. However very satisfactory experimental results are presented next, together with the energy decrease versus iterations.

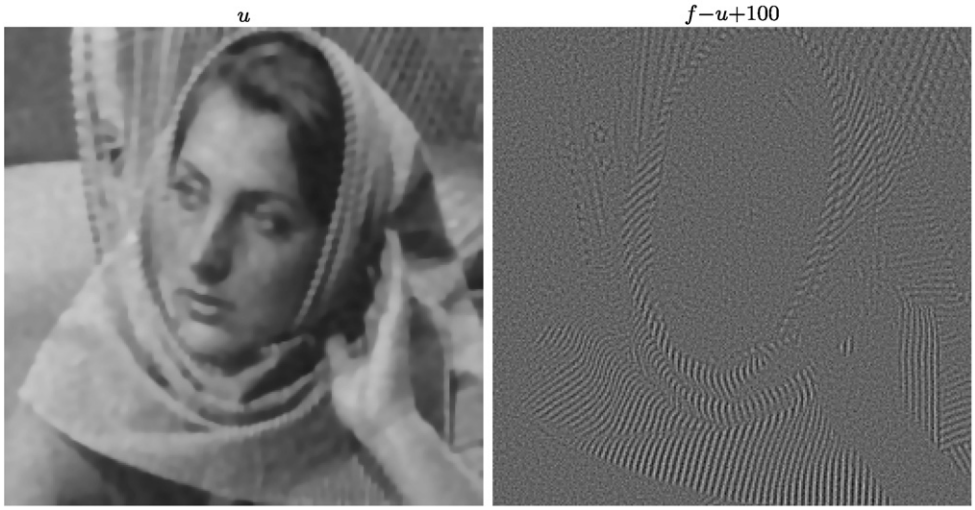

Fig. 17. A decomposition of the noisy image f_2 using the Cauchy–Poisson kernel P_t in (42) with $\varphi(\nabla u) = -\beta + \sqrt{|\nabla u|^2 + \beta^2}$, $\beta = \sqrt{10}$, $\alpha = 1$, $p = 2$, $\mu = 10$, and $\lambda = 0.01$.

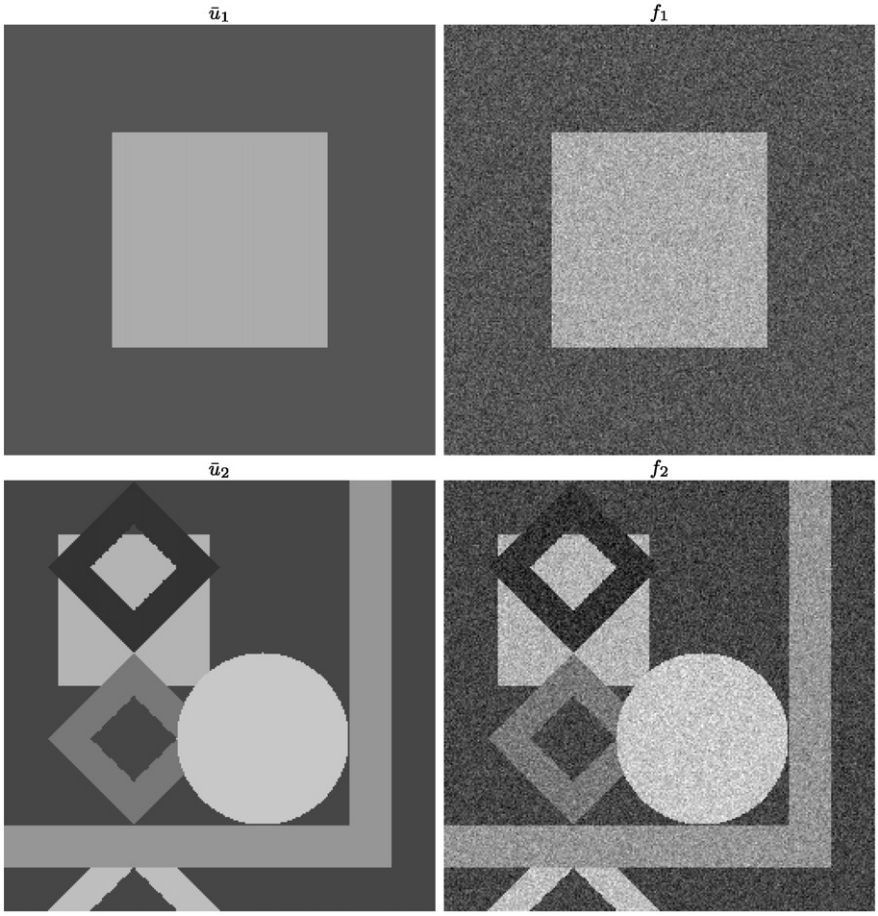

Fig. 18. Exact cartoon images and their noisy versions corrupted by additive Gaussian noise.

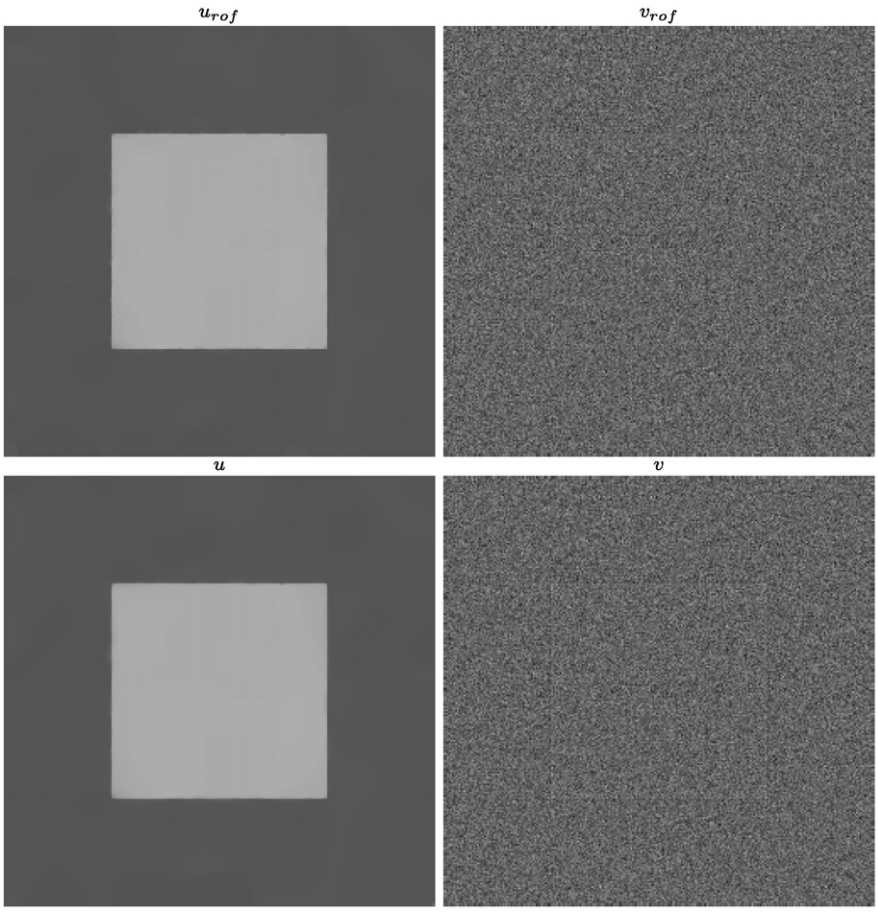

Fig. 19. The top row shows a decomposition of f_1 from Fig. 18 using the ROF model (2) with $rmse = 1.3923$. The bottom row shows a decomposition of the same image using W_t, $p = 2$ and $\alpha = 1$ in (50) with $rmse = 1.7200$.

4. Numerical results

In this section, we would like to show some numerical results of the proposed models applied to images with the presence of noise and/or texture. We also test the proposed models on cartoon images corrupted by additive Gaussian noise. Let $\bar{u}$ and u be the exact image and the reconstructed image, respectively, of size $n \times n$. For denoising, we measure the goodness of the reconstructed image by the following quantity

$$rmse = \left(\frac{\sum_{i,j=1}^{n} |u_{i,j} - \bar{u}_{i,j}|^2}{n^2} \right)^{1/2}. \tag{55}$$

Figures 3–8 show the decompositions of f_1 from Fig. 2 and an energy decrease versus iterations, with $p = 1$ and with different values of α, $0 < \alpha < 2$ using the Cauchy–Poisson kernel P_t in the approximated model (42) with $\varphi(\nabla u) = |\nabla u|$. Notice the improvement in terms of cartoon and texture decomposition of f_1 as $\alpha \to 0$. Local oscillations are well captured in v.

Figures 9–12 show an exact decomposition (50) of f_1 from Fig. 2 with $\varphi(\nabla u) = |\nabla u|$, $p = 1$ and various values of α using the Cauchy–Poisson kernel P_t.

Figure 13 shows an exact decomposition (50) of f_5 from Fig. 2 with $\varphi(\nabla u) = |\nabla u|$, $p = 1$ and $\alpha = 1$ using the Cauchy–Poisson kernel P_t.

Figure 14 shows an exact decomposition (50) of f_3 from Fig. 2 with $\varphi(\nabla u) = |\nabla u|$, $p = 1$ and $\alpha = 1$ using the Cauchy–Poisson kernel P_t.

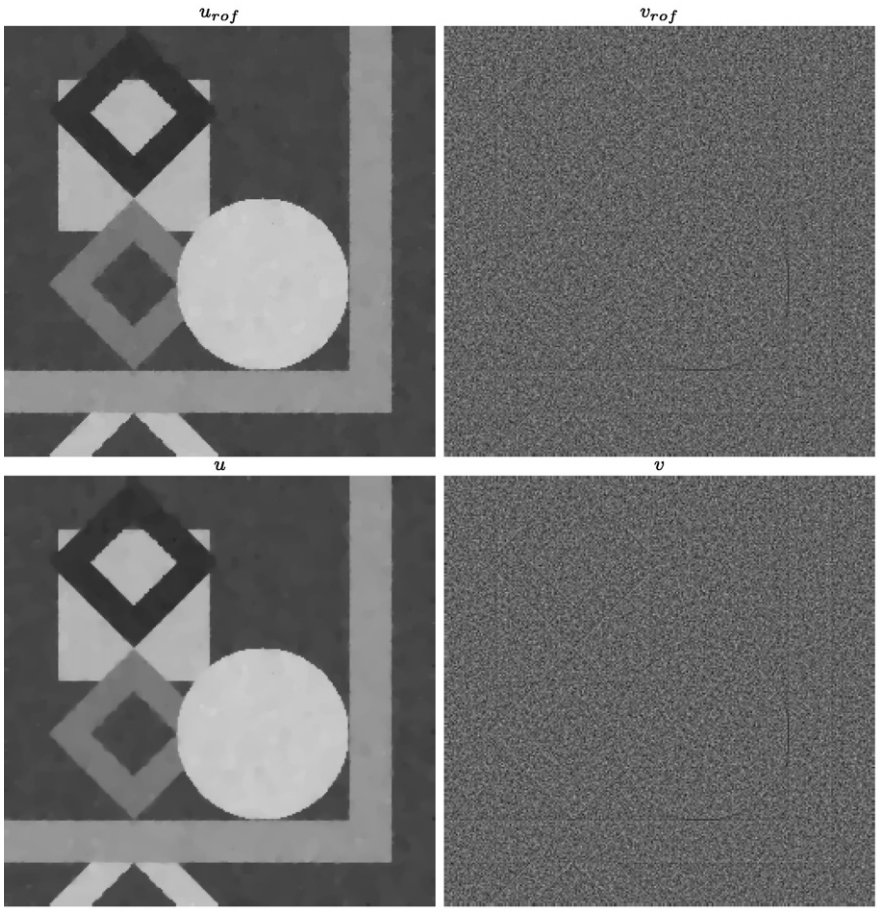

Fig. 20. The top row shows a decomposition of f_2 from Fig. 18 using the ROF model (2) with $rmse = 4.4287$. The bottom row shows a decomposition of the same image using W_t, $p = 2$ and $\alpha = 1$ in (50) with $rmse = 4.7221$.

Fig. 21. A decomposition of f_1 from Fig. 2 using the model (59) with the kernel being the characteristic function of a square centered at 0 having 3-pixel length for the sides, $\varphi(\nabla u) = |\nabla u|$, $p = 1$, and $\lambda = 1.5$.

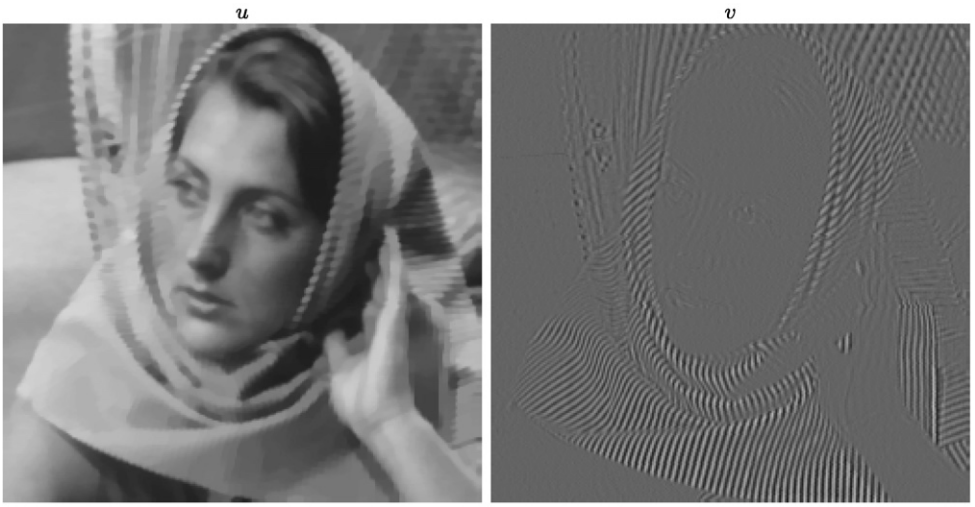

Fig. 22. A decomposition of f_1 from Fig. 2 using the model (59) with the kernel being the characteristic function of a square centered at 0 having 5-pixel length for the sides, $\varphi(\nabla u) = |\nabla u|$, $p = 1$, and $\lambda = 1.5$.

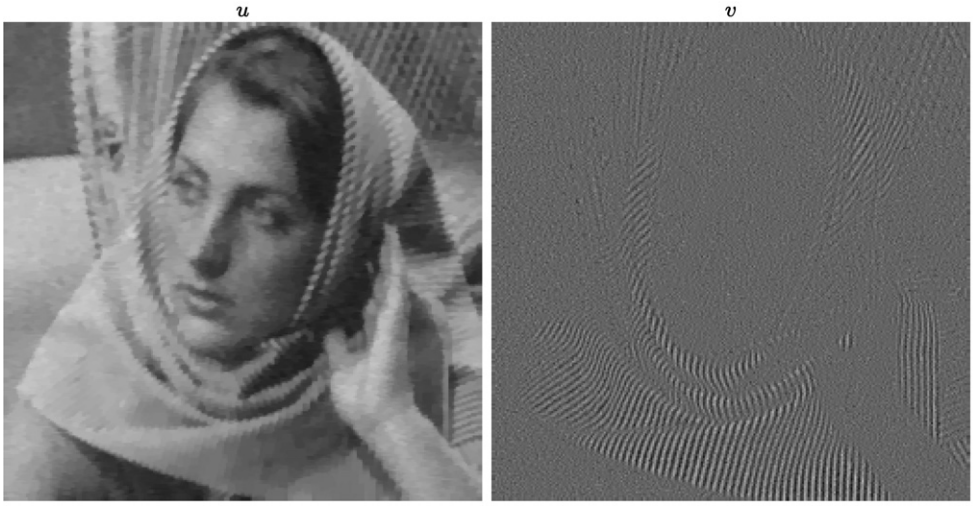

Fig. 23. A decomposition of the noisy image f_2 from Fig. 2 using the model (59) with the kernel being the characteristic function of a square centered at 0 having 3-pixel length for the sides, $\varphi(\nabla u) = |\nabla u|$, $p = 1$, and $\lambda = 2$.

Figure 15 shows an exact decomposition (50) of f_4 from Fig. 2 with $\varphi(\nabla u) = -\beta + \sqrt{|\nabla u|^2 + \beta^2}$, $\beta = \sqrt{10}$, $p = 1$ and $\alpha = 1$ using the Cauchy–Poisson kernel P_t.

Figure 16 shows a decomposition of f_1 from Fig. 2 using the Cauchy–Poisson kernel P_t in (42) with $\varphi(\nabla u) = |\nabla u|$, $p = \infty$, $\alpha = 1$. Here $v \in E = \dot{B}^{-1}_{\infty,\infty}$ (as proposed by Y. Meyer). The oscillatory component v tends to be uniform, and captures other non-repeated patterns as well. This comes from the property of L^∞.

Figure 17 shows a decomposition of f_2 from Fig. 2 using the Cauchy–Poisson kernel in (42) with $\varphi(\nabla u) = -\beta + \sqrt{|\nabla u|^2 + \beta^2}$, $\beta = \sqrt{10}$, $p = 2$, $\alpha = 1$. Here $v \in \dot{B}^{-1}_{2,\infty}$.

In Figs. 19–20, we show the results for image denoising using the model (50), and the ROF model (2) for comparison, applied to the noisy images from Fig. 21. We notice that the proposed new model does not give improvement over the ROF model in this denoising case.

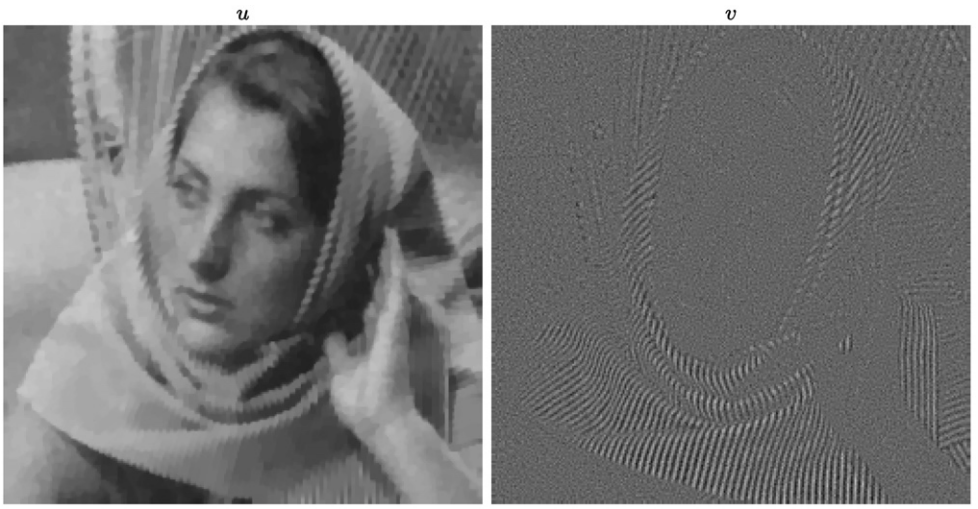

Fig. 24. A decomposition of f_2 from Fig. 2 using the model (59) with the kernel being the characteristic function of a disk centered at 0 having the radius equal to 1-pixel length, $\varphi(\nabla u) = |\nabla u|$, $p = 1$, and $\lambda = 1.5$.

5. Discussion

As seen from Figs. 1–17, the proposed models perform fairly well on texture decompositions. The repeated patterns are well captured in the v component when $p = 1$. For the case with $p = \infty$, other oscillations (not just repeated patterns) are also captured in v. However, for denoising of cartoon images (Figs. 19–20), we do not see an improvement over the ROF model (2).

In the minimization problem (50), choosing a parameter $\alpha > 0$ amounts to finding a $\bar{t} > 0$ such that $\|H^{\alpha}_{\bar{t}} * v\|_{L^p}$ is optimal. Therefore, (50) can be seen as the minimization problem,

$$\inf_{u\in BV}\left\{\int_{\Omega}\varphi(\nabla u)+\lambda\left\|K_{\bar{t}}*(f-u)\right\|_{L^p},\ 1\leqslant p\leqslant\infty\right\}, \tag{56}$$

where $K_{\bar{t}}$ is either $P_{\bar{t}}$ or $W_{\bar{t}}$, for some fixed $\bar{t} = \bar{t}(\alpha) > 0$. Computationally, (56) is less demanding compared to (50). Similarly, we can also consider a more simple problem instead of (42),

$$\inf_{u\in BV}\left\{\int_{\Omega}\varphi(\nabla u)+\mu\|f-u-\Delta g\|^2_{L^2}+\lambda\left\|\frac{\partial^2 K_{\bar{t}}}{\partial t^2}*(f-u)\right\|_{L^p},\ 1\leqslant p\leqslant\infty\right\}. \tag{57}$$

Now, $\bar{t}$ is another parameter to the problem, in addition to μ and λ.

Also note that, given $f \in L^p$ for $1 \leqslant p < \infty$, we have $K_t * f \to f$ strongly in L^p as $t \to 0$, [30]. Therefore, as $\bar{t} \to 0$, (56) becomes the minimization problem,

$$\inf_{u\in BV}\left\{\int_{\Omega}\varphi(\nabla u)+\lambda\left\|(f-u)\right\|_{L^p},\ 1\leqslant p<\infty\right\}. \tag{58}$$

Clearly, the norm $\|K_t * v\|_{L^p}$ is weaker than the norm $\|v\|_{L^p}$, for all $t > 0$.

From the numerical results, we notice that the case $p = 1$ gives better cartoon-texture decomposition, and captures well repeated patterns. M. Green [33] experimentally shows that texture-like natural images, when being convolved with kernels of zero mean, have a Laplacian probability distribution (that would amount to minimize an L^1 norm with kernel convolution). In all the above models, we assume that oscillatory components have zero mean; therefore, P_t or W_t are suitable, and our result is in agreement with M. Green's results.

Remark 7. When smooth kernels are used, for instance the Poisson or Gaussian kernel, then the model in (56) is related to the idea arising in discussions, which at that time were not motivated by Besov spaces, between the fourth author and B. Engquist at Princeton University in March 2004.

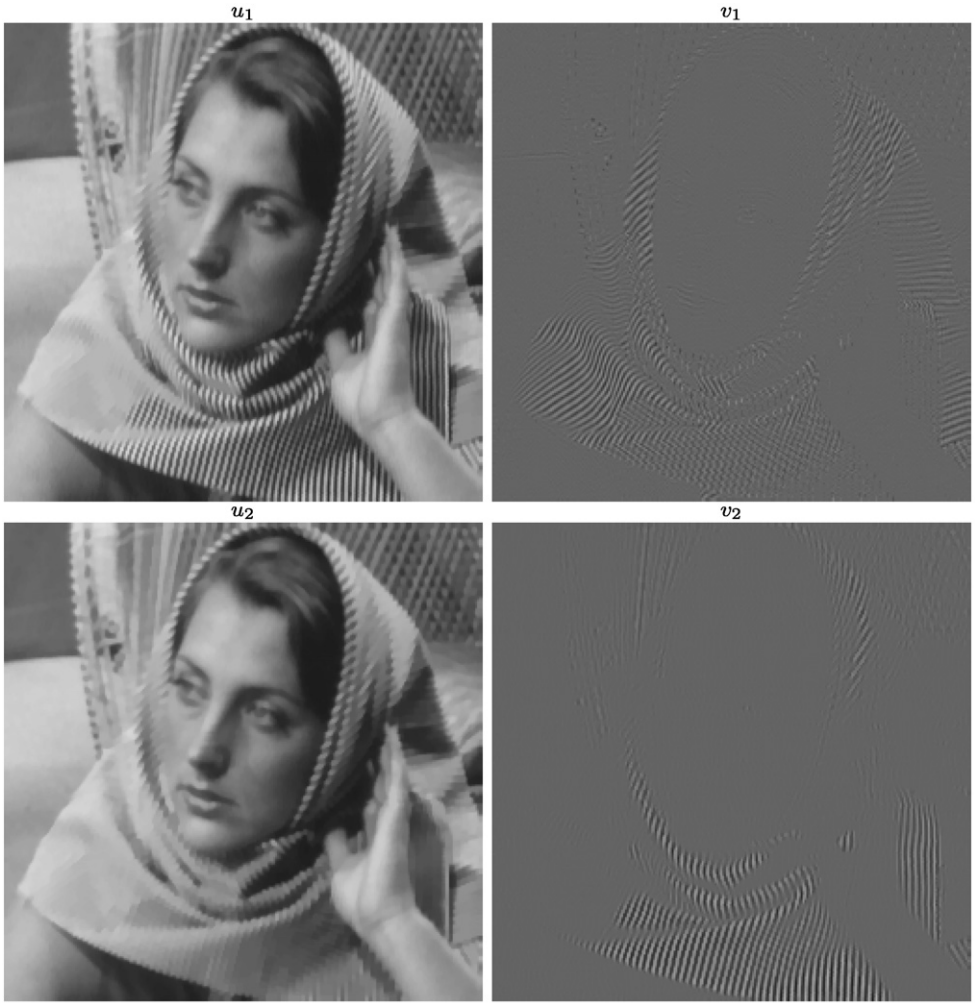

Fig. 25. (u_1, v_1) is a decomposition of $f_1 = u_1 + v_1$ from Fig. 2 using the model (59) with the kernel being the characteristic function of a vertical line centered at 0 having 3-pixel length. u_1 is then decomposed into $u_2 + v_2$ using the same model with the kernel being the characteristic function of a horizontal line centered at 0 having 3-pixel length. The two decompositions have $\varphi(\nabla u) = |\nabla u|$, $p = 1$, and $\lambda = 2$.

Also, Sobolev spaces are particular cases when the kernels are of the form $K_\alpha(x) = ((1 + |\xi|^2)^{(\alpha/2)})^\vee(x)$, for some α, and $p = 2$. These are considered by Lieu and Vese for image denoising and decomposition in [38].

From the image analysis point of view, where images are defined discretely, we consider here other non-smooth kernels as well. An interesting one is the characteristic function of a set, which can be a disk, or a square, etc. To separate oscillations of different orientations, we will also consider anisotropic kernels. We refer the readers to H. Triebel [58] for a discussion of other anisotropic kernels.

Let B be a set containing 0 and $K_B(x) = \frac{1}{|B|}\chi_B(x)$ be the averaging kernel. We have

$$K_B * f(x) = \int_\Omega K_B(x - y) f(y)\, dy = \frac{1}{|B|} \int_B f(x - y)\, dy.$$

Now, (56) can be written as

$$\inf_u \left\{ \mathcal{E}(u) = \int_\Omega \varphi(\nabla u) + \lambda \left\| K_B * (f - u) \right\|_{L^p} \right\}. \tag{59}$$

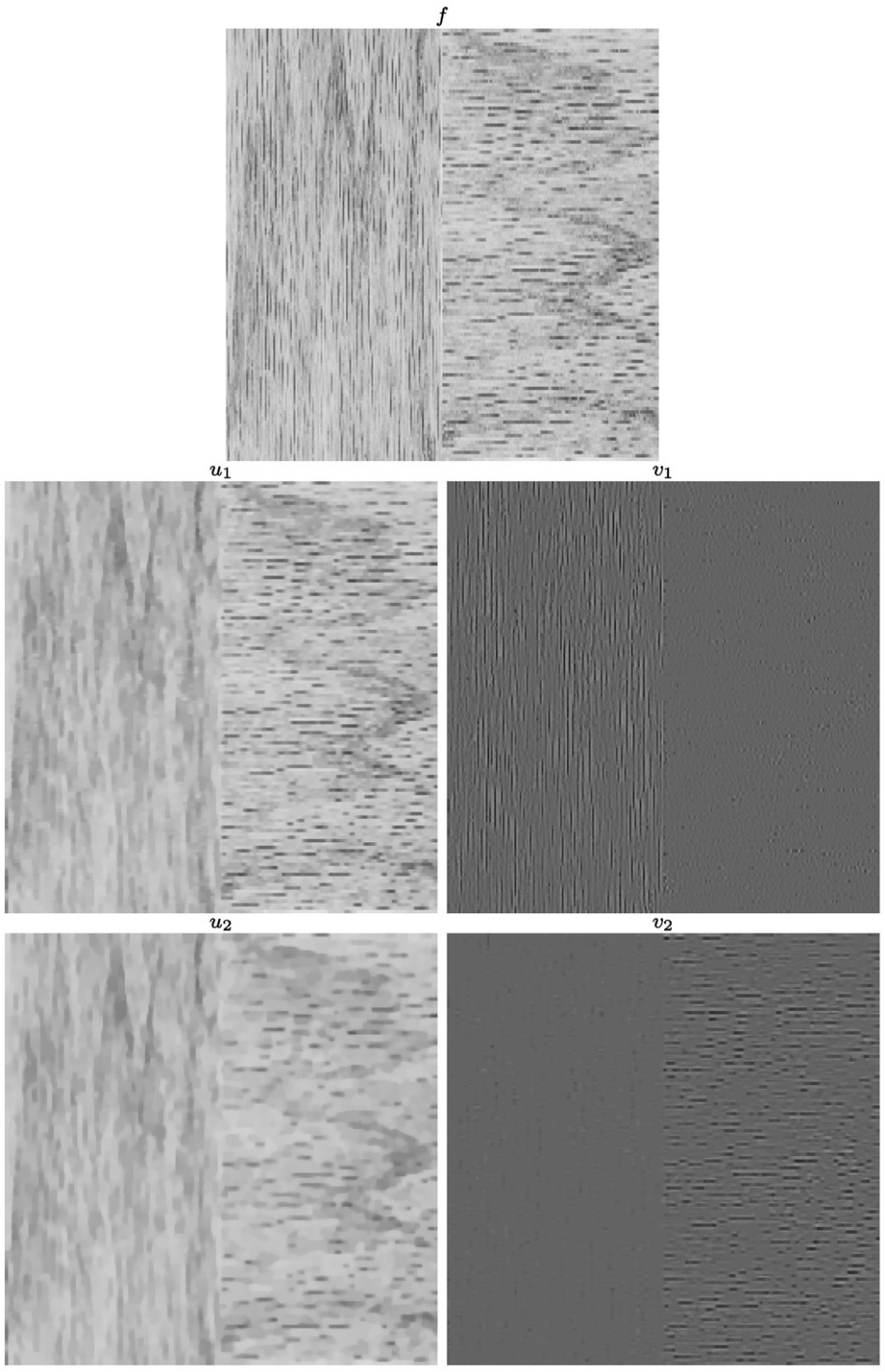

Fig. 26. (u_1, v_1) is a decomposition of $f = u_1 + v_1$ using the model (59) with the kernel being the characteristic function of a horizontal line centered at 0 having 3-pixel length. u_1 is then decomposed into $u_2 + v_2$ using the same model with the kernel being the characteristic function of a vertical line centered at 0 having 3-pixel length. The two decompositions have $\varphi(\nabla u) = |\nabla u|$, $p = 1$, and $\lambda = 2$.

Next, we would like to show some numerical results using the averaging kernels.

Figures 21–22 show decompositions of f_1 from Fig. 2 using the model (59) with a non-smooth averaging kernel K_B, where B is a square centered at 0 with sides parallel to the axis. Both decompositions use $\varphi(\nabla u) = |\nabla u|$, $p = 1$, and $\lambda = 1.5$. However, the decomposition in Fig. 21 uses the square with 3-pixel length for the sides, while the other uses the square with 5 pixel length.

Figure 23 shows a decomposition of the noisy image f_2 from Fig. 2 using the model (59) with B equal to the square of 3 pixel-length centered at 0, $\varphi(\nabla u) = |\nabla u|$, $p = 1$, and $\lambda = 2$.

Figure 24 shows a decomposition of f_2 from Fig. 2 using the model (59) with B equal to the disk centered at 0 and the radius equal to 1-pixel length, $\varphi(\nabla u) = |\nabla u|$, $p = 1$, and $\lambda = 2$.

Figures 25–26 show decompositions using anisotropic kernels, which are suitable for the separation of different orientations of oscillations. Here, we are interested in separating texture that has a vertical orientation from other oscillations in f_1 from Fig. 2 and in f from Fig. 26. We have obtained this result in two steps. In Fig. 25, we first decompose f_1 into $u_1 + v_1$ using the kernel of a characteristic function of a vertical line centered at 0 having 3-pixel length (related with hierarchical multiscale decompositions introduced in [56]). This kernel picks up oscillations in the horizontal direction. We then decompose u_1 into $u_2 + v_2$ using the kernel of a characteristic function of a horizontal line centered at 0 having 3-pixel length. This kernel picks up oscillations in the vertical direction. A similar process is done in Fig. 26, but the kernels are used in reverse order. This process can be done simultaneously in one step via this minimization problem

$$\inf_{u,v}\left\{\int_{\Omega} \varphi(\nabla u) + \lambda_1 \|K_1 * v\|_{L^p} + \lambda_2 \|K_2 * w\|_{L^p},\ f = u + v + w\right\}, \tag{60}$$

where K_1 is the vertical kernel, and K_2 is the horizontal kernel. From Fig. 25, we would have $u = u_2$, $v = v_1$, and $w = v_2$.

Remark 8. Here, we use horizontal and vertical directions. But other directions can also be used to separate different texture orientation, and the directions do not have to be orthogonal.

Acknowledgment

The authors would like to thank Jean-Michel Morel for many fruitful discussions that he had with the authors while visiting UCLA. The authors would also like to thank Mark Green for his valuable discussions and remarks on the statistics of natural images, and Bjorn Engquist for his valuable discussions. Also, the authors would like to thank Michael Frazier for a helpful conversation. Finally, the authors would like to thank the anonymous referees for their useful comments and suggestions, that helped very much to improve the manuscript. The last author would also like to thank David Mumford for suggesting the manuscript by A. Carasso and the mondrian image.

References

[1] R. Acar, C.R. Vogel, Analysis of bounded variation penalty methods for ill-posed problems, Inverse Problems 10 (6) (1994) 1217–1229.
[2] L. Alvarez, Y. Gousseau, J.-M. Morel, Scales in natural images and a consequence on their bounded variation norm, in: Lecture Notes in Comput. Sci., vol. 1682, Springer-Verlag, 1999, pp. 247–258.
[3] S. Alliney, Digital filters as L1-norm regularizers, in: Sixth Multidimensional Signal Processing Workshop, 6–8 Sept. 1989, p. 105.
[4] S. Alliney, Digital filters as absolute norm regularizers, IEEE Trans. on Signal Process. 40 (6) (1992) 1548–1562.
[5] L. Ambrosio, Variational problems in *SBV* and image segmentation, Acta Appl. Math. 17 (1989) 1–40.
[6] L. Ambrosio, N. Fusco, D. Pallara, Functions of Bounded Variation and Free Discontinuity Problems, Oxford University Press, 2000.
[7] F. Andreu-Vaillo, V. Caselles, J.M. Mazon, Parabolic Quasilinear Equations Minimizing Linear Growth Functionals, Progress in Mathematics, vol. 223, Birkhäuser, 2004.
[8] G. Aubert, J.-F. Aujol, Modeling very oscillating signals. Application to image processing, Appl. Math. Optim. 51 (2) (2005) 163–182.
[9] G. Aubert, L. Vese, A Variational method in Image Recovery, SIAM J. Numer. Anal. 34 (5) (1997) 1948–1979.
[10] J.-F. Aujol, Contribution à l'analyse de textures en traitement d'images par méthodes variationnelles et équations aux dérivées partielles, Thèse de Doctorat, University of Nice Sophia Antipolis, France, June, 2004.
[11] J.-F. Aujol, G. Aubert, L. Blanc-Féraud, A. Chambolle, Image decomposition application to SAR images, in: Lecture Notes in Comput. Sci., vol. 2695, Springer-Verlag, 2003, pp. 297–312.
[12] J.-F. Aujol, G. Aubert, L. Blanc-Féraud, A. Chambolle, Image decomposition into a bounded variation component and an oscillating component, J. Math. Imaging Vision 22 (1) (2005) 71–88.

[13] J.-F. Aujol, A. Chambolle, Dual norms and image decomposition models, Internat. J. Comput. Vis. 63 (2005) 85–104.
[14] G. Bouchitté, G. Buttazzo, New lower semi-continuity results for nonconvex functionals defined on measures, Nonlinear Anal. TMA 15 (7) (1990) 679–692.
[15] E. Candes, F. Guo, New multiscale transforms, minimum total variation synthesis: applications to edge-preserving image reconstruction, Signal Process. 82 (11) (2002) 1519–1543.
[16] A.S. Carasso, Singular integrals, image smoothness, and the recovery of texture in image deblurring, SIAM J. Appl. Math. 64 (5) (2004) 1749–1774.
[17] A. Chambolle, R.A. De Vore, N.Y. Lee, B.J. Lucier, Nonlinear wavelet image processing: variational problems, compression, and noise removal through wavelet shrinkage, IEEE Trans. Image Process. 7 (3) (1998) 319–335.
[18] A. Chambolle, P.-L. Lions, Image recovery via total variation minimization and related problems, Numer. Math. 76 (2) (1997) 167–188.
[19] A. Chambolle, B. Lucier, Interpreting translation-invariant wavelet shrinkage as a new image smoothing scale space, IEEE Trans. Image Process. 10 (7) (2001) 993–1000.
[20] T.F. Chan, S. Esedoglu, Aspects of total variation regularized L^1 function approximation, SIAM J. Appl. Math. 65 (5) (2005) 1817–1837.
[21] E. Cheon, A. Paranjpye, L. Vese, S. Osher, Noise removal by total variation minimization, Math 199 REU project, UCLA Department of Mathematics, Spring–Summer 2002.
[22] I. Daubechies, G. Teschke, Wavelet based image decomposition by variational functionals, in: F. Truchetet (Ed.), Wavelets Applications in Industrial Processing, in: Proc. SPIE, vol. 5226, SPIE, Bellingham, WA, 2004, pp. 94–105.
[23] I. Daubechies, G. Teschke, Variational image restoration by means of wavelets: Simultaneous decomposition, deblurring, and denoising, Appl. Comput. Harmon. Anal. 19 (1) (2005) 1–16.
[24] F. Demengel, R. Temam, Convex functions of a measure and applications, Indiana Univ. Math. J. 33 (1984) 673–709.
[25] R.A. DeVore, B.J. Lucier, Fast wavelet techniques for near-optimal image processing, in: IEEE Military Communications Conference Record, San Diego, October 11–14, 1992, IEEE, Piscataway, NJ, 1992, pp. 1129–1135.
[26] D. Donoho, De-noising by soft-thresholding, IEEE Trans. Inform. Theory 41 (1995) 613–627.
[27] D. Donoho, Nonlinear solution of linear inverse problems by wavelet-vaguelette decomposition, Appl. Comput. Harmon. Anal. 2 (1995) 101–126.
[28] I. Ekeland, R. Témam, Convex Analysis and Variational Problems, Classics in Applied Mathematics, vol. 28, SIAM, Philadelphia, 1999.
[29] L.C. Evans, R. F Gariepy, Measure Theory and Fine Properties of Functions, CRC Press, 1991.
[30] G.B. Folland, Modern Analysis: Modern Techniques and Their Applications, second ed., John Willey & Sons, Inc., New York, 1999.
[31] J. Gilles, Décomposition et détection de structures géométriques en imagerie, Thèse de Doctorat, CMLA E.N.S. Cachan, France, June 2006.
[32] Y. Gousseau, J.-M. Morel, Are natural images of bounded variation? SIAM J. Math. Anal. 33 (3) (2001) 634–648.
[33] M. Green, Statistics of images, the TV algorithm of Rudin–Osher–Fatemi for image denoising and an improved denoising algorithm, UCLA CAM Report 02-55, October 2002.
[34] A. Haddad, Méthodes variationnelles en traitement d'image, Thèse de Doctorat, CMLA E.N.S. Cachan, France, June 2005.
[35] T.M. Le, A study of a few image segmentation and decomposition models in a variational approach, Ph.D. Thesis, University of California, Los Angeles, June 2006.
[36] L. Lieu, Contribution to problems in image restoration, decomposition, and segmentation by variational methods and partial differential equations, Ph.D. Thesis, University of California, Los Angeles, June 2006.
[37] T.M. Le, L.A. Vese, Image decomposition using total variation and div(BMO), Multiscale Model. Simul. 4 (2) (2005) 390–423.
[38] L. Lieu, L. Vese, Image restoration and decomposition via bounded variation and negative Hilbert–Sobolev spaces, UCLA CAM Report 05-33, May 2005.
[39] S. Levine, An adaptive variational model for image decomposition, in: Energy Minimization Methods in Computer Vision and Pattern Recognition, in: Lecture Notes in Comput. Sci., vol. 3757, Springer-Verlag, 2005, pp. 382–397.
[40] S. Lintner, F. Malgouyres, Solving a variational image restoration model which involves L^∞ constraints, Inverse Problems 20 (3) (2004) 815–831.
[41] F. Malgouyres, Mathematical analysis of a model which combines total variation and wavelet for image restoration, J. Information Processes 2 (1) (2002) 1–10.
[42] Y. Meyer, Oscillating Patterns in Image Processing and Nonlinear Evolution Equations, University Lecture Series, vol. 22, Amer. Math. Soc., Providence, RI, 2001.
[43] J.-M. Morel, S. Solimini, Variational Methods in Image Segmentation: With Seven Image Processing Experiments, Progress in Nonlinear Differential Equations and Their Applications, Birkhäuser, Boston, 1994.
[44] D. Mumford, B. Gidas, Stochastic models for generic images, Quart. Appl. Math. 59 (1) (2001) 85–111.
[45] D. Mumford, J. Shah, Optimal approximations by piecewise smooth functions and associated variational problems, Comm. Pure Appl. Math. 42 (5) (1989) 577–685.
[46] M. Nikolova, A variational approach to remove outliers and impulse noise, J. Math. Imaging Vision 20 (1–2) (2004) 99–120.
[47] S. Osher, O. Scherzer, G-norm properties of bounded variation regularization, Commun. Math. Sci. 2 (2) (2004) 237–254.
[48] A. Obereder, S. Osher, O. Scherzer, On the use of dual norms in bounded variation type regularization, in: R. Klette, R. Kozera, L. Noakes, J. Weickert (Eds.), Geometric Properties of Incomplete Data, in: Computational Imaging and Vision, vol. 31, Springer, 2006.
[49] S. Osher, A. Solé, L. Vese, Image decomposition and restoration using total variation minimization and the H^{-1} norm, Multiscale Model. Simul. 1 (3) (2003) 349–370.
[50] L. Rudin, S. Osher, E. Fatemi, Nonlinear total variation based noise removal algorithms, Phys. D 60 (1992) 259–268.
[51] H.J. Schmeisser, H. Triebel, Topics in Fourier Analysis and Function Spaces, John Willey & Sons, 1987.

[52] J.-L. Starck, M. Elad, D.L. Donoho, Image decomposition: Separation of texture from piecewise smooth content, in: SPIE Conference on Signal and Image Processing: Wavelet Applications in Signal and Image Processing X, SPIE's 48th Annual Meeting, San Diego, 3–8 August 2003.
[53] E.M. Stein, Harmonic Analysis: Real Variable Methods, Orthogonality, and Oscillatory Integrals, Princeton University Press, Princeton, NJ, 1993.
[54] E.M. Stein, Singular Integrals and Differentiability Properties of Functions, Princeton University Press, Princeton, NJ, 1970.
[55] D. Strong, T. Chan, Edge-preserving and scale-dependent properties of total variation regularization, Inverse Problems 19 (2003) S165–S187.
[56] E. Tadmor, S. Nezzar, L. Vese, A multiscale image representation using hierarchical (BV, L^2) decompositions, Multiscale Model. Simul. 2 (4) (2004) 554–579.
[57] M.H. Taibleson, On the theory of Lipschitz spaces of distributions on Euclidean n-space I. Principal properties, J. Math. Mech. 13 (3) (1964) 407–479.
[58] H. Triebel, Characterizations of Besov–Hardy–Sobolev spaces via harmonic functions, temperatures and related means, J. Approx. Theory 35 (1982) 275–297.
[59] H. Triebel, Theory of Function Spaces II, Monographs in Mathematics, vol. 84, Birkhäuser, 1992.
[60] L. Vese, A study in the BV space of a denoising-deblurring variational problem, Appl. Math. Optim. 44 (2) (September–October 2001) 131–161.
[61] L. Vese, S. Osher, Modeling textures with total variation minimization and oscillating patterns in image processing, J. Sci. Comput. 19 (1–3) (2003) 553–572.
[62] L.A. Vese, S.J. Osher, Image denoising and decomposition with total variation minimization and oscillatory functions, Special Issue on Mathematics and Image Analysis, J. Math. Imaging Vision 20 (2004) 7–18.
[63] S.C. Zhu, D. Mumford, Prior learning and Gibbs reaction-diffusion, IEEE Trans. Pattern Anal. Machine Intelligence 19 (11) (1997) 1236–1250.