

Noisy image decomposition: a new structure, texture and noise model based on local adaptivity

Jérôme Gilles

Centre d'Expertise Parisien
16 bis, avenue Prieur de la Côte d'Or 94114 ARCUEIL Cedex FRANCE
jerome.gilles@etca.fr

The original publication is available at www.springerlink.com.

Abstract

These last few years, image decomposition algorithms have been proposed to split an image into two parts: the structures and the textures. These algorithms are not adapted to the case of noisy images because the textures are corrupted by noise. In this paper, we propose a new model which decomposes an image into three parts (structures, textures and noise) based on a local regularization scheme. We compare our results with the recent work of Aujol and Chambolle. We finish by giving another model which combines the advantages of the two previous ones.

Keywords: image decomposition, BV , texture, noise, oscillating functions, Besov spaces, local adaptivity.

1 Introduction

In 2001, in [3], Y.Meyer investigated the model of image restoration proposed by Rudin-Osher-Fatemi [5]:

$$F^{ROF}(u) = J(u) + (2\lambda)^{-1} \|f - u\|_{L^2}^2 \quad (1)$$

where f is the measured image which implies an estimation:

$$\hat{u} = \arg \inf_{u \in BV} F^{ROF}(u), \quad (2)$$

of the restored image. The main hypothesis is that u belongs to the space BV (the space of bounded variation functions). The quantity $J(u)$ is a semi-norm on BV . It is also knowing as the total variation of the function u and can be expressed by

$$J(u) = \int |\nabla u|. \quad (3)$$

If we write $v = f - u$ and take the point of view of image decomposition (*i.e* f is composed of structures (u) and textures (v)), we can rewrite equation (1) as

$$F^{ROF}(u) = J(u) + (2\lambda)^{-1} \|v\|_{L^2}^2. \quad (4)$$

Y.Meyer proved that this model rejects the oscillatory component of f which is considered to be the textured component. Then, he proposed a new model where the L^2 norm of v is replaced by a norm in a space G close to the dual space of BV and which contains oscillatory functions. The new model introduced by Y.Meyer is then given by

$$F^{YM}(u, v) = J(u) + (2\lambda)^{-1} \|v\|_G, \quad (5)$$

where the G -norm is defined by:

Definition 1 For $v = \partial_1 g_1 + \partial_2 g_2$ where $g_1 \in L^\infty(\mathbb{R}^2), g_2 \in L^\infty(\mathbb{R}^2)$,

$$\|v\|_G = \inf_g \left\| \left(|g_1|^2 + |g_2|^2 \right)^{\frac{1}{2}} \right\|_{L^\infty}. \quad (6)$$

Due to its non-linearity, the G -norm is difficult to compute numerically (see [3]). Then two approaches can be found in the litterature to run Meyer's algorithm.

The first one is the algorithm proposed by L.Vese and S.Osher in [2]. The authors use the following approximation:

$$\forall v \in L^\infty \quad \|v\|_{L^\infty} = \lim_{p \rightarrow +\infty} \|v\|_{L^p}, \quad (7)$$

and the property that any function v in G can be written as $v = \operatorname{div}(g)$ where $g = (g_1, g_2) \in (L^\infty \times L^\infty)$. Then they give a new formulation of the problem:

$$F^{VO}(u, g) = J(u) + (2\lambda)^{-1} \|f - (u + \operatorname{div} g)\|_{L^2}^2 + \mu \left\| \sqrt{g_1^2 + g_2^2} \right\|_{L^p}. \quad (8)$$

The authors seek for

$$(\hat{u}, \hat{g}_1, \hat{g}_2) = \arg \inf_{(u, g_1, g_2) \in (BV \times L^\infty \times L^\infty)} F^{VO}(u, g_1, g_2) \quad (9)$$

The authors deduce the related partial differential equation system and its discrete formulation (see [2] for details). The drawback of their approach is the numerical stability of the algorithm.

The second approach is proposed by Aujol et al. in [1, 11]. Their model is given by

$$F^{AU}(u, v) = J(u) + J^* \left(\frac{v}{\mu} \right) + (2\lambda)^{-1} \|f - u - v\|_{L^2}^2, \quad (10)$$

$$(\hat{u}, \hat{v}) = \arg \inf_{(u, v) \in (BV \times G)} F^{AU}(u, v) \quad (11)$$

where

$$G_\mu = \{v \in G / \|v\|_G \leq \mu\}, \quad (12)$$

and J^* is the adjoint of J ; therefore it is the indicator function defined by

$$J^* \left(\frac{v}{\mu} \right) = \begin{cases} 0 & \text{if } v \in G_\mu, \\ +\infty & \text{if } v \in G \setminus G_\mu. \end{cases} \quad (13)$$



Figure 1: Two part decomposition of noisy image, from left to right: noisy image, object part, texture part corrupted by noise.

The authors prove that the solution of minimizing (10) can be found by an iterative algorithm based on non-linear projectors P_{G_μ} proposed by A.Chambolle (see [4] for details and for a convergence theorem). This algorithm gives good results but is of limited interest in the case of noisy images. Indeed, the noise can be viewed as a very highly oscillatory function (this means that noise is in G_μ). Therefore the algorithm incorporates the noise in the texture component. Then the textures are corrupted by noise (see fig.1 for an example).

In this paper, we propose to extend the two components model to a three components model, $f = u + v + w$, which discriminates between structures (u), textures (v) and noise (w). This work was initiated in [13].

An outline of the paper is given now. In section 2, we describe our new algorithm based on a local regularization principle [7, 10]. We give some numerical aspects of the algorithm and show some results. In section 3 we recall the recent model proposed by Aujol and Chambolle [6] which also decomposes an image into three parts. A comparison between our approach and their model shows the advantages and disadvantages of each algorithm. In section 4, we propose another algorithm which incorporates the advantages of each the previous ones. We conclude and give some perspectives of this work in section 5.

2 Three part image decomposition: a local adaptive algorithm

In this section, we propose a new model to decompose an image into three parts: structures (u), textures (v) and noise (w). As in the $u + v$ model, we consider that structures and textures are modeled by functions in BV and G spaces respectively. We also consider a zero mean gaussian noise added to the image. Let us view noise as a specific very oscillating function. In virtue to the work in [3], where it is shown that the more a function is oscillatory, the smaller its G -norm is, we propose to modelize w as a function in G and consider that its G -norm is much smaller than the norm of the textures ($\|v\|_G \gg \|w\|_G$). These assumptions are equivalent to choose

$$v \in G_{\mu_1}, w \in G_{\mu_2} \quad \text{where} \quad \mu_1 \gg \mu_2 \quad (14)$$

In order to increase the performance, we propose to add a local adaptability behaviour to the algorithm following an idea proposed by G.Gilboa et al. in [7]. The authors investigate the ROF model given by equation (1) and propose a modified version which is able to preserve textures in the denoising process. To do this, they don't choose λ as a constant on the whole image but as a function $\lambda(f)(x, y)$ which represents local properties of the image. In a cartoon-type region, the algorithm enhances the denoising process by increasing the value of λ ; in a texture-type region the algorithm decreases λ to attenuate the regularization in order to preserve the details of textures. So $\lambda(f)(x, y)$ can be viewed as a smoothed partition between textured and untextured regions. Then, in order to decompose an image into three parts, we propose to use the following functional:

$$F_{\lambda, \mu_1, \mu_2}^{JG}(u, v, w) = J(u) + J^* \left(\frac{v}{\mu_1} \right) + J^* \left(\frac{w}{\mu_2} \right) + (2\lambda)^{-1} \|f - u - \nu_1 v - \nu_2 w\|_{L^2}^2, \quad (15)$$

where the functions ν_i represent the smoothed partition of textured and untextured regions (and play the role of λ in Gilboa's paper). The ν_i functions must have the following behaviour:

- if we are in a textured region, we want to favour v instead of w . This is equivalent to ν_1 close to 1 and ν_2 close to 0,
- if we are in an untextured region, we want to favour w instead of v . This is equivalent to ν_1 closed to 0 and ν_2 close to 1.

We see that ν_1 and ν_2 are complementary so it is natural to choose $\nu_2 = 1 - \nu_1 : \mathbb{R}^2 \rightarrow]0; 1[$. The choice of ν_1 and ν_2 is discussed after the following proposition which characterizes the minimizers of $F_{\lambda, \mu_1, \mu_2}^{JG}(u, v, w)$.

Proposition 2.1 *Let $u \in BV$, $v \in G_{\mu_1}$, $w \in G_{\mu_2}$ be the structures, textures and noise parts respectively and f the original noisy image. Let the functions $(\nu_1(f)(\cdot, \cdot), \nu_2(f)(\cdot, \cdot))$ be defined on $\mathbb{R}^2 \rightarrow]0; 1[$, and assume that these functions could be considered as locally constant compared to the variation of v and w (see the proof for details). Then a minimizer defined by*

$$(\hat{u}, \hat{v}, \hat{w}) = \arg_{(u, v, w) \in BV \times G_{\mu_1} \times G_{\mu_2}} \min F_{\lambda, \mu_1, \mu_2}^{JG}(u, v, w), \quad (16)$$

is given by

$$\hat{u} = f - \nu_1 \hat{v} - \nu_2 \hat{w} - P_{G_\lambda}(f - \nu_1 \hat{v} - \nu_2 \hat{w}), \quad (17)$$

$$\hat{v} = P_{G_{\mu_1}} \left(\frac{f - \hat{u} - \nu_2 \hat{w}}{\nu_1} \right), \quad (18)$$

$$\hat{w} = P_{G_{\mu_2}} \left(\frac{f - \hat{u} - \nu_1 \hat{v}}{\nu_2} \right), \quad (19)$$

where P_{G_μ} is the Chambolle's non-linear projectors (see [4]).

The proof of this proposition is given in appendix A. Notice that we have still no results on the uniqueness of the minimizer.

In practice, we propose the following iterative numerical scheme:

1. initialization: $u_0 = v_0 = w_0 = 0$,
2. compute ν_1 and $\nu_2 = 1 - \nu_1$ from f (see below for the choice of ν_i),
3. compute $w_{n+1} = P_{G_{\mu_2}} \left(\frac{f - u_n - \nu_1 v_n}{\nu_2 + \kappa} \right)$, (κ is a small value in order to prevent the division by zero),
4. compute $v_{n+1} = P_{G_{\mu_1}} \left(\frac{f - u_n - \nu_2 w_{n+1}}{\nu_1 + \kappa} \right)$,
5. compute $u_{n+1} = f - \nu_1 v_{n+1} - \nu_2 w_{n+1} - P_{G_\lambda}(f - \nu_1 v_{n+1} - \nu_2 w_{n+1})$,
6. if $\max\{|u_{n+1} - u_n|, |v_{n+1} - v_n|, |w_{n+1} - w_n|\} \leq \epsilon$ or if we did N_{step} iterations then stop the algorithm, else jump to step 3.

Now, let's deal with the choice of the ν_i functions. In [7], the authors choose to compute a local variance on the texture + noise part of the image obtained by the ROF model ($f - u$). In this paper, we use the same strategy but on the v component obtained by the two part decomposition algorithm. This choice is implied by the fact that the additive gaussian noise can be considered orthogonal to the textures. As a consequence, the variance of a textured region is larger than the variance of an untextured region.

So, in practice, we first compute the two part decomposition of the image f . On the v component, $\forall(i, j)$, we compute the local variance on a small window (odd size L) centered on (i, j) . Then $\nu_1(i, j)$ can be expressed by

$$\nu_1(i, j) = \frac{1}{L^2} \sum_{(p,q) \in Z_{ij}} v^2(p, q) - \frac{1}{L^4} \left(\sum_{(p,q) \in Z_{ij}} v(p, q) \right)^2, \quad (20)$$

$$\nu_2(i, j) = 1 - \nu_1(i, j), \quad (21)$$

where $Z_{ij} = [i - (L - 1)/2; i + (L - 1)/2] \times [j - (L - 1)/2; j + (L - 1)/2]$.

At least, we normalize it in order to get the values in $]0; 1[$. Figure 2 shows an example on a noisy version of Barbara ($\sigma = 20$). We see that this image, which represents ν_1 , has larger values in the textured regions. Denote that in [8], the same authors propose to estimate the optimal parameter based on a SNR criteria. At this time, we did not incorporate this method in our algorithm.

We experimented our algorithm on three different images: on a synthetic one (show on the top-left of figure 3), on Barbara (show on the left of figure 2) and on a real outdoor image (show on the top-left of figure 5). The parameters are set to

- Synthetic: $\lambda = 50$, $\mu_1 = 1000$, $\mu_2 = 10$, the size of the window for computing ν is 7 pixels,



Figure 2: On left we have the noisy image (Barbara+gaussian noise, $\sigma = 20$); on right, texture partition ν_1 .

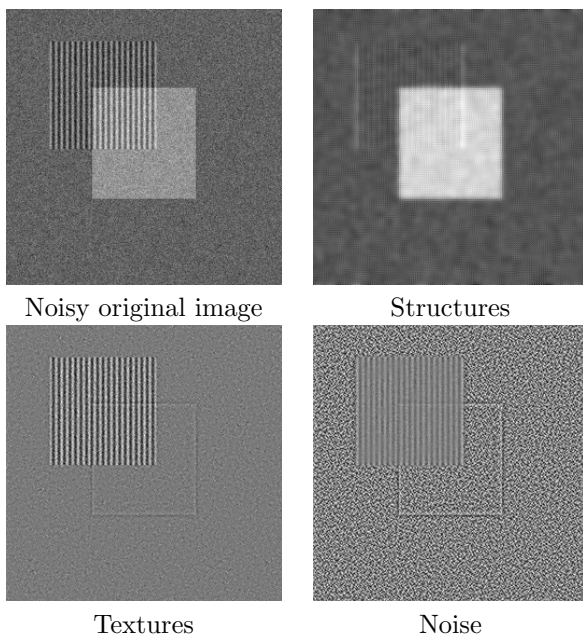


Figure 3: Results given by $F_{\lambda, \mu_1, \mu_2}^{JG}$ applied on the synthetic image.

- Barbara: $\lambda = 10$, $\mu_1 = 1000$, $\mu_2 = 1$, the size of the window for computing ν is 15 pixels,
- Outdoor: $\lambda = 50$, $\mu_1 = 1000$, $\mu_2 = 10$, the size of the window for computing ν is 7 pixels.

The results are given respectively in figures 3, 4 and 5. We can see in the whole cases, that the proposed algorithm successfully separates the noise from textures. In the zoomed images of figure 4, we can see that some residual noise stayed in the textured part and some residual texture is extracted in the noise part. This is inevitable since the ν_i functions made a tradeoff between extracting the information in the texture part or in the noise part.

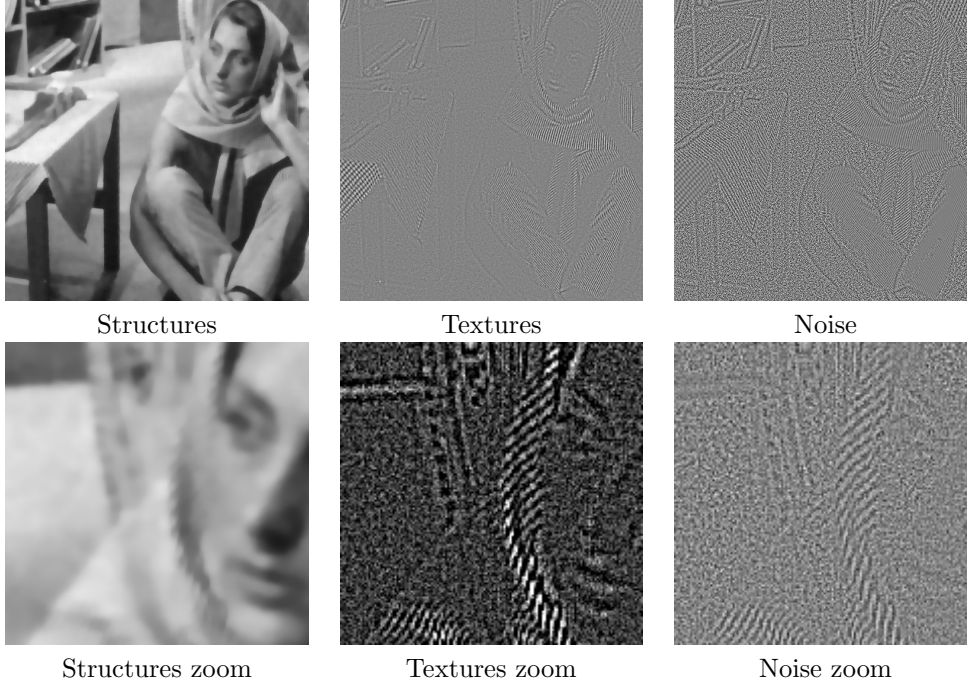


Figure 4: Results given by $F_{\lambda, \mu_1, \mu_2}^{JG}$ applied on Barbara. First row gives the whole images and second row zooms on a section of the images.

3 Aujol-Chambolle's algorithm

In [6], the authors propose another model to decompose an image into three parts. Their assumption is to consider noise as a distribution modeled by the Besov space $\dot{B}_{-1, \infty}^{\infty}$, and the following functional is used:

$$F_{\lambda, \mu, \delta}^{AC}(u, v, w) = J(u) + J^* \left(\frac{v}{\mu} \right) + B^* \left(\frac{w}{\delta} \right) + (2\lambda)^{-1} \|f - u - v - w\|_{L^2}^2, \quad (22)$$

where $u \in BV, v \in G_{\mu}, w \in E_{\delta}$ are defined by (we use the standard setting $s = -1, p = q = +\infty$ for the Besov space $B_{s, q}^p$):

$$E_{\delta} = \left\{ w \in \dot{B}_{-1, \infty}^{\infty} / \|w\|_{\dot{B}_{-1, \infty}^{\infty}} \leq \delta \right\}. \quad (23)$$

The term $B^*(.)$ is the indicator function on E_{δ} defined in the same way as equation (13). Then the minimizer is given by (see [6]):

$$\hat{u} = f - \hat{v} - \hat{w} - P_{G_{\lambda}}(f - \hat{v} - \hat{w}), \quad (24)$$

$$\hat{v} = P_{G_{\mu}}(f - \hat{u} - \hat{w}), \quad (25)$$

$$\hat{w} = P_{E_{\delta}}(f - \hat{u} - \hat{v}) = f - \hat{u} - \hat{v} - WST(f - \hat{u} - \hat{v}, 2\delta), \quad (26)$$

where $WST(f - \hat{u} - \hat{v}, 2\delta)$ is the Wavelet Soft Thresholding operator applied on $f - \hat{u} - \hat{v}$ with a threshold set to 2δ ($\delta = \eta\sigma\sqrt{\log N}$ and N is the number of pixels, η is a tuning parameter, see [12, 9]). A numerical iterative algorithm

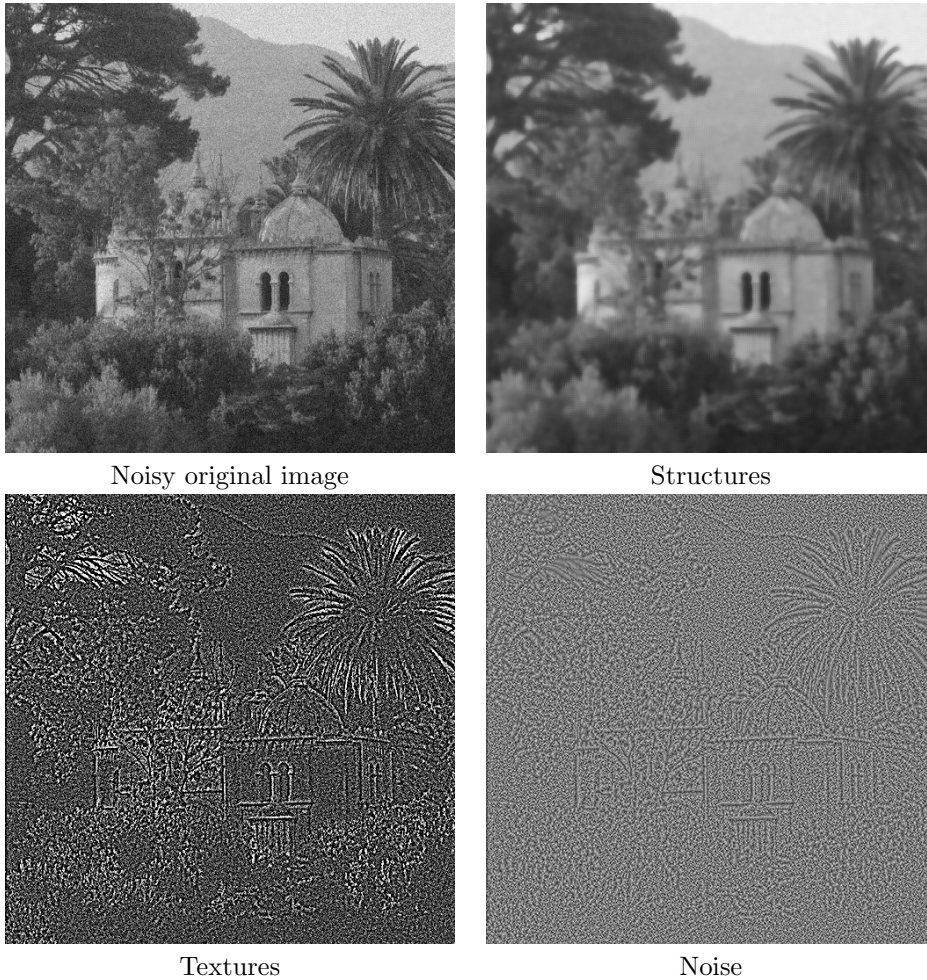


Figure 5: Results given by $F_{\lambda, \mu_1, \mu_2}^{JG}$ applied on the real outdoor image.

can be deduced from these formula (see [6]).

Figures 6, 7, 8 give the results obtained with this model on the same three image test used in the previous section. The parameters are set to

- Synthetic: $\mu = 1000$, $\lambda = 20$, $\sigma = 20$, $\eta = 0.2$,
- Barbara: $\mu = 100$, $\lambda = 1$, $\sigma = 20$, $\eta = 0.1$,
- Outdoor: $\mu = 5000$, $\lambda = 20$, $\sigma = 20$, $\eta = 0.1$.

We can see that the textures are better denoised by this model. This is a consequence of a better noise modeling by distributions in the Besov space. But the residual texture is more important than the one given by our algorithm in the noise part. This behavior is certainly implied by the local adaptivity capabilities of our algorithm. In the next section, we propose to merge the use of Besov space and the local adaptivity in a unique algorithm.

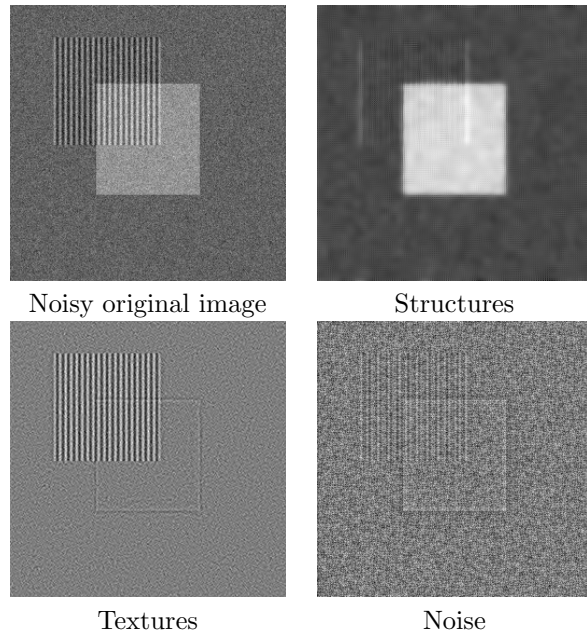


Figure 6: Results given by $F_{\lambda, \mu, \delta}^{AC}$ applied on the synthetic image.

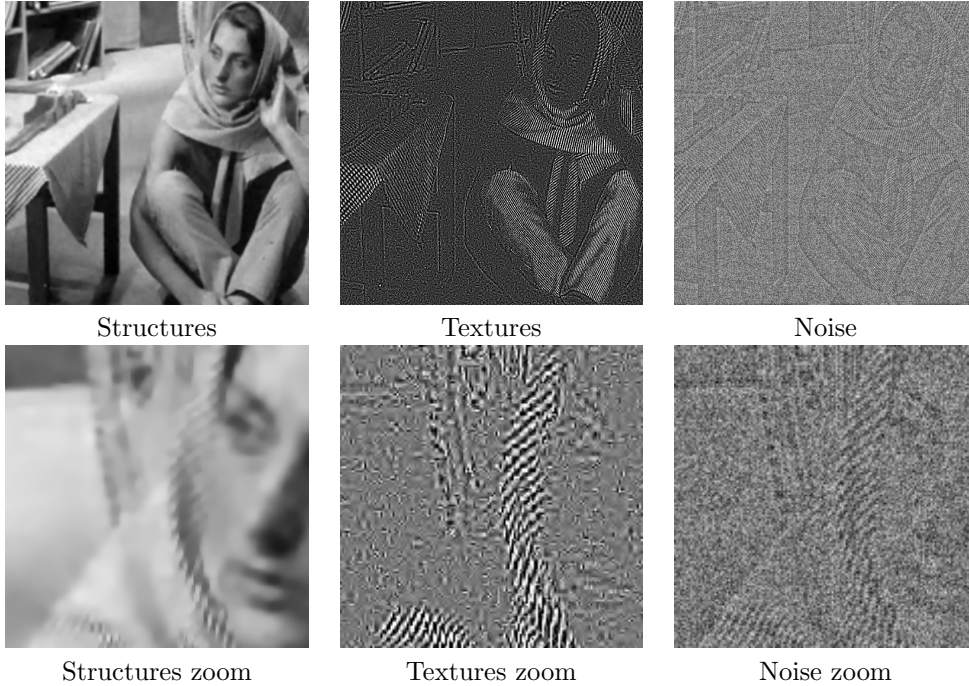


Figure 7: Results given by $F_{\lambda, \mu, \delta}^{AC}$ applied on Barbara.

4 Adaptive wavelet algorithm

We saw in the previous sections that the Besov space $\dot{B}_{-1, \infty}^{\infty}$ is better adapted to modelize the noise than the set G_{μ} (where μ is very small). In the same

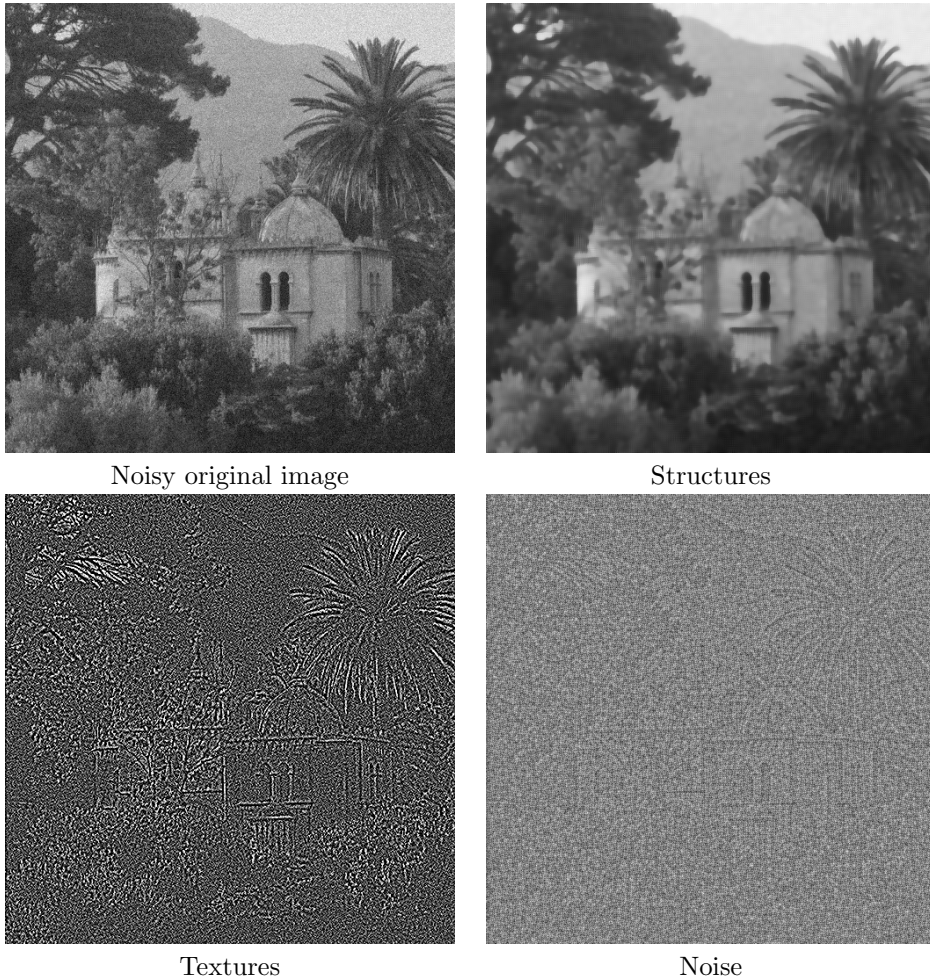


Figure 8: Results given by $F_{\lambda, \mu, \delta}^{AC}$ applied on the outdoor.

way, we saw that local adaptivity implies a decrease of the residual quantity of texture in the noise part. So the question is: how to merge these advantages in a unique algorithm? We propose the following model as an answer:

$$F_{\lambda, \mu, \delta}^{JG^2}(u, v, w) = J(u) + J^* \left(\frac{v}{\mu} \right) + B^* \left(\frac{w}{\delta} \right) + (2\lambda)^{-1} \|f - u - \nu_1 v - \nu_2 w\|_{L^2}^2, \quad (27)$$

where $u \in BV, v \in G, w \in E_\delta$ (as defined in the previous section) and the ν_i functions are the ones defined in section 2. A minimizer is given by the following proposition.

Proposition 4.1 *Let $u \in BV, v \in G_\mu, w \in E_\delta$ be the structures, textures and noise parts respectively and $(\nu_1(f)(x, y), \nu_2(f)(x, y))$ be defined as in proposition 2.1. Then a minimizer of:*

$$(\hat{u}, \hat{v}, \hat{w}) \in \underset{(u, v, w) \in BV \times G_\mu \times E_\delta}{\arg \min} F_{\lambda, \mu, \delta}^{JG^2}(u, v, w) \quad (28)$$



Figure 9: Local smoothed partition ν and its pyramidal version $\tilde{\nu}$.

is given by:

$$\begin{aligned}\hat{u} &= f - \nu_1 \hat{v} - \nu_2 \hat{w} - P_{G_\lambda}(f - \nu_1 \hat{v} - \nu_2 \hat{w}), \\ \hat{v} &= P_{G_\mu} \left(\frac{f - \hat{u} - \nu_2 \hat{w}}{\nu_1 + \kappa} \right), \\ \hat{w} &= \frac{f - \hat{u} - \nu_1 \hat{v}}{\nu_2 + \kappa} - \frac{\lambda}{\delta \nu_2^2 + \kappa} \widehat{WST} \left(\frac{\delta \nu_2}{\lambda} (f - \hat{u} - \nu_1 \hat{v}); \frac{2\delta^2 \tilde{\nu}_2^2}{\lambda} \right),\end{aligned}$$

where P_{G_λ} is the Chambolle's nonlinear projector defined in [4] and $\widehat{WST}(f, \delta)$ is a modified Wavelet Soft Thresholding of f . The threshold is dependent on the location in each wavelet coefficient, $\tilde{\nu}_2$ is a pyramidal version of ν_2 (see fig 9).

A proof of this proposition is given in appendix B.

This “merged-algorithm” gives the results showed in figure 10, 11. and 12. The parameters are set to

- Synthetic: $\mu = 1000$, $\lambda = 50$, $\sigma = 20$, $\eta = 0.7$, window's size for $\nu_i = 7$,
- Barbara: $\mu = 1000$, $\lambda = 30$, $\sigma = 20$, $\eta = 0.5$, window's size for $\nu_i = 7$,
- Outdoor: $\mu = 1000$, $\lambda = 30$, $\sigma = 20$, $\eta = 0.3$, window's size for $\nu_i = 5$.

As we expected, the combination of the two properties (Besov space and local adaptivity) increases the quality of the decomposition. Textures are better denoised and the noise part contains less residual texture. However, the noise is “damaged” by the local behaviour of the algorithm.

5 Conclusion

In this paper, we presented an original extension of the two components image decomposition algorithm of Aujol et al [1] in order to deal with the case of noisy images. This new algorithm uses the assumption that noise is a highly oscillatory function and uses a local adaptivity principle. We compared our results with ones obtained by the Aujol-Chambolle's algorithm and we then proposed a “merged-algorithm” which incorporates the advantages of the two

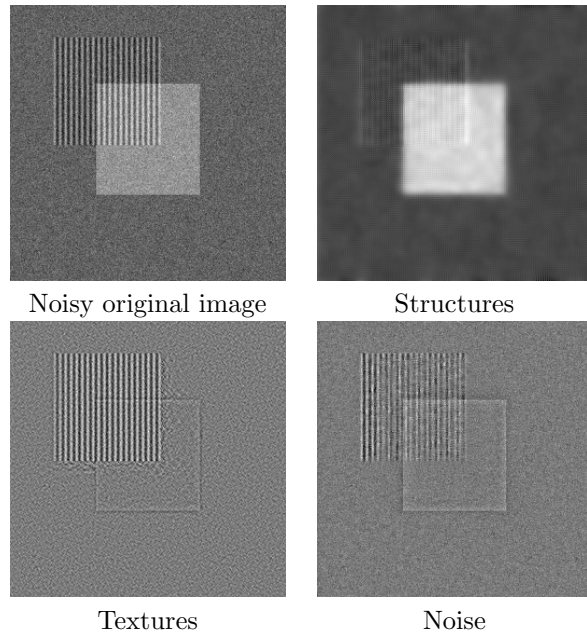


Figure 10: Results given by $F_{\lambda, \mu, \delta}^{JG2}(u, v, w)$ applied on the synthetic image.

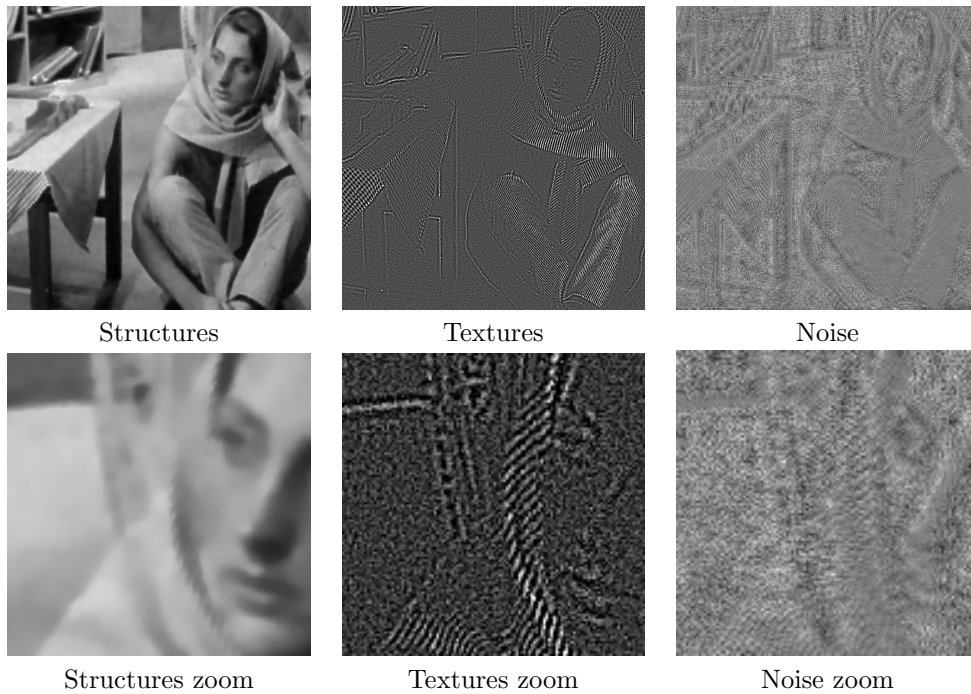


Figure 11: Results given by $F_{\lambda, \mu, \delta}^{JG2}(u, v, w)$ applied on Barbara.

preceding models. The results are quite promising. Futur work will focus on

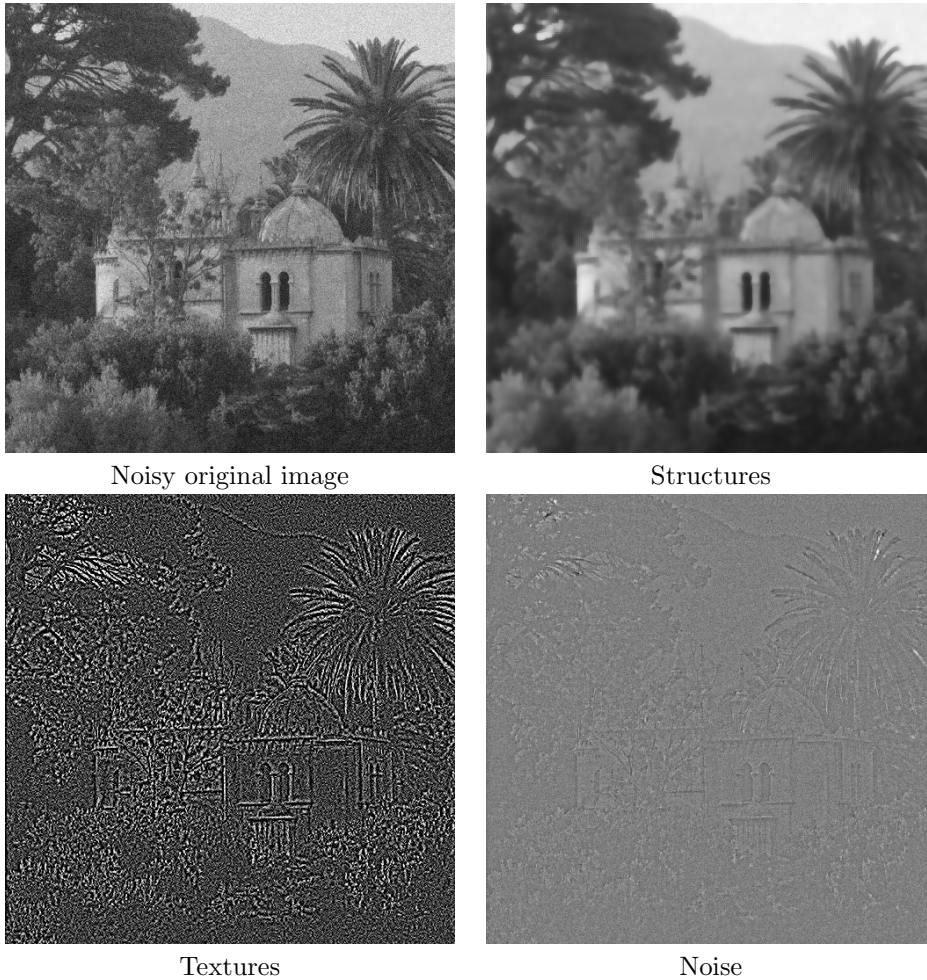


Figure 12: Results given by $F_{\lambda,\mu,\delta}^{JG^2}(u, v, w)$ applied on the outdoor image.

the choice of the different parameters. This choice is still not automated, and a manual choice of the good parameters is not an easy task. Another line of research is on the choice of the functions ν_i . Another modeling of the local behavior of the image may be obtained from a local approximation of oscillation quantity in term of the G -norm. Some other interesting works are done in [14, 18, 16, 17, 15] where different functional spaces are tested and may be an alternative way to the space G . It is another important issue to built objective comparison criteria, for comparing the different decomposition algorithms. We developed metrics adapted to each image component. These metrics will be describe in a futur paper.

Acknowledgements

The author would like to thank Prof. Yves Meyer for his salient remarks in the redaction of this paper. We also wants to thank the reviewers for their remarks

which resulted in a great improvement of the paper.

A Appendix A

The proof of proposition 2.1 is largely inspired from [4], but is detailed now for the reader convenience.

The Euler-Lagrange equations with respect to u are

$$-\frac{1}{\lambda}(f - u - \nu_1 v - \nu_2 w) + \partial J(u) \ni 0 \quad (29)$$

$$\Leftrightarrow u \in \partial J^* \left(\frac{1}{\lambda}(f - u - \nu_1 v - \nu_2 w) \right). \quad (30)$$

By adding $f - u - \nu_1 v - \nu_2 w$ on each side and multiplying by $\frac{1}{\lambda}$, it is deduced:

$$\frac{1}{\lambda}(f - \nu_1 v - \nu_2 w) \in \frac{1}{\lambda}(f - u - \nu_1 v - \nu_2 w) + \frac{1}{\lambda} \partial J^* \left(\frac{1}{\lambda}(f - u - \nu_1 v - \nu_2 w) \right). \quad (31)$$

By setting $\eta = (f - u - \nu_1 v - \nu_2 w)/\lambda$, it comes:

$$\frac{1}{\lambda}(f - \nu_1 v - \nu_2 w) \in \eta + \frac{1}{\lambda} \partial J^*(\eta). \quad (32)$$

Thus, using the results in [4], we get

$$\eta = P_G \left(\frac{1}{\lambda}(f - \nu_1 v - \nu_2 w) \right) = \frac{1}{\lambda} P_{G_\lambda}(f - \nu_1 v - \nu_2 w). \quad (33)$$

Using the expression of η we conclude that

$$\hat{u} = f - \nu_1 v - \nu_2 w - P_{G_\lambda}(f - \nu_1 v - \nu_2 w). \quad (34)$$

The u part is now completed. To compute the v and w parts, we use the following lemma.

Lemma A.1 *Let be $f \in L^2(\mathbb{R}^2)$, $v \in G_\mu$ and $\nu : \mathbb{R}^2 \rightarrow]0; 1[$. As ν is more regular than v , we assume ν as locally constant compared to v (this assumption is like a first-order assumption). Then the solution:*

$$\hat{v} = \arg \inf_{v \in G_\mu} \left\{ (2\lambda)^{-1} \|f - \nu v\|_{L^2}^2 + J^* \left(\frac{v}{\mu} \right) \right\} \quad (35)$$

can be approximated by:

$$\hat{v} = P_{G_\mu} \left(\frac{f}{\nu} \right) \quad (36)$$

By setting $\eta = \frac{v}{\mu}$, we get

$$F(\eta) = (2\lambda)^{-1} \|f - \mu \nu \eta\|_{L^2}^2 + J^*(\eta) \quad (37)$$

and now we seek for

$$\hat{\eta} = \arg \inf_{\eta} F(\eta). \quad (38)$$

First, let's show that it exists some $F_m(\eta)$ and $F_M(\eta)$ such that

$$F_m(\eta) \leq F(\eta) \leq F_M(\eta). \quad (39)$$

Let's begin by $F_M(\eta)$,

$$F(\eta) = (2\lambda)^{-1} \left\| \mu\nu \left(\frac{f}{\mu\nu} - \eta \right) \right\|_{L^2}^2 + J^*(\eta) \quad (40)$$

$$\leq (2\lambda)^{-1} \|\mu\nu\|_{L^2}^2 \left\| \frac{f}{\mu\nu} - \eta \right\|_{L^2}^2 + J^*(\eta) \quad (41)$$

then we choose

$$F_M(\eta) = (2\lambda)^{-1} \|\mu\nu\|_{L^2}^2 \left\| \frac{f}{\mu\nu} - \eta \right\|_{L^2}^2 + J^*(\eta). \quad (42)$$

Now let's deal with $F_m(\eta)$, if we choose

$$F_m(\eta) = (2\lambda)^{-1} \frac{\mu}{\|\nu\|_{L^2}^2} \left\| \frac{f}{\mu\nu} - \eta \right\|_{L^2}^2 + J^*(\eta). \quad (43)$$

We have

$$\left\| \frac{f}{\mu\nu} - \eta \right\|_{L^2}^2 \leq \|\nu\|_{L^2}^2 \left\| \frac{1}{\nu} \left(\frac{f}{\mu\nu} - \eta \right) \right\|_{L^2}^2 \quad (44)$$

then

$$\implies \frac{1}{\|\nu\|_{L^2}^2} \left\| \frac{f}{\mu\nu} - \eta \right\|_{L^2}^2 \leq \left\| \frac{1}{\nu} \left(\frac{f}{\mu\nu} - \eta \right) \right\|_{L^2}^2 \quad (45)$$

$$= \left\| \frac{1}{\mu\nu^2} (f - \mu\nu\eta) \right\|_{L^2}^2 \quad (46)$$

$$\leq \left\| \frac{1}{\mu\nu^2} \right\|_{L^2}^2 \|f - \mu\nu\eta\|_{L^2}^2. \quad (47)$$

We suppose that $\left\| \frac{1}{\nu^2} \right\|_{L^2}^2 \lesssim 1$ (we can verify it in all our experiments). Then

$$\frac{1}{\|\nu\|_{L^2}^2} \left\| \frac{f}{\mu\nu} - \eta \right\|_{L^2}^2 \leq \frac{1}{\mu} \|f - \mu\nu\eta\|_{L^2}^2. \quad (48)$$

This implies

$$F_m(\eta) \leq F(\eta) \quad (49)$$

Now we want to characterize $F(\eta)$ by studying the behaviour of $F_m(\eta)$ and $F_M(\eta)$. First, as the three energy are convex, we have

$$\inf F_m(\eta) \leq \inf F(\eta) \leq \inf F_M(\eta). \quad (50)$$

Let calculate $F_m^{min} = \inf F_m(\eta)$ and $F_M^{min} = \inf F_M(\eta)$. First, we have

$$\hat{\eta}_M = \arg \inf_{\eta} F_M(\eta). \quad (51)$$

$$\iff -(2\lambda)^{-1}\|\mu\nu\|_{L^2}^2 \left(\frac{f}{\mu\nu} - \eta \right) + \partial J^*(\eta) \ni 0 \quad (52)$$

$$\iff \eta + \frac{2\lambda}{\|\mu\nu\|_{L^2}^2} \partial J^*(\eta) \ni \frac{f}{\mu\nu} \quad (53)$$

Chambolle [4] shows that

$$\hat{\eta}_M = P_{G_1} \left(\frac{f}{\mu\nu} \right) = \frac{1}{\mu} P_{G_\mu} \left(\frac{f}{\nu} \right) \quad (54)$$

and then

$$F_M^{min} = (2\lambda)^{-1} \|\nu\|_{L^2}^2 \left\| \frac{f}{\nu} - P_{G_\mu} \left(\frac{f}{\nu} \right) \right\|_{L^2}^2. \quad (55)$$

By the same way, we find that

$$\hat{\eta}_m = P_{G_1} \left(\frac{f}{\mu\nu} \right) = \frac{1}{\mu} P_{G_\mu} \left(\frac{f}{\nu} \right) \quad (56)$$

and

$$F_m^{min} = \frac{(2\lambda)^{-1}}{\|\nu\|_{L^2}^2} \left\| \frac{f}{\nu} - P_{G_\mu} \left(\frac{f}{\nu} \right) \right\|_{L^2}^2. \quad (57)$$

Then if we set $B = (2\lambda)^{-1} \left\| \frac{f}{\nu} - P_{G_\mu} \left(\frac{f}{\nu} \right) \right\|_{L^2}^2$, we get

$$\frac{B}{\|\nu\|_{L^2}^2} \leq F^{min} = F(\hat{\eta}) \leq B \|\nu\|_{L^2}^2 \quad (58)$$

Remind that we suppose that $\left\| \frac{1}{\nu^2} \right\|_{L^2}^2 \lesssim 1$, in practice we can easily check that $\|\nu\|_{L^2} \approx 1$ and as $\hat{\eta}_m = \hat{\eta}_M$, a good approximation of $\hat{\eta}$ is given by

$$\hat{\eta} = \frac{1}{\mu} P_{G_\mu} \left(\frac{f}{\nu} \right) \quad (59)$$

Since $\hat{\eta} = \frac{\hat{v}}{\mu}$, we conclude the proof of the lemma that \hat{v} is well approximated by:

$$\hat{v} = P_{G_\mu} \left(\frac{f}{\nu} \right). \quad (60)$$

We apply this lemma two times to $F_{\lambda, \mu_1, \mu_2}^{JG}(u, v, w)$. Then \hat{v} is obtained from u and w , while \hat{w} is obtained from u and v . This concludes the proof of proposition 2.1.

B Appendix B

Now, we give the proof of proposition 4.1 of section 4. First, it is easy to see that for the components u and v , the proof is the same as the one of proposition 2.1 given in appendix A. So we will focus on w and then suppose the components u and v are fixed. We need to find

$$\hat{w} = \arg \min_{w \in E_\delta} \left\{ (2\lambda)^{-1} \|f - u - \nu_1 v - \nu_2 w\|_{L^2}^2 + B^* \left(\frac{w}{\delta} \right) \right\}. \quad (61)$$

To get this solution, the following lemma will be used

Lemma B.1 *Let $f \in L^2$, $w \in E_\delta$ and $\nu : \mathbb{R}^2 \rightarrow]0; 1[$. Assume that, the function ν can be considered as locally constant compared to the variation of v . Then the minimizer:*

$$\hat{w} = \arg \min_{w \in E_\delta} \left\{ (2\lambda)^{-1} \|f - \nu w\|_{L^2}^2 + B^* \left(\frac{w}{\delta} \right) \right\} \quad (62)$$

is given by:

$$\hat{w} = \frac{f}{\nu} - \frac{\lambda}{\delta \nu^2} WST \left(\frac{\delta f \nu}{\lambda}; \frac{2\delta^2 \tilde{\nu}^2}{\lambda} \right), \quad (63)$$

where $WST(f, \delta)$ is the Wavelet Soft Thresholding operator (see [9, 12]).

The proof is based on the differentiation of (62) with respect to w . Using the Euler-Lagrange equation, we get

$$0 \in -\frac{\nu}{\lambda}(f - \nu w) + \frac{1}{\delta} \partial B^* \left(\frac{w}{\delta} \right) \quad (64)$$

$$\Leftrightarrow \frac{w}{\delta} \in \partial B \left(\frac{\delta \nu (f - \nu w)}{\lambda} \right). \quad (65)$$

Let $\eta = \frac{\delta \nu (f - \nu w)}{\lambda}$. Then

$$0 \in \frac{\lambda}{\delta^2 \nu^2} \eta - \frac{f}{\delta \nu} + \partial B(\eta) \quad (66)$$

$$\Leftrightarrow 0 \in \eta - \frac{\delta f \nu}{\lambda} + \frac{\delta^2 \nu^2}{\lambda} \partial B(\eta). \quad (67)$$

This comes from the Euler-Lagrange equation of

$$\arg \inf_{\eta \in L^2} \left\{ \frac{1}{2} \left\| \eta - \frac{\delta f \nu}{\lambda} \right\|_{L^2}^2 + \frac{\delta^2 \nu^2}{\lambda} \|\eta\|_{B_{1,1}^1} \right\}. \quad (68)$$

Based on the result of Chambolle et al. [9] and because we suppose the ν behaves locally as a constant compared to η , we get

$$\hat{\eta} = WST \left(\frac{\delta f \nu}{\lambda}; \frac{2\delta^2 \tilde{\nu}^2}{\lambda} \right) \quad (69)$$

where $\tilde{\nu}$ is the pyramidal version of ν defined in section 4. Now, we get from the expression of $\hat{\eta}$:

$$\frac{\delta f \nu}{\lambda} - \frac{\delta \nu^2 \hat{w}}{\lambda} = WST \left(\frac{\delta f \nu}{\lambda}; \frac{2\delta^2 \tilde{\nu}^2}{\lambda} \right) \quad (70)$$

$$\Rightarrow \hat{w} = \frac{f}{\nu} - \frac{\lambda}{\delta \nu^2} WST \left(\frac{\delta f \nu}{\lambda}; \frac{2\delta^2 \tilde{\nu}^2}{\lambda} \right). \quad (71)$$

In order to end the proof of proposition 4.1, we just have to substitute f by $f - u - \nu_1 v$ and ν by ν_2 in the previous lemma. Then we get:

$$\hat{w} = \frac{f - u - \nu_1 v}{\nu_2} - \frac{\lambda}{\delta \nu_2^2} WST \left(\frac{\delta \nu_2}{\lambda} (f - u - \nu_1 v); \frac{2\delta^2 \tilde{\nu}_2^2}{\lambda} \right). \quad (72)$$

References

- [1] J.F. Aujol and G. Aubert and L. Blanc-Féraud and A. Chambolle, Decomposing an image: Application to SAR images, Scale-Space'03, Lecture Notes in Computer Science, Vol 1682, 2003.
- [2] L. Vese and S. Osher, Modeling textures with total variation minimization and oscillating patterns in image processing, *Journal of Scientific Computing*, Vol 19, No 1-3, 553–572, 2002.
- [3] Y. Meyer, Oscillating patterns in image processing and in some nonlinear evolution equations, The Fifteenth Dean Jacqueline B. Lewis Memorial Lectures, American Mathematical Society, 2001.
- [4] A. Chambolle, An algorithm for total variation minimization and applications, *Journal of Mathematical Imaging and Vision*, Vol 20, No 1-2, 89–97, 2004.
- [5] L. Rudin and S. Osher and E. Fatemi, Nonlinear total variation based noise removal algorithms, *Physica D*, Vol 60, 259–268, 1992.
- [6] J.F. Aujol and A. Chambolle, Dual norms and image decomposition models, *International Journal of Computer Vision*, Vol 63, No 1, 85-104, 2005.
- [7] G. Gilboa and Y. Zeevi and N. Sochen, Texture preserving variational denoising using an adaptive fidelity term, *IEEE Trans. Image Proc*, Vol 15(8), 2281–2289, 2006.
- [8] G. Gilboa and N. Sochen and Y. Zeevi, Estimation of optimal PDE-based denoising in the SNR sense, *IEEE Trans. Image Proc*, Vol 15(8), 2269–2280, 2006.
- [9] A. Chambolle and R.A. DeVore and N. Lee and B.J. Lucier, Nonlinear wavelet image processing: variational problems, compression and noise removal through wavelet shrinkage, *IEEE Trans. on Image Processing*, Vol 7, 319–335, 1998.
- [10] J.F. Aujol and G. Gilboa and T. Chan and S. Osher, Structure-texture image decomposition-modeling, algorithms and parameter selection, *International Journal of Computer Vision*, Vol 67, No 1, 111–136, 2006.
- [11] J.F. Aujol, Contribution à l'analyse de textures en traitement d'images par méthodes variationnelles et équations aux dérivées partielles, PhD, Université de Nice Sophia Antipolis, 2004.
- [12] D. Donoho, De-noising by soft-thresholding, *IEEE Trans. Inform. Theory*, Vol 41, 613–627 1995.
- [13] J. Gilles, Décomposition et détection de structures géométriques en imagerie, PhD, Ecole Normale Supérieure de Cachan, <http://jerome.gilles91.free.fr>, 2006.
- [14] A. Haddad, Méthodes variationnelles en traitement d'image, PhD, Ecole Normale Supérieure de Cachan, 2005.

- [15] L. Lieu, Contribution to Problems in Image Restoration, Decomposition, and Segmentation by Variational Methods and Partial Differential Equations, Ph.D. Thesis, CAM Report 06-46, 2006.
- [16] J.B. Garnett and T.M. Le and Y. Meyer and L. Vese, Image Decompositions Using Bounded Variation and Homogeneous Besov Spaces, CAM Report 05-57, 2005.
- [17] T.M. Le and L. Vese, Image decomposition using total variation and $\text{div}(\text{BMO})$, CAM Report 04-36, 2004.
- [18] A. Haddad and Y. Meyer, Variational Methods in Image Processing, CAM Report 04-52, 2004.