Pré-Publicações do Departamento de Matemática
Universidade de Coimbra
Preprint Number 20–28

VARIATIONAL MULTI-CHANNEL INPAINTING AND DENOISING MODEL: EXISTENCE OF SOLUTION AND NUMERICAL APPROXIMATION

ISABEL N. FIGUEIREDO AND MAHDI DODANGEH

Abstract: In this paper we propose a variational "total variation like" inpainting and denoising model, for multi-channel images, prove the existence and uniqueness of its solution, and define and implement a numerical scheme for its solution. This variational model includes, besides the data-fidelity term, an extension of the total variation regularizer, appropriate for vector-valued images, aiming at reconstructing sharp image edges, and also a smooth regularizer for removing noise. The proposed numerical algorithm is an instance of the so-called alternating direction method of multipliers, for fast image recovery, which transforms the discrete version of the variational model into a constrained optimization problem, by using variable splitting. This constrained problem is subsequently solved by an augmented Lagrangian approach. For assessing the method, tests are performed on real-world and good quality color images, that are artificially damaged by adding noise and removing randomly pixel values in the different image channels. In addition an experiment on *ab initio* degraded medical image, corrupted with specular highlights, is also carried out.

Keywords: Variational Model, Numerical Solution, Image Processing.
Math. Subject Classification (2010): 49J27; 46B25; 65Q10; 65K10; 68U10.

1. Introduction

Image inpainting is a kind of interpolation procedure, which consists in filling in missing parts of a degraded image, by using the information available from the parts of the image that are complete and not damaged. Denoising is an image processing technique, that aims to reduce noise (an imperfection that degrades images and occurs during the acquisition of images), by means of a smoothing procedure.

In this paper we propose a variational "total variation like" model for recovering incomplete and noisy multi-channel images. It is a minimization problem, and like the total variation (TV) inpainting model of [13], it includes

Received October 06, 2020.
This work was supported by the FCT (Fundação para a Ciência e a Tecnologia, Portugal) research project PTDC/EMD- EMD/28960/2017, and also partially by the Centre for Mathematics of the University of Coimbra (CMUC) - UIDB/00324/2020, funded by the Portuguese Government through FCT/MCTES..

a data fidelity term, that adjusts the "searched good image" to the original one, in the intact parts, plus a regularizing term which is an extension of TV to vector-valued functions. In addition, in our model there is a second regularizer for reducing noise (see the definition of the proposed model in equation (1)).

There are several methods for restoring degraded images, and those based on TV (see [6, 12] for seminal works on TV) have proved to lead to superior image restoration quality, in particular by reducing noise and blur, without smearing the edges.

Here we use the "Color TV" regularizer proposed in [3], for vector-valued images, herein denoted by TV^{ext} (see the definition in (5)). It is an extension of TV. It reduces to the usual total variation, TV, in the case of a scalar image, and it keeps the main properties of TV, namely preservation of edges and rotation invariance. It also acts as a coupling among the channels, that, and according to [3, formulas (7) and (8)] *"takes the form of a global channel-wise scaling of the diffusion coefficient, such that a channel with larger TV will be smoothed more than a channel with with smaller TV.... and retains more detail in the weaker channels"* (see the last equation in (25) of this paper, where this statement becomes evident). Moreover, other works existing on the literature report on the ability of this regularizer to recover multi-channel images (for example, in remote sensing images cf. [7]).

The numerical method chosen to find the solution of the proposed model relies on a suitable modification of a fast algorithm, introduced in [1], which itself is an instance of the so-called alternating method of multipliers [9]. This fast algorithm is appropriate for a variety of image restoration/reconstruction problems, which are formulated by unconstrained optimization problems, where the objective functional includes a data-fidelity term and a non-smooth regularizer, as for instance the total variation. Basically, this algorithm depends on a variable splitting technique, that transforms the unsconstrained problem into a constrained one, which is afterwards solved with an augmented Lagrangian technique.

We remark that in [7] the authors define a multi-channel nonlocal total variation model for inpainting remote sensing images, which is solved by a numerical procedure relying on a Split-Bregman type algorithm. However no proof is provided for the existence and uniqueness of solution of the model, neither for the convergence of the numerical method. We also refer to [15] for the description and implementation of a fast algorithm for multi-channel

[9] Eckstein, J., and Bertsekas, D. P. On the Douglas - Rachford splitting method and the proximal point algorithm for maximal monotone operators. *Mathematical Programming 55*, 1-3 (1992), 293–318.
[10] Ekeland, I., and Temam, R. *Convex analysis and variational problems*, vol. 28. Siam, 1999.
[11] Liu, C., Ng, M. K.-P., and Zeng, T. Weighted variational model for selective image segmentation with application to medical images. *Pattern Recognition 76* (2018), 367–379.
[12] Rudin, L. I., Osher, S., and Fatemi, E. Nonlinear total variation based noise removal algorithms. *Physica D: nonlinear phenomena 60*, 1-4 (1992), 259–268.
[13] Shen, J., and Chan, T. F. Mathematical models for local nontexture inpaintings. *SIAM Journal on Applied Mathematics 62*, 3 (2002), 1019–1043.
[14] Vese, L. A., and Le Guyader, C. *Variational methods in image processing*. CRC Press, 2015.
[15] Yang, J., Yin, W., Zhang, Y., and Wang, Y. A fast algorithm for edge-preserving variational multichannel image restoration. *SIAM Journal on Imaging Sciences 2*, 2 (2009), 569–592.

Isabel N. Figueiredo
University of Coimbra, CMUC, Department of Mathematics, Faculty of Sciences and Technology, Portugal.
E-mail address: isabelf@mat.uc.pt
URL: http://http://www.mat.uc.pt/∼isabelf

Mahdi Dodangeh
University of Coimbra, CMUC, Department of Mathematics, Faculty of Sciences and Technology, Portugal
E-mail address: dodangeh@mat.uc.pt

in the images, they can be missed by the doctor who is analysing the video, and this is a risk for the patient. For example, in the medical image shown in Figure 5-a, we can perceive, in the middle top, a small polyp covered with specular highlights.

Clearly, the method we propose in this paper is suited for inpainting these specular highlights. Figure 5-a displays the original image, that is already degraded with the bright white spots. The regions to be inpainted are precisely these bright white spots. The identification of these regions can be done by using a simple threshold method based on the intensity of the pixels. In this way, we obtain the characteristic functions γ_i (3), whose graphic is displayed in Figure 5-b. For this image we consider the three functions γ_1, γ_2, and γ_3 to be equal for the three RGB channels. Finally, we obtain the restored image, shown in Figure 5-c, with the proposed algorithm for model (6) described in subsection 3.3.3. As can be observed, all the specular highlights have disappeared and we have a better visualization of the mucosa and the small colonic polyp.

For all the experiments, a computer with Intel(R) Core(TM) processor i7-6700 CPU@3.40 GHz was used, and the algorithms were implemented with MATLAB® 2018b (The MathWorks, Inc., Natick, Massachusetts, United States).

References

[1] Afonso, M. V., Bioucas-Dias, J. M., and Figueiredo, M. A. Fast image recovery using variable splitting and constrained optimization. *IEEE Transactions on Image Processing 19*, 9 (2010), 2345–2356.

[2] Attouch, H., Buttazzo, G., and Michaille, G. *Variational analysis in Sobolev and BV spaces: applications to PDEs and optimization*. SIAM, 2014.

[3] Blomgren, P., and Chan, T. F. Color TV: total variation methods for restoration of vector-valued images. *IEEE Transactions on Image Processing 7*, 3 (1998), 304–309.

[4] Cai, X., Chan, R., Nikolova, M., and Zeng, T. A three-stage approach for segmenting degraded color images: Smoothing, lifting and thresholding (SLaT). *Journal of Scientific Computing 72*, 3 (2017), 1313–1332.

[5] Cai, X., Chan, R., and Zeng, T. A two-stage image segmentation method using a convex variant of the Mumford-Shah model and thresholding. *SIAM Journal on Imaging Sciences 6*, 1 (2013), 368–390.

[6] Chambolle, A., and Lions, P.-L. Image recovery via total variation minimization and related problems. *Numerische Mathematik 76*, 2 (1997), 167–188.

[7] Cheng, Q., Shen, H., Zhang, L., and Li, P. Inpainting for remotely sensed images with a multichannel nonlocal total variation model. *IEEE Transactions on Geoscience and Remote Sensing 52*, 1 (2013), 175–187.

[8] Ciarlet, P. G. *Introduction à l'analyse numérique matricielle et à l'optimisation*. Masson, 1982.

image deblurring and denoising, where it is adopted another multi-channel total variation regularizer, different from the one used in the current paper.

We finish this introduction with an outline of the paper. After this introduction, in Section 2 we describe the model, introduce basic notations and give the proof of existence and uniqueness of solution, in the continuous setting. We also prove the existence of solution, for the same model but without the regularizer for reducing noise, in a broader functional space. A comparison of the two models, in the last Section, highlights the beneficial importance of the regularizer for reducing noise. In Section 3 we recall the background numerical methods, define the approximate problem in a finite dimensional setting, and describe in detail the proposed numerical method. Finally, in Section 4 we present the experiments performed on a variety of color (RGB) images in order to evaluate the quality of the proposed model and method.

2. Proposed Variational Model

Let $f = (f_1, \dots, f_m)$ be a given multichannel image with channels $f_i : \Omega \to \mathbb{R}$, $i = 1, \dots, m$, where $\Omega \subset \mathbb{R}^2$, is a square or a rectangle representing the pixel domain. For an RGB image, $m = 3$ and $f = (f_1, f_2, f_3)$. The given image f is in general noisy and incomplete. Let Ω_i, for $i = 1, \dots, m$, be the part of the domain Ω where the given channel f_i is known.

The proposed variational multi-channel model intends to restore the given image, by simultaneously performing image denoising (a technique that consists in improving the quality of the image by noise removal) and image inpainting (a procedure which consists to filling the missing information in $\Omega \setminus \Omega_i$, for each channel i). Its definition is

$$\min_{u=(u_i)_{i=1}^m} E(u), \quad \text{with}$$

$$E(u) = \underbrace{\frac{\beta}{2} \sum_{i=1}^{m} \int_\Omega \omega_i^2 \gamma_i |u_i - f_i|^2 dx}_{\text{fitting term}} + \underbrace{\mu\, TV^{ext}(u) + \frac{\alpha}{2} \sum_{i=1}^{m} \int_\Omega |\nabla u_i|^2 dx}_{\text{regularizing terms}}. \quad (1)$$

2.1. Basic notations. Before explaining the meaning of each term in (1), we firstly recall the notations, basic terminology and definition for Sobolev and BV (bounded variation) spaces. For the complete theory and proofs associated to these preliminaries we refer for example to the books [2, 14].

We denote $W^{m,p}(\Omega)$ the Sobolev spaces ($1 \leq p < \infty$, $m = 0, 1, 2, \ldots$) whose definition is $W^{m,p}(\Omega) = \{v \in L^p(\Omega) : \ \partial_j^k v \in L^p(\Omega), \ k = 1, \ldots, m\}$, where $\partial_j^k v$ means the weak partial derivative of v of order k, with respect to x_j, and $x = (x_1, x_2)$ an arbitrary point in Ω. The space $L^p(\Omega) = \{v : \Omega \to \mathbb{R}, v \text{ is Lesbegue measurable and} \|v\|_{L^p(\Omega)} < \infty\}$, with $\|v\|_{L^p(\Omega)} = (\int_\Omega v^p \, dx)^{\frac{1}{p}}$. We denote by $BV(\Omega)$ the space of functions v of bounded variation in Ω, that is, the functions $v \in L^1(\Omega)$, such that the total variation of v in Ω denoted by $|Dv|(\Omega)$ and defined by $|Dv|(\Omega) = \sup\left\{\int_\Omega v \, div\phi \, dx : \ \phi \in C_c^1(\Omega; \mathbb{R}^2), \ |\phi| < 1\right\}$ is finite, where $C_c^1(\Omega; \mathbb{R}^2)$ is the space of functions ϕ with continuous derivatives until the order 1 and with compact support. We remark that $v \in BV(\Omega)$, if and only if, $v \in L^1(\Omega)$ and Dv is a $\mathbb{R}^2$-valued Borel measure, where Dv is defined by $< Dv, \phi >= \int_\Omega v \, div\phi \, dx$, that is, the gradient of v in the distributional sense.

We recall that for a function $v : \Omega \to \mathbb{R}$ of bounded variation belonging to $L^1(\Omega)$, and denoting by Dv its gradient in the distributional sense, then "$\int_\Omega |\nabla v|$" is the usual notation for the total mass $|Dv|(\Omega) = \int_\Omega |Dv|$ of the total variation $|Dv|$, also denoted by "$TV(v)$", of the measure Dv. This notation "$\int_\Omega |\nabla v|$" is generally adopted, because whenever the gradient of $\nabla v = (\partial_1 v, \partial_2 v)$ exists, or more precisely, when $v \in W^{1,1}(\Omega)$ then, $|Dv|(\Omega) = \int_\Omega |\nabla v| dx$ (see [14, Definition 11, pages 26-27] or [2, Definition 10.1.1, pages 371-372]).

Moreover, we also recall that in the numerical approximations, a discrete form of $\int_\Omega |\nabla v|$ is always used. It is generally defined by $\|\nabla v\|_1$, that is,

$$\int_\Omega |\nabla v| \approx \|\nabla v\|_1 = \sum_{j \in \Omega} \sqrt{(\partial_1 v)_j^2 + (\partial_2 v)_j^2}, \tag{2}$$

in which $v \in \mathbb{R}^{n \times n}$ is the matrix representing the scalar image v at $n \times n$ points in the pixel domain, the pair $\left(\partial_1 v, \partial_2 v\right)_j \in \mathbb{R}^2$ represents certain first-order finite differences of v at pixel j in horizontal ($(\partial_1 v)_j$) and vertical ($(\partial_2 v)_j$) directions, and the summation is taken over all pixels $j = 1, \ldots, , n \times n$.

2.2. Explanation of the Proposed Model. In formula (1),

- the unknown restored image is the vector-valued function $u = (u_1, \ldots, u_m) : \Omega \subset \mathbb{R}^2 \to \mathbb{R}^m$,
- α, β and μ are positive constants representing weights,

(a) Medical image with specular highlights (b) Binary mask of the specular highlights $\equiv$ function γ_i (c) Restored image

FIGURE 5. Original *in vivo* medical image (a) and restored image (c) obtained with model (6). In (b) the white spots correspond to the graphic of the characteristic function in (3) ($\gamma_1 = \gamma_2 = \gamma_3$).

and four different percentages of information loss (respectively, 5%, 10%, 15% and 20%). On can see that the curves diverge slightly and become parallel to each other when the percentage of information loss is fixed. This means that the initial added noise level has a negligible impact on the final result. In contrast, when the noise level is fixed, the corresponding curves depending on the amount of information loss converge to parallel lines and tend to decrease their distance. This also means that the proposed algorithm effectively recovers/overcomes the lost pixel value information.

In the previous experiments the original good quality images are artificially corrupted in order to evaluate the performance and quality of the proposed method. In this last test we consider an original image, which is already naturally degraded due to the acquisition procedure. Figure 5 shows the results obtained for a medical image, shown in Figure 5-a, with the proposed algorithm for model (6). It is an *in vivo* image extracted from a colonoscopy video, acquired during a medical exam to a patient (courtesy of the Department of Gastroenterology, CHUC - Centro Hospitalar e Universitário de Coimbra, Coimbra, Portugal). This type of colonoscopy images is in general degraded. In fact, very often these images exhibit specular highlights, which are bright white spots, originated by reflection of the colonoscope light source. This is a serious drawback for doctors who search to detect different abnormalities in the images, mainly colonic polyps, precursors of colon cancer. When very small polyps are covered or hidden by specular highlights

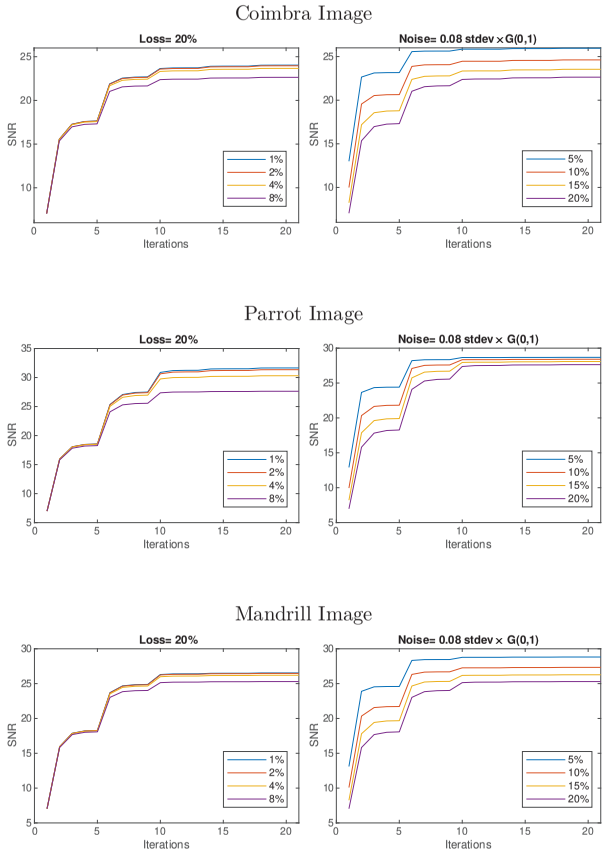


FIGURE 4. Evolution of SNR when the percentages of noise level or, vice-versa, information loss are fixed, during the execution of the algorithm for model (6).

of SNR measure, along the execution of the proposed algorithm for model (6). Similarly to the previous Figure 3, on the left column of Figure 4, each graphic exhibits four curves representing the SNR function for a fixed 20% of information loss, and four percentages of noise level (1%, 2%, 4% and 8%). On the right column the graphics show the SNR curves for a fixed noise level

- γ_i is the characteristic function of Ω_i, a subdomain of Ω,

$$\gamma_i = \begin{cases} 1, & x \in \Omega_i \subset \Omega, \\ 0, & x \in \Omega \setminus \Omega_i, \end{cases} \qquad i = 1, \ldots, m, \tag{3}$$

- ω_i, for $i = 1, \ldots, m$, is another weighting function, that involves an edge detector function g_i defined by

$$\omega_i^2(x) = 1 - g_i(x) \quad \in]0, 1] \qquad \text{and} \qquad g_i(x) = \frac{1}{1 + \varepsilon |\nabla f_i(x)|^2} < 1, \tag{4}$$

  where ∇ is the gradient operator, and $\varepsilon > 0$ a parameter that adjusts the strength of the edge,
- and finally $TV^{ext}(u)$ is the extended total variational (TV) norm for color and other vector-valued images (or equivalently multi-channel images), defined by

$$TV^{ext}(u) = \sqrt{\sum_{i=1}^{m} \Big(TV(u_i)\Big)^2} = \sqrt{\sum_{i=1}^{m} \left[\int_\Omega |\nabla u_i|\right]^2}. \tag{5}$$

The fitting term in (1) obliges the restored image $u = (u_1, \ldots, u_m)$ to be close to its observation (the data $f = (f_1, \ldots, f_m)$ in the possible different subdomains Ω_i). The weighting function ω_i is large near the edges (because the edge detector g is large) and becomes smaller away from the edges (where the edge detector is small). Consequently, with this definition for the fitting term, in regions near the edges details are kept while regions far away from the edges are smoothed more (see [11], where a similar weighting function was used for defining a selective image segmentation model).

The two regularizing terms impose constraints to the solution u. Mainly they play a vital role in reducing noise and in restoring the image in the missing domains $\Omega \setminus \Omega_i$, for $i = 1, \ldots, m$, by diffusion. The first regularizer, the extended total variation, proposed in [3, Definition 2], couples the different channels, has the property of not penalizing discontinuities (edges) in each channel, and is rotationally invariant (it is also a form of anisotropic diffusion). The second regularizing term ensures also more smoothness of the minimizer u.

2.3. Existence and Uniqueness of Solution. Notice that if, for each i, $f_i \in [L^2(\Omega)]^m$, the integrand of the fitting term in (1) can be written as $|\mathcal{A}_i u_i - \bar{f}_i|^2$, with $\bar{f}_i = \omega_i^2 \gamma_i f_i$ and $\mathcal{A}_i = \omega_i^2 \gamma_i I : L^2(\Omega) \to L^2(\Omega)$ a linear

and bounded operator defined by $\mathcal{A}_i(v) = \omega_i^2 \gamma_i v$ (I represents the identity operator). Then, by defining $\bar{f} = (\bar{f}_1, \ldots, \bar{f}_m)$, $\mathcal{A} = (\mathcal{A}_1, \ldots, \mathcal{A}_m)$ and $\overrightarrow{\nabla} u = (\nabla u_1, \ldots, \nabla u_m)$ for any $u \in [W^{1,2}(\Omega)]^m$ or $u \in [BV(\Omega)]^m$ (where in this latter case the notation $\overrightarrow{\nabla} u$ should be interpreted as $|\overrightarrow{D} u| = (|Du_1|, \ldots, |Du_m|)$, as explained before in section 2.1), there are two equivalent and alternative forms for defining the energy function $E(u)$ of the variational model (1), which lead to the following equivalent formulations for model (1) :

$$\begin{aligned} &\min_{u=(u_i)_{i=1}^m} \left[\frac{\beta}{2} \sum_{i=1}^m \int_\Omega |\mathcal{A}_i u_i - \bar{f}_i|^2 dx + \frac{\alpha}{2} \sum_{i=1}^m \int_\Omega |\nabla u_i|^2 dx + \mu\, TV^{ext}(u) \right] \\ &\equiv \min_{u=(u_i)_{i=1}^m} \left[\frac{\beta}{2} \int_\Omega \|\mathcal{A} u - \bar{f}\|_{\mathbb{R}^m}^2 dx + \frac{\alpha}{2} \int_\Omega \|\overrightarrow{\nabla} u\|_{\mathbb{R}^m}^2 dx + \mu\, TV^{ext}(u) \right], \end{aligned} \tag{6}$$

where $\|.\|_{\mathbb{R}^m}$ is the usual Euclidean norm in $\mathbb{R}^m$.

Now, we prove firstly the existence and uniqueness of solution in the space $[W^{1,2}(\Omega)]^m$. Afterwards, by keeping only the regularizer involving $TV^{ext}(u)$ in the definition of $E(u)$, we also prove the existence and uniqueness of solution in the space $[BV(\Omega)]^m$

Theorem 2.1 (*Minimizer in $[W^{1,2}(\Omega)]^m$*)**.** *Let Ω be a bounded connected open subset of $\mathbb{R}^2$ with a Lipschitz boundary. Let $\mathcal{A}_i : L^2(\Omega) \to L^2(\Omega)$ be bounded and linear, and, for each $i \in \{1, \ldots, m\}$, assume that*
i) $\bar{f}_i \in L^2(\Omega)$,
ii) $\inf_{x \in \omega} \omega_i(x) > 0$ and $\omega_i^{min} \leq \omega_i(x) \leq \omega_i^{max}$, with ω_i^{min} and ω_i^{max} two positive constants,
iii) $\mathcal{K}ern(\mathcal{A}_i) \cap \mathcal{K}ern(\nabla) = \{0\}$, where $\mathcal{K}ern(...)$ denotes the kernel of the operator between brackets.

Then, the variational model (1) has a unique minimizer $u \in [W^{1,2}(\Omega)]^m$.

Proof: The proof relies on arguments of the same type of those used in [5, Theorem 2.4] (see also [4]).

If the infimum of E is finite because $E(u) \geq 0$ on $[W^{1,2}(\Omega)]^m$ and by choosing $u = 0$, then

$$0 \leq \inf_{u \in [W^{1,2}(\Omega)]^m} E(u) \leq E(0) = \frac{\beta}{2} \sum_{i=1}^m \int_\Omega |\bar{f}_i|^2 dx < \infty.$$

For proving the existence of a minimum of E we just need to show that $E(u)$ is coercive in $[W^{1,2}(\Omega)]^m$, that is, $E(u) \to +\infty$, when $\|u\|_{[W^{1,2}(\Omega)]^m} \to +\infty$,

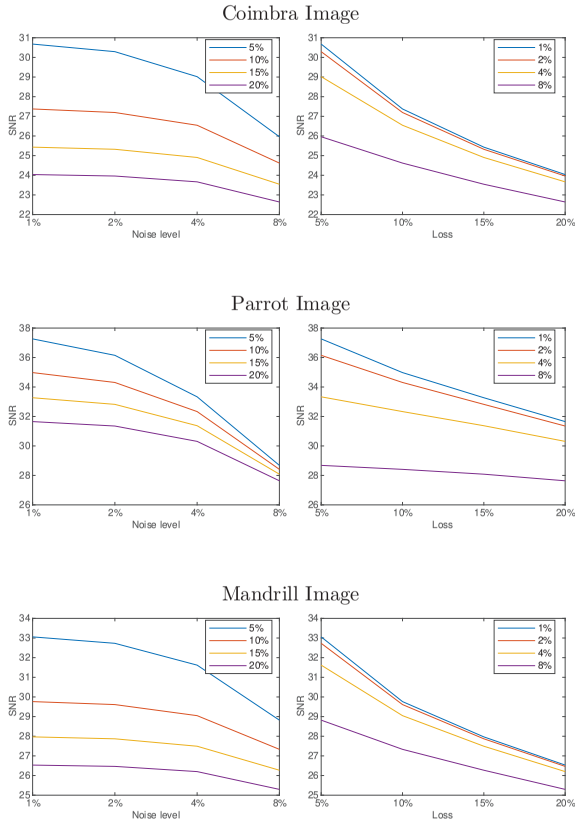
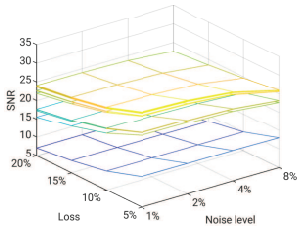


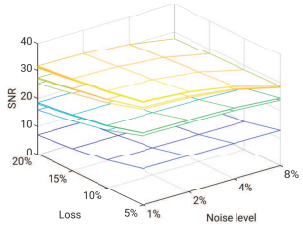
FIGURE 3. Effects of the percentages of Gaussian noise level and information loss, separately, on the final result in terms of SNR, with model (6). On the left : SNR curves for four fixed information loss and varied noise level. On the right : SNR curves for four fixed noise level and varied information loss.

the proposed method/model is slightly more robust against the Gaussian noise rather than information loss.

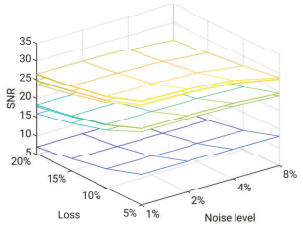
Figure 4 tracks the change of the degraded image (with added Gaussian noise and information loss) towards the final restored image, again in terms



(a) Coimbra Image



(b) Parrot Image



(c) Mandrill Image

FIGURE 2. SNR of six iterations of the proposed algorithm for model (6), as a scalar function of the two degradation variables: information loss (variable denoted by "Loss" in the horizontal plane), and added Gaussian noise (variable denoted by "Noise level" in the horizontal plane).

In Figure 3 the two variables were fixed one by one. On the left column the curves represent SNR for four fixed information loss (with percentages 5%, 10%, 15% and 20%) and four varied percentages of noise level equal to 1%, 2%, 4% and 8%. Vice-versa, on the right column, the curves represent SNR for the four fixed noise levels (with percentages 1%, 2%, 4% and 8%) and varied information loss equal to 5%, 10%, 15% and 20%. As can be seen, when the percentage of noise level or information loss increases, the final error also increases, because SNR decreases. However, we see that the SNR curves are converging to each other. This simply means that the proposed method is robust against noise and information loss. In particular, the convergence of the curves is more clear when the noise level increases. The later means that

with

$$\|u\|_{[W^{1,2}(\Omega)]^m} = \underbrace{\left[\sum_{i=1}^{m}\|u_i\|^2_{L^2(\Omega)}\right]^{\frac{1}{2}}}_{\|u\|_{[L^2(\Omega)]^m}} + \underbrace{\left[\sum_{i=1}^{m}\|\nabla u_i\|^2_{L^2(\Omega)}\right]^{\frac{1}{2}}}_{\|\overrightarrow{\nabla} u\|_{[L^2(\Omega)]^m}} \tag{7}$$

where $\|.\|_{L^2(\Omega)}$ is the usual norm in $L^2(\Omega)$. This statement is consequence of the fact that $[W^{1,2}(\Omega)]^m$ is a reflexive Banach space, and $E(u)$ is convex and lower semi-continuous (see for instance [10, Proposition 1.2] or [8, Théorème 8.2.-2, Remarques (1), (2)]). Clearly we have

$$\|\overrightarrow{\nabla} u\|_{[L^2(\Omega)]^m} = \left[\sum_{i=1}^{m}\|\nabla u_i\|^2_{L^2(\Omega)}\right]^{\frac{1}{2}} = \left[\sum_{i=1}^{m}\int_\Omega |\nabla u_i|^2 dx\right]^{\frac{1}{2}} \leq \left[\frac{2}{\alpha}E(u)\right]^{\frac{1}{2}}, \tag{8}$$

so it remains to show that $\|u\|_{[L^2(\Omega)]^m}$ is also bounded by $[E(u)]^{\frac{1}{2}}$.

Using the Poincaré inequality on $W^{1,2}(\Omega)$, and (8), we have for each i,

$$\begin{aligned} &\|u_i - u_i^\Omega\|_{L^2(\Omega)} \leq C_i^\Omega \|\nabla u_i\|_{L^2(\Omega)} \leq C_i^\Omega \left[\frac{2}{\alpha}E(u)\right]^{\frac{1}{2}}, \quad \text{with} \\ &u_i^\Omega = \frac{1}{|\Omega|}\int_\Omega u_i(x)dx. \end{aligned} \tag{9}$$

Therefore setting $u^\Omega = (u_1^\Omega, \ldots, u_m^\Omega)$

$$\|u - u^\Omega\|_{[L^2(\Omega)]^m} \leq [CE(u)]^{\frac{1}{2}}, \quad \text{where} \quad C = \frac{2}{\alpha}(\max_i C_i^\Omega)^2 m. \tag{10}$$

In addition we have that

$$\|\mathcal{A}_i u_i^\Omega\|_{L^2(\Omega)} = \|\omega_i^2\gamma_i I u_i^\Omega\|_{L^2(\Omega)} = |u_i^\Omega|\left[\int_{\Omega_i} w_i^4 dx\right]^{\frac{1}{2}} \geq |u_i^\Omega|(\omega_i^{min})^2|\Omega_i|^{\frac{1}{2}} \tag{11}$$

and

$$0 \leq \frac{\beta}{2}\int_\Omega |\mathcal{A}_i u_i - \bar{f}_i|^2 dx \leq E(u) \Longrightarrow \|\mathcal{A}_i u_i - \bar{f}_i\|_{L^2(\Omega)} \leq \left[\frac{2}{\beta}E(u)\right]^{\frac{1}{2}}. \tag{12}$$

So, from (11), (12) and (10)

$$\begin{aligned} |u_i^\Omega|(\omega_i^{min})^2|\Omega_i|^{\frac{1}{2}} &\leq \|\mathcal{A}_i u_i^\Omega\|_{L^2(\Omega)} \\ &\leq \|\mathcal{A}_i u_i - \bar{f}_i\|_{L^2(\Omega)} + \|\bar{f}_i - \mathcal{A}_i(u_i - u_i^\Omega)\|_{L^2(\Omega)} \\ &\leq \left[\tfrac{2}{\beta}E(u)\right]^{\frac{1}{2}} + \|\bar{f}_i\|_{L^2(\Omega)} + [CE(u)]^{\frac{1}{2}} \end{aligned} \tag{13}$$

Finally from the triangular inequality

$$\|u\|_{[L^2(\Omega)]^m} \;\leq\; \|u-u^\Omega\|_{[L^2(\Omega)]^m} + \|u^\Omega\|_{[L^2(\Omega)]^m}, \tag{14}$$

noticing that $\|u_i^\Omega\|_{L^2(\Omega)} = \left[\int_\Omega |u_i^\Omega|^2 dx\right]^{1/2} = \left[|u_i^\Omega|^2 \int_\Omega 1 dx\right]^{1/2} = |u_i^\Omega||\Omega|^{\frac{1}{2}}$, and using the estimates (10) and (13), we conclude by (14), that $\|u\|_{[L^2(\Omega)]^m}$ is also bounded by a constant plus $[E(u)]^{\frac{1}{2}}$ times a constant. Therefore the coercivity of $E(u)$ is proved, as well as the existence of minimizer.

For proving the uniqueness we suppose that u and v are two different minimizers. Due to the convexity of E we have the equality

$$E(\theta u + (1-\theta)v) = \theta E(u) + (1-\theta)E(v), \qquad \forall \theta \in (0,1). \tag{15}$$

In addition, because of the strict convexity of the two first terms of E, as defined in (6), the following two equalities hold

$$\begin{aligned}
&\sum_{i=1}^m \frac{\beta}{2}\int_\Omega |\mathcal{A}_i(\theta u_i + (1-\theta)v_i) - \bar{f}_i|^2 dx = \\
&\qquad \sum_{i=1}^m \left[\frac{\theta\beta}{2}\int_\Omega |\mathcal{A}_i u_i - \bar{f}_i|^2 dx + \frac{(1-\theta)\beta}{2}\int_\Omega |\mathcal{A}_i v_i - \bar{f}_i|^2 dx\right] \\
&\sum_{i=1}^m \frac{\alpha}{2}\int_\Omega |\nabla(\theta u_i + (1-\theta)v_i)|^2 dx = \\
&\qquad \sum_{i=1}^m \left[\frac{\theta\alpha}{2}\int_\Omega |\nabla u_i|^2 dx + \frac{(1-\theta)\alpha}{2}\int_\Omega |\nabla v_i|^2 dx\right].
\end{aligned} \tag{16}$$

These two equations impose that $\mathcal{A}_i u_i = \mathcal{A}_i v_i$ and $\nabla u_i = \nabla v_i$, respectively (choosing for example $\theta = \frac{1}{2}$). Consequently, due to the assumption $\mathcal{K}ern(\mathcal{A}_i) \cap \mathcal{K}ern(\nabla) = \{0\}$, which is obviously verified because $\mathcal{A}_i = \omega_i^2 \gamma_i I$, we conclude that $u = v$. ∎

For information loss a total of 30% of pixel values are modified randomly, by changing to zero, indiscriminately, 10% of the pixel values, per channel.

To assess (and compare) the quality of the restored images with the two models, the signal-to-noise-ratio (SNR) is used. For a scalar image $\bar{u}$ its definition is

$$\text{SNR}(\bar{u}, u) = 10 \times \log_{10} \frac{\|\bar{u} - E(\bar{u})\|_{L^2}}{\|\bar{u} - u\|_{L^2}},$$

with $E(\bar{u})$ the average intensity of the original image $\bar{u}$, and u the restored image.

Figure 1 shows the recovered images with the proposed model (6) and Table 1 gives the performance comparison among the two models (6) and (17), in terms of the SNR measure. From this Table 1, we see that model (6) outperforms clearly model (17).

Model	Coimbra	Parrot	Mandrill	Synthesis
(6)	27.38	34.98	29.76	36.18
(17)	21.48	27.65	24.47	29.35

TABLE 1. SNR of the restored images with models (6) and (17)

To better analyse the effect of the added Gaussian noise and information loss on the performance of the proposed algorithm for model (6), we vary the magnitude of these two variables and computed SNR, in each iteration during the execution of the algorithm. Therefore, instead of only one noise level, we take four levels, by considering stdev in (34) multiplied by

$$n_1 = 0.01 = 1\%,\ n_2 = 0.02 = 2\%,\ n_3 = 0.04 = 4\%, \text{ and } n_4 = 0.08 = 8\%. \tag{35}$$

Similarly, we also consider four levels of information loss, by changing to zeo, randomly

$$l_1 = 5\%,\ l_2 = 10\%,\ l_3 = 15\%, \text{ and } l_4 = 20\%, \tag{36}$$

pixel values, from each image-channel.

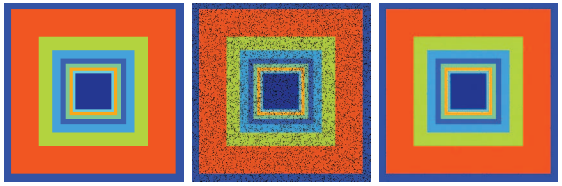
For each value of the pair (n_i, l_j), with $i, j = 1, 2, 3, 4$, the algorithm was executed and the SNR measure computed: the results are displayed in Figure 2.

The next two Figures 3 and 4 allow a better interpretation of Figure 2, as well as better study of the influence of the two variables (the percentages of added Gaussian noise level and information loss), separately, on the performance of the proposed method.

(a) Original (b) Degraded (c) Restored

(d) Original (e) Degraded (f) Restored

(g) Original (h) Degraded (i) Restored



(j) Original (k) Degraded (l) Restored

FIGURE 1. Original, degraded and restored images obtained with model (6).

by 1% of the standard deviation of the channel (denoted by "stdev"), that is

$$\text{Noise} = 0.01\,\text{stdev} \times G(0,1). \tag{34}$$

Theorem 2.2 (*Minimizer in* $[BV(\Omega)]^m$). *Under the same hypothesis of the previous Theorem 2.1, the following variational model with only one regularizer, the total variation* $TV^{ext}(u)$*, defined by*

$$\begin{aligned} &\min_{u\in[BV(\Omega)^m]} \left[\frac{\beta}{2}\sum_{i=1}^{m}\int_\Omega |\mathcal{A}_i u_i - \bar{f}_i|^2 dx + \mu\, TV^{ext}(u)\right] \\ &\equiv \min_{u\in[BV(\Omega)^m]} \left[\frac{\beta}{2}\int_\Omega \|\mathcal{A}u - \bar{f}\|^2_{\mathbb{R}^m} dx + \mu\, TV^{ext}(u)\right], \end{aligned} \tag{17}$$

has a unique minimizer $u \in [BV(\Omega)]^m$.

Proof: - The statement follows by remarking that, the assumptions and proof of existence and uniqueness of minimizer in the space $BV(\Omega)$ of [14, Theorem 26, page 54], can be extended to the cartesian space $[BV(\Omega)]^m$. In particular, for our problem the function ϕ in [14, Theorem 26] is now defined by $\phi : \mathbb{R}^m \to \mathbb{R}^+$, with $\phi(z) = (|z_1|^2 + \ldots + |z_m|^2)^{\frac{1}{2}}$, and satisfies assumption H1 of that Theorem 26. ■

3. Numerical Solutions and Convergence

We firstly recall in Section 3.1 some background methods, afterwards we present the discrete setting of (6) and (17) in Section 3.2, and then we describe the numerical method for the approximate solution. We present in detail all the steps of the proposed algorithm, in Sections 3.3 and Subsections 3.3.1 and 3.3.2, and finally in the last Subsection 3.3.3 we sum up the global numerical method for the solution of (6) and (17).

3.1. Background Methods - ADMM and SALSA. The alternating direction method of multipliers (ADMM) [9], appropriate to address the following abstract problem

$$\begin{aligned} &\min_{u\in\mathbb{R}^n} f_1(u) + f_2(Gu) \\ &\text{with } G \in \mathbb{R}^{m\times n},\ f_1 : \mathbb{R}^n \to \mathbb{R},\ f_2 : \mathbb{R}^m \to \mathbb{R}. \end{aligned} \tag{18}$$

Problem (18) is clearly equivalent to the following constrained optimization reformulation

$$\begin{aligned} &\min_{u\in\mathbb{R}^n,\, v\in\mathbb{R}^m} f_1(u) + f_2(v) \\ &\text{subject to } Gu = v. \end{aligned} \tag{19}$$

The ADMM applied to (19) is the following:

Algorithm ADMM

(1) Set $k = 0$, choose $\rho > 0$, v^0 and d^0.
(2) **repeat**
(3) $\quad u^{k+1} \in argmin_u \left(f_1(u) + \frac{\rho}{2}\|Gu - v^k - d^k\|_2^2\right)$
(4) $\quad v^{k+1} \in argmin_v \left(f_2(v) + \frac{\rho}{2}\|Gu^{k+1} - v - d^k\|_2^2\right)$
(5) $\quad d^{k+1} = d^k - (Gu^{k+1} - v^{k+1})$
(6) $\quad k \leftarrow k + 1$
(7) **until** stopping criterion is satisfied.

Above and hereafter, $\|.\|_2$ denotes the usual Euclidean norm, or $2-norm$, in the Euclidean space $\mathbb{R}^n$.

The convergence of algorithm ADMM to problems of the form (18) is demonstrated in [9]. We recall here that statement.

Theorem 3.1. [9, Theorem 8] *Consider problem* (18)*, where f_1 and f_2 are closed, proper convex functions, and $G \in \mathbb{R}^{m\times n}$ has full column rank. Consider arbitrary $\rho > 0$ and v_0, d_0 in $\mathbb{R}^m$. Let $\{\eta_k \geq 0,\ k = 0, 1, \ldots\}$ and $\{\nu_k \geq 0,\ k = 0, 1, \ldots\}$ be two sequences such that*

$$\sum_{k=0}^{\infty} \eta_k < \infty \quad and \quad \sum_{k=0}^{\infty} \nu_k < \infty.$$

Consider the sequences $\{u^k \in \mathbb{R}^n,\ k = 0, 1, \ldots\}$, $\{v^k \in \mathbb{R}^m,\ k = 0, 1, \ldots\}$, and $\{d^k \in \mathbb{R}^m,\ k = 0, 1, \ldots\}$ that satisfy

$$\begin{aligned} \eta_k &\geq \left\|u^{k+1} - argmin_u \left(f_1(u) + \tfrac{\rho}{2}\|Gu - v^k - d^k\|_2^2\right)\right\|_2 \\ \nu_k &\geq \left\|v^{k+1} - argmin_v \left(f_2(v) + \tfrac{\rho}{2}\|Gu^{k+1} - v - d^k\|_2^2\right)\right\|_2 \\ d^{k+1} &= d^k - (Gu^{k+1} - v^{k+1}). \end{aligned}$$

Then, if (18) *has a solution, the sequence $\{u_k\}$ converges, $u_k \to u^*$, where u^* is a solution of* (18)*. If* (18) *does not have a solution, then at least one of the sequences $\{u_k\}$ or $\{d_k\}$ diverges.*

The split augmented Lagrangian shrinkage algorithm (SALSA) proposed in [1] is an instance of ADMM, with $G = I$, where I is the identity map. We refer to [1] for further properties and analysis of SALSA algorithm.

3.2. Discretization of (6) **and** (17)**.** The discrete setting of (6) is

$$\min_u E(u), \quad \text{with} \quad E(u) = \frac{\beta}{2}\|\mathcal{A}u - \bar{f}\|_2^2 + \frac{\alpha}{2}\|\overrightarrow{\nabla} u\|_2^2 + \mu\,\|\overrightarrow{\nabla} u\|_1, \qquad (20)$$

(c) $\quad \theta^k = \mu \dfrac{\sum_{l=1}^m \|\nabla v_l^k\|_1}{\|\overrightarrow{\nabla} v^k\|_1}$

$$\begin{aligned} v_i^{p+1} &\in argmin_{v_i} \left(\frac{\rho}{2\theta^k}\|v_i - (u_i^{k+1} - d_i^k)\|_2^2 + \frac{\zeta}{2}\|\nabla v_i - w_i^p - e_i^p\|_2^2\right) \\ &= (\frac{\rho}{\theta^k} I + \zeta \nabla^T \nabla)^{-1} \left(\frac{\rho}{\theta^k}(u_i^{k+1} - d_i^k) + \zeta \nabla^T (w_i^p + e_i^p)\right). \end{aligned}$$

(d)

$$\begin{aligned} \omega_i^{p+1} &\in argmin_{\omega_i} \left(\|\nabla \omega_i\|_2 + \tfrac{\zeta}{2}\|\nabla v_i^{p+1} - \omega_i - e_i^p\|_2^2\right) \\ &= \frac{\nabla v_i^{p+1} - e_i^p}{\|\nabla v_i^{p+1} - e_i^p\|_2} \max\left\{(\|\nabla v_i^{p+1} - e_i^p\|_2 - \frac{1}{\zeta}, 0\right\} \end{aligned}$$

(e) $\quad e_i^{p+1} = e_i^k - (\nabla v_i^{p+1} - \omega_i^{p+1})$
(f) $\quad p \leftarrow p + 1$
(g) **until** stopping criterion is satisfied.
(5) $\quad d^{k+1} = d^k - (u^{k+1} - v^{k+1})$
(6) $\quad k \leftarrow k + 1$
(7) **until** stopping criterion is satisfied.

4. Experimental results

In this section we describe the results obtained with the proposed algorithm, defined in subsection 3.3.3, for the model (6). Some tests are also performed with the model (17), which does not contain the smooth regularizer, for comparison purposes.

We present the results firstly for four RGB color images : three real world images - an image of Coimbra city, a mandrill image, and a parrot image, and one synthesis image - all shown in the first column of Figure 1. Then, the results for a damaged *in vivo* medical image exhibiting specular highlights (see Figure 5-a) are also shown.

The different parameters included in these two models were chosen by trial and error until finding the most successful results for each model.

In the first experiments the four good quality original images (Coimbra city, mandrill, parrot and synthesis images) are artificially corrupted with added Gaussian noise and information loss. In each channel, the added Gaussian noise is of mean 0 and variance 1 (we adopt the notation $G(0, 1)$), multiplied

Both points 3 and 4, above, admit explicit solutions. The solution v_i^{p+1} in point 3 is defined by

$$\frac{\rho}{\theta^k}(v_i^{p+1} - (u_i^{k+1} - d_i^k)) + \zeta\nabla^T(\nabla v_i^{p+1} - w_i^p - e_i^p) = 0 \tag{29}$$

where ∇^T is again the transpose of the discrete gradient $\nabla = (\partial_1, \partial_2)$, and consequently

$$v_i^{p+1} = (\frac{\rho}{\theta^k}I + \zeta\nabla^T\nabla)^{-1}\left(\frac{\rho}{\theta^k}(u_i^{k+1} - d_i^k) + \zeta\nabla^T(w_i^p + e_i^p)\right). \tag{30}$$

The solution ω_i^{p+1} in point 4 is defined by

$$w_i^{p+1} = \frac{\nabla v_i^{p+1} - e_i^p}{\|\nabla v_i^{p+1} - e_i^p\|_2}\max\left\{(\|\nabla v_i^{p+1} - e_i^p\|_2 - \frac{1}{\zeta}, 0\right\}, \tag{31}$$

as a direct consequence of a generalized shrinkage formula, see for instance [15], whose result we recall here.

Lemma 3.2. [15, Lemma 3.3] *For any α, $\beta > 0$ and $t \in \mathbb{R}^q$, the minimizer of*

$$\min_{s\in\mathbb{R}^q}\left\{\alpha\|s\|_2 + \frac{\beta}{2}\|s - t\|_2^2\right\} \tag{32}$$

is given by

$$s(t) = \max\left\{\|t\|_2 - \frac{\alpha}{\beta}, 0\right\}\frac{t}{\|t\|_2} \tag{33}$$

where we follow the convention $0.(0/0) = 0$.

3.3.3. *Pesudocode for the solution of models* (6) *and* (17). We can now summarize the proposed numerical method, whose steps are the following :

(1) Set $k = 0$, choose $\rho > 0$, v^0 and d^0.
(2) **repeat**
(3) $u^{k+1} \in argmin_u \left(\frac{\beta}{2}\|\mathcal{A}u - \bar{f}\|_2^2 + \frac{\alpha}{2}\|\overrightarrow{\nabla}u\|_2^2 + \frac{\rho}{2}\|u - v^k - d^k\|_2^2\right)$
$[\alpha = 0$, in the case of model (17)$]$
(4) $v^{k+1} \in argmin_u \left(\mu\|\overrightarrow{\nabla}v\|_1 + \frac{\rho}{2}\|u^{k+1} - v - d^k\|_2^2\right)$
(a) Set $p = 0$, choose $\zeta > 0$, ω_i^0 and e_i^0.
(b) **repeat**

and of (17) is

$$\min_u E(u), \quad \text{with} \quad E(u) = \frac{\beta}{2}\|\mathcal{A}u - \bar{f}\|_2^2 + \mu\|\overrightarrow{\nabla}u\|_1, \tag{21}$$

where,

$$\begin{aligned}
\|\mathcal{A}u - \bar{f}\|_2^2 &= \sum_{i=1}^m\sum_{j\in\Omega}(\mathcal{A}_i u_i - \bar{f}_i)_j^2,\\
\|\overrightarrow{\nabla}u\|_2^2 &= \sum_{i=1}^m\|\nabla u_i\|_2^2 = \sum_{i=1}^m\sum_{j\in\Omega}\left((\partial_1 u_i)^2 + (\partial_2 u_i)^2\right)_j,\\
\|\overrightarrow{\nabla}u\|_1 &= \sqrt{\sum_{i=1}^m\|\nabla u_i\|_1^2} = \left(\sum_{i=1}^m\left(\sum_{j\in\Omega}\sqrt{(\partial_1 u_i)_j^2 + (\partial_2 u_i)_j^2}\right)^2\right)^{\frac{1}{2}},
\end{aligned} \tag{22}$$

j is a general grid point in the pixel domain Ω, $\nabla = (\partial_1, \partial_2)$ is the discrete gradient operator defined by using backward finite differences (with periodic boundary conditions), and $\|\overrightarrow{\nabla}u\|_1$ is the discrete extended total variational $TV^{ext}(u)$, that follows the usual discretization of the $TV(u_i)$ semi-norm, denoted by $\|\nabla u_i\|_1$ (see (2)).

Then rewriting (20) in the form (19), it reads as

$$\begin{cases} f_1(u) &= \frac{\beta}{2}\|\mathcal{A}u - \bar{f}\|_2^2 + \frac{\alpha}{2}\|\overrightarrow{\nabla}u\|_2^2\\ f_2(u) &= \mu\|\overrightarrow{\nabla}u\|_1 \qquad \text{[discrete problem (6)]},\\ G &= I \end{cases} \tag{23}$$

and doing the same for (21), it reads as

$$\begin{cases} f_1(u) &= \frac{\beta}{2}\|\mathcal{A}u - \bar{f}\|_2^2\\ f_2(u) &= \mu\|\overrightarrow{\nabla}u\|_1 \qquad \text{[discrete problem (17)]}.\\ G &= I \end{cases} \tag{24}$$

3.3. Algorithm Proposed for the Solutions of (6) **and** (17)**.** It essentially consists in SALSA algorithm, but with suitable modifications of some of its steps, which are explained hereafter.

Algorithm SALSA for problems (23)/(24) *(equivalently, discrete problem* (6)/(17)*)*:

(1) Set $k=0$, choose $\rho>0$, v^0 and d^0.
(2) **repeat**
(3) $u^{k+1} \in argmin_u \left(\frac{\beta}{2}\|\mathcal{A}u-\bar{f}\|_2^2+\frac{\alpha}{2}\|\overrightarrow{\nabla}u\|_2^2+\frac{\rho}{2}\|u-v^k-d^k\|_2^2\right)$
[in the case of model (24), or equivalently (17), $\alpha=0$]
(4) $v^{k+1} \in argmin_v \left(\mu\,\|\overrightarrow{\nabla}v\|_1+\frac{\rho}{2}\|u^{k+1}-v-d^k\|_2^2\right)$
(5) $d^{k+1}=d^k-(u^{k+1}-v^{k+1})$
(6) $k \leftarrow k+1$
(7) **until** stopping criterion is satisfied.

In order to solve point 3 and point 4 in algorithm SALSA, either for problem (23) or (24), we formally compute the Gâteaux derivative of the associated functionals.

This leads to the following optimality conditions,

$$\begin{aligned}
&\textit{for point 3:}\quad (\beta\mathcal{A}^T\mathcal{A}-\alpha\overrightarrow{\Delta}+\rho I)u^{k+1}=\beta\mathcal{A}^T\bar{f}+\rho(v^k+d^k),\\
&\qquad\text{[in the case of model (24), or equivalently (17), } \alpha=0]\\
&\textit{for point 4:}\qquad \textit{for } \ i=1,\dots,m,\\
&\mu\frac{\sum_{l=1}^m\|\nabla v_l^{k+1}\|_1}{\|\overrightarrow{\nabla}v^{k+1}\|_1}\,\nabla^T\left(\frac{\nabla v_i^{k+1}}{\|\nabla v_i^{k+1}\|_2}\right)+\rho\Big(v_i^{k+1}-(u_i^{k+1}-d_i^k)\Big)=0.
\end{aligned} \tag{25}$$

where $\mathcal{A}^T$ is the transpose of $\mathcal{A}$, I is the identity, ∇^T is the transpose of the discrete gradient $\nabla=(\partial_1,\partial_2)$, and $\overrightarrow{\Delta}u=(\Delta u_1,\dots,\Delta u_m)$ represents the Laplace operator, with $\Delta=-(\partial_1^T\partial_1+\partial_2^T\partial_2)$.

3.3.1. *Numerical Solution for Point 3 of Algorithm SALSA.* Concerning problem (23), or equivalently (6), in order to determine u^{k+1} we apply the Gauss-Seidel iterative method, which is convergent because of the properties of the matrix $(\beta\mathcal{A}^T\mathcal{A}-\alpha\overrightarrow{\Delta}+\rho I)$, that is symmetric and positive definite (cf. Ostrowski-Reich theorem, for instance in [8, Théorème 5.3-2, page 103]).

Regarding problem (24), or equivalently (17), the same theorem [8, Théorème 5.3-2, page 103] also applies to $(\beta\mathcal{A}^T\mathcal{A}+\rho I)$.

3.3.2. *Numerical Solution for Point 4 of Algorithm SALSA.* Point 4 of the algorithm SALSA is equal for both problems (23) and (24) (that is, for problems (6) and (17)). To solve it we use a numerical procedure, that consists firstly in replacing, the unknown v^{k+1} by the previous iterate v^k in the first term (it then acts as a constant, denoted by θ^k, in the current iterate $k+1$). That is, for $i=1,\dots,m$,

$$\begin{aligned}
&\theta^k\,\nabla^T\left(\frac{\nabla v_i^{k+1}}{\|\nabla v_i^{k+1}\|_2}\right)+\rho\Big(v_i^{k+1}-(u_i^{k+1}-d_i^k)\Big)=0,\\
&\text{with}\qquad \theta^k=\mu\frac{\sum_{l=1}^m\|\nabla v_l^k\|_1}{\|\overrightarrow{\nabla}v^k\|_1}\quad \text{a constant.}
\end{aligned} \tag{26}$$

Then, by remarking that the above equation, multiplied by $\frac{1}{\theta^k}$, is exactly the optimality condition of the following minimization problem

$$\min_{v_i}\left(\frac{\rho}{2\theta^k}\|v_i-(u_i^{k+1}-d_i^k)\|_2^2+\|\nabla v_i\|_2\right),\quad \text{for } i=1,\dots,m, \tag{27}$$

and by doing a variable splitting, that is by creating a new variable $\omega_i=\nabla v_i$, we are led to the following equivalent constrained problem

$$\begin{aligned}
&\min_{v_i}\left(\frac{\rho}{2\theta^k}\|v_i-(u_i^{k+1}-d_i^k)\|_2^2+\|\nabla v_i\|_2\right),\quad \text{for } i=1,\dots,m,\\
&\text{subject to } \nabla v_i=\omega_i.
\end{aligned} \tag{28}$$

Therefore, the determination for each $i=1,\dots,m$, of v_i^{k+1} in (27) can be done by applying ADMM to (28). It works as follows:

Algorithm ADMM for (28) *(equivalently point 4 in SALSA)*

(1) Set $p=0$, choose $\zeta>0$, ω_i^0 and e_i^0.
(2) **repeat**
(3) $v_i^{p+1} \in argmin_{v_i}\left(\frac{\rho}{2\theta^k}\|v_i-(u_i^{k+1}-d_i^k)\|_2^2+\frac{\zeta}{2}\|\nabla v_i-w_i^p-e_i^p\|_2^2\right)$
(4) $\omega_i^{p+1} \in argmin_{\omega_i}\left(\|\nabla\omega_i\|_2+\frac{\zeta}{2}\|\nabla v_i^{p+1}-\omega_i-e_i^p\|_2^2\right)$
(5) $e_i^{p+1}=e_i^k-(\nabla v_i^{p+1}-\omega_i^{p+1})$
(6) $p \leftarrow p+1$
(7) **until** stopping criterion is satisfied.