HAL
open science

Simultaneous deconvolution and denoising using a second order variational approach applied to image super resolution

Amine Laghrib, Mahmoud Ezzaki, Mohammed El Rhabi, Abdelilah Hakim, Pascal Monasse, Said Raghay

To cite this version:

Amine Laghrib, Mahmoud Ezzaki, Mohammed El Rhabi, Abdelilah Hakim, Pascal Monasse, et al.. Simultaneous deconvolution and denoising using a second order variational approach applied to image super resolution. Computer Vision and Image Understanding, 2017, 10.1016/j.cviu.2017.08.007 . hal-01616743

HAL Id: hal-01616743

https://enpc.hal.science/hal-01616743v1

Submitted on 14 Oct 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Noname manuscript No.
(will be inserted by the editor)

Simultaneous deconvolution and denoising using a second order variational approach applied to image super resolution

Amine Laghrib · Mahmoud Ezzaki · Mohammed El Rhabi · Abdelilah Hakim · Pascal Monasse · Said Raghay.

Received: date / Accepted: date

Abstract The aim of a Super Resolution (SR) technique is to construct a high-resolution image from a sequence of observed low-resolution ones of the same scene. One major roadblock of an SR reconstitution is removing noise and blur without destroying edges. We propose a novel multiframe image SR algorithm based on a convex combination of Bilateral Total Variation and a non-smooth second order variational regularization, using a controlled weighting parameter. We prove the existence of a minimizer of the proposed SR model in the space of functions of bounded Hessian, and we confirm the success of this approach in avoiding undesirable artifacts. The simulation results show the efficient performance of the proposed algorithm compared to other methods in the literature using two criteria, PSNR and SSIM.

Keywords Multiframe Super Resolution · Bilateral TV filter · Bounded Hessian space · Second order regularization · Relaxed function.

1 Introduction

Image super resolution (SR) reconstruction is a challenging problem and an active research area in image processing. The principle of this technique is to reconstruct a high-resolution (HR) image by fusing degraded low-resolution (LR) ones. The SR is used in many applications, such as medical diagnostics [1], video surveillance [2], satellite imaging, etc.

After the first seminal work proposed by Tsai and al. [3], several techniques of the muliframe SR [4–6] have been proposed to enhance the quality of the obtained HR image. Since the SR approach is an ill-posed problem, many SR methods are based on a regularization framework [2,7,6,8]. Farsiu et al. proposed a fast and robust SR [9]; this technique reduces the complexity of the SR method by decomposing it into two steps: first, finding a blurred HR image from the LR measurements, then, deblurring and denoising it. Both steps are critical to the quality of the HR image. We concentrate here on the second one; we impose some prior on the HR image in a Bayesien framework. Some of the widely-used prior functions are Tikhonov-type regularizers [6, 10] and Total Variation (TV)-type regularizers [11]. Another successful regularization is the bilateral total variation (BTV) with the L^1 norm proposed by Farsiu et al. [9] and an adaptive norm with BTV regularizer [12]. These approaches are successful in recovering images with sharp edges but fail on images with smooth surface, suffering from the staircasing effect. To overcome this defect, a second order variational approach [13–15] can be used. Despite its success in numerous denoising problems, it suffers from blur and sometimes does not preserve edges correctly. To perform the process of simultaneous deblurring and denoising, a combined first and second order regularization [16,17] has been used and proved its robustness in image restoration, but no such attempts have been proposed to increase the performance of SR methods with the BTV regularizer. An alternative to BTV is the total generalized

Amine Laghrib, Abdelilah Hakim, Said Raghay
LAMAI, Faculty of Science and Technology, University Cadi Ayyad, Morocco.
E-mail: laghrib.amine@gmai.com

Mahmoud Ezzaki, Mohammed El Rhabi
Applied Mathematics and Computer Science department, Ecole des Ponts ParisTech (ENPC), Paris, France.

Pascal Monasse
Université Paris-Est, LIGM (UMR CNRS 8049), ENPC, F-77455 Marne-la-Vallée, France.

variation (TGV) [18], TGV^2_α, known for its effectiveness in avoiding the staircasing effect near discontinuity regions. However, Papafitsoros and Schönlieb [17] demonstrate that the combined first and second order regularization yields sometimes better results than TGV^2_α at a much lower computation cost. The best performing regularization terms are learned from training data in the work of Chen et al. [19,20]. Nevertheless, the mathematical well-posedness of the problems using these priors is not ascertained, and we did not investigate those. Motivated by the success of the combination models in restoration problem, we propose in this paper a new combination of the BTV regularizer and a second order differential regularization in the choice of the prior function.

The main idea behind this combination is to use a large weight for the BTV regularization to enhance sharp edges and also to correct the misregistration errors, and use the second order regularizer to eliminate the staircasing caused by the BTV regularization without blurring the image. In the following, we will prove existence of the minimization deblurring and denoising problem (second step of SR), which is a combination of the BTV and second order variational regularization. We use as functional space the bounded Hessian space $HB(\Omega)$ [21] and we use the classical steps based on relaxed functions [22,23].

In Section 2, we present the general SR problem and describe the proposed SR algorithm steps. In Section 3, we introduce the variational problem and the space $HB(\Omega)$ and its essential properties. Hence, we will prove the existence of a minimum of the relaxed functional associated to the deblurring and denoising problem, using standard techniques from calculus of variations. In Section 4, we present some experimental results using both synthetic and real images, and we compare our model with some existing algorithms in the SR literature.

2 Problem formulation

The observed images of a real scene are usually in low resolution due to some degradation operators. In practice, the acquired image is corrupted by noise, blur and sampling [24,3,5]. We assume that the LR images are taken under the same environmental conditions using the same sensor. The relationship between an ideal HR image X of size $M = rN_1 \times rN_2$ denoted by a column vector $X = [x_1, x_2 \ldots, x_M]^\top$, where $r \geq 1$ is the downsampling factor and the LR images Y_k of size $N_1 \times N_2$, represented also by a column vector $Y_k = [y_{k,1}, y_{k,2} \ldots, y_{k,N_1N_2}]^\top$, where $k = 1 \ldots n$ and n is the number of the LR images, is given by the relation

$$Y_k = DF_kHX + e_k, \qquad \forall k = 1, 2, \ldots, n, \tag{1}$$

where H represents the linear blur operator of kernel size $M \times M$, D represents the sampling operator of size $N_1N_2 \times M$, F_k is a geometric warp matrix of size $M \times M$ representing a non-parametric transformation that differs in all frames, and e_k a vector of size N_1N_2 represents the additive noise for each image, assumed to follow a zero mean Gaussian distribution.

In the presence of different operators of degradation (sampling, blur and noise), the problem becomes very unstable. To deal with it, we use the same approach as in [9] that suggests to separate it into three steps:

1. Computing the warp matrix F_k for each image.
2. Fusing the low-resolution images Y_k into a blurred HR version $B = HX$.
3. Finding the estimation of the HR image X from B.

To compute the warp matrix F_k, there are many approaches in the literature (for more details see [24]). In this paper, we use a non parametric registration [25]. For the fusion step, we use the technique of [9], recalled in the following section.

2.1 The fusion step

The first part of our algorithm is to compute the blurred HR version $B = HX$. We assume that the additive noise is Gaussian distributed and follows the same distribution for all low resolution images. The blurred image $\widehat{B}$ can be found via the principle of maximum likelihood estimator (ML) [24]. The ML suggests the choice of $\widehat{B}$ that maximizes the likelihood function, which also minimizes the negative log-likelihood function

$$\begin{aligned}\widehat{B} &= \arg\max_B \{p(Y_k|B)\} \\ &= \arg\min_B \{-\log(p(Y_k|B))\} \\ &= \arg\min_B \sum_{k=1}^{n} \|Y_k - DF_kB\|^2_{L^2}.\end{aligned} \tag{2}$$

The steepest descent algorithm [26] can be used to resolve this minimization problem, or, even better, the primal-dual algorithm [27].

2.2 Deconvolution and denoising step

In this step we try to find the HR image X by deblurring and denoising the image $\widehat{B}$. Unfortunately we are facing an unstable inverse problem due to the presence of noise and blur at the same time. To overcome this difficulty,

we impose some prior knowledge about the image in a Bayesian framework. Since $\widehat{B}$ has white Gaussian noise, the measured vectors Y_k also undergo a Gaussian blur. Via the Bayes rule, finding the HR image X leads us to look for the solution of the following minimization problem (3) using the maximum a posteriori (MAP) estimator:

$$\begin{aligned}\widehat{X}_{MAP} &= \arg\max_X \{p(X|\widehat{B})\} \\ &= \arg\max_X \left\{\frac{p(\widehat{B}|X)p(X)}{p(\widehat{B})}\right\} \\ &= \arg\min_X \left\{-\log(p(\widehat{B}|X)) - \log(p(X))\right\}, \end{aligned} \tag{3}$$

where $p(\widehat{B}|X)$ represents the likelihood term (data attachment term) and $p(X)$ denotes the prior knowledge on the HR image. To formulate precisely this problem, we need to assume a prior Gibbs function (PGF).

2.3 The prior Gibbs function

A well-known manner to represent the image prior PGF function p is the Gibbs function represented by

$$p(x) = c_G \cdot \exp\{-G(x)\}, \tag{4}$$

where c_G is a normalization constant and $G(x)$ is a non negative energy function. Since we know the robustness of the BTV regularizer to remove noise and preserve edges, and the success of the variational second order regularization to remove the blocky effect, we propose to use a combination of these two regularizers in the choice of the Gibbs function, controlled by a weight parameter that depends on the gradient of the image:

$$\begin{aligned}p(X) = &\exp\left\{-\delta \sum_{i=-p}^{p}\sum_{j=-p}^{p} \alpha^{|i|+|j|}\|X - S_x^i S_y^j X\|_1\right\} \\ &\cdot \exp\left\{-(1-\delta)\|f(\nabla^2 X)\|_1\right\},\end{aligned}$$

where the operators S_x^i and S_y^j shift X by i and j pixels in horizontal and vertical directions respectively:

$$S_x^i X(x,y) = X(x+i,y) \text{ and } S_y^j X(x,y) = X(x,y+j).$$

This formulation depends on the following parameters:

- α: a scalar weight, applied to give a spatially decaying effect to the summation of the regularization terms, $0 < \alpha < 1$.
- p: the spatial window size, $p \geq 1$.
- f: a function defined from $\mathbb{R}^4$ to $\mathbb{R}^+$ verifying some assumptions that we define below.
- δ: the weight that controls the regularization combination. We define in the following how it is calculated from the gradient of the image.

The first term is a measure of bilateral variation, containing a spatial decaying influence and a color difference in the neighborhood.

Notice that in the discrete setting, the BTV regularizer

$$\sum_{i=-p}^{p}\sum_{j=-p}^{p} \alpha^{|i|+|j|}\|X - S_x^i S_y^j X\|_1$$

can be translated into

$$\sum_{i=-p}^{p}\sum_{j=-p}^{p} \omega_{i,j} \sum_{k=1}^{M} |(A_{i,j}X)_k|,$$

with weights $\omega_{i,j} = \alpha^{|i|+|j|}$; the linear operators $A_{i,j} = I - S_x^i S_y^j$ correspond to convolutions with a linear kernel. Under this point of view, the BTV regularization is a simple version of the learned regularizers of Chen et al. [20], but with the ℓ^1 norm.

3 Resolution of the final SR problem

In this section, we rewrite the problem (3) with the defined PGF function and the likelihood terms, which will constitute the final problem. To do so, we use the operators F_k defined for each image and the computed blurred HR image $\widehat{B}$ obtained in (2).

$$\widehat{X} = \arg\min_X F(X), \tag{5}$$

with

$$\begin{aligned}F(X) = &\|HX - \widehat{B}\|_1 \\ &+ \delta \sum_{i=-p}^{p}\sum_{j=-p}^{p} \alpha^{|i|+|j|}\|X - S_x^i S_y^j X\|_1 \\ &+ (1-\delta)\|f(\nabla^2 X)\|_1.\end{aligned} \tag{6}$$

The norm of the Lebesgue space $L^1(\Omega)$ in the expression $\|HX - \widehat{B}\|_1$, where Ω contains all the pixels on the HR grid X, is used because it is robust against outliers [9]. The choice of L^1 instead of the L^2 norm is essentially due to the better effectiveness of the former in removing impulse noise. Also, it is contrast invariant in a weak sense: if $C \in \mathbb{R}^+$, CX is a solution for $C\hat{B}$ in (6). Before solving the minimization problem (5), we have to check the existence of a solution. In the following, we construct the functional framework and we ensure the existence of a solution.

3.1 Functional framework

We consider a compact set $\Omega \subset \mathbb{R}^2$ smooth enough. We construct the functional framework where the function F is well defined. In other words, we look for a Banach space of functions X defined from Ω to $\mathbb{R}$, which ensures the existence of the different norms in (6). We suppose that $X \in L^1(\Omega)$; this ensures that $\|HX(x) - \widehat{B}(x)\|_1$ is well defined. Moreover, the operator $(I - S_x^{\cdot} S_y^{\cdot})$ is defined such that $I - S_x^{\cdot} S_y^{\cdot} : L^1(\Omega) \longrightarrow L^1(\Omega)$ which implies the existence of $\|X - S_x^{\cdot} S_y^{\cdot} X\|_1$. Moreover, we assume that f is convex and satisfies the following condition:

$$\exists k, k' > 0,\ \forall x \in \mathbb{R}^4,\ k|x|_2 - k' \leq f(x) \leq k|x|_2 + k', \quad (7)$$

where $|\,.\,|_2$ is the Euclidean norm in $\mathbb{R}^4$. The condition (7) ensures that if $\nabla^2 X \in \left[L^1(\Omega)\right]^4$ then $\|f(\nabla^2 X)\|_1$ is well defined. From the conditions $X \in L^1(\Omega)$, $(X - S_x^{\cdot} S_y^{\cdot} X) \in L^1(\Omega)$ and $\nabla^2 X \in \left[L^1(\Omega)\right]^4$, we recognize as an adequate functional framework, for $q = 1$, the Sobolev space [28]

$$W^{2,q} = \{X \in L^q(\Omega) \,|\, D^\alpha X \in L^q(\Omega), [\alpha] \leq 2\}, \quad (8)$$

where

$$D^\alpha = \frac{\partial^{\alpha_1}\partial^{\alpha_2}}{\partial x^{\alpha_1}\partial y^{\alpha_2}} \text{ and } [\alpha] = \alpha_1 + \alpha_2.$$

Unfortunately, this space is not reflexive [28] and we cannot apply directly the following theorem of optimization:

Theorem 1 *We consider a reflexive space V, and a function $F : V \mapsto \mathbb{R}$ such that F is coercive, lower semi-continuous (l.s.c) and convex. Then the minimum of F exists. Moreover if F is strictly convex, then this minimum is unique.*

We cannot say anything about a bounded minimizing sequence in $W^{2,1}(\Omega)$. To overcome the ill-posedness of this problem, we use the procedure of relaxation [29]. A typical choice of the space that guarantees the compactness results is the space of Bounded Hessian $HB(\Omega)$ noted also $BV^2(\Omega)$, introduced for the first time by F. Demengel in 1985 [21]. We present next the space $HB(\Omega)$ and its useful properties.

3.2 The space $HB(\Omega)$

We define the space $HB(\Omega)$ of bounded Hessian functions. First, we define the space of bounded variation functions $BV(\Omega)$ as

$$BV(\Omega) = \left\{X \in L^1(\Omega) \text{ such that } \nabla X \in \mathcal{M}_b(\Omega, \mathbb{R}^2)\right\},$$

where $\mathcal{M}_b(\Omega, \mathbb{R}^2)$ is the space of bounded measures from Ω to $\mathbb{R}^2$. We define the space $HB(\Omega)$ using the space $BV(\Omega)$ by

$$HB(\Omega) = \left\{X \in W^{1,1}(\Omega), \nabla X \in (BV(\Omega))^2\right\} \quad (10)$$

It is clear that $W^{2,1}(\Omega) \subset HB(\Omega)$. We define now a norm of the space $HB(\Omega)$.

Theorem 2 *The space $HB(\Omega)$ is a Banach space, when endowed with the norm $\|.\|_{HB}$ defined as*

$$\|X\|_{HB} = \|X\|_{W^{1,1}} + |DX|,$$

where

$$\|X\|_{W^{1,1}} = \|X\|_1 + \|\nabla X\|_1,$$

and

$$|DX| = \sup\{\int_\Omega X \operatorname{div}^2(\phi) : \phi \in \mathcal{C}_0^1(\Omega)^2, \|\phi\|_{L^\infty(\Omega)} \leq 1\};$$

div^2 *is the second order divergence operator with the adjointness property*

$$\operatorname{div}^2 X \cdot Y = X \cdot \nabla^2 Y \quad \forall Y \in \mathbb{R}^M,\ X \in \left(\mathbb{R}^M\right)^4,$$

and the dot "$\cdot$" denotes the Euclidean inner product. $\mathcal{C}_0^1(\Omega)$ represents the space of continuously differentiable functions with compact support in Ω.

We have also the embedding results proved by Demengel [21] as well:

Theorem 3 *Assuming that the dimension $n > 1$, we have*

$$HB(\Omega) \hookrightarrow W^{1,q}(\Omega), \text{ when } q \leq \frac{n}{n-1},$$

with continuous embedding. Moreover the embedding is compact if $q < \dfrac{n}{n-1}$. In particular, we have

$$HB(\Omega) \hookrightarrow L^q(\Omega), \text{ for } q \leq \frac{n}{n-2}, \text{ if } n > 2,$$
$$HB(\Omega) \hookrightarrow L^q(\Omega), \text{ for } q \geq 1, \text{ if } n = 2.$$

Since in our problem we treat the case $n = 2$ and Ω is a subset of $\mathbb{R}^2$, we have in fact $HB(\Omega) \hookrightarrow W^{1,2}(\Omega)$ with continuous embedding and $HB(\Omega) \hookrightarrow W^{1,1}(\Omega)$ with compact embedding.

Definition 1 *(Weak * convergence in $HB(\Omega)$)*
Let $(X_k)_{k\in\mathbb{N}}$ and $X \in HB(\Omega)$. We say that (X_k) converges to X weakly if*

$$X_k \overset{L^1(\Omega)}{\longrightarrow} X \quad \text{and} \quad \nabla X_k \underset{BV-\omega*}{\rightharpoonup} \nabla X,$$

where $BV-\omega*$ is the weak topology in $BV(\Omega)$ defined as:*

$$X_k \underset{BV-\omega*}{\rightharpoonup} X \Longleftrightarrow \begin{cases} X_k \underset{L^1}{\longrightarrow} X \\ DX_k \underset{M}{\overset{*}{\rightharpoonup}} DX, \end{cases}$$

and $DX_k \underset{M}{\overset{*}{\rightharpoonup}} DX$ means

$$\int_\Omega \varphi\, DX_k \longrightarrow \int_\Omega \varphi\, DX \qquad \forall \varphi \in \mathcal{C}_0^1(\Omega).$$

We also need this compactness result in $BV(\Omega)$:

Theorem 4 *(Compactness in $BV(\Omega)$)*
Every uniformly bounded sequence $(X_k)_{k\in\mathbb{N}}$ in $BV(\Omega)$ is relatively compact in $L^q(\Omega)$, for $1 \le q < \frac{n}{n-1}, n \ge 1$. Moreover, there exist a subsequence $(X_{kj})_{j\in\mathbb{N}}$ and $X \in BV(\Omega)$ such as $X_{kj} \underset{BV-\omega*}{\rightharpoonup} X$.

Another very useful theorem to demonstrate the l.s.c is the Buttazo and Freddi theorem defined as

Theorem 5 *Let Ω be an open subset of $\mathbb{R}^n$, ν, $(\nu_k)_{k\in\mathbb{N}}$ be $\mathbb{R}^m$-valued finite Radon measures and μ, $(\mu_k)_{k\in\mathbb{N}}$ be positive Radon measure in Ω. Let $g : \mathbb{R}^m \to \mathbb{R}$ be a convex function and suppose that $\mu_k \to \mu$ and $\nu_k \to \nu$ weakly* in Ω. Consider the Lebesgue decompositions:*

$$\nu = \left(\frac{\nu}{\mu}\right)\mu + \nu_s \quad , \quad \nu_k = \left(\frac{\nu_k}{\mu_k}\right)\mu_k + \nu_{s_k}, \quad k \in \mathbb{N}.$$

If $\mu = \mu_k = dx^n$, then

$$g(\nu)(\Omega) \le \liminf_{k\to\infty} g(\nu_k)(\Omega).$$

3.3 The Variational Problem

In this subsection, we prove the existence of a solution to the problem (5) by using the direct method in calculus of variations.

Theorem 6 *Under the hypothesis (7), and assuming that the operators $(I - S_x^{\cdot} S_y^{\cdot})$ and H defined: $L^1(\Omega) \longrightarrow L^1(\Omega)$ are continuous; assuming also that H does not annihilate the constants (in particular $H.1 \neq 0$); then, the optimization problem*

$$\min_X \overline{F}_{ext}(X), \tag{11}$$

has a solution in $HB(\Omega)$, where $\overline{F}_{ext}$ is the lower semi-continuous envelope of F defined in the proof.

Proof First of all, we extend the function F in $HB(\Omega)$ by :

$$F_{ext}(X) = \begin{cases} F(X) & \text{if } X \in W^{2,1}(\Omega) \\ +\infty & \text{if } X \in HB(\Omega) \setminus W^{2,1}(\Omega) \end{cases} \tag{12}$$

We can check easily that F_{ext} is not lower semi continuous for the weak* topology defined in $HB(\Omega)$ following the same steps as in [17], so we should determine its l.s.c envelope defined in the whole space $HB(\Omega)$ by

$$\begin{aligned} \overline{F}_{ext}(X) = &\|HX - \widehat{B}\|_1 \\ &+ \delta \sum_{i=-p}^{p} \sum_{j=-p}^{p} \alpha^{|i|+|j|} \|X - S_x^i S_y^j X\|_1 \\ &+ (1-\delta)\|f(D^2X)\|_1, \end{aligned} \tag{13}$$

where $D^2X = D\nabla X$ and D is the distributional gradient of X. With these notations $\overline{F}_{ext}$ coincides with F on $W^{2,1}(\Omega)$. Let us prove now the coercivity of $\overline{F}_{ext}$.

- Let $(X_n)_{n\in\mathbb{N}}$ be a minimizing sequence of the function F_{ext} such that

  $$\lim_{n\to+\infty} \overline{F}_{ext}(X_n) = \inf_{X\in HB(\Omega)} \overline{F}_{ext}(X).$$

  Using the hypothesis (7) on F_{ext}, we can deduce that

  $$\begin{cases} |D^2X_n|(\Omega) \le M_1, \\ \|HX_n - \widehat{B}\|_1 \le M_2, \\ \sum_{i=-p}^{p}\sum_{j=-p}^{p} \alpha^{|i|+|j|}\|X_n - S_x^i S_y^j X_n\|_1 \le M_3. \end{cases} \tag{14}$$

  We prove now that $\|X_n\|_1$ is also bounded. We use a classical approach [30]. We construct two sequences $Y_n = \frac{1}{|\Omega|}\int_\Omega X_n\, dx$ and $Z_n = X_n - Y_n$; then

  $$\int_\Omega Z_n\, dx = 0, \quad \text{and} \quad \nabla Z_n = \nabla X_n. \tag{15}$$

  Using the generalized Poincaré-Wirtinger inequality [28] for the norm $\|.\|_2$ of the Lebesgue space $L^2(\Omega)$, there is a universal constant C that we will use in the following, such that

  $$\|Z_n\|_2 \le C\|\nabla Z_n\|_2. \tag{16}$$

  Based on the approximation given to the TV prior suggested in [31], we can deduce that

  $$\|\nabla X_n\|_1 \simeq \|Q_x X_n\|_1 + \|Q_y X_n\|_1,$$

  This equivalence is obtained computationally, where the operators Q_x and Q_y can be defined in the regularization BTV such that $i = 1, j = 0$ and $i = 0, j = 1$ respectively, which is equivalent to the choice $Q_x = (I - S_x^1)$ and $Q_y = (I - S_y^1)$. Since we

choose $p \geq 1$ in the BTV regulizer, we can deduce that

$$\|\nabla X_n\|_1 \leq \sum_{i=-p}^{p} \sum_{j=-p}^{p} \alpha^{|i|+|j|} \|X_n - S_x^i S_y^j X_n\|_1. \quad (17)$$

Using the third inequality in (14) and relations (15), (16), we have

$$\|Z_n\|_2 \leq C, \quad (18)$$

then

$$\begin{aligned} \|X_n\|_2 &= \|X_n - Y_n + Y_n\|_2 \\ &= \|Z_n + Y_n\|_2 \\ &\leq \|Z_n\|_2 + \|Y_n\|_2 \\ &\leq C + \|Y_n\|_2 \end{aligned} \quad (19)$$

with $\|Y_n\|_{L^2(\Omega)} = \| \int_\Omega X_n \, dx\|$. We also have

$$\begin{aligned} \left\| H\left(\frac{1}{|\Omega|}\int_\Omega X_n \, dx\right)\right\|_1 &\leq \|HY_n - HX_n\|_1 \\ &+ \|HX_n - \widehat{B}\|_1 + \|\widehat{B}\|_1 \\ &\leq \|H\|_{L^\infty(\Omega)} \|Z_n\|_1 + C + \|\widehat{B}\|_1 \\ &\leq \|H\|_{L^\infty(\Omega)} \|Z_n\|_2 + C \\ &\leq C, \end{aligned} \quad (20)$$

and finally

$$\begin{aligned} \left\| H\left(\frac{1}{|\Omega|}\int_\Omega X_n \, dx\right)\right\|_1 &= \left|\frac{1}{|\Omega|}\int_\Omega X_n \, dx\right| \|H.1\|_1 \\ &\leq C. \end{aligned} \quad (21)$$

We can deduce that the sequence $(X_n)_{n\in\mathbb{N}}$ is bounded in $L^2(\Omega)$, and Ω is bounded so it is also bounded in $L^1(\Omega)$. In addition we have that, $\forall X_n \in HB(\Omega)$:

$$\int_\Omega |\nabla X_n| \, dx \leq C_1 |D^2 X_n|(\Omega) + C_2 \|X_n\|_1. \quad (22)$$

Since $(X_n)_{n\in\mathbb{N}}$ is bounded in $L^1(\Omega)$, $(\nabla X_n)_{n\in\mathbb{N}}$ is bounded in $L^1(\Omega)$ and $(D^2 X_n)_{n\in\mathbb{N}}$ is bounded in $L^1(\Omega)$, we deduce finally that (X_n) is bounded in $HB(\Omega)$. Using the embedding results in Theorem 3, the sequence $(X_n)_{n\in\mathbb{N}}$ is also bounded in $W^{1,1}(\Omega)$ and $HB(\Omega) \hookrightarrow W^{1,1}(\Omega)$ compactly. Thus, we can extract a subsequence $(X_{n_k})_{k\in\mathbb{N}}$ such that

$$X_{n_k} \xrightarrow{W^{1,1}(\Omega)} X.$$

In addition $(\nabla X_n)_{n\in\mathbb{N}}$ is also bounded in $[BV(\Omega)]^2$, so that, based on the result of compactness in theorem 4, we can extract a subsequence $(\nabla X_{n_k})_{k\in\mathbb{N}}$ satisfying

$$\nabla X_{n_k} \overset{w*}{\rightharpoonup} Y.$$

Using the weak* convergence in $BV(\Omega)$ we have

$$\begin{cases} \nabla X_{n_k} & \underset{L^1}{\longrightarrow} Y, \\ D\nabla X_{n_k} & \underset{M}{\overset{*}{\rightharpoonup}} DY, \end{cases}$$

where $D\nabla X_{n_k} \underset{M}{\overset{*}{\rightharpoonup}} DY$ is equivalent to

$$\int_\Omega \varphi \, D\nabla X_{n_k} \longrightarrow \int_\Omega \varphi \, DY \qquad \forall \varphi \in \mathcal{C}_0^1(\Omega).$$

Therefore $\nabla X = Y$, and we finally get

$$\begin{cases} X_{n_k} \longrightarrow X & \text{in } W^{1,1}(\Omega) \\ \nabla X_{n_k} \overset{w*}{\rightharpoonup} \nabla X & \text{in } [BV(\Omega)]^2 \end{cases}$$

$$\Longleftrightarrow X_{n_k} \overset{*}{\rightharpoonup} X \text{ in } HB(\Omega).$$

To finish this proof it remains to show that F_{ext} is weak sequentially l.s.c in $HB(\Omega)$.

- To prove that $\overline{F}_{ext}$ is weak sequentially l.s.c, it suffices to establish the following inequality

$$\liminf_{X_{n_k} \rightharpoonup X} \overline{F}_{ext}(X_n) \geq \overline{F}_{ext}(X).$$

Let X and $(X_n)_{n\in\mathbb{N}}$ be functions taken from $HB(\Omega)$ such that

$$X_n \underset{HB(\Omega)}{\overset{*}{\rightharpoonup}} X.$$

This implies

$$\|X_n - X\|_1 \longrightarrow 0.$$

Since H and $(I - S_x^i S_y^j)$ are continuous operators from L^1 to itself we can check that

$$\|HX_n - B\|_1 \longrightarrow \|HX - B\|_1,$$

and

$$\sum_{i=-p}^{p} \sum_{j=-p}^{p} \alpha^{|i|+|j|} \|X_n - S_x^i S_y^j X_n\|_1 \longrightarrow \sum_{i=-p}^{p} \sum_{j=-p}^{p} \alpha^{|i|+|j|} \|X - S_x^i S_y^j X\|_1. \quad (23)$$

Finally we use theorem 5 of Buttazo and Freddi

$$\nu_n = dx^2, \ \mu = D^2 X \quad \text{and} \quad \mu_n = D^2 X_n$$

to deduce that

$$\liminf_{X_n \rightharpoonup X} f(D^2 X_n)(\Omega) \geq f(D^2 X)(\Omega),$$

which gives

$$\liminf_{X_n \rightharpoonup X} \overline{F}_{ext}(X_n) \geq \overline{F}_{ext}(X),$$

i.e., X is the minimum of $\overline{F}_{ext}$ in $HB(\Omega)$.

We cannot say anything about the uniqueness of the solution since $\|HX_n - B\|_1$ is not strictly convex, but if we replace this term by $\|HX_n - B\|_2$, the solution is unique since the blur operator H is injective. In the next section, we define the proposed iterative SR algorithm.

3.4 Proposed SR algorithm

Let us describe briefly the minimization algorithm used in our SR approach. We can simply use the alternative approach proposed in [32], based on the discretization of the gradient descent PDE associated to the minimizing function defined in (6). Unfortunately this approach is very slow and may sometimes generate some errors, hence a faster and more consistent algorithm is needed. We use an efficient primal-dual algorithm [27], well adapted to our non-smooth convex optimization problem. Let us detail the proposed algorithm.

First, we define the function f in (6) using the second order differential operators noted ∇^2 as

$$f(\nabla^2 X) = \nabla_{xx}X + \nabla_{xy}X + \nabla_{xy}X + \nabla_{yy}X,$$

where

$$\nabla^2 X = \nabla_{xx}X + \nabla_{xy}X + \nabla_{xy}X + \nabla_{yy}X.$$

We start with the following notation:

$$K_1 = \delta \left.\begin{pmatrix} \alpha^{2p}(I - S_x^{-p}S_y^{-p}) \\ \alpha^{2p-1}(I - S_x^{-p}S_y^{-p+1}) \\ \vdots \\ \alpha^{2p}(I - S_x^{p}S_y^{p}) \end{pmatrix}\right\} (2p+1)^2, \tag{24}$$

$$K_2 = (1-\delta) \sum_{a,b\in\{x,y\}} \nabla_{ab}, \tag{25}$$

and

$$K = \left.\begin{pmatrix} K_1 \\ K_2 \end{pmatrix}\right\} (2p+1)^2 + 1, \tag{26}$$

and we define also the function $\mathtt{F}_1$ as follows

$$\begin{aligned} \mathtt{F}_1(KX) &= \|K_1X\|_1 + \|K_2X\|_1 \\ &= \delta \sum_{i=-p}^{p}\sum_{j=-p}^{p} \alpha^{|i|+|j|}\|(I - S_x^i S_y^j)X\|_1 \\ &+ (1-\delta)\,\|\sum_{a,b\in\{x,y\}} \nabla_{ab}X\|_1. \end{aligned} \tag{27}$$

Also, the function $\mathtt{F}_2$ is defined as

$$\mathtt{F}_2(\tilde{K}X) = \|HX - \widehat{B}\|_1, \tag{28}$$

where $\mathtt{F}_2(X) = \|X\|_1$. Using the notations above, the minimization problem (5) becomes

$$\widehat{X} = \arg\min_X \{\ \mathtt{F}_1(KX) + \mathtt{F}_2(\tilde{K}X)\}. \tag{29}$$

Now, we can apply the primal-dual algorithm [27] to minimize the general problem (29), where $\mathtt{F}_1$ and $\mathtt{F}_2$ are convex functions. $\tilde{K}$ is a linear operator defined from $L^1(\Omega) \to L^1(\Omega)$, and K is a linear operator defined from $L^1(\Omega) \to [L^1(\Omega)]^{(2p+1)^2+1}$. Thus, using the saddle point problem [27], we get the equivalent primal-dual problem

$$\inf_X \sup_{Y,Z} \{< KX, Y > -\mathtt{F}_1^*(Y) + < HX - \widehat{B}, Z > -\mathtt{F}_2^*(Z)\}, \tag{30}$$

where $Y = \begin{pmatrix} Y^1 \\ Y^2 \\ \vdots \\ Y^{(2p+1)^2+1} \end{pmatrix} \in [L^\infty(\Omega)]^{(2p+1)^2+1}$ and $Z \in L^\infty(\Omega)$ denoting the dual variables, $\mathtt{F}_1^*$ and $\mathtt{F}_2^*$ are the dual functions of respectively $\mathtt{F}_1$ and $\mathtt{F}_2$ defined in a similar way as

$$\mathtt{F}_i^*(Y) = \iota_P(Y) = \begin{cases} 0 & Y \in P \\ +\infty & Y \notin P \end{cases} \tag{31}$$

where $P = \{Y : \|Y\|_\infty \leq 1\}$, and $\|\cdot\|_\infty$ denotes the discrete maximum norm defined as

$$\|Y\|_\infty = \max_i |Y_i|.$$

For the primal-dual algorithm, we have to define the proximal operator functions $(I + \sigma\partial\mathtt{F}_1^*)^{-1}$ and $(I + \sigma\partial\mathtt{F}_2^*)^{-1}$ using the projection on a convex set P (noted Π_P) as follows

$$Y = (I + \sigma\partial\mathtt{F}_1^*)^{-1}(\widehat{Y}) = \Pi_P(\widehat{Y}), \tag{32}$$

and,

$$Z = (I + \sigma\partial\mathtt{F}_2^*)^{-1}(\widehat{Z}) = \Pi_P(\widehat{Z}), \tag{33}$$

where

$$\Pi_P(\widehat{W}_i) = \frac{\widehat{W}_i}{\max(\|\widehat{W}_i\|_\infty, 1)}.$$

T complete the primal-dual algorithm, we have to compute the resolvent operator $(I + \tau\partial\mathtt{G})^{-1}$ for the function $G(X) = 0$ (in our case) as follows

$$X = (I + \tau\partial G)^{-1}(\widehat{X}) = \widehat{X}. \tag{34}$$

Now, we can apply the iterations of the primal-dual algorithm [33] to our saddle-point problem (30), we obtain the following iterative algorithm:

$$\begin{cases} \widehat{Y}^{n+1} = (I + \sigma\partial \mathtt{F}_1^*)^{-1}(\widehat{Y}^n + \sigma K X^n), \\ \widehat{Z}^{n+1} = (I + \sigma\partial \mathtt{F}_2^*)^{-1}(\widehat{Z}^n + \sigma \tilde{K} X^n), \\ \widehat{X}^{n+1} = \widehat{X}^n - \tau(K^*\widehat{Y}^{n+1} + \tilde{K}^*\widehat{Z}^{n+1}), \\ \widehat{X}^{n+1} = 2\widehat{X}^{n+1} - \widehat{X}^n. \end{cases} \tag{35}$$

Using the expressions of $\mathtt{F}_1^*$ and $\mathtt{F}_2^*$, the solution of the iterative problem (35) is given as

$$\begin{cases} \widehat{Y}^{n+1} = \Pi_P(\widehat{Y}^n + \sigma K X^n), \\ \widehat{Z}^{n+1} = \Pi_P(\widehat{Z}^n + \sigma \tilde{K} X^n), \\ \widehat{X}^{n+1} = \widehat{X}^n - \tau(K^*\widehat{Y}^{n+1} + \tilde{K}^*\widehat{Z}^{n+1}), \\ \widehat{X}^{n+1} = 2\widehat{X}^{n+1} - \widehat{X}^n. \end{cases} \tag{36}$$

where Π_P is the projection on the convex set P defined above. We need also to define the operators K^* and $\tilde{K}^*$, which are respectively the adjoint of the operators K and $\tilde{K}$.

$$K^*Y = \sum_{k=1}^{(2p+1)^2} \left(\delta \sum_{i=-p}^{p} \sum_{j=-p}^{p} \alpha^{|i|+|j|}(I - S_y^{-j} S_x^{-i}) Y_k \right) + (1-\delta) \sum_{a,b\in\{x,y\}} \overline{\nabla_{ab}} Y_{(2p+1)^2+1}, \tag{37}$$

where the operators S_x^{-i} and S_y^{-j} define the transpose of matrices S_x^i and S_y^j respectively and have a shifting effect in the opposite directions. $\overline{\nabla_{ab}}$ is the adjoint operator of ∇_{ab} that we will define later in a discrete form. We also have

$$\tilde{K}^* = H^\intercal.$$

This provides all the ingredients needed to implement the primal-dual algorithm associated to the problem (29). Let us introduce the discrete setting.

We define the second order differential operators ∇_{xx}, ∇_{yy}, and ∇_{xy} using convolutions with the kernels

$$k_{xx} = \begin{pmatrix} 0 & 0 & 0 \\ 1 & -2 & 1 \\ 0 & 0 & 0 \end{pmatrix} \; k_{yy} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & -2 & 0 \\ 0 & 1 & 0 \end{pmatrix} \; k_{xy} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & -1 & 1 \end{pmatrix}$$

As a consequence, the matrix K_2 in (27) can be interpreted as the kernel k_2 defined as:

$$k_2 = k_{xx} + k_{yy} + 2k_{xy} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & -2 & -1 \\ 0 & -1 & 2 \end{pmatrix}.$$

Therefore, the result of K_2X is merely the convolution of X with the above linear kernel k_2. Moreover, the transpose of K_2^* is given by the convolution with the rotated kernel $\bar{k}_2$ in the case of periodic boundary condition exploited in this paper. Algorithm 1 is used to perform the deconvolution and denoising of the proposed SR method.

In this algorithm, we choose carefully the weighting parameter δ. Indeed, we propose a new manner to control this parameter, using a function depending on the gradient of the image. Since we know the success of the BTV regularization in preserving sharp edges and of the second variational approach in recovering smoother surfaces, the ideal choice of the weighting parameter is in the surrounding of 1 along edges; and to avoid the artifacts due to the BTV regularization, the choice of the weighting parameter δ is $0 \leq \delta < 1$ in smooth regions and small jumps. Thus, we use the following formula, which yields the parameter δ giving usually the best quality of the reconstitution HR image:

$$\delta = \begin{cases} \exp(-\frac{\|\nabla X\|_1^2}{2M^2}) & \text{if } \|\nabla X\|_1/M \leq T \\ \exp(-\frac{T^2}{2}) & \text{if } \|\nabla X\|_1/M > T \end{cases} \tag{38}$$

where T is the Otsu binarization threshold [34]. To approve this choice, we used several test images; while sometimes we find a better result for a hand tuned parameter, this parameter is very close to the controlled weighting parameters calculated by the function.

To compare the performance of the proposed algorithm to other regularization methods, we set up the following deconvolution and denoising experiment. First, we added a blur of $\sigma = 5 \times 3$, then, we added Gaussian white noise of zero mean and standard deviation $\sigma = 30$ to the original *Lena* image. Figure 1 shows the obtained image using different regularizations. We compare our method with TV [11], BTV [9], 2nd-order regularization [14], and also the combined first order and second order (TV + TV^2) regularizations [17]. We select the regularization parameters according to the best visual result. This example demonstrates that the proposed regularization functional not only produces sharp edges but also avoids the staircasing effect.

4 Experimental results

In this section, both synthetic and real images are tested to judge the performance of our proposed method, and compare it to some popular approaches. Some of the used benchmark images are listed in Figure 2. The chosen images are often used in image processing applications, since they differ in texture, smoothness and gray level histogram. To generate the LR images from these HR images, we follow the classical steps: First, we define the sub-pixel motion, after, we align the images using the registration process [35], in the third and fourth step we blur and decimate the image respectively and

Data: Choose τ, $\sigma > 0$, $(\widehat{Y}^0, \widehat{Z}^0, \widehat{X}^0) \in \mathbb{R}^M \times \mathbb{R}^M$ and set $\widehat{Y}^0 = \widehat{Z}^0 = \widehat{X}^0$.
Result: the restored HR image $\widehat{X}$
Iterate until negligible variation of $\widehat{X}^n$:
begin

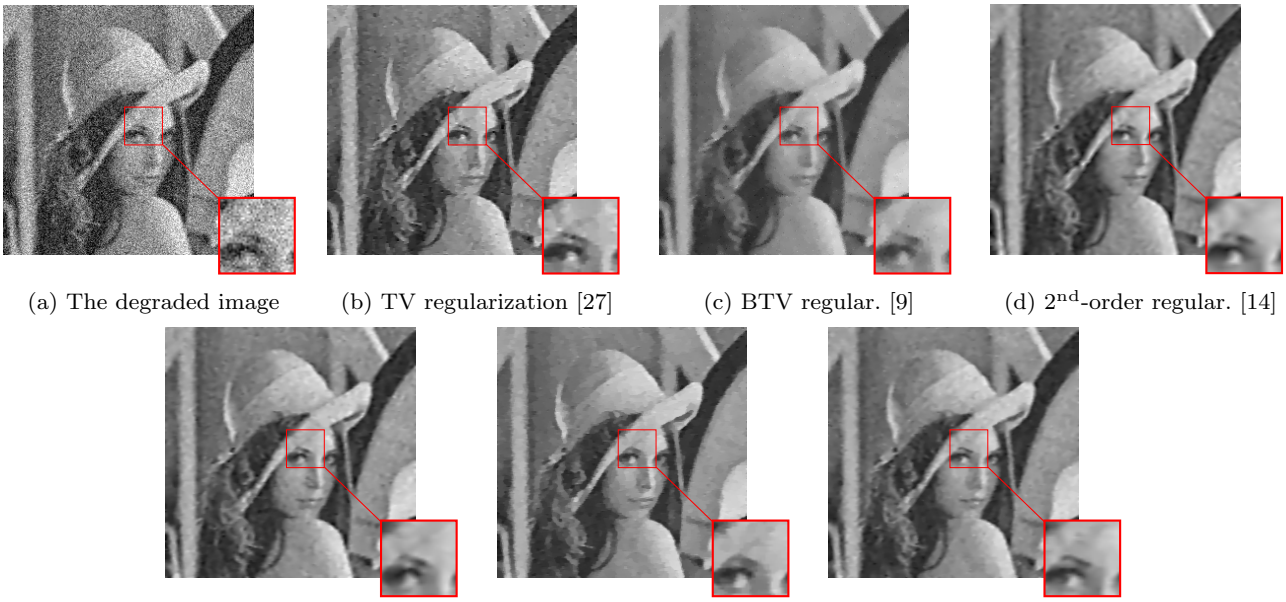
$\widehat{Y}^{n+1} = \Pi_P(\widehat{Y}^n + \sigma K X^n)$ // Use (32)

$\widehat{Z}^{n+1} = \Pi_P(\widehat{Z}^n + \sigma \tilde{K} X^n)$ // Use (33)

$\widehat{X}^{n+1} = \widehat{X}^n - \tau(K^* \widehat{Y}^{n+1} + \tilde{K}^* \widehat{Z}^{n+1})$ // Use (37)

$\widehat{X}^{n+1} = 2\widehat{X}^{n+1} - \widehat{X}^n$

Algorithm 1: Primal-Dual algorithm



(a) The degraded image (b) TV regularization [27] (c) BTV regular. [9] (d) 2nd-order regular. [14]

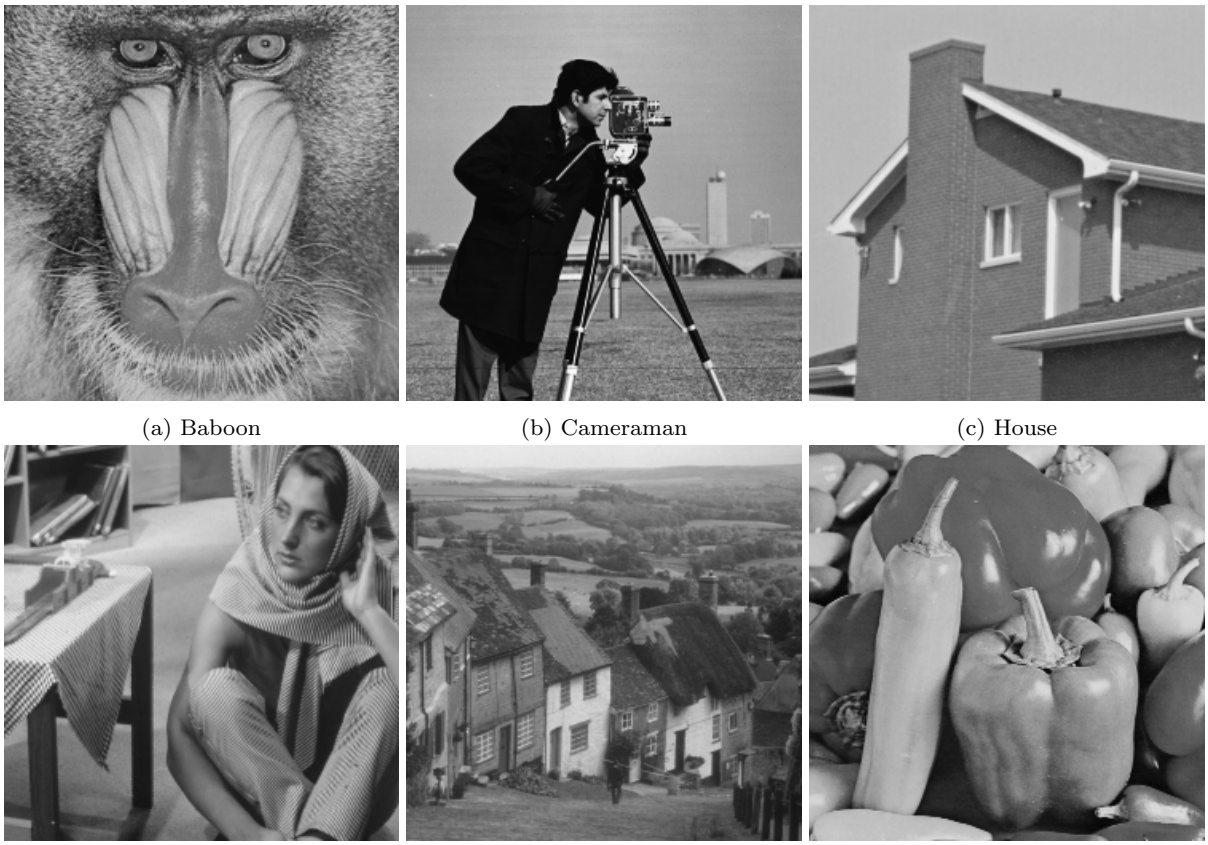
(e) TV + TV^2 regular. [17] (f) Proposed method ($p = 2$) (g) Proposed method ($p = 3$)

Fig. 1: Simulation results of deconvolution and denoising of *Lena* using different regularization methods.

finally we add an additive zero-mean white Gaussian noise with a chosen Signal-to-Noise Ratio level (SNR). We can now justify the contribution of this proposed combination in noise, blur and misregistration errors removing using different generated LR images. Hence, to evaluate the performance of the SR process, we use two measures such as peak signal to noise ratio (PSNR) and the structural similarity (SSIM). The PSNR is a popular metric used to measure the quality of the estimated HR image, while the SSIM is a complementary measure, which gives an indication of image quality based on known characteristics of the human visual system. In the next section, we describe the parameters chosen in the simulation results.

4.1 Simulation results

We use the six image of Figure 2, all these images have the same size 248×248. Then, we generate 120 LR images for each of the six images: we use a Gaussian kernel with standard deviation $\sigma = 1.5$ truncated in a 3×3 window to blur it, and subsample it by a factor $r = 4$. Notice that this introduces severe aliasing (a blur of $\sigma = 4 \times 0.8 = 3.2$ would be necessary according to [36]) and ringing (small window size). In addition to measure the robustness of our proposed method against noise, we add an additive white Gaussian noise with standard deviation $\sigma = 10$ to all the generated LR obtained from the six benchmark images. The initial HR image X^0 is obtained by a bicubic interpolation of the LR image Y_1, the scalar weight $\alpha = 0.7$, the spatial window size $p = 2$. We stop at the first iteration n where $\|\widehat{X}^{n+1} - \widehat{X}^n\|_1 < 10^{-4}\|\widehat{X}^n\|_1$. We justify the effectiveness of our method by comparing it with popular methods used in SR, such as bicubic, TV [11], BTV [9], BEP regularizer [12], the second order variational regularization used in denoising problem [14], and also the combined first order and second order (TV + TV^2) regularization [17]. We show the reconstructed HR images $\widehat{X}$ and also the obtained images $\widehat{B}$ from the fusion step



(a) Baboon (b) Cameraman (c) House

(d) Barbara (e) Goldhill (f) Peppers

Fig. 2: Set of benchmark images used in tests

in figures 3 to 8. By a visual evaluation, we can see the effectiveness of the proposed method in removing jagged artifacts. Note that we select the regularization parameters according to the best visually pleasant result and the highest PSNR in all the experiments for the other methods. In Table 1, the SSIM and PSNR values are computed, with two standard deviations σ for noise. The best results are represented by a bold number. Usually, the proposed method outperforms the others in terms of both PSNR and SSIM. Even if the result obtained by the BEP method exceeds ours in the *Goldhill* figure in terms of both PSNR and SSIM, we can find a better result if we change the regularization parameter's value δ.

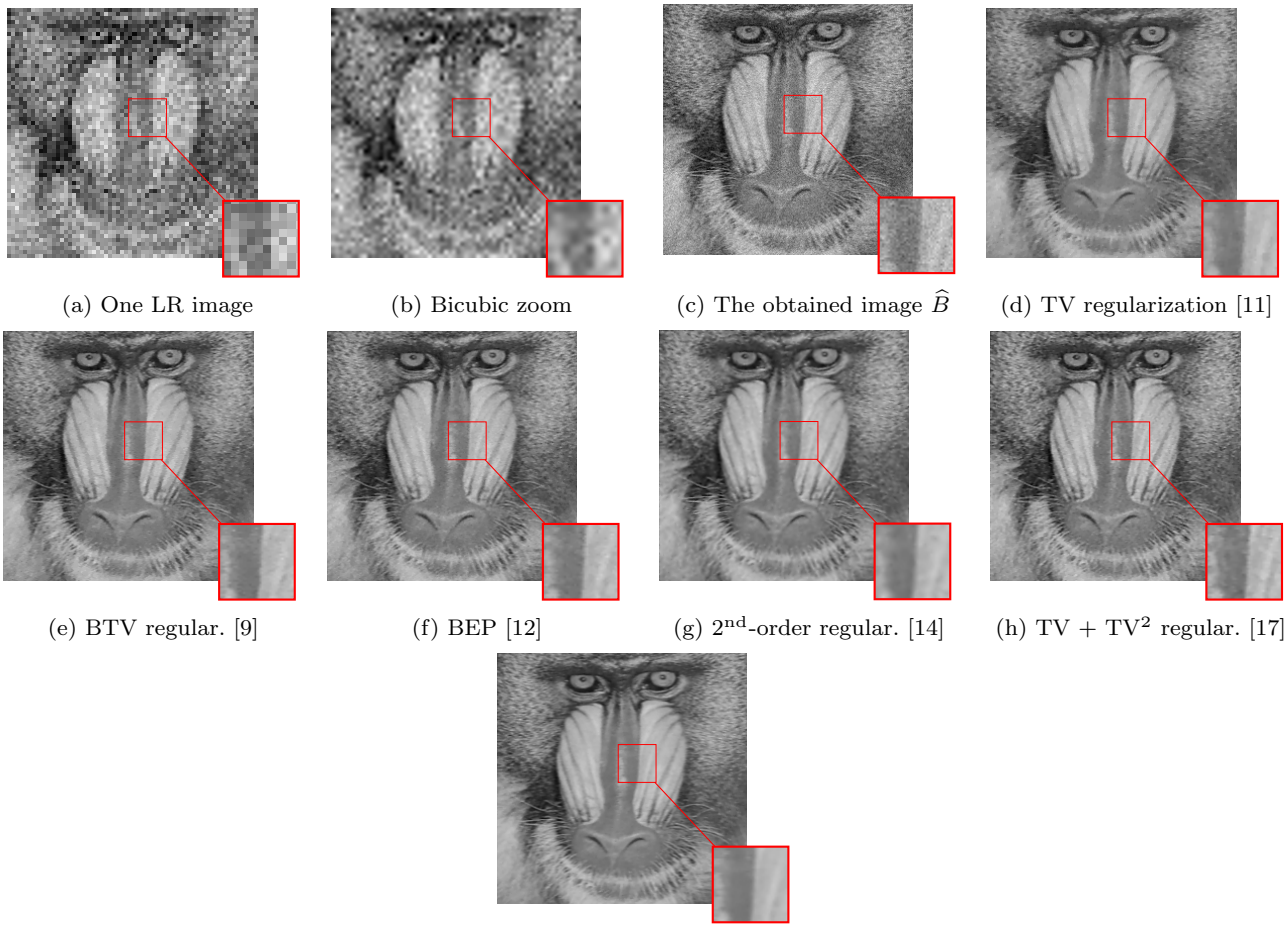
4.2 Real experiment

In this section, we discuss the performance of our proposed algorithm on natural data. We use in the first experiment the "Alpaca" video sequence of 55 compressed grayscale frames of size 96×128, downloaded from Milanfar's web site[1]. Without clue of the camera's PSF, we assume that it is a 11×11 Gaussian kernel with standard deviation equal to 1.5. In figure 9, we show the reconstituted HR image from 4 LR frames taken from the video. The resolution is augmented by a factor $r = 2$ and we compare our algorithm with the other SR method used in the previous tests. If we focus on what happens around the edges of the obtained HR image, we can see that our proposed method can restore the characters in the images at least as well as the best other SR algorithms.

In the second experiment, we use 16 low-resolution images obtained from the EIA video sequence[2] of size 90×90. We assume that the PSF is a 3×3 Gaussian kernel with standard deviation equal to 1.5. In figure 10, we show the result obtained of different SR algorithm HR image from 16 LR frames taken from the EIA video. The resolution is augmented by a factor $r = 3$ and we compare also our algorithm with the other SR method

[1] https://users.soe.ucsc.edu/~milanfar/software/sr-datasets.html

[2] https://users.soe.ucsc.edu/~milanfar/software/sr-datasets.html



(a) One LR image (b) Bicubic zoom (c) The obtained image $\widehat{B}$ (d) TV regularization [11]

(e) BTV regular. [9] (f) BEP [12] (g) 2nd-order regular. [14] (h) TV + TV^2 regular. [17]

(i) Proposed method

Fig. 3: Super resolution of *Baboon* by different methods.

used in the previous tests. Since the video of EIA has a lot more noise compared with Alpaca, we can see the robustness of the proposed algorithm against noise compared with the other SR methods.

5 Conclusion

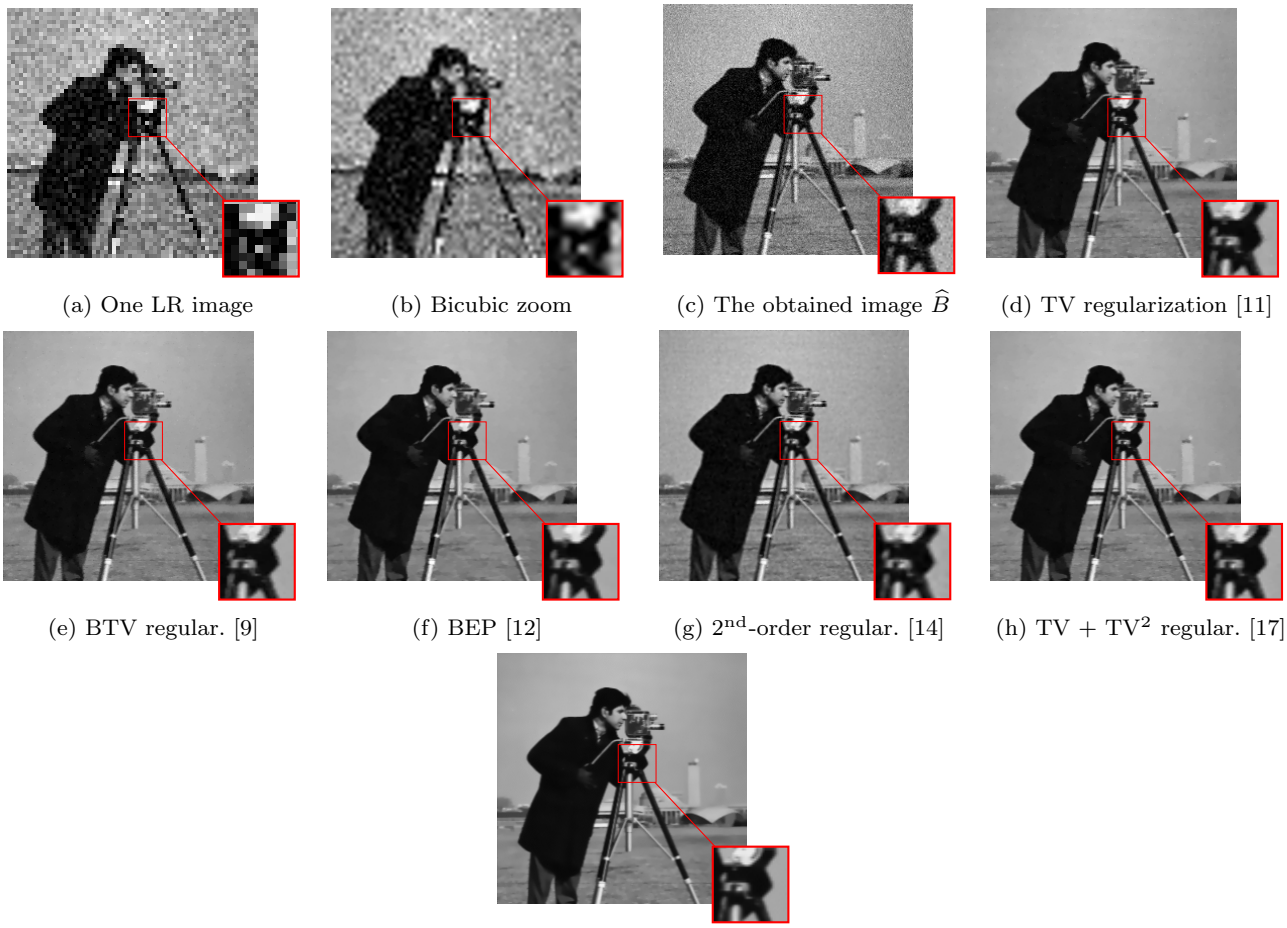
In this paper, a novel energy has been proposed for multiframe super-resolution image reconstruction problem. We presented a new combination of BTV regulizer and a second order variational problem using a controlled weighting parameter. Since the BTV term is effective in preserving sharp edges in images when removing blur and noise, the second order term has been introduced to avoid the undesirable staircasing effect. To confirm the choice of this combination, the existence of a minimizer has been proved using a relaxation technique. Also, the results have been performed on both the simulated and real images, and the proposed method has confirmed its effectiveness visually and quantitatively using the PSNR and SSIM criteria. One remaining open question is the uniqueness of the minimizer. We could neither prove nor discard this hypothesis. Another interesting path of investigation is the use of a non-local term instead of the bilateral term.

Acknowledgments

The authors are grateful to the anonymous reviewers for their insightful remarks and corrections. Their feedback had a great influence on the improved quality compared to the first draft version of this paper.

Compliance with Ethical Standards

– Funding: This research was entirely funded by the respective institutions of the authors.
– Conflict of interest: The authors declare that they have no conflict of interest.



(a) One LR image (b) Bicubic zoom (c) The obtained image $\widehat{B}$ (d) TV regularization [11]

(e) BTV regular. [9] (f) BEP [12] (g) 2nd-order regular. [14] (h) TV + TV^2 regular. [17]

(i) Proposed method

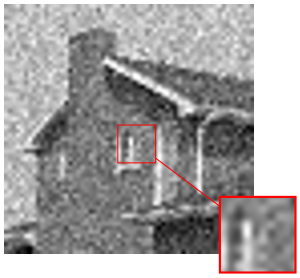
Fig. 4: Super resolution of *Cameraman* by different methods.

– Neither human participants nor animals are involved in this research.

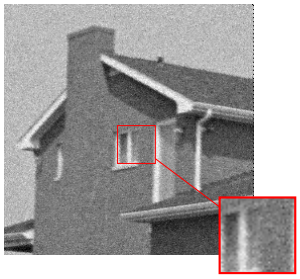
References

1. D. L. G. Hill, P. G. Batchelor, M. Holden, D. J. Hawkes, Medical image registration, Physics in medicine and biology 46 (3) (2001) R1.
2. Q. H. Luong, Advanced image and video resolution enhancement techniques, Ph.D. thesis, Faculty of Engineering Ghent University (2009).
3. R. Y. Tsai, T. S. Huang, Multiframe image restoration and registration, Advances in Computer Vision and Image Processing 1, chap. 7. (JAI Press, Greenwich, Conn, USA) (1984) 317–339.
4. H. Su, L. Tang, Y. Wu, D. Tretter, J. Zhou, Spatially adaptive block-based super-resolution, IEEE Transactions on Image Processing 21 (3) (2012) 1031–1045.
5. S. C. Park, M. K. Park, M. G. Kang, Super-resolution image reconstruction: a technical overview, IEEE Signal Processing Magazine 20 (3) (2003) 21–36.
6. E. S. Lee, M. G. Kang, Regularized adaptive high-resolution image reconstruction considering inaccurate subpixel registration, IEEE Transactions on Image Processing 12 (7) (2003) 806–813.
7. S. Villena, M. Vega, S. Babacan, R. Molina, A. K. Katsaggelos, Bayesian combination of sparse and non-sparse priors in image super resolution, Digital Signal Processing 23 (2) (2013) 530–541.
8. V. Caselles, G. Sapiro, D. H. Chung, Vector median filters, inf-sup operations, and coupled PDEs: Theoretical connections, Journal of Mathematical Imaging and Vision 12 (2) (2000) 109–119.
9. S. Farsiu, M. D. Robinson, M. Elad, P. Milanfar, Fast and robust multiframe super resolution, IEEE Transactions on Image Processing 13 (10) (2004) 1327–1344.
10. V. Patanavijit, S. Jitapunkul, A robust iterative multiframe superresolution reconstruction using a Huber Bayesian approach with Huber Tikhonov regularization, in: International Symposium on Intelligent Signal Processing and Communications, Yonago, Japan, 2006, pp. 13–16.
11. M. K. Ng, H. Shen, E. Y. Lam, L. Zhang, A total variation regularization based superresolution reconstruction algorithm for digital video, EURASIP Journal on Advances in Signal Processing (2007) 1–16.
12. X. Zeng, L. Yang, A robust multiframe super-resolution algorithm based on half-quadratic estimation with modi-

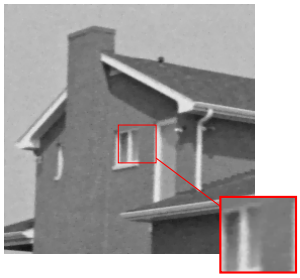
(a) One LR image



(b) Bicubic zoom

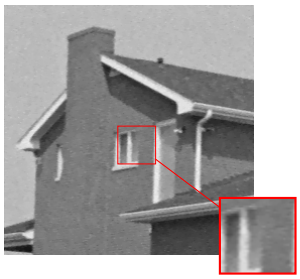


(c) The obtained image $\widehat{B}$

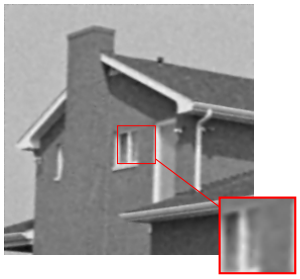


(d) TV regularization [11]

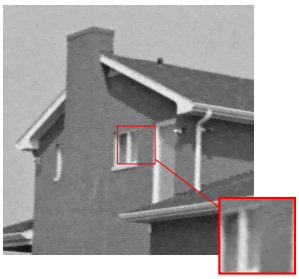
(e) BTV regular. [9]



(f) BEP [12]



(g) 2nd-order regular. [14]



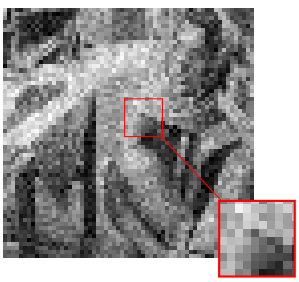
(h) TV + TV^2 regular. [17]

(i) Proposed method

Fig. 5: Super resolution of *House* by different methods.

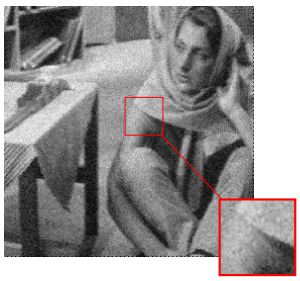
fied BTV regularization, Digital Signal Processing 23 (1) (2013) 98–109.

13. M. Lysaker, A. Lundervold, X.-C. Tai, Noise removal using fourth-order partial differential equation with applications to medical magnetic resonance images in space and time, IEEE Transactions on Image Processing 12 (12) (2003) 1579–1590.
14. M. Bergounioux, L. Piffet, A second-order model for image denoising, Set-Valued and Variational Analysis. 18 (3-4) (2010) 277–306.
15. Y.-L. You, M. Kaveh, Fourth-order partial differential equations for noise removal, IEEE Transactions on Image Processing 9 (10) (2000) 1723–1730.
16. M. Lysaker, X. C. Tai, Iterative image restoration combining total variation minimization and a second-order functional, International Journal of Computer Vision 66 (1) (2006) 5–18.
17. K. Papafitsoros, C. B. Schönlieb, A combined first and second order variational approach for image reconstruction, Journal of Mathematical Imaging and Vision 48 (2) (2014) 308–338.
18. K. Bredies, K. Kunisch, T. Pock, Total generalized variation, SIAM Journal on Imaging Sciences 3 (3) (2010) 492–526.
19. Y. Chen, R. Ranftl, T. Pock, Insights into analysis operator learning: From patch-based sparse models to higher order MRFs, IEEE Transactions on Image Processing 23 (3) (2014) 1060–1072.
20. Y. Chen, T. Pock, R. Ranftl, H. Bischof, Revisiting loss-specific training of filter-based MRFs for image restoration, in: Pattern Recognition, Springer, 2013, pp. 271–281.
21. F. Demengel, Fonctions à hessien borné, Annales de l'Institut Fourier 34 (2) (1985) 155–190.
22. G. Aubert, P. Kornprobst, Mathematical Problems in Image Processing Partial Differential Equations and the Calculus of Variations, Vol. 147, Springer Science, New York, 2006.
23. G. Demengel, F. Demengel, Espaces fonctionnels. Utilisation dans la résolution des équations aux dérivées partielles, EDP Sciences, 2012.
24. P. Milanfar, Super-Resolution Imaging, CRC Press, 2010.
25. J. Modersitzki, Numerical Methods for Image Registration, Oxford University Press, USA, 2003.
26. J. Nocedal, S. Wright, Numerical Optimization, Springer Science, 2006.
27. A. Chambolle, T. Pock, A first-order primal-dual algorithm for convex problems with applications to imaging, Journal of Mathematical Imaging and Vision 40 (1) (2011) 120–145.

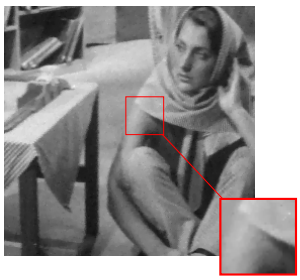


(a) One LR image

(b) Bicubic zoom

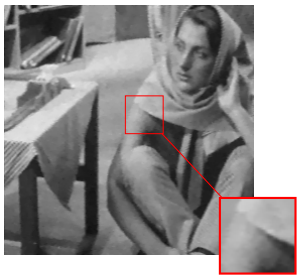


(c) The obtained image $\widehat{B}$

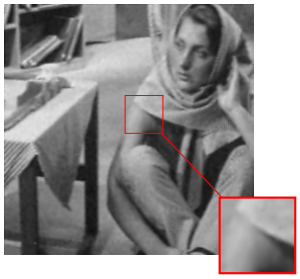


(d) TV regularization [11]

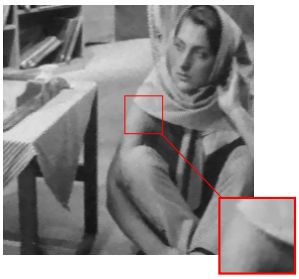
(e) BTV regular. [9]



(f) BEP [12]



(g) 2nd-order regular. [14]

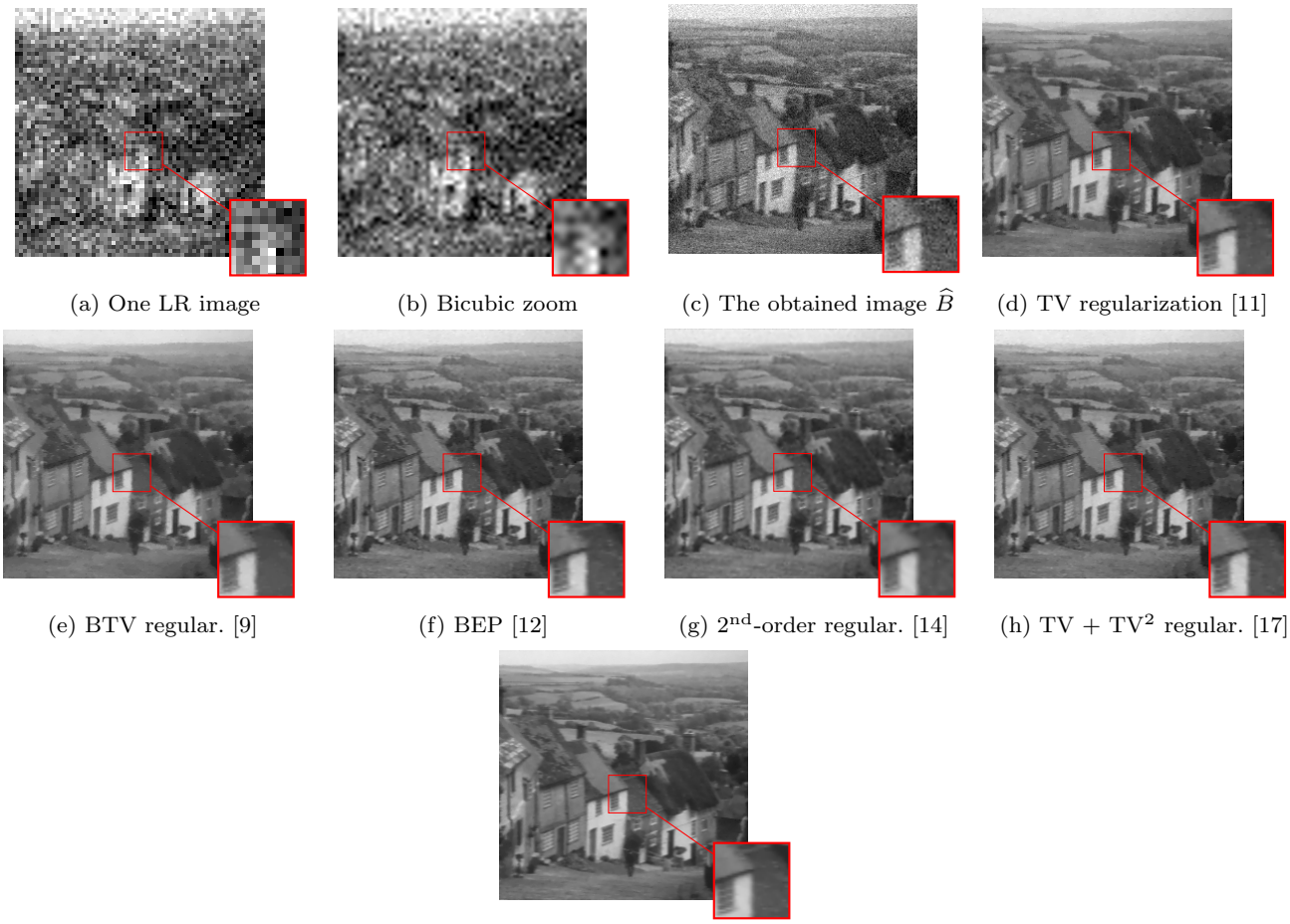


(h) TV + TV^2 regular. [17]

(i) Proposed method

Fig. 6: Super resolution of *Barbara* by different methods.

28. H. Brezis, Functional Analysis, Sobolev Spaces and Partial Differential Equations, Springer Science, New York, 2010.
29. F. Demengel, R. Temam, Convex functions of a measure and applications, Indiana University Mathematics Journal 33 (5) (1984) 673–709.
30. L. Vese, Problémes variationnels et EDP pour lanalyse dimages et lévolution de courbes, Ph.D. thesis, Université de Nice Sophia-Antipolis (1996).
31. Y. Li, F. Santosa, A computational algorithm for minimizing total variation in image restoration, IEEE Transactions on Image Processing 5 (6) (1996) 987–995.
32. L. Rudin, S. Osher, E. Fatemi, Nonlinear total variation based noise removal algorithms, Physica D 60 (1992) 259–268.
33. M. Unger, T. Pock, M. Werlberger, H. Bischof, A convex approach for variational super-resolution, in: Pattern Recognition, Springer, 2010, pp. 313–322.
34. N. Otsu, A threshold selection method from gray-level histograms, Automatica 11 (285-296) (1975) 23–27.
35. A. Laghrib, A. Hakim, S. Raghay, M. El Rhabi, Robust super resolution of images with non-parametric deformations using an elastic registration, Applied Mathematical Sciences 8 (179) (2014) 8897–8907.
36. J.-M. Morel, G. Yu, Is SIFT scale invariant?, Inverse Problems and Imaging 5 (1) (2011) 115–136.

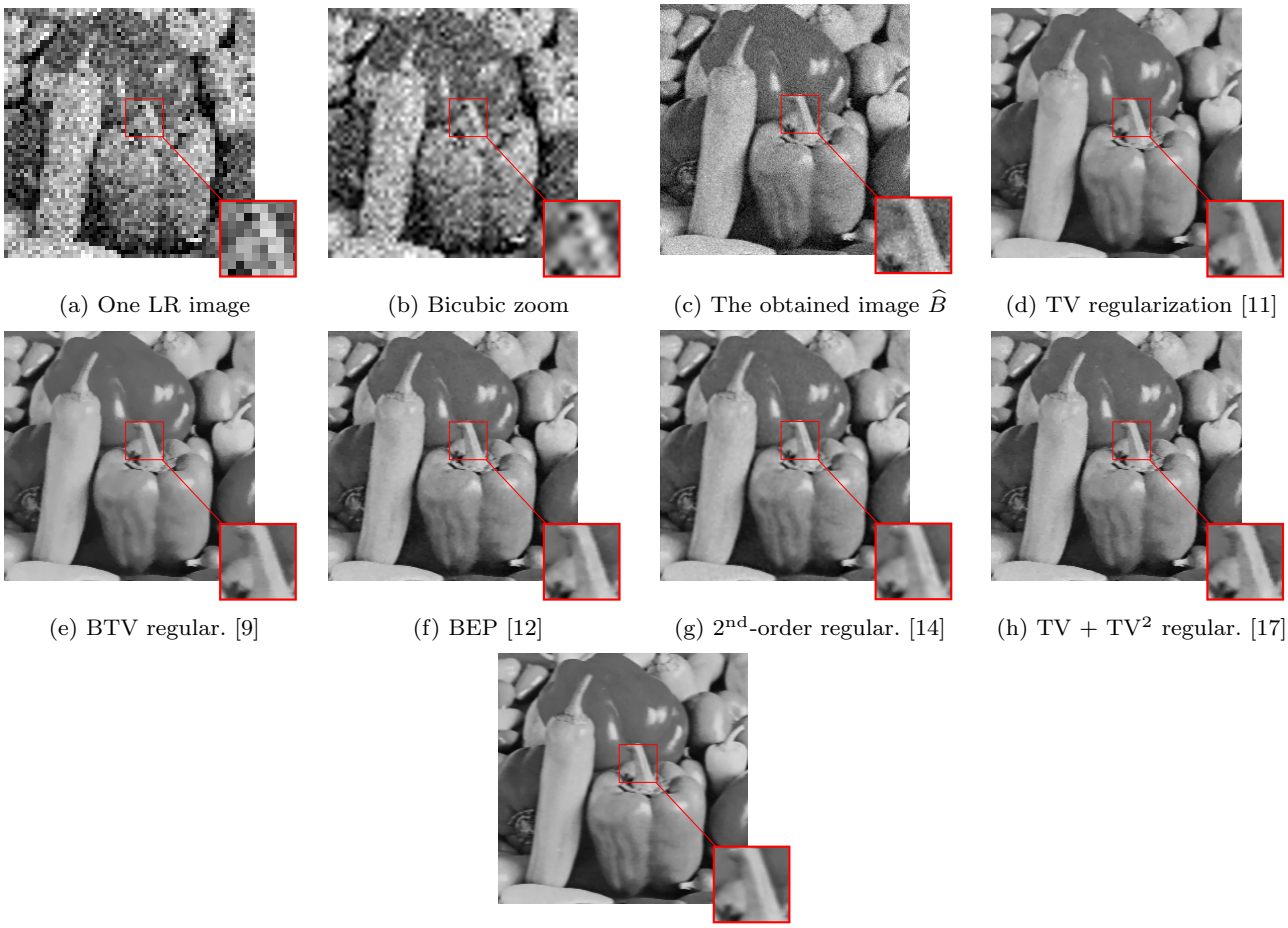


(a) One LR image (b) Bicubic zoom (c) The obtained image $\widehat{B}$ (d) TV regularization [11]

(e) BTV regular. [9] (f) BEP [12] (g) 2nd-order regular. [14] (h) TV + TV^2 regular. [17]

(i) Proposed method

Fig. 7: Super resolution of *Goldhill* by different methods.



(a) One LR image (b) Bicubic zoom (c) The obtained image $\widehat{B}$ (d) TV regularization [11]

(e) BTV regular. [9] (f) BEP [12] (g) 2nd-order regular. [14] (h) TV + TV^2 regular. [17]

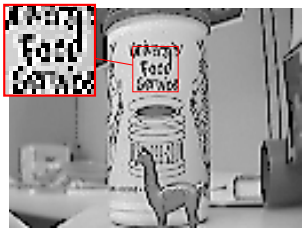
(i) Proposed method

Fig. 8: Super resolution of *Peppers* by different methods.

Table 1: PSNR and SSIM results obtained by applying different methods to the benchmark images. In *bold* the best (highest) score of each row is shown.

Image			Method						
	σ noise	Metric	Bicubic	TV	second order	BEP	BTV	TV+TV^2	proposed
Baboon	20	PSNR	17.95	25.69	25.79	26.50	26.45	26.50	**26.74**
		SSIM	0.229	0.738	0.744	0.788	0.781	0.789	**0.801**
	30	PSNR	16.98	24.22	24.57	25.11	24.96	25.09	**25.77**
		SSIM	0.201	0.685	0.688	0.713	0.699	0.704	**0.746**
Cameraman	20	PSNR	18.22	26.48	26.52	27.52	26.63	26.93	**27.58**
		SSIM	0.288	0.852	0.837	0.848	0.855	0.864	**0.892**
	30	PSNR	17.88	25.43	25.49	26.48	26.04	26.18	**26.89**
		SSIM	0.200	0.783	0.768	0.796	0.795	0.802	**0.827**
House	20	PSNR	18.62	33.27	32.45	33.53	33.32	33.90	**34.40**
		SSIM	0.326	0.870	0.858	0.875	0.872	0.878	**0.900**
	30	PSNR	17.55	31.82	31.49	32.88	32.05	32.92	**33.21**
		SSIM	0.256	0.838	0.833	0.814	0.810	0.838	**0.876**
Barbara	20	PSNR	18.83	26.97	27.31	27.68	27.02	27.34	**27.87**
		SSIM	0.412	0.772	0.787	0.796	0.770	0.787	**0.813**
	30	PSNR	17.41	26.15	26.93	26.94	26.55	26.85	**27.03**
		SSIM	0.328	0.709	0.750	0.752	0.750	0.754	**0.763**
Goldhill	20	PSNR	20.19	29.16	29.15	**29.99**	29.02	29.24	29.98
		SSIM	0.487	0.832	0.829	0.847	0.820	0.836	**0.868**
	30	PSNR	19.45	28.00	28.11	**28.68**	28.22	28.29	28.66
		SSIM	0.338	0.771	0.763	0.809	0.779	0.790	**0.820**
Peppers	20	PSNR	18.39	30.74	30.58	31.40	30.81	31.37	**31.65**
		SSIM	0.438	0.889	0.885	0.893	0.889	0.883	**0.920**
	30	PSNR	17.88	29.73	29.80	30.15	29.24	30.08	**30.72**
		SSIM	0.318	0.855	0.856	0.853	0.839	0.835	**0.896**

(a) First LR frame



(b) Fourth LR frame

(c) Bicubic zoom

(d) TV regular. [11]

(e) BTV regular. [9]

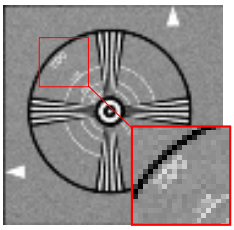
(f) BEP [12]

(g) 2nd-order regular. [14]

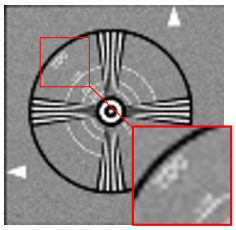
(h) TV + TV^2 regular. [17]

(i) Our method

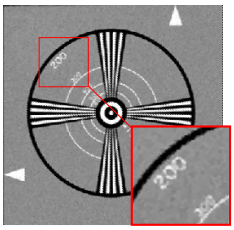
Fig. 9: Results on the first four frames of the *Alpaca* sequence.



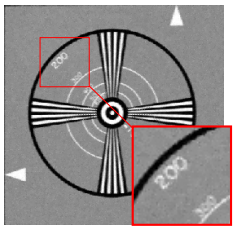
(a) First LR frame



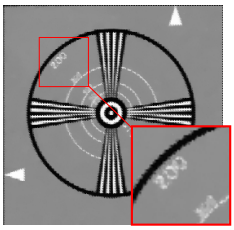
(b) Bicubic zoom



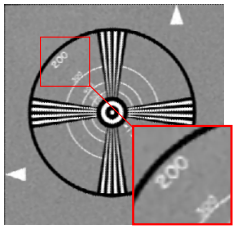
(c) TV regular. [11]



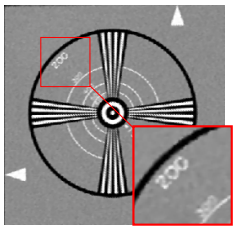
(d) BTV regular. [9]



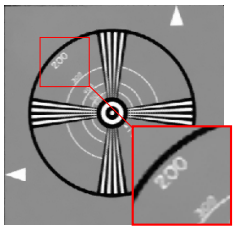
(e) BEP [12]



(f) 2nd-order regular. [14]



(g) TV + TV^2 regular. [17]



(h) Our method

Fig. 10: Results on the *EIA* sequence.