An Alternating Proximal Approach for Blind Video Deconvolution

Feriel Abboud, Emilie Chouzenoux, Jean-Christophe Pesquet, Jean-Hugues Chenot, Louis Laborelli

An Alternating Proximal Approach for Blind Video Deconvolution

Feriel Abboud[a], Émilie Chouzenoux[b,c,*], Jean-Christophe Pesquet[c], Jean-Hugues Chenot[d], Louis Laborelli[d]

[a] *WITBE, Les Collines de l'Arche, Immeuble Opéra. 92800 Puteaux, France*
[b] *LIGM, UMR CNRS 8049, Université Paris-Est. 77454 Champs-sur-Marne, France*
[c] *Center for Visual Computing, CentraleSupélec, INRIA Saclay. 91190 Gif-sur-Yvette, France*
[d] *INA, Institut National de l'Audiovisuel. 94366 Bry sur Marne, France*

Abstract

Blurring occurs frequently in video sequences captured by consumer devices, as a result of various factors such as lens aberrations, defocus, relative camera-scene motion, and camera shake. When it comes to the contents of archive documents such as old films and television shows, the degradations are even more serious due to several physical phenomena happening during the sensing, transmission, recording, and storing processes. We propose in this paper a versatile formulation of blind video deconvolution problems that seeks to estimate both the sharp unknown video sequence and the underlying blur kernel from an observed video. This inverse problem is *ill-posed*, and an appropriate solution can be obtained by modeling it as a nonconvex minimization problem. We provide a novel iterative algorithm to solve it, grounded on the use of recent advances in convex and nonconvex optimization techniques, and having the ability of including numerous well-known regularization strategies.

Keywords: Blind deconvolution, video processing, regularization, nonconvex optimization, proximal algorithms

[*] Corresponding author
Email address: emilie.chouzenoux@u-pem.fr (Émilie Chouzenoux)

1. Introduction

Video processing has been growing in popularity for the last decades and plays henceforth a prominent role in many application fields, such as telecommunication, video surveillance, microscopy, medical imaging, astronomy and so on. Real-life video sequences are usually blurred due to the overall effect of different factors such as defocus, motion blur, and optical blur. These degraded videos can typically be modeled as the noisy convolution of original ones with the impulse response of some blur kernel, also called point spread function (PSF) [1, 2, 3]. Thereby, a deconvolution process becomes mandatory for retrieving a visually sharp video [4].

Video deconvolution problems can be categorized into two types: non-blind deconvolution problem where the blur kernel is assumed to be known, and blind deconvolution problem where one has to estimate both the video and the blur. The blind scheme is more realistic, and it is frequently encountered, for instance in optics due to imperfect optical instruments, and in photography due to misfocusing or camera shake, resulting in blurry images with unknown PSF. Blind video deconvolution is an *ill-posed* inverse problem since an infinity of pairs (images/blur kernel) can lead to the same observed video. Therefore, the use of some prior knowledge on the sought video and kernels is required to compute a stable solution to the problem and circumvent its *ill-posedness.* This usually comes with the formulation of the problem as the optimization of a nonconvex cost function accounting for the observation model, through a data fidelity term, and prior information thanks to regularization functions.

A large amount of works have been dedicated to the search for optimal regularization strategies for 2D blind deconvolution. Sparsity is often favored and exploited in such context, with the use of regularization terms based on ℓ_1/ℓ_2-norm [5, 6], ℓ_1-norm [7], or ℓ_0-norm [8], applied to linear transforms of the image. A very popular approach is the Total Variation (TV) regularization that has been proposed in [9] in order to penalize small discontinuities while preserving the edges in the image. Another interesting regularization approach consists of resorting to wavelet/frame decompositions [10, 11, 12, 13], where the sparsity of the frame coefficients of the sought image is enforced. Regularization and hard constraints are also applied to the blur kernel. They usually model the physical properties of the imaging system by imposing the positivity of the kernel coefficients, their smoothness or sparsity, and a mean or energy preservation property through a sum-to-one constraint [5, 14, 15]. Note that this last condition is particularly useful in the blind case since it allows to overcome the scaling ambiguity issue.

The extension of these approaches to the case of video processing can be performed in a straightforward manner, by considering each frame in an independent manner. However, such naive strategies may be detrimental to the restored video quality as they do not account for dependencies between consecutive video frames. Moreover, in some specific contexts of video restoration, such as television archive processing, the blur kernel is assumed to be constant along time, which again should be taken into account in order to improve the blind deconvolution procedure. Up to our knowledge, the aforementioned issues have not been treated in the literature and this paper aims at addressing them by considering a more versatile formulation of the video blind deconvolution problem.

An important challenge in the context of blind deconvolution with variational techniques lies in the nonconvexity of the objective function to be minimized. The nonconvexity is actually intrinsic to the blind model. It can also arise from the use of sparsity enhancement priors, such as the ℓ_0 quasi-norm and its approximations [16]. Therefore, there is a crude need for efficient optimization tools dealing with such a class of complicated problems. In the context of 2D blind deconvolution, most techniques rely on an alternating minimization strategy, consisting in updating, at each iteration of the process, either the image or the blur kernel, by minimizing the cost function with respect to each of both variables [5, 4]. However, the convergence guarantees for this scheme require very restrictive assumptions [17] that are usually not met in the context of blind deconvolution. Therefore, improved optimization methods have been proposed, based on the use of proximal steps, that benefit from convergence guarantees even in this challenging nonconvex context [18], and only require mild changes in the minimization process. In this paper, we pursue this promising direction by extending these ideas to the context of blind video deconvolution.

In a nutshell, the contributions of our paper are:

- the proposition of a versatile regularized formulation of the blind video deconvolution problem, that accounts in particular for temporal correlations between consecutive video frames,
- the proposition of a new alternating proximal algorithm to solve the underlying nonconvex minimization problem with sound convergence guarantees on the produced iterates,
- extensive numerical comparisons of several spatial gradient-based regularizers for blind deconvolution of synthetic and real video sequences.

The reminder of this paper is structured as follows: we introduce in Section 2 the formulation of the blind video deconvolution problem as a minimization problem and present a number of regularization strategies that are encompassed by our model. Afterwards, we present our minimization approach in Section 3, which allows us to solve efficiently the resulting nonconvex problem. Section 4 provides some illustrative experimental results on synthetic and real video sequences. Finally, some conclusions are given in Section 5.

2. Problem statement

2.1. Observation model

Blind video deconvolution amounts to inferring an original sharp video sequence $\boldsymbol{x} = (x_t)_{1 \leqslant t \leqslant T} \in \mathbb{R}^{TN}$ and a spatial convolution kernel $h \in \mathbb{R}^P$ from an observed degraded video sequence $\boldsymbol{y} = (y_t)_{1 \leqslant t \leqslant T} \in \mathbb{R}^{TN}$, satisfying the following degradation model:

$$(\forall t \in \{1, \dots, T\}) \qquad y_t = h * x_t + w_t, \tag{1}$$

where T denotes the number of frames included in the video sequence, $*$ stands for the 2D convolution operator, and $(w_t)_{1\leqslant t\leqslant T} \in \mathbb{R}^{TN}$ represents an additive noise. (Throughout the paper, images of size $N_1 \times N_2$ are reshaped as column vectors of dimension $N = N_1N_2$).

Note that, for simplicity, we will consider throughout the paper, circulant boundary padding of the images. Let us however emphasize that our approach can encompass any type of boundary assumptions, as soon as all the involved operators and their respective adjoints are computed in a consistent manner. Furthermore, we assume that the kernel is temporally invariant, which is a realistic model for our target application, that is the restoration of television archives. However, here again, it must be emphasized that the methodology described in this paper could be easily extended to the case of temporal varying kernels such as those related to camera motion.

If no additional information is supplied, the problem of estimating the images and the kernel from the noisy and blurry observations is *ill-posed* and its resolution may lead to unstable and unsatisfactory results. Thus, we resort to the following penalized formulation in order to solve it:

$$\text{Find } (\widehat{\boldsymbol{x}}, \widehat{h}) \in \mathbb{R}^{TN} \times \mathbb{R}^{P} \text{ such that } F(\widehat{\boldsymbol{x}}, \widehat{h}) = \inf F \tag{2}$$

with

$$(\forall \boldsymbol{x} \in \mathbb{R}^{TN})(\forall h \in \mathbb{R}^{P}) \quad F(\boldsymbol{x}, h) = \Phi(\boldsymbol{x}, h) + \Psi(\boldsymbol{x}) + \Theta(h) \tag{3}$$

The cost function (3) is composed of a least squares data fidelity term Φ which ensures the compliance with Model (1), and is given by

$$(\forall \boldsymbol{x} \in \mathbb{R}^{TN})(\forall h \in \mathbb{R}^{P}) \quad \Phi(\boldsymbol{x}, h) = \frac{1}{2}\sum_{t=1}^{T} \|h * x_t - y_t\|^2, \tag{4}$$

and of two regularization functions Ψ and Θ that incorporate prior information on the sought images and kernel, respectively. The objective function F

is nonconvex due to the coupling existing in the data fidelity term between the variables $\boldsymbol{x}$ and h. This suggests the use of an optimization method that alternates between the estimation of the images composing the sequence $\boldsymbol{x}$, and the identification of the PSF h in order to reach a critical point of (2). Besides, it is worth noticing that the choice of the regularization functions Ψ and Θ plays a prominent role in the quality of the restored video and the identified kernel. A number of spatial regularization strategies has been proposed in the context of image processing. We propose to combine them with a temporal regularizer to account for the correlation between consecutive video frames. The adopted optimization method heavily depends on the mathematical properties of the retained penalty functions. The main contribution of this paper is to propose a unique and versatile optimization framework that can handle a wide class of regularization functions, as detailed in the following.

2.2. Video estimation

Let us first focus on the simpler problem of estimating the video sequence while assuming a known PSF h. The images composing the video sequence can be inferred by solving the following problem

$$\underset{\boldsymbol{x}\in\mathbb{R}^{TN}}{\operatorname{minimize}}\ \Phi(\boldsymbol{x},h)+\Psi(\boldsymbol{x}). \tag{5}$$

Here, we propose to define Ψ as follows:

$$(\forall \boldsymbol{x}\in\mathbb{R}^{TN})\quad \Psi(\boldsymbol{x})=\sum_{t=1}^{T}\big(\eta\,\psi(x_t)+\iota_{[x_{\min},x_{\max}]^N}(x_t)\big)+\mathcal{M}(\boldsymbol{x}), \tag{6}$$

where ψ is a spatial regularization function handling each frame x_t separately, $\iota_{[x_{\min},x_{\max}]^N}$ denotes an indicator function that sets a range on the pixel values of each image. Moreover, we propose to rely on the following temporal regularization term:

$$(\forall \boldsymbol{x}\in\mathbb{R}^{TN})\quad \mathcal{M}(\boldsymbol{x})=\frac{1}{2}\sum_{t=1}^{T}\sum_{\ell\in\mathcal{V}_t}\beta_{\ell,t}\|x_t-M_{\ell\to t}x_\ell\|^2, \tag{7}$$

where, for every t and ℓ, $\beta_{\ell,t}$ are positive weights selected proportionally to the distance $|t-\ell|$ between the frame index of images x_t and x_ℓ, the index set $\mathcal{V}_t$ defines the neighborhood of the current image x_t (i.e., $\ell \in \mathcal{V}_t$ is such that $|\ell - t|$ is small and nonzero), and $M_{\ell\to t} \in \mathbb{R}^{N\times N}$ is a linear operator modeling the motion fields between the current image x_t and the neighboring image x_ℓ. These motion fields can be efficiently estimated from the data, as detailed in the experimental part of the paper. Note that $\mathcal{M}$ can be viewed as a smooth version of the temporal regularization we proposed in our previous work [19] with the advantage of a reduced computational cost.

Various choices for spatial regularization term ψ can be adopted in Model (6). In this work, we will consider several of them that are listed herebelow:

- *Total Variation (TV)* is one of the most popular regularization method in image restoration. It has been initially introduced for image denoising and reconstruction problems [9], and reads:

  $$(\forall z \in \mathbb{R}^N) \quad \psi(z) = \chi_2\left(Dz\right), \tag{8}$$

  where $D \in \mathbb{R}^{2N\times N}$ is the discrete gradient operator defined as the concatenation of the horizontal and vertical gradient operators:

  $$D = \begin{bmatrix} \nabla_{\mathsf{H}} \\ \nabla_{\mathsf{V}} \end{bmatrix}, \quad \text{with} \quad \nabla_{\mathsf{H}} \in \mathbb{R}^{N\times N}, \quad \nabla_{\mathsf{V}} \in \mathbb{R}^{N\times N}, \tag{9}$$

  and for every $q \in \mathbb{N}^*$, $\chi_q \colon \mathbb{R}^{qN} \to \mathbb{R}$ is a sparsity promoting function given by

  $$\big(\forall (z_1,\ldots,z_q) \in (\mathbb{R}^N)^q\big) \quad \chi_q(z_1,\ldots,z_q) = \sum_{k=1}^{N} \sqrt{\left(z_{1,k}\right)^2 + \cdots + \left(z_{q,k}\right)^2}. \tag{10}$$

  The total variation promotes the sparsity of the image derivatives, which has the advantage of reducing the noise and preserving sharp edges. However, it may lead to piecewise constant images and induce staircase artifacts [20].

- *Semi-Local Total Variation (SLTV)* that has been proposed in [21], is a variant of the classical total variation that attenuates the staircase artifacts, and is defined as

$$(\forall z \in \mathbb{R}^N) \quad \psi(z) = \sum_{\ell \in \Omega} \chi_2\left(Dz - V_\ell Dz\right), \tag{11}$$

  where $D \in \mathbb{R}^{2N \times N}$ is the linear operator introduced in (9), $\Omega = \{1, \ldots, 6\}$ and $(V_\ell)_{1 \leqslant \ell \leqslant 6} \in (\mathbb{R}^{2N \times 2N})^6$ represent shift operators as illustrated in Figure 1.

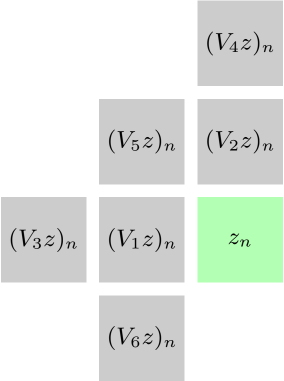


Figure 1: Shift operators $(V_\ell)_{\ell \in \{1,\ldots,6\}}$ applied to a given pixel position $n \in \{1, \ldots, N\}$.

- *Total Generalized Variation (TGV)* is a high order total variation regularization with acknowledged good performance in the context of image restoration [22, 23]. We use here the definition given in [24, 25]:

$$(\forall z \in \mathbb{R}^N) \quad \psi(z) = \min_{q \in \mathbb{R}^{2N}} \; \alpha_0 \, \chi_2\left(Dz - q\right) + \alpha_1 \, \chi_3\left(\mathsf{G}q\right),$$

  with $(\alpha_0, \alpha_1) \in]0, +\infty[^2$, $D \in \mathbb{R}^{2N \times N}$ and χ_q are defined in (9) and (10) respectively. Finally $\mathsf{G} \in \mathbb{R}^{3N \times N}$ can be viewed as a further derivative operator given by

$$\mathsf{G} = \begin{bmatrix} -\nabla_{\mathsf{H}}^{\top} & 0 \\ -\nabla_{\mathsf{V}}^{\top} & -\nabla_{\mathsf{H}}^{\top} \\ 0 & -\nabla_{\mathsf{V}}^{\top} \end{bmatrix}. \tag{12}$$

- *Total Variation on a Staggered Grid (TVSG)* that has been recently proposed in [26], introduces a new formulation of the total variation with a more accurate adaptation of its continuous definition to the discrete domain, instead of the one based on the classical finite differences in (9). It resorts to a sophisticated gradient operator which is defined as

$$(\forall z \in \mathbb{R}^N) \quad \psi(z) = \min_{(v_1,v_2,v_3)\in\mathbb{R}^{(2N)^3}} \Big\{ \chi_2\,(v_1) + \chi_2\,(v_2) + \chi_2\,(v_3) \;\; | \\ L_1^{\top} v_1 + L_2^{\top} v_2 + L_3^{\top} v_3 = Dz \Big\}, \tag{13}$$

where $L_1^{\top}, L_2^{\top}, L_3^{\top}$ denote the adjoint operators of L_1, L_2, L_3 respectively, defined as follows. Let $u \in \mathbb{R}^{2N} = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$, then

$L_1 u \in \mathbb{R}^{2N} = \begin{bmatrix} q_{1,1} \\ q_{1,2} \end{bmatrix}$ is such that

$q_{_1,1}(n,m) = u_1(n,m),$

$q_{_1,2}(n,m) = \left(u_2(n,m) + u_2(n,m-1) + u_2(n+1,m) + u_2(n+1,m-1)\right)/4,$

$L_2 u \in \mathbb{R}^{2N} = \begin{bmatrix} q_{2,1} \\ q_{2,2} \end{bmatrix}$ is such that

$q_{2,1}(n,m) = \left(u_1(n,m) + u_1(n-1,m) + u_1(n,m+1) + u_1(n-1,m+1)\right)/4,$

$q_{2,2}(n,m) = u_2(n,m),$

$L_3 u \in \mathbb{R}^{2N} = \begin{bmatrix} q_{3,1} \\ q_{3,2} \end{bmatrix}$ is such that

$q_{3,1}(n,m) = \left(u_1(n,m) + u_1(n-1,m)\right)/2,$

$q_{3,2}(n,m) = \left(u_2(n,m) + u_2(n,m-1)\right)/2,$

where $n \in \{1, \ldots, N_1\}$ and $m \in \{1, \ldots, N_2\}$ are vertical and horizontal pixel indices, with $N = N_1 N_2$.

This new definition of gradient fields leads to a regularized approach that improves the sharpness of the edges, and presents a better isotropy compared to the standard total variation.

- *Smoothed One Over Two-Total Variation (SOOT-TV)* is a nonconvex sparsity promoting function combining the ℓ_1/ℓ_2 norm and the total variation operator. ℓ_1/ℓ_2 can be viewed as a more accurate approximation to ℓ_0 compared with the convex ℓ_1 norm, as shown in Figure 2.

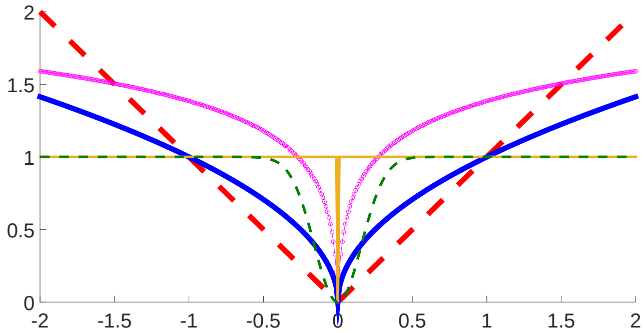


Figure 2: Sparsity promoting norms: ℓ_0 norm (thin solid yellow), ℓ_1 norm (thick dashed red), ℓ_1/ℓ_2 norm (thick solid blue), log-ℓ_1 norm (thin magenta '$\circ$'), Welsch penalty (thin dashed green).

Here, we will focus on the log-smoothed version of the ℓ_1/ℓ_2 norm called "SOOT" introduced in [6]. The prior then reads:

$$(\forall z \in \mathbb{R}^N) \quad \psi(z) = \log\left(\frac{\ell_{1,\alpha}(Dz) + \beta}{\ell_{2,\lambda}(Dz)}\right), \tag{14}$$

where

$$\ell_{1,\alpha}(Dz) = \sum_{i=1}^{2N}\left(\sqrt{(D^{(i)}z)^2 + \alpha^2} - \alpha\right), \quad \ell_{2,\lambda}(Dz) = \sqrt{\sum_{i=1}^{2N}(D^{(i)}z)^2 + \lambda^2}, \tag{15}$$

$D \in \mathbb{R}^{2N\times N}$ is the discrete gradient operator defined in (9), $D^{(i)} \in \mathbb{R}^{1\times N}$ denotes the i-th row of D, and α, β, λ are positive parameters.

- *Smoothed log-Total Variation (log-TV)* is a nonconvex smooth sparsity promoting regularization function from [27] defined as follows

$$(\forall z \in \mathbb{R}^N) \quad \psi(z) = \frac{1}{2}\sum_{i=1}^{N}\log\left((\nabla_{\mathsf{H}}^{(i)}z)^2 + (\nabla_{\mathsf{V}}^{(i)}z)^2 + \alpha^2\right), \tag{16}$$

where $\alpha > 0$, and $\nabla_{\mathsf{H}}^{(i)} \in \mathbb{R}^{1\times N}$ (resp. $\nabla_{\mathsf{V}}^{(i)} \in \mathbb{R}^{1\times N}$) denotes the i-th row of ∇_{H} (resp. ∇_{V}). Similarly to the ℓ_1/ℓ_2 norm, the log-based penalty used in (16) can be viewed as a nonconvex approximation to ℓ_0.

- *Welsch-Total Variation (Welsch-TV)* is based on the so-called "Welsch function" [28] defined as

  $$\phi\colon \mathbb{R} \to \mathbb{R} \,:\, t \to 1 - \exp\left(-t^2/(2\sigma^2)\right). \tag{17}$$

  The Welsch function is bounded and approaches 1 exponentially fast as $|t| \to +\infty$, as shown by Figure 2. It is convex near the origin, for $t^2 < \sigma^2$ and nonconvex elsewhere. Its adaptation to the context of image and video deconvolution is realized by applying it to the image gradients in order to measure their sparsity:

  $$(\forall z \in \mathbb{R}^N) \quad \psi(z) = \sum_{i=1}^{N} \phi\big(\sqrt{(\nabla_{\mathsf{H}}^{(i)} z)^2 + (\nabla_{\mathsf{V}}^{(i)} z)^2}\big). \tag{18}$$

2.3. Kernel identification

The spatial convolution kernel can be estimated by solving the minimization problem (2) with respect to h while keeping the images $(x_t)_{1\leqslant t\leqslant T}$ fixed, which reduces to

$$\underset{h\in\mathbb{R}^P}{\text{minimize}} \;\; \Phi(\boldsymbol{x}, h) + \Theta(h), \tag{19}$$

where Θ stands for an indicator function of a set $\mathcal{H}$ representing a constrained set, so that *a priori* information on the sought kernel are satisfied. In the proposed method, the following constraints are considered:

$$(\forall h \in \mathbb{R}^P) \quad \Theta(h) = \iota_{\mathcal{H}}(h) = \begin{cases} 0 & \text{if } h \in \mathcal{H} \\ +\infty & \text{otherwise,} \end{cases} \tag{20}$$

with

$$\mathcal{H} = \Big\{ h = (h_p)_{1\leqslant p\leqslant P} \in \mathbb{R}^P \mid \sum_{p=1}^{P} h_p = 1, \tag{21}$$

$$(\forall p \in \{1,\ldots,P\})\; h_{\min,p} \leqslant h_p \leqslant h_{\max,p} \Big\}. \tag{22}$$

The first constraint (21) is used to circumvent the so-called scaling ambiguity. In fact, let $(\widehat{x}, \widehat{h})$ be a solution to (2), then each pair $(\alpha\widehat{x}, \frac{1}{\alpha}\widehat{h})$ with $\alpha \neq 0$ is also a solution satisfying Model (1). This ambiguity is avoided by imposing Constraint (21). The second constraint (22) is adjusted regarding prior information on the physical properties of the sought convolution kernel. As an example, for old television archive contents, the kernel h may have a narrow spike and small (possibly negative) components.

3. Optimization method

3.1. Minimization strategy

The objective function F is nonconvex, yet has a simple structure with respect to each image variable x_t, $t \in \{1, \dots, T\}$, and with respect to the kernel h. A standard resolution approach is thus to adopt an alternating minimization strategy, where, at each iteration, F is minimized with respect to one variable while the others remain fixed. However, its convergence is only guaranteed under restrictive assumptions [17]. Therefore, alternative strategies based on proximal tools have been proposed which benefit from sounder convergence properties, particularly in the nonconvex setting. They consist of replacing, at each iteration, the minimization step by either a (single) proximal step [29, 18] or a forward-backward step [30, 31, 32], giving rise, respectively, to the so-called proximal (resp. forward-backward) alternating algorithms. Here, we propose a novel alternating algorithm for the resolution of Problem (2) that can be viewed as a hybrid solution between proximal and forward-backward alternating methods.

In order to present our alternating minimization strategy, let us rewrite the cost function (3) as follows:

$$(\forall \boldsymbol{x} \in \mathbb{R}^{TN})(\forall h \in \mathbb{R}^{P}) \quad F(\boldsymbol{x}, h) = f_1(\boldsymbol{x}, h) + f_2(\boldsymbol{x}) + \Theta(h), \tag{23}$$

where f_1 represents the smooth part of $\Phi(\boldsymbol{x}, h) + \Psi(\boldsymbol{x})$ and f_2 its nonsmooth part taking the separable form:

$$(\forall \boldsymbol{x} \in \mathbb{R}^{TN}) \quad f_2(\boldsymbol{x}) = \sum_{t=1}^{T} f_{2,t}(x_t). \tag{24}$$

Two cases arise depending on the selected regularization function ψ:

- ψ is nonsmooth, e.g., in case of TV, SLTV, TGV and TVSG, then

$$\begin{aligned} (\forall \boldsymbol{x} \in \mathbb{R}^{TN}), (\forall h \in \mathbb{R}^{P}) \quad & f_1(\boldsymbol{x}, h) = \Phi(\boldsymbol{x}, h) + \mathcal{M}(\boldsymbol{x}), \\ (\forall \boldsymbol{x} \in \mathbb{R}^{TN})(\forall t \in \{1, \ldots, T\}) \quad & f_{2,t}(x_t) = \eta\, \psi(x_t) + \iota_{[x_{\min}, x_{\max}]^N}(x_t). \end{aligned}$$

- ψ is smooth, e.g., it corresponds to the nonconvex regularizations, SOOT-TV, log-TV and Welsch-TV, then

$$\begin{aligned} (\forall \boldsymbol{x} \in \mathbb{R}^{TN}), (\forall h \in \mathbb{R}^{P}) \quad & f_1(\boldsymbol{x}, h) = \Phi(\boldsymbol{x}, h) + \eta \sum_{t=1}^{T} \psi(x_t) + \mathcal{M}(\boldsymbol{x}), \\ (\forall \boldsymbol{x} \in \mathbb{R}^{TN})(\forall t \in \{1, \ldots, T\}) \quad & f_{2,t}(x_t) = \iota_{[x_{\min}, x_{\max}]^N}(x_t). \end{aligned}$$

Function (23) is then minimized with the proximal-based alternating minimization strategy presented in [32]. Before stating the algorithm, let us first define the notion of proximity operator relative to a metric.

Definition 1 *Let $f\colon \mathbb{R}^N \to]-\infty, +\infty]$ be a convex, proper, lower semicontinuous function, let $A \in \mathbb{R}^{N \times N}$ be a symmetric positive definite matrix. For every $x \in \mathbb{R}^N$, the minimization problem*

$$\underset{z \in \mathbb{R}^N}{\text{minimize}} \;\; f(z) + \frac{1}{2}\|x - z\|_A^2, \tag{25}$$

admits a unique solution, which is denoted by $\mathrm{prox}_{A,f}(x)$. The so-defined operator $\mathrm{prox}_{A,f}\colon \mathbb{R}^N \to \mathbb{R}^N$ is the proximity operator *of f relative to the metric induced by A.*

Hereabove, $\|\cdot\|_A$ denotes the weighted norm defined by $\|.\|_A^2 = \langle\cdot, A\cdot\rangle$, where $\langle\cdot,\cdot\rangle$ is the usual Euclidean scalar product. Note that, if A is the identity matrix, one recovers the usual proximity operator $\mathrm{prox}_f \colon \mathbb{R}^N \to \mathbb{R}^N$, which is at the core of numerous convex optimization algorithms (see [33, 34, 35] for tutorials and use for multicomponent image processing).[1] We are now ready to provide Algorithm 1 for the minimization of function F:

Algorithm 1 Blind video deconvolution

```
**Initialization:**
For every $k \in \mathbb{N}$, $\left(\gamma_t^k\right)_{1\leqslant t\leqslant T} \in ]0,2[$ and $\mu^k \in ]0,+\infty[$
**for** $k = 0, 1, \ldots$ **do**
    **for** $t = 1, \ldots, T$ **do**
        $\check{\boldsymbol{x}}^{t,k} = \left(x_1^{k+1}, \ldots, x_{t-1}^{k+1},\ x_t^k\ , x_{t+1}^k \ldots, x_T^k\right)$
        $\widetilde{x}_t^k = x_t^k - \gamma_t^k A_{t,k}^{-1}\left(\nabla_{x_t} f_1(\check{\boldsymbol{x}}^{t,k}, h^k)\right)$
        $x_t^{k+1} = \mathrm{prox}_{(\gamma_t^k)^{-1}A_{t,k}, f_{2,t}}\left(\widetilde{x}_t^k\right)$
    **end for**
    $h^{k+1} = \mathrm{prox}_{\mu^k\,(\Theta+\Phi(\boldsymbol{x}^{k+1},\cdot))}(h^k)$
**end for**
```

Hereabove, $\nabla_{x_t} f_1$ denotes the gradient of f_1 with respect to frame x_t. The update rule on each image x_t with $t \in \{1, \ldots, T\}$ corresponds to a forward-backward iteration, alternating a partial gradient descent step with respect to x_t on the smooth part f_1 and a proximal step on the nonsmooth part $f_{2,t}$, preconditioned by a positive definite matrix $A_{t,k} \in \mathbb{R}^{N\times N}$ at iteration k. This matrix is chosen so as to satisfy a majorizing condition for f_1 at $\check{\boldsymbol{x}}^{t,k}$, i.e. [36]:

$$\begin{aligned}\left(\forall x \in \mathbb{R}^N\right) \quad Q(x, \check{\boldsymbol{x}}^{t,k}) &= f_1(\check{\boldsymbol{x}}^{t,k}) + \langle x - x_t^k \mid \nabla_{x_t} f_1(\check{\boldsymbol{x}}^{t,k})\rangle + \frac{1}{2}\|x - x_t^k\|_{A_{t,k}}^2, \\ &\geqslant f_1(x_1^{k+1}, \ldots, x_{t-1}^{k+1}, x, x_{t+1}^k, \ldots, x_T^k). \end{aligned} \tag{26}$$

After each image has been updated, a proximal step is applied on the restriction of F to the kernel variable h while all the images of the sequence $\boldsymbol{x}$ are kept

[1] See also http://proximity-operator.net.

unchanged. The proposed hybrid scheme handles the minimization with respect to the kernel in a simple and stable manner via a proximal step. The update rule for the images accounts explicitly for the presence of smooth and non-smooth terms in the regularization function, and is expected to have a good numerical efficiency thanks to the preconditioning strategy based on the aforementioned majorization principle. We will show in our convergence analysis in Section 3.5 that Algorithm 1 can actually be viewed as a special instance of the alternating proximal scheme from [32] for an appropriate choice of the underlying metrics.

3.2. Construction of the majorant

The choice of a good majorant function $Q(\cdot, \check{\boldsymbol{x}}^{t,k})$ of the restriction of f_1 to the image x_t at each iteration $k \in \mathbb{N}$ has a strong leverage on the numerical performance of the proposed method. Thereby, one has to favor curvature matrices $(A_{t,k})_{1 \leqslant t \leqslant T, k \in \mathbb{N}}$ that are easy to handle.
Depending on the choice of the spatial regularization ψ, $(A_{t,k})_{1 \leqslant t \leqslant T, k \in \mathbb{N}}$ are defined as described below (proofs are given in appendix).

1. *ψ is non smooth* e.g., in case of TV, SLTV, TGV, and TVSG

   $$(\forall t \in \{1,\ldots,T\})(\forall k \in \mathbb{N}) \quad A_{t,k} = \upsilon_{t,k}\,\mathrm{Id}_N, \tag{27}$$

   where $\upsilon_{t,k}$ is a Lipschitz constant of the gradient of f_1 with respect to x_t^k:

   $$\nabla\left(\frac{1}{2}\|h^k * \cdot - y_t\|^2 + \mathcal{M}(x_1^{k+1},\ldots,x_{t-1}^{k+1},\cdot,x_{t+1}^k,\ldots,x_T^k)\right).$$

   According to (7), such a Lipschitz constant is thus expressed as

   $$\upsilon_{t,k} = \|H_k\|^2 + \sum_{\ell \in \mathcal{V}_t} \beta_{\ell,t} + \sum_{\ell \in \{1,\ldots,T\}: t \in \mathcal{V}_\ell} \beta_{t,\ell}\|M_{t\to\ell}\|^2. \tag{28}$$

   where $\|H_k\|$ is maximum magnitude of the frequency response of the blur filter estimate at iteration k and $(\|M_{t\to\ell}\|)_{1 \leqslant \ell \leqslant T, t \in \mathcal{V}_\ell}$ denote the spectral norms of the operators used for motion compensation based on x_t^k. Expressions of these norms have been derived in [19, Section 5.2.2] for motion estimation with fractional accuracy. Note that, for every $k \in \mathbb{N}$, $\upsilon_{t,k}$ is lower bounded by a positive constant, namely $\sum_{\ell \in \mathcal{V}_t} \beta_{\ell,t}$.[2]

2. ψ *is SOOT-TV*

$$(\forall t \in \{1,\dots,T\})(\forall k \in \mathbb{N}) \;\; A_{t,k} = \left(\upsilon_{t,k} + \frac{9\eta \, \|D\|^2}{8\lambda^2}\right) \mathrm{Id}_N + \frac{\eta \, A_{\ell_{1,\alpha}}(\check{\boldsymbol{x}}^{t,k})}{\ell_{1,\alpha}(D\check{\boldsymbol{x}}^{t,k}) + \beta}, \tag{29}$$

Hereabove,

$$(\forall z \in \mathbb{R}^N) \quad A_{\ell_{1,\alpha}}(z) = \mathrm{Diag}\left(\Omega^\top s(z)\right) \tag{30}$$

where, for every $z \in \mathbb{R}^N$, $s(z) \in \mathbb{R}^{2N}$ is such that, for every $i \in \{1,\dots,2N\}$, its i-th component is given by

$$s^{(i)}(z) = ((D^{(i)}z)^2 + \alpha^2)^{-1/2}, \tag{31}$$

and the $(i,j) \in \{1,\dots,2N\}^2$ element of Ω is

$$\Omega^{(i,j)} = \left|D^{(i,j)}\right| \sum_{l=1}^{N} \left|D^{(i,l)}\right|. \tag{32}$$

3. ψ *is log-TV*

$$(\forall t \in \{1,\dots,T\})(\forall k \in \mathbb{N}) \quad A_{t,k} = \upsilon_{t,k}\mathrm{Id}_N + \eta \, A_{\log}(\check{\boldsymbol{x}}^{t,k}), \tag{33}$$

where

$$(\forall z \in \mathbb{R}^N) \quad A_{\log}(z) = \mathrm{Diag}\left(\Omega^\top s(z)\right), \tag{34}$$

[2] If no motion compensation is performed, a positive constant $\epsilon > 0$ must be added to the right-hand side of (28) to guarantee the existence of such a bound independent of k. The same correction should be made in subsequent expressions of $A_{t,k}$.

for every $z \in \mathbb{R}^N$, $s(z) = \big(s_i(z)\big)_{1 \leqslant i \leqslant N} \in \mathbb{R}^N$ is such that

$$(\forall i \in \{1, \dots, N\}) \qquad s^{(i)}(z) = \big((\nabla_{\mathsf{H}}^{(i)} z)^2 + (\nabla_{\mathsf{V}}^{(i)} z)^2 + \alpha^2\big)^{-1}, \tag{35}$$

and Ω is the matrix $\Omega_{\mathsf{H}} + \Omega_{\mathsf{V}}$ whose $(i,j) \in \{1, \dots, N\}^2$ elements are expressed as

$$\Omega_{\mathsf{H}}^{(i,j)} = \left|\nabla_{\mathsf{H}}^{(i,j)}\right| \sum_{l=1}^{N} \left|\nabla_{\mathsf{H}}^{(i,l)}\right| \;,\; \Omega_{\mathsf{V}}^{(i,j)} = \left|\nabla_{\mathsf{V}}^{(i,j)}\right| \sum_{l=1}^{N} \left|\nabla_{\mathsf{V}}^{(i,l)}\right|. \tag{36}$$

4. *ψ is Welsch-TV*

$$(\forall t \in \{1, \dots, T\})(\forall k \in \mathbb{N}) \quad A_{t,k} = \upsilon_{t,k} \mathrm{Id}_N + \eta\, A_w(\breve{\boldsymbol{x}}^{t,k}), \tag{37}$$

where

$$(\forall z \in \mathbb{R}^N) \quad A_w(z) = \sigma^{-2} \mathrm{Diag}\,\big(\Omega^\top s(z)\big) \tag{38}$$

and, for every $z \in \mathbb{R}^N$, $s(z) \in \mathbb{R}^N$ is such that

$$(\forall i \in \{1, \dots, N\}) \qquad s^{(i)}(z) = \exp\Big(-\big((\nabla_{\mathsf{H}}^{(i)} z)^2 + (\nabla_{\mathsf{V}}^{(i)} z)^2\big)/(2\sigma^2)\Big)\,, \tag{39}$$

and $\Omega = \Omega_{\mathsf{H}} + \Omega_{\mathsf{V}}$, the (i,j) elements of Ω_{H} and Ω_{V} being expressed by (36).

3.3. Implementation of the proximity operator of $f_{2,t}$

The retained metric matrices $(A_{t,k})_{1 \leqslant t \leqslant T, k \in \mathbb{N}}$ being diagonal, the proximity operators involved in Algorithm 1 may have a closed form expression when $(f_2)_{1 \leqslant t \leqslant T}$ are “simple” functions. However, when the latter are more sophisticated, for example when they represent sum of functions possibly composed with linear operators, we have to resort to iterative strategies in order to evaluate it. In our framework, the computation of the proximity operator of $f_{2,t}$ for each $t \in \{1, \ldots, T\}$ (i.e., $\mathrm{prox}_{(\gamma_t^k)^{-1} A_{t,k}, f_{2,t}}$) depends on the choice of the spatial regularization function ψ. In some instances, it has an explicit form while in others, we must use specific algorithms to evaluate it, namely:

- for smooth nonconvex regularization functions ψ (i.e., SOOT-TV, log-TV and Welsch-TV), we have

  $$(\forall t \in \{1, \ldots, T\})(\forall z \in \mathbb{R}^N) \quad f_{2,t}(z) = \iota_{[x_{\min}, x_{\max}]^N}(z),$$

  so that the proximity operator has an explicit expression, since it reduces to compute scalar projections onto $[x_{\min}, x_{\max}]$.

- for nonsmooth convex regularization functions ψ (i.e., TV, SLTV, TGV, TVSG), we have

  $$(\forall t \in \{1, \ldots, T\})(\forall z \in \mathbb{R}^N) \quad f_{2,t}(z) = \psi(z) + \iota_{[x_{\min}, x_{\max}]^N}(z),$$

  and the proximity operator has been evaluated by using the following algorithms:

TV	SLTV	TGV	TVSG
Dual forward-backward [37]	Dual block-coordinate forward-backward [19]	Primal-dual splitting [38]	Alternating proximal gradient [39]

Table 1: List of optimization algorithms used for computing the proximity operator with respect to the different convex regularization functions.

3.4. Implementation of the proximity operator for kernel estimation

The blur kernel h is estimated in Algorithm 1 by computing the proximity operator of the sum of the data fidelity term and regularization function Θ (i.e., $\mathrm{prox}_{\mu^k\,(\Theta+\Phi(\boldsymbol{x}^{k+1},\cdot))}$). Since there is no closed form expression for the latter proximity operator, we resort to the parallel proximal algorithm (PPXA) in [40] to evaluate it.

3.5. Convergence analysis

The convergence properties of Algorithm 1 depend on the settings of parameters $(\gamma_t^k,\, \mu^k)_{1\leqslant t\leqslant T, k\in\mathbb{N}}$ and on the choice for the preconditioning matrices $(A_{t,k})_{1\leqslant t\leqslant T, k\in\mathbb{N}}$. First, let us state the following proposition related to the quadratic form of the data fidelity term Φ.

Proposition 3.1 *Let us define the symmetric positive definite matrix*

$$B = \mu X^\top X + \mathrm{Id}_P, \tag{40}$$

where Id_P *is the identity matrix of* $\mathbb{R}^P$*,* $\mu \in]0, +\infty[$ *and* $X \in \mathbb{R}^{TN\times P}$ *is such that, for every* $h \in \mathbb{R}^P$*,* $Xh = (h * x_t)_{1\leqslant t\leqslant T}$*. Then, for every* $h \in \mathbb{R}^P$ *and* $\boldsymbol{x} \in \mathbb{R}^{TN}$*,*

$$\mathrm{prox}_{\mu\,(\Theta+\Phi(\boldsymbol{x},\cdot))}\,(h) = \mathrm{prox}_{\mu^{-1}B,\Theta}(h - \mu B^{-1}\nabla_h \Phi(\boldsymbol{x}, h)), \tag{41}$$

and, for every $h' \in \mathbb{R}^P$*,*

$$\Phi(\boldsymbol{x}, h') + \nabla_{h'}\Phi(\boldsymbol{x}', h')^\top (h - h') + \frac{1}{2}\|h - h'\|^2_{\mu^{-1}B} \geqslant \Phi(\boldsymbol{x}, h). \tag{42}$$

Proof. Let $q \in \mathbb{R}^P$ be the value of the proximity operator of $\Phi(\boldsymbol{x}, \cdot) + \Theta$ at h, i.e., $q = \mathrm{prox}_{\mu(\Theta+\Phi(\boldsymbol{x},\cdot))}\,(h)$. We have the following subdifferential inclusion:

$$\begin{aligned}
&h - q \in \mu\,(\partial\,\Theta(q) + \nabla_q \Phi(\boldsymbol{x}, q))\\
\Leftrightarrow\ &h - q \in \mu\,\partial\,\Theta(q) + \mu\,X^\top(Xq - \boldsymbol{y})\\
\Leftrightarrow\ &h - (\mathrm{Id}_P + \mu\,X^\top X)q + \mu\,X^\top \boldsymbol{y} \in \mu\,\partial\,\Theta(q)\\
\Leftrightarrow\ &(\mathrm{Id}_P + \mu\,X^\top X)^{-1}(h + \mu\,X^\top \boldsymbol{y}) - q \in \mu\,(\mathrm{Id}_P + \mu\,X^\top X)^{-1}\partial\,\Theta(q) \qquad (43)
\end{aligned}$$

where $\partial\,\Theta(q)$ denotes the subdifferential of Θ at q. [3] Thus, by setting $B =$

[3] In our case, it is the normal cone to $\mathcal{H}$ at q, but the result is valid for any proper convex lower-semicontinuous function Θ.

$\mathrm{Id}_P + \mu X^\top X$,

$$
\begin{aligned}
& q = \mathrm{prox}_{\mu^{-1}\,B,\Theta}\,(B^{-1}(h + \mu\, X^\top \boldsymbol{y})) \\
\Leftrightarrow\ & q = \mathrm{prox}_{\mu^{-1}\,B,\Theta}\,(h - \mu\, B^{-1} X^\top (Xh - \boldsymbol{y})) \\
\Leftrightarrow\ & q = \mathrm{prox}_{\mu^{-1}\,B,\Theta}\,(h - \mu\, B^{-1} \nabla_h \Phi(x,h)). \qquad (44)
\end{aligned}
$$

Because of the quadratic form of Φ,

$$\Phi(\boldsymbol{x},h') + \nabla_{h'}\Phi(\boldsymbol{x}',h')^\top (h-h') + \frac{1}{2}(h-h')^\top \nabla^2_{h'}\Phi(\boldsymbol{x}',h')(h-h') = \Phi(\boldsymbol{x},h), \quad (45)$$

where the Hessian of $\Phi(\boldsymbol{x}',\cdot)$ is

$$\nabla^2_{h'}\Phi(\boldsymbol{x}',h') = X^\top X \preceq X^\top X + \mu^{-1}\mathrm{Id}_P = \mu^{-1} B \quad (46)$$

($\preceq$ stands for the Loewner order). This yields (42). □

This allows us to derive the following convergence result:

Theorem 3.2 *Let us consider Algorithm 1. Assume that*

$$
\begin{aligned}
& (\forall t \in \{1,\ldots,T\}) \qquad 0 < \inf_{k\in\mathbb{N}} \gamma_t^k \quad \text{and} \quad \sup_{k\in\mathbb{N}} \gamma_t^k < 2, \qquad (47) \\
& 0 < \inf_{k\in\mathbb{N}} \mu^k \quad \text{and} \quad \sup_{k\in\mathbb{N}} \mu^k < +\infty. \qquad (48)
\end{aligned}
$$

Then, the sequence $(\boldsymbol{x}^k, h^k)_{k\in\mathbb{N}}$ converges to a critical point $(\widehat{\boldsymbol{x}}, \widehat{h})$ of F. Moreover, $\big(F(\boldsymbol{x}^k,h^k)\big)_{k\in\mathbb{N}}$ is a nonincreasing sequence converging to $F(\widehat{\boldsymbol{x}}, \widehat{h})$.

Proof. It follows from Proposition 3.1 that the proximal step for kernel estimation in Algorithm 1 at iteration $k \in \mathbb{N}$ reduces to a preconditioned forward-backward iteration with the preconditioning matrix $(\mu^k)^{-1} B_k$ where $B_k = \mu^k X_k^\top X_k + \mathrm{Id}_P$ and $X_k \in \mathbb{R}^{TN\times P}$ is such that, for every $h \in \mathbb{R}^P$, $X_k h = (h * x_t^k)_{1\leqslant t\leqslant T}$. Algorithm 1 thus appears as a special case of the block-coordinate variable metric forward-backward algorithm studied in [32] where $T+1$ blocks of variables are involved (corresponding to the T frames and the kernel to be estimated). Indeed, the cost function (23) satisfies the assumptions required in [32]:

- it is a coercive function (since both variables $\boldsymbol{x}$ and h are constrained to belong to compact sets) and it satisfies Kurdyka-Łojasiewicz inequality[4];
- f_1 is a function with a Lipschitz continuous gradient;
- the function $(\boldsymbol{x}, h) \mapsto f_2(\boldsymbol{x}) + \Theta(h)$ is a proper convex lower-semicontinuous function which is separable with respect to the blocks of variables;
- according to Section 3.2 and Equation (42), the curvature matrices $(A_{t,k})_{1 \leqslant t \leqslant T}$ and $(\mu^k)^{-1} B_k$ used at each iteration $k \in \mathbb{N}$ provide quadratic majorant approximations to the restriction of f_1 to the current activated block.

In addition, since $(x_t^k)_{1 \leqslant t \leqslant T, k \in \mathbb{N}}$ and $(h^k)_{k \in \mathbb{N}}$ are constrained to belong to bounded sets, it follows from the expressions derived in Section 3.2 and the positive lower bound already exhibited on $(\upsilon_{t,k})_{1 \leqslant t \leqslant T, k \in \mathbb{N}}$ that, for every $t \in \{1, \ldots, T\}$, there exists $(a_{t,\min}, a_{t,\max}) \in]0, +\infty[^2$ such that

$$(\forall k \in \mathbb{N}) \qquad a_{t,\min} \mathrm{Id}_N \preceq A_{t,k} \preceq a_{t,\max} \mathrm{Id}_N. \tag{49}$$

According to (48), there also exists $(b_{\min}, b_{\max}) \in]0, +\infty[^2$ such that

$$(\forall k \in \mathbb{N}) \qquad b_{\min} \mathrm{Id}_P \preceq (\mu^k)^{-1} B_k = X_k^\top X_k + (\mu^k)^{-1} \mathrm{Id}_P \preceq b_{\max} \mathrm{Id}_P. \tag{50}$$

The convergence result then follows from [32, Theorem 3.1]. □

[4] KL inequality allows the regularity of a function to be quantified around its stationary points [41]. It is satisfied for a wide class of functions, in particular real analytic and semi-algebraic functions.

4. Experimental results

We assess in this section the performance of the proposed approach on synthetic and real video sequences. We first start our evaluation on $T = 20$ frames of the synthetic video sequences **Foreman** and **Claire** of sizes $N = 352 \times 288$ and $N = 360 \times 288$, sourced from http://media.xiph.org/video/derf/. The sequences have been blurred using the four convolution kernels displayed in Figure 3, and degraded by white Gaussian noise with zero mean, and variance equal to 2. We also apply the proposed method to the $T = 20$ frames of the real blurred video sequences **Tachan** and **Au théâtre ce soir** of size $N = 720 \times 576$ supplied by the French National Audiovisual Institute (INA).

Since it is usually challenging to develop a blind deconvolution method that achieves satisfactory results for recovering both the blur kernel and the video contents in a single step, we propose to apply successively our method in two stages: First, a blind deconvolution step is performed which aims at identifying the blur kernel from the input degraded sequence using Algorithm 1. Second, we apply a non-blind (supervised) deconvolution step where the observed degraded sequence and the identified kernel from the first step are both employed to reach an improved estimate for the sought restored video sequence. To this aim, we resort to the same optimization strategy than in Algorithm 1 (possibly with other regularizers), where we omit the kernel updates, which results in the simplified Algorithm 2.

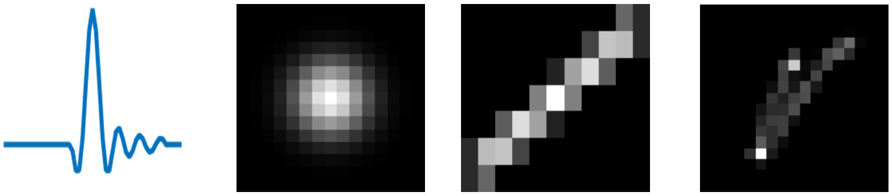
The algorithm parameters are set as follows: for every $t \in \{1, \dots, T\}$ and $k \in \mathbb{N}$, $\gamma_t^k = 1.9$ and $\mu^k = 1$. In all tested methods, the inner loops for computing the proximity operator with respect to each image (resp. to the kernel) are stopped as soon as the relative difference in norms between consecutive iterates is below 10^{-4} (resp. 10^{-6}). Finally, Algorithm 1 (or its simplified version Algorithm 2) is run until the following relative decrease condition on the cost function holds:

$$|F(\boldsymbol{x}^k, h^k) - F(\boldsymbol{x}^{k-1}, h^{k-1})| \leqslant 10^{-5} |F(\boldsymbol{x}^k, h^k)|.$$

Algorithm 2 Non-blind video deconvolution

```
Initialization:
Let $h \in \mathbb{R}^P$ be the identified kernel
For every $k \in \mathbb{N}$, $\left(\gamma_t^k\right)_{1 \leqslant t \leqslant T} \in ]0,2[$
for $k = 0, 1, \dots$ do
    for $t = 1, \dots, T$ do
        $\breve{\boldsymbol{x}}^{t,k} = \left(x_1^{k+1}, \dots, x_{t-1}^{k+1},\ x_t^k\ , x_{t+1}^k \dots, x_T^k\right)$
        $\widetilde{x}_t^k = x_t^k - \gamma_t^k A_{t,k}^{-1} \left(\nabla_{x_t} f_1(\breve{\boldsymbol{x}}^{t,k}, h)\right)$
        $x_t^{k+1} = \operatorname{prox}_{(\gamma_t^k)^{-1} A_{t,k}, f_{2,t}} \left(\widetilde{x}_t^k\right)$
    end for
end for
```

Note that, although our convergence analysis does not deal with the case of an inexactly computed proximity operator, for such small tolerance, a good convergence of the approach was observed in practice. Regarding the temporal regularization term $\mathcal{M}$, the closest neighboring frames such that $|\ell - t| = 1$ have been taken into account. The motion matrices $(M_{\ell \to t})_{\ell,t}$ have been estimated from the degraded sequence $\boldsymbol{y}$, using the optical flow estimation algorithm from [42]. It should be noted that all the experiments for both blind and non-blind steps are initialized with the Dirac delta function for the kernel identification step, and the input degraded video sequence.



(a) $P = 1 \times 53$ (b) $P = 15 \times 15$ (c) $P = 7 \times 11$ (d) $P = 17 \times 17$

Figure 3: Synthetic convolution kernels.

4.1. Blind video deconvolution step

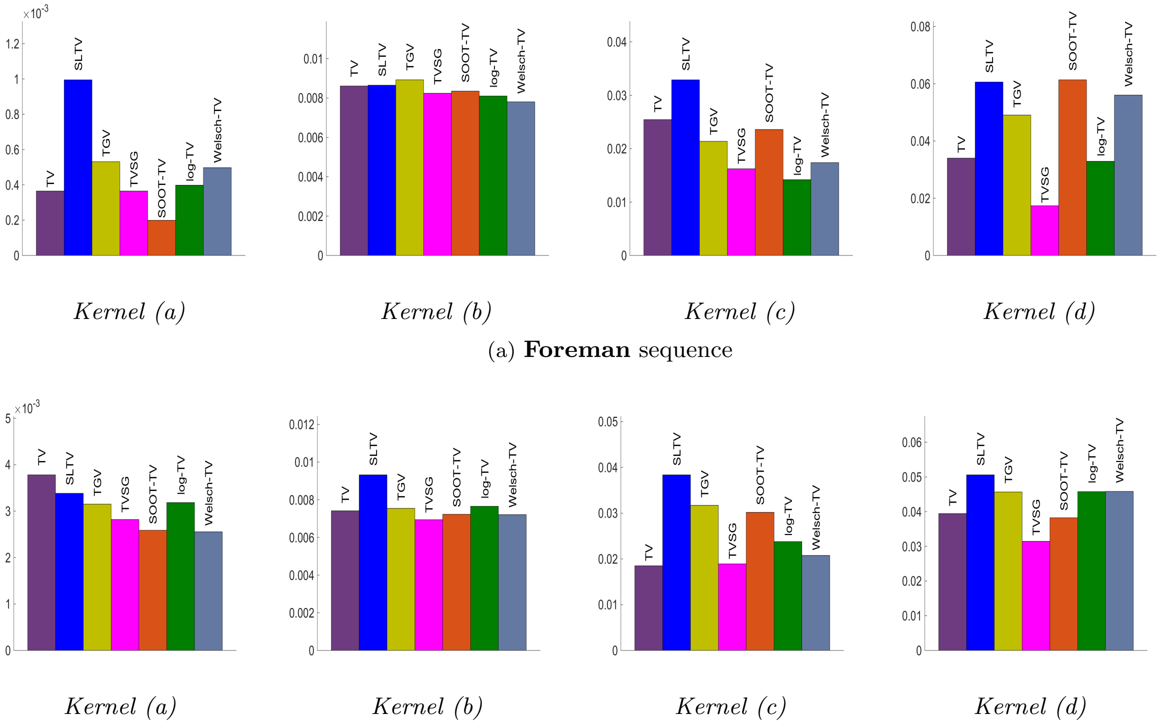
Figure 4 shows the quadratic error in kernel identification for the seven regularization approaches presented in Section 2.2, and the four tested convolution kernels. This error is evaluated as follows

$$\text{Error} = \|h - \widehat{h}\|^2, \tag{51}$$

where $\widehat{h} \in \mathbb{R}^P$ denotes the ground truth kernel and $h \in \mathbb{R}^P$ is the estimated one. Since the ground truth is available in this set of experiments, we are able to adjust the parameters involved in the regularization terms in order to obtain the smallest possible error.

We observe in Figure 4 that the results vary slightly depending on the kernels and video sequences. We can notice that the TVSG achieves low errors regardless of the kernel and the video sequence. The nonconvex regularizations also produce low errors on kernel identification in certain cases (e.g., kernels (a) and (c) with **Foreman** sequence, and kernel (a) with **Claire** sequence). Nevertheless, they can also fail in identifying the correct kernel possibly because of the existence of spurious local minima. In addition, they are harder to adjust since they involve multiple parameters, that may be quite different from a degraded sequence to another (see Table 3). Note that the worst results in terms of kernel identification are usually achieved by the SLTV regularization. This may be due to the fact that the latter relies on second order derivative operators, that tend to over-smooth the images and particularly the edges, resulting in an unsatisfactory kernel identification.
It should be noted that the convex regularizations SLTV and TVSG are approximately 5 times slower than TV, while TGV is about 10 times slower. Besides, the nonconvex regularizations are comparable with TV in terms of computational cost by iteration. Moreover, SOOT-TV leads to a slower convergence compared with the other nonconvex penalties.

Table 2 illustrates the gap between the best and worst identification qual-



(a) **Foreman** sequence

(b) **Claire** sequence

Figure 4: Performance in terms of error on kernel identification with respect to the different regularizations and blur kernels, from left to right: TV, SLTV, TGV, TVSG, SOOT-TV, log-TV, Welsch-TV.

ity scores with respect to the two synthetic sequences and the four convolution kernels. We also provide the results obtained by applying the 2D blind deconvolution approach from [15] on each frame of the degraded video, with parameters optimized manually in order to get the best kernel estimate. The averaged value for the kernel estimation error is provided for each tested scenario. One can notice that our strategy allows us to obtain very competitive results, in particular for the challenging kernels (a) and (d).

Sequences		Kernel (a)	Kernel (b)	Kernel (c)	Kernel (d)
Foreman	Worst	9.94×10^{-4} *(SLTV)*	8.92×10^{-3} *(TGV)*	3.28×10^{-2} *(SLTV)*	6.13×10^{-2} *(SOOT-TV)*
	Best	1.98×10^{-4} *(SOOT-TV)*	7.8×10^{-3} *(Welsch-TV)*	1.41×10^{-2} *(log-TV)*	1.73×10^{-2} *(TVSG)*
	[15]	6.1×10^{-1}	9×10^{-2}	1.9×10^{-1}	7.5×10^{-1}
Claire	Worst	3.78×10^{-3} *(TV)*	9.32×10^{-3} *(SLTV)*	3.83×10^{-2} *(SLTV)*	5.05×10^{-2} *(SLTV)*
	Best	2.55×10^{-3} *(Welsch-TV)*	6.93×10^{-3} *(TVSG)*	1.84×10^{-2} *(TV)*	3.14×10^{-2} *(TVSG)*
	[15]	5.9×10^{-1}	5.1×10^{-2}	1.8×10^{-1}	3.6×10^{-1}

Table 2: Kernel identification scores using our method (for worst and best regularization choices) and the approach from [15].

We would like to point out that in the considered nonconvex context, the performance of a method may result from an interplay between the variational model, the optimization algorithm as well as the regularizer.

4.2. Non-blind video deconvolution step

Let us now investigate the performance of the various tested regularization strategies in the second stage based on non-blind video deconvolution. We make use of the estimated kernel with the lowest error at the output of the blind step, to perform a non-blind deconvolution of the input degraded sequence. Figures 5 and 6 illustrate the restoration quality in terms of SNR and MOVIE, for the different regularization approaches. The MOVIE indicator, introduced in [43], is a perceptual measure mimicking the human subjective judgment to assess the visual quality of a natural video, the smaller value MOVIE is, the better the video restoration result is.

The performance in terms of images quality are more stable than those concerning the errors on kernel identification. High SNR scores are usually obtained using the TGV and TVSG regularizations, and in some cases by TV and log-TV. Moreover, the lowest values of SNR are usually obtained with the

	Foreman				Claire			
	Kernel (a)	Kernel (b)	Kernel (c)	Kernel (d)	Kernel (a)	Kernel (b)	Kernel (c)	Kernel (d)
TV	$\eta = 6.24 \times 10^{-4}$ $\beta_{\ell\to t} = 6.76 \times 10^{-4}$	$\eta = 9.2 \times 10^{-3}$ $\beta_{\ell\to t} = 1.53 \times 10^{-3}$	$\eta = 1.52 \times 10^{-2}$ $\beta_{\ell\to t} = 1.33 \times 10^{-3}$	$\eta = 2.82 \times 10^{-2}$ $\beta_{\ell\to t} = 8.05 \times 10^{-4}$	$\eta = 7.53 \times 10^{-3}$ $\beta_{\ell\to t} = 3.13 \times 10^{-3}$	$\eta = 1.60 \times 10^{-2}$ $\beta_{\ell\to t} = 7.02 \times 10^{-3}$	$\eta = 1.52 \times 10^{-2}$ $\beta_{\ell\to t} = 1.10 \times 10^{-3}$	$\eta = 2.90 \times 10^{-1}$ $\beta_{\ell\to t} = 9.32 \times 10^{-2}$
SLTV	$\eta = 6.17 \times 10^{-5}$ $\beta_{\ell\to t} = 2.17 \times 10^{-3}$	$\eta = 3.99 \times 10^{-4}$ $\beta_{\ell\to t} = 2.01 \times 10^{-2}$	$\eta = 2.83 \times 10^{-3}$ $\beta_{\ell\to t} = 6.97 \times 10^{-3}$	$\eta = 2.08 \times 10^{-3}$ $\beta_{\ell\to t} = 7.48 \times 10^{-3}$	$\eta = 3.51 \times 10^{-4}$ $\beta_{\ell\to t} = 1.57 \times 10^{-3}$	$\eta = 2.34 \times 10^{-3}$ $\beta_{\ell\to t} = 7.59 \times 10^{-3}$	$\eta = 1.44 \times 10^{-3}$ $\beta_{\ell\to t} = 3.89 \times 10^{-2}$	$\eta = 2.34 \times 10^{-3}$ $\beta_{\ell\to t} = 1.04 \times 10^{-2}$
TGV $\alpha_0 = 1 - \alpha_1$	$\eta = 1.99 \times 10^{-3}, \alpha_1 = 0.63$ $\beta_{\ell\to t} = 4.95 \times 10^{-3}$	$\eta = 5.50 \times 10^{-3}, \alpha_1 = 0.77$ $\beta_{\ell\to t} = 2.08 \times 10^{-2}$	$\eta = 3.32 \times 10^{-1}, \alpha_1 = 0.68$ $\beta_{\ell\to t} = 2.12 \times 10^{-3}$	$\eta = 1.34, \alpha_1 = 0.51$ $\beta_{\ell\to t} = 8.65 \times 10^{-4}$	$\eta = 9.36 \times 10^{-3}, \alpha_1 = 0.79$ $\beta_{\ell\to t} = 1.10 \times 10^{-2}$	$\eta = 5.46 \times 10^{-2}, \alpha_1 = 0.92$ $\beta_{\ell\to t} = 1.32 \times 10^{-2}$	$\eta = 8.13 \times 10^{-2}, \alpha_1 = 0.32$ $\beta_{\ell\to t} = 1.70 \times 10^{-3}$	$\eta = 2.44 \times 10^{-2}, \alpha_1 = 0.86$ $\beta_{\ell\to t} = 2.16 \times 10^{-2}$
TVSG	$\eta = 2.52 \times 10^{-3}$ $\beta_{\ell\to t} = 1.18 \times 10^{-4}$	$\eta = 3.30 \times 10^{-2}$ $\beta_{\ell\to t} = 6.53 \times 10^{-4}$	$\eta = 4.36 \times 10^{-2}$ $\beta_{\ell\to t} = 3.03 \times 10^{-4}$	$\eta = 2.55 \times 10^{-2}$ $\beta_{\ell\to t} = 8.26 \times 10^{-4}$	$\eta = 3.16 \times 10^{-3}$ $\beta_{\ell\to t} = 1.78 \times 10^{-9}$	$\eta = 2.93 \times 10^{-2}$ $\beta_{\ell\to t} = 5.58 \times 10^{-4}$	$\eta = 2.73 \times 10^{-2}$ $\beta_{\ell\to t} = 4.44 \times 10^{-4}$	$\eta = 1.84 \times 10^{-2}$ $\beta_{\ell\to t} = 6.20 \times 10^{-4}$
SOOT-TV $\beta = 1.26 \times 10^{-5}$ $\alpha = 1.2 \times 10^{-3}$	$\eta = 1.48, \lambda = 541.43$ $\beta_{\ell\to t} = 3.55 \times 10^{-4}$	$\eta = 13.65, \lambda = 13$ $\beta_{\ell\to t} = 5.58 \times 10^{-3}$	$\eta = 23.48, \lambda = 7.94$ $\beta_{\ell\to t} = 8.94 \times 10^{-6}$	$\eta = 30.63, \lambda = 11.79$ $\beta_{\ell\to t} = 1.15 \times 10^{-5}$	$\eta = 4.61, \lambda = 11.40$ $\beta_{\ell\to t} = 2.31 \times 10^{-7}$	$\eta = 9.80, \lambda = 14.97$ $\beta_{\ell\to t} = 1.66 \times 10^{-4}$	$\eta = 29.78, \lambda = 297.78$ $\beta_{\ell\to t} = 2.09 \times 10^{-4}$	$\eta = 6.48, \lambda = 9.50$ $\beta_{\ell\to t} = 5.18 \times 10^{-3}$
log-TV	$\eta = 2.59 \times 10^{-5}$ $\alpha = 1.32 \times 10^{-2}$ $\beta_{\ell\to t} = 1.50 \times 10^{-3}$	$\eta = 4.94 \times 10^{-5}$ $\alpha = 1.32 \times 10^{-2}$ $\beta_{\ell\to t} = 1.12 \times 10^{-3}$	$\eta = 7.38 \times 10^{-4}$ $\alpha = 2.79 \times 10^{-3}$ $\beta_{\ell\to t} = 1.59 \times 10^{-5}$	$\eta = 2.41 \times 10^{-4}$ $\alpha = 3.35 \times 10^{-3}$ $\beta_{\ell\to t} = 3.26 \times 10^{-4}$	$\eta = 5.61 \times 10^{-5}$ $\alpha = 1.03 \times 10^{-2}$ $\beta_{\ell\to t} = 1.35 \times 10^{-3}$	$\eta = 7.24 \times 10^{-5}$ $\alpha = 8.44 \times 10^{-3}$ $\beta_{\ell\to t} = 1.08 \times 10^{-3}$	$\eta = 4.17 \times 10^{-4}$ $\alpha = 3.04 \times 10^{-3}$ $\beta_{\ell\to t} = 2.33 \times 10^{-4}$	$\eta = 5.63 \times 10^{-5}$ $\alpha = 1.03 \times 10^{-2}$ $\beta_{\ell\to t} = 1.47 \times 10^{-3}$
Welsch-TV	$\eta = 7.21 \times 10^{-5}$ $\sigma = 5.66 \times 10^{-2}$ $\beta_{\ell\to t} = 3.59 \times 10^{-4}$	$\eta = 1.69 \times 10^{-4}$ $\sigma = 1.61 \times 10^{-2}$ $\beta_{\ell\to t} = 3.11 \times 10^{-4}$	$\eta = 5.05 \times 10^{-4}$ $\sigma = 3.66 \times 10^{-2}$ $\beta_{\ell\to t} = 1.64 \times 10^{-4}$	$\eta = 5.52 \times 10^{-4}$ $\sigma = 3.80 \times 10^{-2}$ $\beta_{\ell\to t} = 1.61 \times 10^{-4}$	$\eta = 2.21 \times 10^{-4}$ $\sigma = 1.92 \times 10^{-2}$ $\beta_{\ell\to t} = 2.92 \times 10^{-4}$	$\eta = 2.04 \times 10^{-4}$ $\sigma = 2.16 \times 10^{-2}$ $\beta_{\ell\to t} = 3.10 \times 10^{-4}$	$\eta = 4.21 \times 10^{-4}$ $\sigma = 1.19 \times 10^{-2}$ $\beta_{\ell\to t} = 4.34 \times 10^{-4}$	$\eta = 2.04 \times 10^{-4}$ $\sigma = 2.46 \times 10^{-2}$ $\beta_{\ell\to t} = 3.20 \times 10^{-4}$

Table 3: Regularization parameters used in the blind deconvolution step.

SLTV regularization and by some nonconvex regularization such as the Welsch-TV (e.g., kernel (a) for **Foreman** and **Claire** sequences). The regularizations achieving the best scores in terms of restoration quality with respect to SNR and MOVIE are displayed in Tables 4 and 5, respectively. The latter emphasizes the good performance of some nonconvex regularizations. Figures 8 and 9 show images extracted from **Foreman** and **Claire** sequences respectively. We present some frames from the degraded sequences and the corresponding restored frames with the best regularizations in terms of SNR. We can observe from the above figures and tables that the proposed method achieves good performance in terms of video restoration quality, where a gain up to 7 dB is obtained.

Sequences		Kernel (a)	Kernel (b)	Kernel (c)	Kernel (d)
Foreman	Degraded	28.72 dB	22.72 dB	21.58 dB	19.88 dB
	Restored	33.60 dB	26.90 dB	26.64 dB	24.83 dB
		(TVSG)	*(TVSG)*	*(TGV)*	*(TV)*
Claire	Degraded	26.99 dB	22.40 dB	20.49 dB	20.04 dB
	Restored	30.84 dB	28.35 dB	27.67 dB	27.17 dB
		(TVSG)	*(log-TV)*	*(log-TV)*	*(TVSG)*

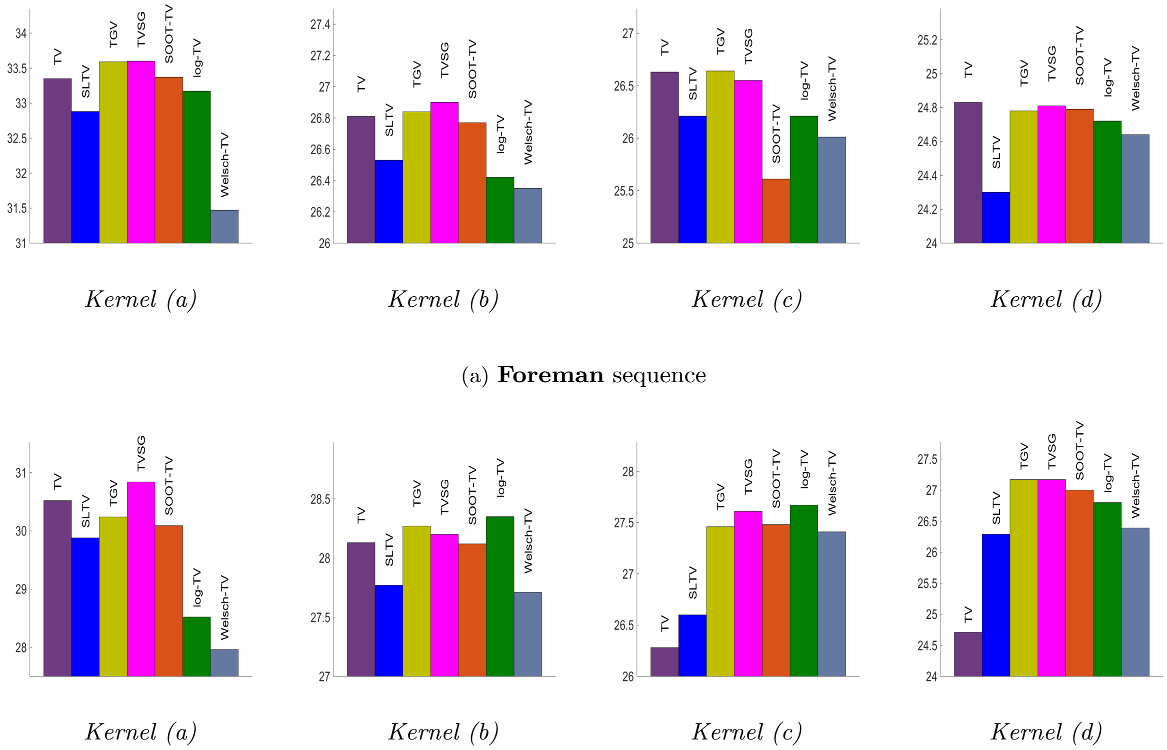
Table 4: Performance of the best non-blind deconvolution methods in terms of SNR.

Sequences		Kernel (a)	Kernel (b)	Kernel (c)	Kernel (d)
Foreman	Degraded	2.03×10^{-4}	$1.58\ \times10^{-3}$	2.41×10^{-3}	3.12×10^{-3}
	Restored	7.7×10^{-5}	4.82×10^{-4}	1.17×10^{-3}	1.83×10^{-3}
		(SOOT-TV)	*(SOOT-TV)*	*(TV)*	*(SOOT-TV)*
Claire	Degraded	2.04×10^{-3}	9.52×10^{-3}	1.18×10^{-2}	1.19×10^{-2}
	Restored	5.28×10^{-4}	2.9×10^{-3}	3.32×10^{-3}	3.67×10^{-3}
		(TVSG)	*(Welsch-TV)*	*(log-TV)*	*(TGV)*

Table 5: Performance of the best non-blind deconvolution methods in terms of Movie.

4.3. Real Data

We have applied our blind deconvolution method to the interlaced real sequences **Tachan** and **Au théâtre ce soir** provided by INA. The odd and even



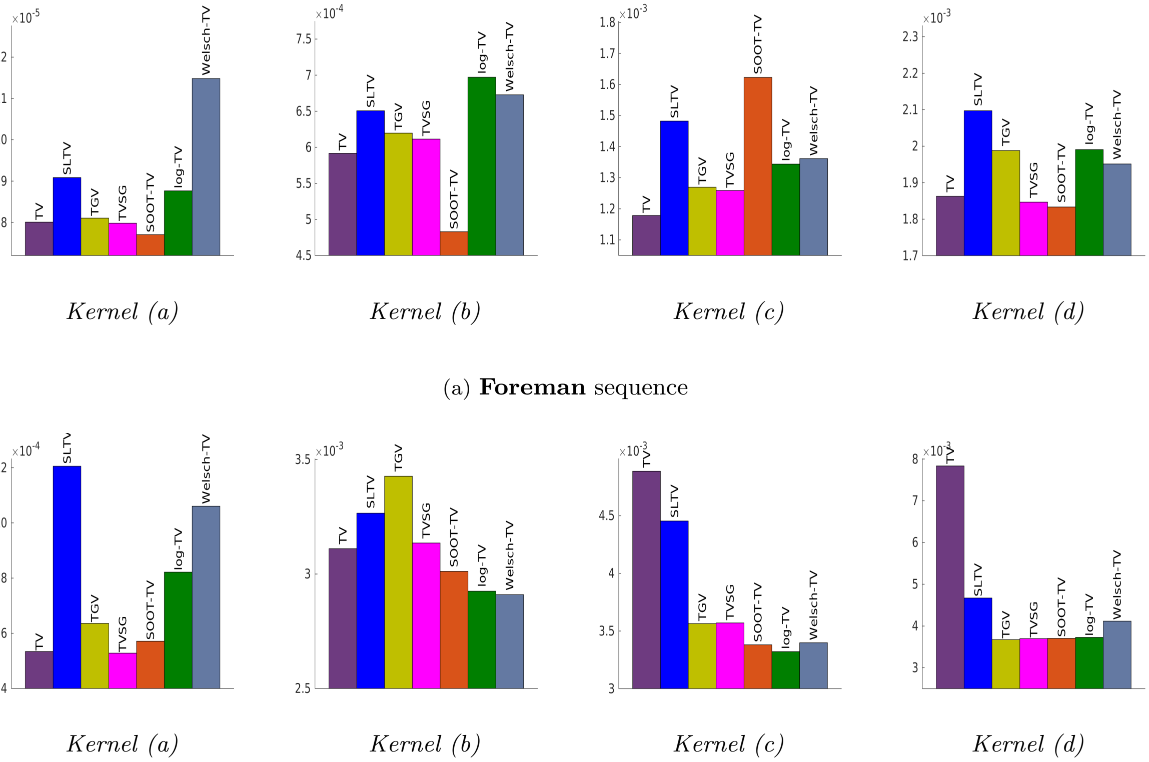
(a) **Foreman** sequence

(b) **Claire** sequence

Figure 5: Performance in terms of SNR with respect to the different regularizations and blur kernels, from left to right: TV, SLTV, TGV, TVSG, SOOT-TV, log-TV, and Welsch-TV.

fields of each frame are extracted and both blind and non-blind deconvolution stages are performed on them. Once the restored fields are obtained, they are merged in order to reconstruct a deblurred interlaced sequence. The estimated kernels with respect to the different regularization approaches are displayed in Figure 7.

Since no ground truth is available for the real sequences, the kernels retained after the blind step, and the associated regularization parameters are selected based on visual inspection on the restored videos. In practice, we selected the kernels estimated with SOOT-TV and log-TV regularizations for **Tachan** and



(a) **Foreman** sequence

(b) **Claire** sequence

Figure 6: Performance in terms of MOVIE with respect to the different regularizations and blur kernels, from left to right: TV, SLTV, TGV, TVSG, SOOT-TV, log-TV, and Welsch-TV.

Au théâtre ce soir sequences respectively. Moreover, we have tested in the second non-blind step the spatial regularization functions that achieved the best performance on synthetic data, namely TGV, TVSG and log-TV. Figures 10-13 illustrate images taken from the input degraded sequences and the restored ones with the above-listed regularizations. One can notice the enhancement of the sharpness and the visual quality of the restored images, and the attenuation of several artifacts such as the echoes and overshoot effects (i.e., weak replicas of the transmitted image) observed on the sequence **Au théâtre ce soir**. Echoes and overshoot are typical defects in analog television records [44], which

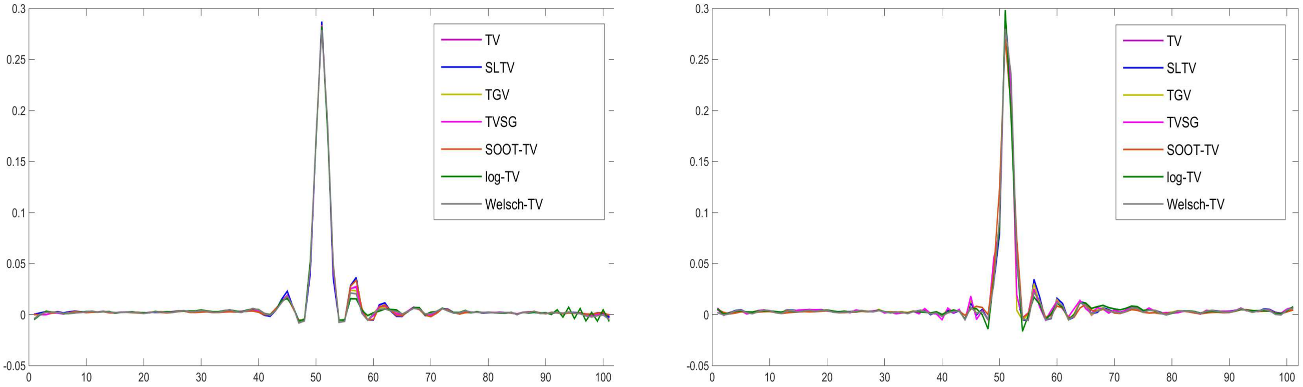
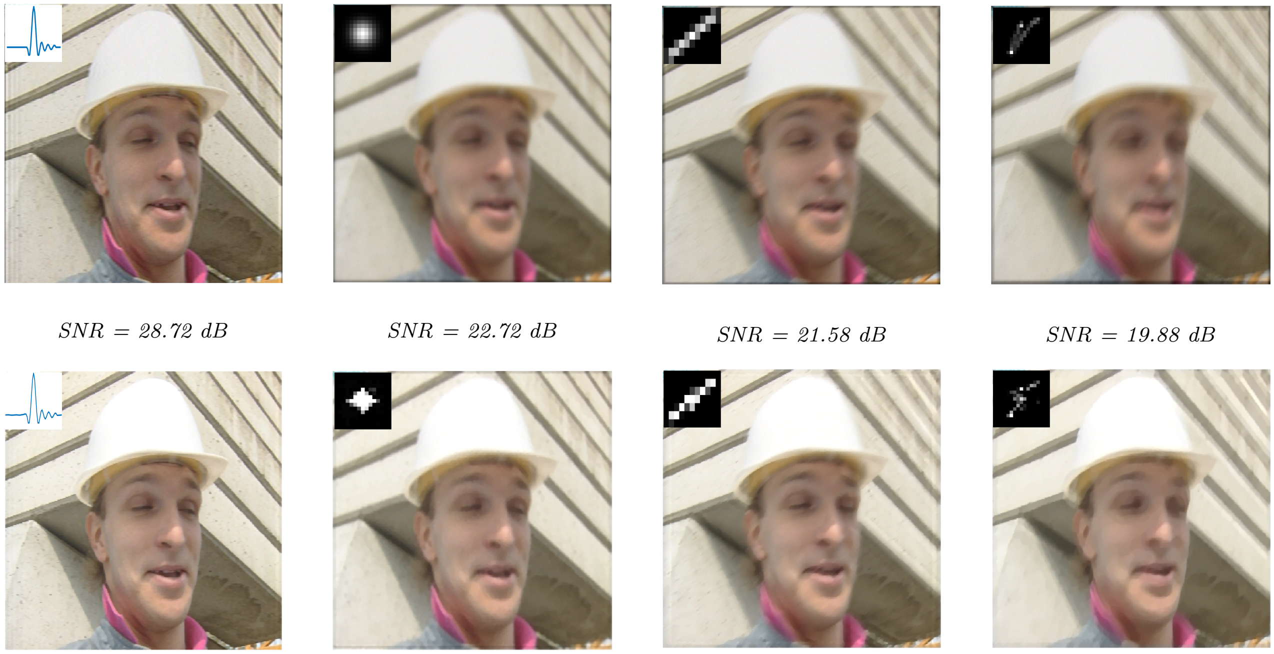


Figure 7: Identified blur kernels (P = 101) with the different regularization approaches: **Tachan** (left), **Au théâtre ce soir** (right).

result from the fusion of television signal waves, some of them being attenuated and delayed by reflections before reaching the receiver. By visual inspection of the resulting sequences at video rate, it appears that the TVSG and log-TV regularizations are more effective in reducing these undesirable effects.



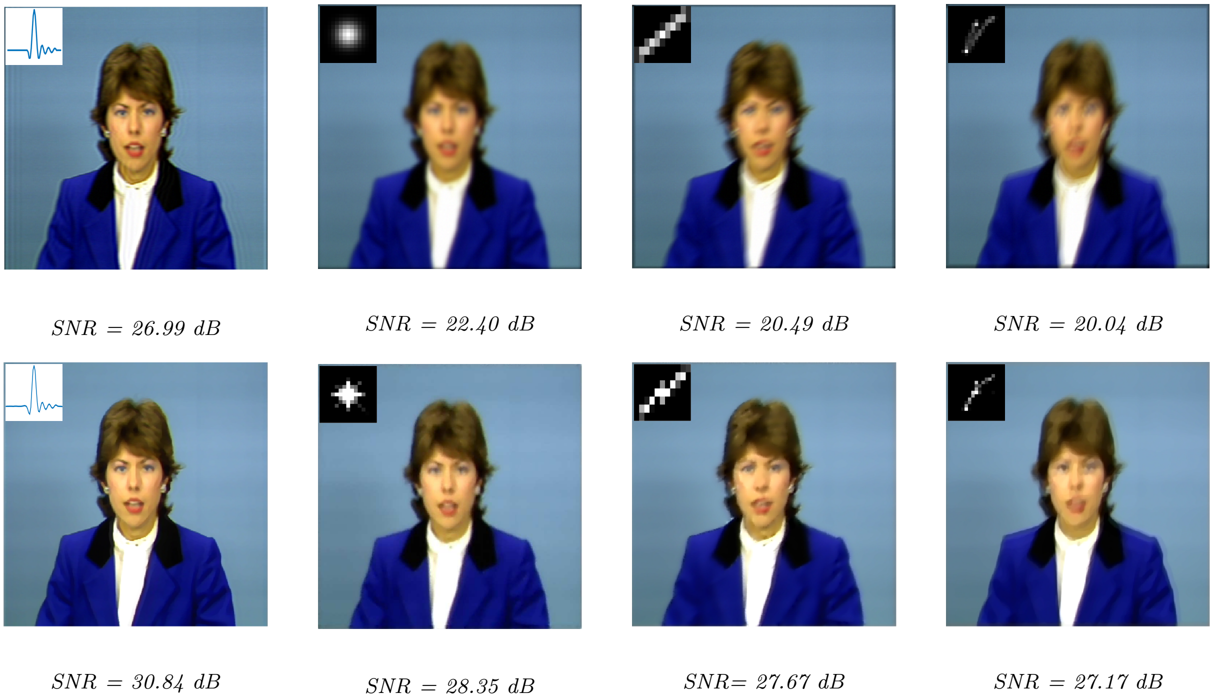
SNR = 28.72 dB *SNR = 22.72 dB* *SNR = 21.58 dB* *SNR = 19.88 dB*

SNR = 33.60 dB *SNR = 26.90 dB* *SNR = 26.64 dB* *SNR = 24.83 dB*

Figure 8: **Foreman** sequence: images from the degraded sequence (top), corresponding restored images with the best choice of spatial regularizations in terms of SNR (bottom).

5. Conclusion

In this paper, we have presented a new variational method for blind video deconvolution. Our approach relies on the minimization of a penalized criterion to enhance the restoration quality. Our iterative algorithm alternates between two proximal steps, namely a video estimation step followed by a kernel identification stage, and we showed that its convergence is guaranteed. The versatility of the proposed method allows us to consider a temporal regularization associated with various convex and nonconvex spatial regularization strategies that are usually employed for solving image and video restoration problems. The experimental results on both synthetic and real data revealed that our method achieves good results in video deconvolution problems, depending on the chosen



SNR = 26.99 dB *SNR = 22.40 dB* *SNR = 20.49 dB* *SNR = 20.04 dB*

SNR = 30.84 dB *SNR = 28.35 dB* *SNR= 27.67 dB* *SNR = 27.17 dB*

Figure 9: **Claire** sequence: images from the degraded sequence (top), corresponding restored images with the best choice of spatial regularizations in terms of SNR (bottom).

spatial regularization and the considered problem (blind/non-blind deconvolution). This approach could be further accelerated through preconditioning and/or a suitable implementation on parallel computing architectures.

Appendix A. Derivation of the majorants in Section 3.2

Appendix A.1. SOOT-TV

Let us set

$$\left(\forall z \in \mathbb{R}^N\right) \quad \eta\, \psi(z) = \eta \log\left(\frac{\ell_{1,\alpha}(Dz) + \beta}{\ell_{2,\lambda}(Dz)}\right) = \psi_1(z) + \psi_2(z), \tag{A.1}$$

where $\psi_1(z) = \eta \log(\ell_{1,\alpha}(Dz) + \beta)$ and $\psi_2(z) = -\eta \log(\ell_{2,\lambda}(Dz))$. We shall prove that

1. $A_{\ell_{1,\alpha}}(z)$ satisfies the majoration condition for $\ell_{1,\alpha}$ at z,

2. $\frac{A_{\ell_{1,\alpha}}(z)}{\ell_{1,\alpha}(Dz)+\beta}$ satisfies the majoration condition for ψ_1 at z,

3. $\frac{9\eta \|D\|^2}{8\lambda^2}$ is a Lipschitz constant of ψ_2.

Proving Statements 2 and 3 is similar to the proof provided in [6]. Let us now consider Statement 1, by first defining

$$\ell_{1,\alpha}(D\,\cdot) = \sum_{i=1}^{2N} \phi\left(D^{(i)}\,\cdot\right), \tag{A.2}$$

where

$$(\forall v \in \mathbb{R}) \qquad \phi(v) = \sqrt{v^2+\alpha^2} - \alpha. \tag{A.3}$$

We have [45]

$$(\forall u \in \mathbb{R}) \quad \phi\left(u\right) \leqslant \phi\left(v\right) + (u-v)\dot{\phi}\left(v\right) + \frac{\kappa(v)}{2}\left(u-v\right)^2, \tag{A.4}$$

with, for every $v \in \mathbb{R}$,

$$\dot{\phi}(v) = \frac{v}{\sqrt{v^2+\alpha^2}} \quad \text{and} \quad \kappa(v) = \frac{1}{\sqrt{v^2+\alpha^2}}. \tag{A.5}$$

Thus, for every $(\omega, z) \in (\mathbb{R}^N)^2$,

$$\phi\left(D^{(i)}\omega\right) \leqslant \phi\left(D^{(i)}z\right) + (\omega-z)^\top D^{(i)\top}\dot{\phi}\left(D^{(i)}z\right) + \frac{\kappa(D^{(i)}z)}{2}\left(D^{(i)}(\omega-z)\right)^2, \tag{A.6}$$

By combining (A.2) and (A.6), we deduce that

$$\begin{aligned}\ell_{1,\alpha}(D\omega) \leqslant \ell_{1,\alpha}(Dz) &+ \langle \omega - z, \nabla\left(\ell_{1,\alpha}\circ D\right)(z)\rangle \\ &+ \frac{1}{2}\left(D(\omega-z)\right)^\top \mathrm{Diag}(s(z))\, D(\omega-z),\end{aligned} \tag{A.7}$$

where $s(z) \in \mathbb{R}^{2N}$ is such that its i-th component with $i \in \{1,\dots,2N\}$ is given by (31).

Let us define $(\sigma_j)_{1\leqslant j\leqslant N} \in]0,+\infty[^N$ such that $\sum_{j=1}^N \sigma_j = 1$, so that, for every

$i \in \{1, \ldots, 2N\}$,

$$\begin{aligned}
\left(D^{(i)}(\omega - z)\right)^2 &= \left(\sum_{j=1}^{N} D^{(i,j)}(\omega^j - z^j)\right)^2, \\
&= \left(\sum_{j=1}^{N} \sigma_j \frac{D^{(i,j)}(\omega^j - z^j)}{\sigma_j}\right)^2. \qquad \text{(A.8)}
\end{aligned}$$

According to the Jensen's inequality, we get

$$\begin{aligned}
\left(\sum_{j=1}^{N} \sigma_j \frac{D^{(i,j)}(\omega^j - z^j)}{\sigma_j}\right)^2 &\leqslant \sum_{j=1}^{N} \sigma_j \left(\frac{D^{(i,j)}(\omega^j - z^j)}{\sigma_j}\right)^2, \\
&= \sum_{j=1}^{N} \frac{\left(D^{(i,j)}(\omega^j - z^j)\right)^2}{\sigma_j}. \qquad \text{(A.9)}
\end{aligned}$$

Setting, for every $j \in \{1, \ldots, N\}$, $\sigma_j = \frac{\left|D^{(i,j)}\right|}{\sum_{l=1}^{N}\left|D^{(i,l)}\right|}$, leads to

$$\begin{aligned}
\left(D^{(i)}(\omega - z)\right)^2 &= \sum_{j=1}^{N} \left(\frac{\left(D^{(i,j)}(\omega^j - z^j)\right)^2}{\frac{\left|D^{(i,j)}\right|}{\sum_{l=1}^{N}\left|D^{(i,l)}\right|}}\right), \\
&= \sum_{j=1}^{N} \left|D^{(i,j)}\right| \sum_{l=1}^{N} \left|D^{(i,l)}\right| (\omega^j - z^j)^2. \qquad \text{(A.10)}
\end{aligned}$$

This yields

$$\|D(\omega - z)\|^2 \leqslant \|\omega - z\|^2_{A_{\ell_{1,\alpha}}(z)}, \qquad \text{(A.11)}$$

with $A_{\ell_{1,\alpha}}(z)$ given by (30) and (32).

Appendix A.2. log-TV

The process of constructing the majorant when ψ stands for log-TV regularization, is similar to the one of SOOT-TV, by setting [32]

$$\begin{aligned}(\forall z \in \mathbb{R}^N) \quad \psi(z) &= \frac{1}{2}\sum_{i=1}^{N} \log\left((\nabla_{\mathsf{H}}^{(i)} z)^2 + (\nabla_{\mathsf{V}}^{(i)} z)^2 + \alpha^2\right), \\ &= \sum_{i=1}^{N} \phi\left(\sqrt{(\nabla_{\mathsf{H}}^{(i)} z)^2 + (\nabla_{\mathsf{V}}^{(i)} z)^2}\right), \end{aligned} \tag{A.12}$$

where

$$(\forall v \in \mathbb{R}) \qquad \phi(v) = \frac{1}{2}\log\left(v^2 + \alpha^2\right). \tag{A.13}$$

By proceeding similarly to the derivation of (A.6) we obtain, for every $(\omega, z) \in (\mathbb{R}^N)^2$,

$$\begin{aligned}\phi\left(\sqrt{(\nabla_{\mathsf{H}}^{(i)} \omega)^2 + (\nabla_{\mathsf{V}}^{(i)} \omega)^2}\right) \leqslant\ & \phi\left(\sqrt{(\nabla_{\mathsf{H}}^{(i)} z)^2 + (\nabla_{\mathsf{V}}^{(i)} z)^2}\right) \\ &+ \kappa\left(\sqrt{(\nabla_{\mathsf{H}}^{(i)} z)^2 + (\nabla_{\mathsf{V}}^{(i)} z)^2}\right)(\omega - z)^\top((\nabla_{\mathsf{H}}^{(i)})^\top \nabla_{\mathsf{H}}^{(i)} + (\nabla_{\mathsf{V}}^{(i)})^\top \nabla_{\mathsf{V}}^{(i)})z \\ &+ \frac{\kappa\left(\sqrt{(\nabla_{\mathsf{H}}^{(i)} z)^2 + (\nabla_{\mathsf{V}}^{(i)} z)^2}\right)}{2}\left((\nabla_{\mathsf{H}}^{(i)}(\omega - z))^2 + (\nabla_{\mathsf{V}}^{(i)}(\omega - z))^2\right), \end{aligned} \tag{A.14}$$

where

$$(\forall v \in \mathbb{R}) \qquad \kappa(v) = \frac{\dot{\phi}(v)}{v} = \frac{1}{v^2 + \alpha^2}. \tag{A.15}$$

By combining (A.12) and (A.14), we obtain

$$\psi(\omega) \leqslant \psi(z) + \langle \omega - z, \nabla\psi(z)\rangle + \frac{1}{2}(\omega - z)^\top \operatorname{Diag}(s(z))\,(\omega - z), \tag{A.16}$$

where the components of $s(z) \in \mathbb{R}^N$ are given by (35). Therefore, for every $\omega \in \mathbb{R}^N$,

$$\psi(\omega) \leqslant \psi(z) + \langle \omega - z, \nabla\psi(z)\rangle + \frac{1}{2}\|\omega - z\|^2_{A_{\log}(z)}, \tag{A.17}$$

where, for every $z \in \mathbb{R}^N$, matrix $A_{\log}(z)$ is expressed by (34) where $\Omega = \Omega_{\mathsf{H}} + \Omega_{\mathsf{V}}$

and the elements of Ω_{H} and Ω_{V} are given by (36).

Appendix A.3. Welsch-TV

The construction of this majorant is analogous to the one of log-TV regularization, by taking

$$(\forall z \in \mathbb{R}^N) \quad \psi(z) = \sum_{i=1}^{N} \left(1 - \exp\left(-\left((\nabla_{\mathsf{H}}^{(i)} z)^2 + (\nabla_{\mathsf{V}}^{(i)} z)^2\right)/(2\sigma^2)\right)\right),$$

$$= \sum_{i=1}^{N} \phi\left(\sqrt{(\nabla_{\mathsf{H}}^{(i)} z)^2 + (\nabla_{\mathsf{V}}^{(i)} z)^2}\right), \qquad \text{(A.18)}$$

where ϕ is given by (17). Inequality (A.14) still holds where

$$(\forall v \in \mathbb{R}) \qquad \kappa(v) = \frac{1}{\sigma^2} \exp\left(-v^2/(2\sigma^2)\right). \qquad \text{(A.19)}$$

By combining this inequality with (A.18) and following the same approach as in the previous proof, we have, for every $(\omega, z) \in (\mathbb{R}^N)^2$,

$$\psi(\omega) \leqslant \psi(z) + \langle \omega - z, \nabla\psi(z) \rangle + \frac{1}{2}\|\omega - z\|^2_{A_w(z)}, \qquad \text{(A.20)}$$

where $A_w(z)$ is given by (38), (39), and (36).

References

[1] G. Chantas, N. Galatsanos, A. Likas, M. Saunders, Variational Bayesian image restoration based on a product of t-distributions image prior, IEEE Trans. Image Process. 7 (10) (2008) 1795–1805.

[2] A. Danielyan, V. Katkovnik, K. Egiazarian, BM3D frames and variational image deblurring, IEEE Trans. Image Process. 21 (4) (2012) 1715–1728.

[3] M. A. T. Figueiredo, R. D. Nowak, An EM algorithm for wavelet-based image restoration, IEEE Trans. Image Process. 12 (8) (2003) 906–916.

[4] R. Fergus, B. Singh, A. Hertzmann, S. T. Roweis, F. W. T., Removing camera shake from a single photograph, ACM Trans. Graph. 25 (3) (2006) 787–794.

[5] D. Krishnan, T. Tay, R. Fergus, Blind deconvolution using a normalized sparsity measure, in: IEEE Conf. Comput. Vis. Pattern Recognt. (CVPR 2011), Colorado Springs, CO, USA, 2011, pp. 233–240.

[6] A. Repetti, M. Q. Pham, L. Duval, E. Chouzenoux, J.-C. Pesquet, Euclid in a taxicab: Sparse blind deconvolution with smoothed l1/l2 regularization, Signal Process. Lett. 22 (5) (2015) 539–543.

[7] J. Zhang, D. Zhao, R. Xiong, S. Ma, M. Gao, Image restoration using joint statistical modeling in a space-transform domain, IEEE Trans. Circuits Syst. Video Technol. 24 (6) (2014) 915–928.

[8] L. Xu, S. Zheng, J. Jia, Unnatural l0 sparse representation for natural image deblurring, in: IEEE Conf. Comput. Vis. Pattern Recognt. (CVPR 2013), Ohio, USA, 2013, pp. 1107–1114.

[9] L. I. Rudin, S. Osher, E. Fatemi, Nonlinear total variation based noise removal algorithms, Phys. D 60 (1) (1992) 259–268.

[10] S. Mallat, A Wavelet Tour of Signal Processing, 3rd Edition, Academic Press, 2008.

[11] N. Pustelnik, A. Benazza-Benhayia, Y. Zheng, J.-C. Pesquet, Wavelet-based image deconvolution and reconstruction, Wiley Encyclopedia of Electrical and Electronics Engineering.

[12] M. Elad, P. Milanfar, R. Ron, Analysis versus synthesis in signal priors, Inverse Problems 23 (3) (2007) 947–968.

[13] N. Pustelnik, C. Chaux, J.-C. Pesquet, C. Comtat, Parallel algorithm and hybrid regularization for dynamic PET reconstruction, in: IEEE Med. Imaging Conf., Knoxville, Tennessee, USA, 2010, pp. 2423–2427.

[14] N. Komodakis, N. Paragios, MRF-based blind image deconvolution, in: Asian Conf. Comput. Vision (ACCV 2012), Daejeon, North Korea, 2012, pp. 361–374.

[15] J. Kotera, F. Śroubek, P. Milanfar, Blind deconvolution using alternating maximum a posteriori estimation with heavy-tailed priors, Comput. Anal. Images Patterns 8048 (2013) (2013) 59–66.

[16] E. Chouzenoux, A. Jezierska, J.-C. Pesquet, H. Talbot, A majorize-minimize subspace approach for l2-l0 image regularization, SIAM J. Imag. Sci. 6 (1) (2013) 563–591.

[17] Z. Q. Luo, P. Tseng, On the convergence of the coordinate descent method for convex differentiable minimization, J. Optim. Theory App. 72 (1) (1992) 7–35.

[18] J. Bolte, P. L. Combettes, J.-C. Pesquet, Alternating proximal algorithm for blind image recovery, in: IEEE International Conf. Image Processing (ICIP), Honk Kong, 2010, pp. 1673–1676.

[19] F. Abboud, E. Chouzenoux, J.-C. Pesquet, J.-H. Chenot, L. Laborelli, Dual block coordinate forward-backward algorithm with application to deconvolution and deinterlacing of video sequences, J. Math. Imaging Vision 59 (3) (2017) 415–431.

[20] M. Nikolova, Weakly constrained minimization: Application to the estimation of images and signals involving constant regions, J. Math. Imaging Vision 21 (2) (2009) 155–175.

[21] L. Condat, Semi-local total variation for regularization of inverse problems, in: IEEE Eur. Signal Process. Conf. (EUSIPCO 2014), Lisbon, Portugal, 2014, pp. 1806–1810.

[22] K. Bredies, K. Kunisch, T. Pock, Total generalized variation, SIAM J. Imaging Sci. 3 (3) (2010) 492–526.

[23] F. Balle, T. Beck, D. Eigler, J. H. Fitschen, S. Schuff, G. Steidl, Strain analysis by a total generalized variation regularized optical flow model, Inverse Problems in Science and Engineering (to appear), https://arxiv.org/pdf/1704.06028.pdf.

[24] S. Setzer, G. Steidl, Variational methods with higher order derivatives in image processing, in: Approximation XII, San Antonio, USA, 2008, pp. 360–386.

[25] S. Setzer, G. Steidl, T. Teuber, Infimal convolution regularizations with discrete ℓ_1-type functionals, Commun. Math. Sci. 9 (3) (2011) 797–827.

[26] L. Condat, Discrete Total Variation: New definition and minimization, SIAM J. Imaging Sci. 10 (3) (2017) 1258–1290.

[27] D. Perrone, P. Favaro, A logarithmic image prior for blind deconvolution, Int. J. Comput. Vis. 17 (2) (2016) 159–172.

[28] J. E. Dennis, R. E. Welsch, Techniques for nonlinear least squares and robust regression, Comm. Statist. Simulation Comput. 7 (4) (1978) 345–359.

[29] H. Attouch, J. Bolte, P. Redont, A. Soubeyran, Proximal alternating minimization and projection methods for nonconvex problems. An approach based on the Kurdyka-Łojasiewicz inequality, Math. Oper. Res. 35 (2) (2010) 438–457.

[30] H. Attouch, J. Bolte, B. F. Svaiter, Convergence of descent methods for semi-algebraic and tame problems: proximal algorithms, forward-backward splitting, and regularized Gauss-Seidel methods, Math. Program. 137 (2011) 91–129.

[31] J. Bolte, S. Sabach, M. Teboulle, Proximal alternating linearized minimization for nonconvex and nonsmooth problems, Math. Program. 146 (1-2) (2014) 459–494.

[32] E. Chouzenoux, J.-C. Pesquet, A. Repetti, A block coordinate variable metric forward-backward algorithm, J. Global Optim. 66 (3) (2016) 457–485.

[33] P. L. Combettes, D. Dung, B. C. Vũ, Dualization of signal recovery problems, Set-Valued Var. Anal. 18 (3) (2010) 373–404.

[34] L. M. Briceño Arias, P. L. Combettes, J.-C. Pesquet, N. Pustelnik, Proximal algorithms for multicomponent image recovery problems, J. Math. Imaging Vision 41 (1) (2011) 3–22.

[35] N. Parikh, S. Boyd, Proximal algorithms, Found. Trends Optim. 1 (3) (2014) 127–239.

[36] E. Chouzenoux, J.-C. Pesquet, A. Repetti, Variable metric forward-backward algorithm for minimizing the sum of a differentiable function and a convex function, J. Optim. Theory Appl. 162 (1) (2014) 107–132. doi:10.1007/s10957-013-0465-7.

[37] P. L. Combettes, D. Dũng, B. C. Vũ, Proximity for sums of composite functions, J. Math. Anal. Appl. 380 (2) (2011) 680–688.

[38] L. Condat, A primal-dual splitting method for convex optimization involving Lipschitzian, proximable and linear composite terms, J. Optim. Theory App. 158 (2) (2013) 460–479.

[39] S. Ma, Alternating proximal gradient method for convex minimization, J. Sci. Comput. 68 (2) (2016) 546–572.

[40] P. L. Combettes, J.-C. Pesquet, A proximal decomposition method for solving convex variational inverse problems, Inverse Problems 24 (8) (2008) 065014.

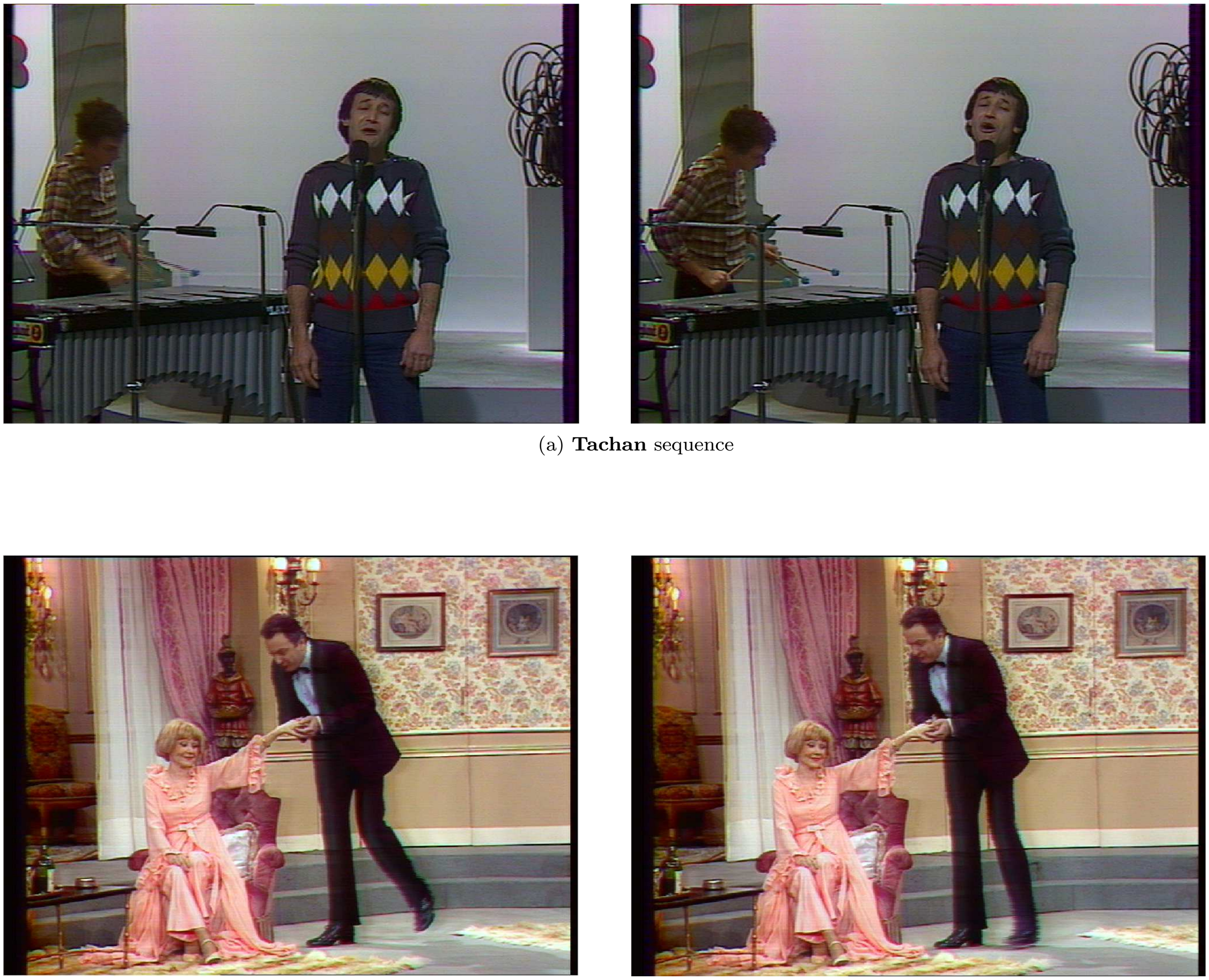
[41] H. Attouch, J. Bolte, On the convergence of the proximal algorithm for nonsmooth functions involving analytic features, Math. Program. 116 (2009) 5–16.

[42] C. Liu, W. T. Freeman, E. H. Adelson, Y. Weiss, Human-assisted motion annotation, in: IEEE Conf. Comput. Vis. Pattern Recogn. (CVPR 2008), Anchorage, Alaska, USA, 2008, pp. 1–8.

[43] K. Seshadrinathan, A. C. Bovik, Motion tuned spatio-temporal quality assessment of natural videos, IEEE Trans. Image Process. 19 (2) (2010) 335–350.

[44] A. C. Kokaram, On missing data treatment for degraded video and film archives: A survey and a new bayesian approach, IEEE Trans. Image Process. 13 (3) (2004) 397–415.

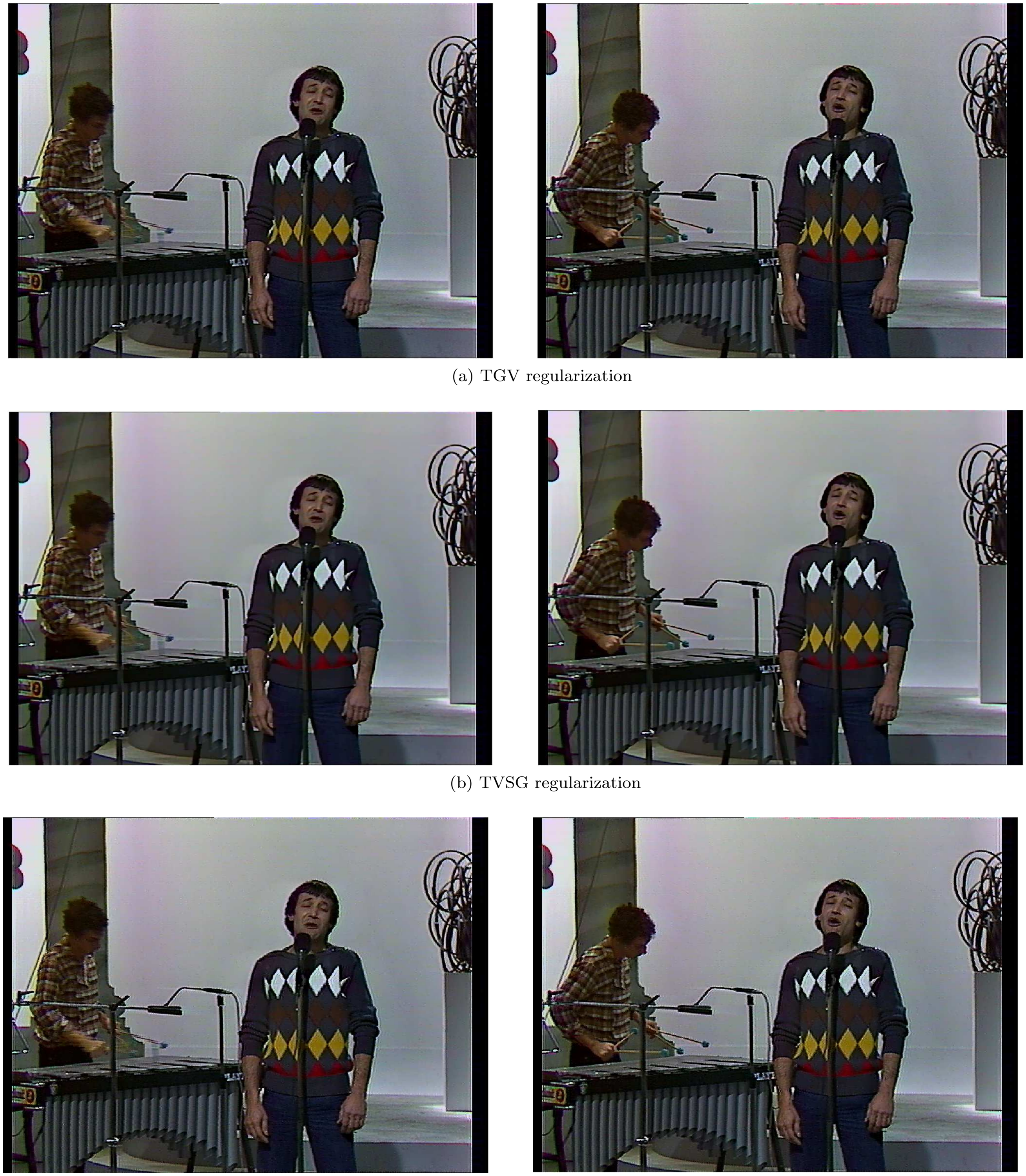
[45] M. Allain, J. Idier, Y. Goussard, On global and local convergence of half quadratic algorithms, IEEE Trans. Image Process. 15 (5) (2006) 1130–1142.



(a) **Tachan** sequence

(b) **Au théâtre ce soir** sequence

Figure 10: 4-*th* and 9-*th* frames from the input degraded and interlaced sequences.

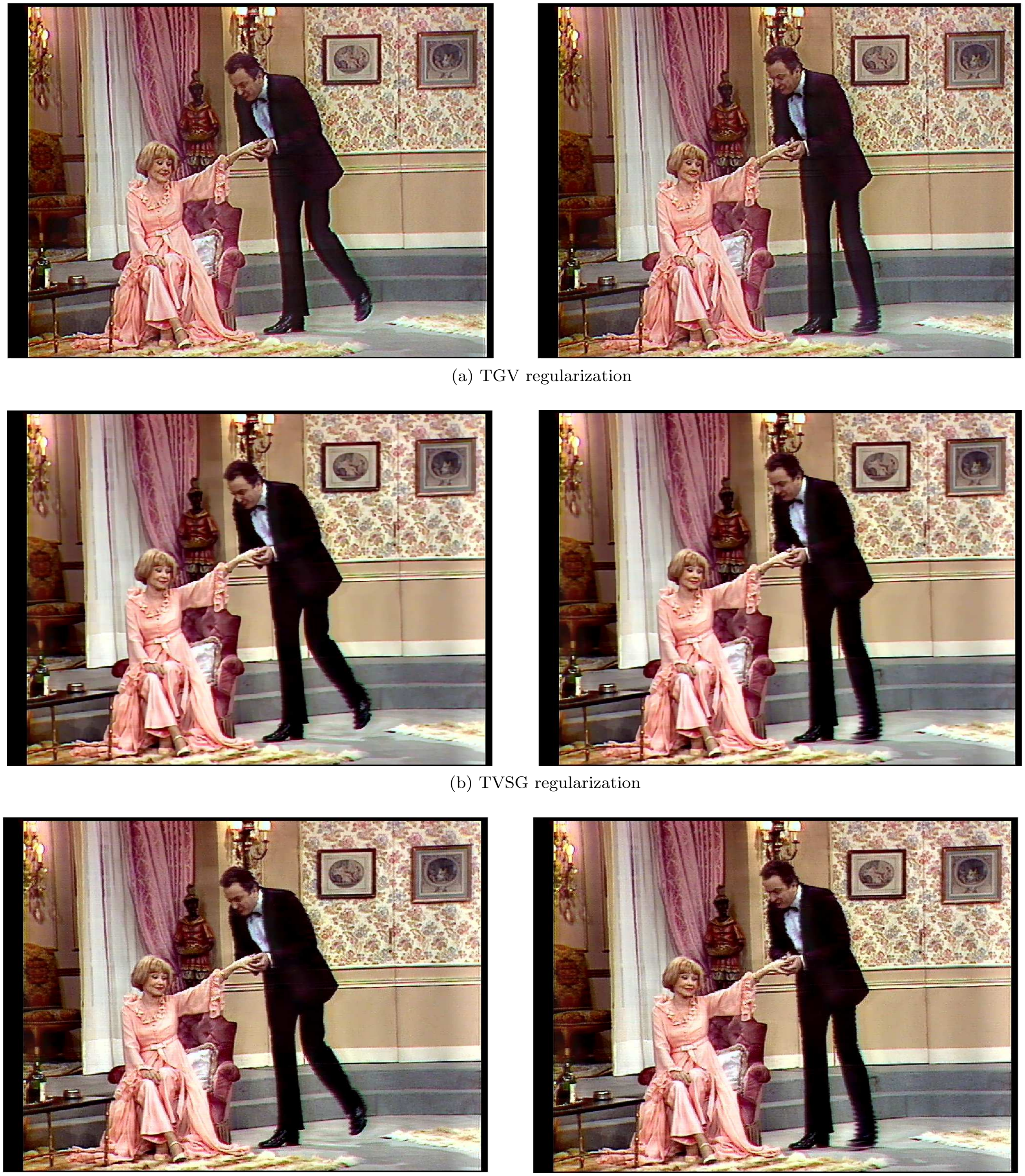


(a) TGV regularization

(b) TVSG regularization

(c) log-TV regularization

Figure 11: **Tachan** sequence: 4-*th* and 9-*th* frames from the restored sequences with the best spatial regularizations in non-blind deconvolution.

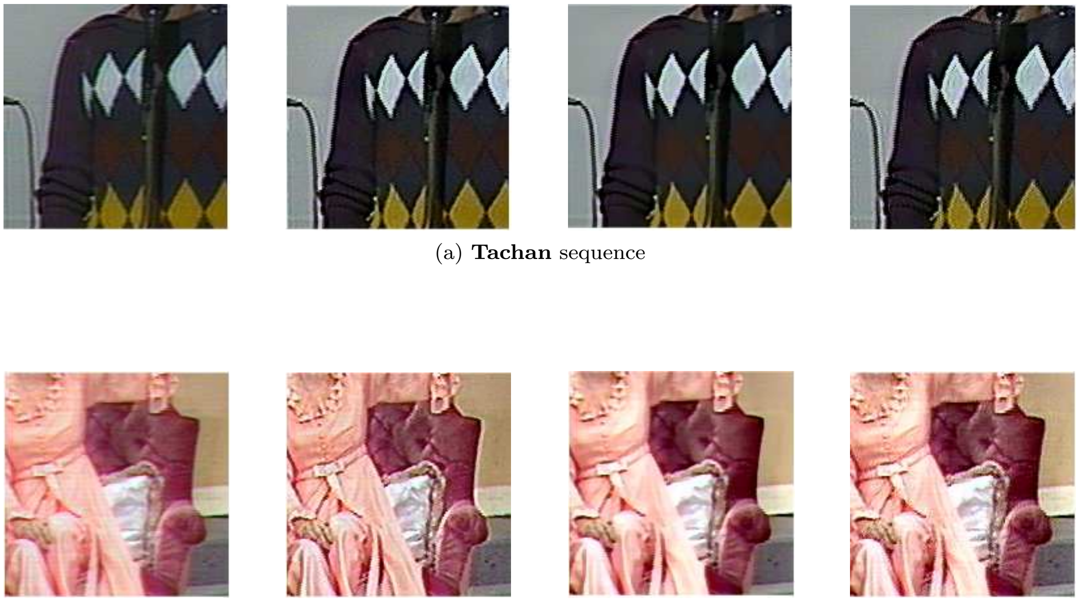


(a) TGV regularization

(b) TVSG regularization

(c) log-TV regularization

Figure 12: **Au théâtre ce soir** sequence: 4-*th* and 9-*th* frames from the restored sequences with the best spatial regularizations in non-blind deconvolution.



(a) **Tachan** sequence

(b) **Au théâtre ce soir** sequence

Figure 13: Zoom on part of images, from left to right: degraded sequence, restored sequence with TGV, restored sequence with TVSG, restored sequence with log-TV.