A HYBRID ALTERNATING PROXIMAL METHOD FOR BLIND VIDEO RESTORATION

F. Abboud[1,2], *E. Chouzenoux*[1], *J.-C. Pesquet*[1], *J.-H. Chenot*[2], *and L. Laborelli*[2]

[1] Université Paris-Est, LIGM, UMR CNRS 8049, Champs sur Marne, France
[2] INA, Institut National de l'Audiovisuel. 94366 Bry sur Marne, France

ABSTRACT

Old analog television sequences suffer from a number of degradations. Some of them can be modeled through convolution with a kernel and an additive noise term. In this work, we propose a new blind deconvolution algorithm for the restoration of such sequences based on a variational formulation of the problem. Our method accounts for motion between frames, while enforcing some level of temporal continuity through the use of a novel penalty function involving optical flow operators, in addition to an edge-preserving regularization. The optimization process is performed by a proximal alternating minimization scheme benefiting from theoretical convergence guarantees. Simulation results on synthetic and real video sequences confirm the effectiveness of our method.

Index Terms— Blind deconvolution, video processing, optical flow, proximal algorithms, convex optimization, regularization.

1. INTRODUCTION

The visual contents of videos from the previous decades are often in a degraded state so that they no longer correspond to the actual public demand. The observed degradations result from several physical phenomena which happened during the sensing, transmission, and recording processes [1–6]. Most of them can be summarized into two main categories. The first type of degradation is of random nature and appears in images as noise, mainly caused by electronic devices. The second one is deterministic and results in blur and oscillations, whose common causes are camera defocus, lens aberrations, relative camera-scene motion, diffusion in sensors, physical or electronic transmission problems. In this paper, we address the problem of restoring a video sequence $(\boldsymbol{y}_t)_{1\leqslant t\leqslant T}$, composed of T successive frames having N pixels, related to an original unknown sequence $(\overline{\boldsymbol{x}}_t)_{1\leqslant t\leqslant T}$ through the degradation model

$$(\forall t \in \{1,\dots,T\}) \quad \boldsymbol{y}_t = S(\overline{\boldsymbol{h}})\overline{\boldsymbol{x}}_t + \boldsymbol{w}_t, \tag{1}$$

where $\overline{\boldsymbol{h}} \in \mathbb{R}^P$ is the vectorized version of an unknown convolution kernel, S is the linear operator which maps the kernel to its associated Hankel-block Hankel matrix form, and $\boldsymbol{w}_t \in \mathbb{R}^N$ represents an unknown additive noise. The objective of video restoration is to recover $(\overline{\boldsymbol{x}}_t)_{1\leqslant t\leqslant T}$ from $(\boldsymbol{y}_t)_{1\leqslant t\leqslant T}$, while $\overline{\boldsymbol{h}}$ is unknown. When $T = 1$, one recovers the well known still image blind deconvolution problem, that has received much attention [7–12]. However, the specific multi-frame case when $T > 1$ remains scarcely studied in the literature. Among existing works, variational approaches based on alternating minimization strategies have been proposed in [13, 14]. Both methods deal with the case of a time-varying positive blur operator, possibly combined with a decimation [14]. The video restoration process then consists of minimizing a least-squares criterion augmented with regularization terms on the sought images and kernels. In [13], the time variations are assumed to be small, which allows the use of penalty functions on the difference of consecutive frames, in addition to a total variation-like prior on each image. Moreover, a parametric model is used on the convolution operator, so as to enforce a specific form of the kernel, defined a priori (e.g. a Lorentzian function). In [14], motion effects are taken into account through a modification of the observation model, while quadratic penalty functions are chosen for the images and the kernels with the aim to simplify the minimization process. It is also worth mentioning the work in [15] which deals with a variant of Problem (1) where the convolution kernel is time-varying while the same scene is observed (up to some motion effects) in a given group of frames.

The contributions of this paper are (i) a versatile formulation of the blind video sequence restoration problem that includes an original strategy to account for motion effects, and (ii) the proposition of a novel efficient optimization algorithm with proven convergence to solve this problem. Our developments are grounded on the use of recent advances in nonsmooth optimization, and more specifically on a new hybrid proximal alternating method. The paper is organized as follows: Section 2 presents the penalized criterion we minimize. In Section 3 we introduce the proposed algorithm. Then, Section 4 illustrates the applicability of our method through a set of experiments for video archive contents restoration.

2. PROBLEM FORMULATION

2.1. Minimization problem

Let us define $\boldsymbol{x} = (\boldsymbol{x}_t)_{1\leqslant t\leqslant T} \in \mathbb{R}^{TN}$. Estimates of $\overline{\boldsymbol{h}}$ and $\overline{\boldsymbol{x}}$ are obtained by solving the minimization problem:

The research leading to these results has received funding from the EC Seventh Framework Programme (FP7/2007-2013) under Grant Agreement n^{o}600827

$$\underset{\boldsymbol{x}\in\mathbb{R}^{TN},\boldsymbol{h}\in\mathbb{R}^{P}}{\text{minimize}} \quad \left(G(\boldsymbol{x},\boldsymbol{h}) = \Phi(\boldsymbol{x},\boldsymbol{h}) + \Theta(\boldsymbol{h}) + \Psi(\boldsymbol{x})\right). \quad (2)$$

Hereabove, Φ is a data fidelity term, which is chosen in accordance with the noise statistics. In this paper, we consider the case when $\boldsymbol{w} = (\boldsymbol{w}_t)_{1\leqslant t\leqslant T}$ is an additive Gaussian noise, so that Φ is equal to a least squares data criterion, defined, for every $\boldsymbol{x} \in \mathbb{R}^{TN}$ and $\boldsymbol{h} \in \mathbb{R}^{P}$ as

$$\Phi(\boldsymbol{x},\boldsymbol{h}) = \frac{1}{2}\sum_{t=1}^{T} \|S(\boldsymbol{h})\boldsymbol{x}_t - \boldsymbol{y}_t\|^2. \quad (3)$$

Functions Θ and Ψ are regularizing terms acting, respectively, on the kernel and the image sequence. The design of suitable penalization functions plays a major role on the quality of the video restoration process. This will be discussed in the next two sections.

2.2. Regularization strategy for the kernel

Prior information on the kernel $\boldsymbol{h}$ is incorporated by choosing:

$$(\forall \boldsymbol{h} \in \mathbb{R}^{P}) \quad \Theta(\boldsymbol{h}) = \iota_{\mathcal{H}}(\boldsymbol{h}) + \rho\|\boldsymbol{h}\|, \quad (4)$$

where $\rho > 0$, and $\iota_{\mathcal{H}}$ is the indicator function of a closed convex set $\mathcal{H}$, i.e.,

$$\iota_{\mathcal{H}}(\boldsymbol{h}) = \begin{cases} 0 & \text{if } \boldsymbol{h} \in \mathcal{H} \\ +\infty & \text{otherwise.} \end{cases} \quad (5)$$

Here, the constrained set is the polyhedral set defined as

$$\mathcal{H} = \Big\{ \boldsymbol{h} = (h_p)_{1\leqslant p\leqslant P} \in \mathbb{R}^{P} \ : \ \sum_{p=1}^{P} h_p = 1, \\ (\forall p \in \{1,\dots,P\}) \qquad h_{\min,p} \leqslant h_p \leqslant h_{\max,p} \Big\}. \quad (6)$$

2.3. Regularization strategy for the video sequence

We consider a composite penalization function on the sequence $\boldsymbol{x}$, whose first term introduces an a priori knowledge on each image independently, while the second term penalizes the whole sequence, i.e.

$$\Psi(\boldsymbol{x}) = \sum_{t=1}^{T} \Psi_t(\boldsymbol{x}_t) + \mathrm{M}(\boldsymbol{x}). \quad (7)$$

For all $t \in \{1,\dots,T\}$, Ψ_t is defined as

$$(\forall \boldsymbol{x}_t \in \mathbb{R}^{N}) \quad \Psi_t(\boldsymbol{x}_t) = \eta\,\mathrm{tv}(\boldsymbol{x}_t) + \iota_{[x_{\min},x_{\max}]^N}(\boldsymbol{x}_t), \quad (8)$$

where η is a positive regularization constant, and tv denotes the total variation function from [16]. Furthermore, in order to favor the similarity between the successive video frames while taking into account existing motions, we define

$$\mathrm{M}(\boldsymbol{x}) = \frac{1}{2}\sum_{t=1}^{T}\sum_{\ell\in\mathcal{V}_t} \beta_{\ell,t}\|\boldsymbol{x}_t - \boldsymbol{M}_{\ell\to t}\boldsymbol{x}_\ell\|^2, \quad (9)$$

where $(\beta_{\ell,t})_{\ell,t}$ are positive weights, the index set $\mathcal{V}_t$ defines the neighborhood of t (i.e. a set of indices $\ell \in \mathcal{V}_t$ for which $|\ell - t|$ is small), and $\boldsymbol{M}_{\ell\to t} \in \mathbb{R}^{N\times N}$ is the linear operator modeling the motion compensation process between the reference frame t and a neighboring frame ℓ [17].

3. PROPOSED ALGORITHM

3.1. Optimization tools

Let us first recall 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 $\boldsymbol{U} \in \mathbb{R}^{N\times N}$ be a symmetric positive definite matrix. For every $\boldsymbol{x} \in \mathbb{R}^N$, the variational problem*

$$\underset{\boldsymbol{z}\in\mathbb{R}^{N}}{\text{minimize}} \ \ f(\boldsymbol{z}) + \frac{1}{2}\|\boldsymbol{x} - \boldsymbol{z}\|_{\boldsymbol{U}}^2, \quad (10)$$

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

Hereabove, $\|\cdot\|_{\boldsymbol{U}}$ denotes the weighted norm defined by $\|.\|_{\boldsymbol{U}} = \langle\cdot,\boldsymbol{U}\cdot\rangle$, where $\langle\cdot,\cdot\rangle$ is the usual Euclidean scalar product. Note that, if $\boldsymbol{U}$ 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 [18, 19] for a tutorial).

We also recall the following characterization of smooth functions:

Definition 2 *Let $f\colon \mathbb{R}^N \to \mathbb{R}$ be a differentiable function. Its gradient ∇f is said L-Lipschitzian continuous if, for every $(\boldsymbol{x},\boldsymbol{z}) \in (\mathbb{R}^N)^2$,*

$$\|\nabla f(\boldsymbol{x}) - \nabla f(\boldsymbol{z})\| \leqslant L\|\boldsymbol{x} - \boldsymbol{z}\|. \quad (11)$$

3.2. Hybrid proximal alternating strategy

Although the objective function G is nonconvex because of the coupling in the data fidelity term (3), it remains yet convex with respect to each variable $\boldsymbol{x}_1,\dots,\boldsymbol{x}_T$ and $\boldsymbol{h}$, as the regularization terms are convex. A standard approach for solving (2) is thus to adopt an alternating minimization strategy, where, at each iteration, G is minimized with respect to one variable while the others remain fixed. This approach, sometimes referred to as Block Coordinate Descent [20] or nonlinear Gauss-Seidel method [21], has been widely used in the context of blind deconvolution [7, 10, 13–15]. However, its convergence is only guaranteed under restrictive assumptions [20]. Therefore, alternative strategies based on proximal tools have been proposed in [9, 21–26]. They consist of replacing, at each iteration, the minimization step by either a (single) proximal step [21, 24] or a Forward-Backward step [22, 23, 25, 26], giving rise, respectively, to the so-called proximal (resp. forward-backward) alternating algorithms. In this work, we propose to adopt the following hybrid alternating algorithm:

Algorithm 1 Video blind deconvolution algorithm

For every $k \in \mathbb{N}$, $(\lambda_t^1, \ldots, \lambda_T^k, \mu^k) \in (0, +\infty)^{T+1}$.
Initialize with $\boldsymbol{x}_1^0, \ldots, \boldsymbol{x}_T^0 \in \mathbb{R}^N$, and $\boldsymbol{h}^0 \in \mathbb{R}^P$.
Iterations:
For $k = 0, 1, \ldots$
For $t = 1, \ldots, T$
$\check{\boldsymbol{x}}^{t,k} = \left(\boldsymbol{x}_1^{k+1}, \ldots, \boldsymbol{x}_{t-1}^{k+1}, \boldsymbol{x}_t^k, \ldots, \boldsymbol{x}_T^k\right)$,
$\widetilde{\boldsymbol{x}}_t^k = \boldsymbol{x}_t^k - \lambda_t^k \left(\nabla_{\boldsymbol{x}_t}\Phi(\check{\boldsymbol{x}}^{t,k}, \boldsymbol{h}^k) + \nabla_{\boldsymbol{x}_t}\mathrm{M}(\check{\boldsymbol{x}}^{t,k})\right)$,
$\boldsymbol{x}_t^{k+1} = \mathrm{prox}_{\lambda_t^k\Psi_t}\left(\widetilde{\boldsymbol{x}}_t^k\right)$,
$\boldsymbol{h}^{k+1} = \mathrm{prox}_{\mu^k(\Theta+\Phi(\boldsymbol{x}^{k+1},\cdot))}(\boldsymbol{h}^k)$.

In the above algorithm, for every $k \in \mathbb{N}$, for every $t \in \{1, \ldots, T\}$, $\nabla_{\boldsymbol{x}_t}\Phi(\check{\boldsymbol{x}}^{t,k}, \boldsymbol{h}^k)$ is the partial gradient of Φ with respect to $\boldsymbol{x}_t$ computed at $(\check{\boldsymbol{x}}^{t,k}, \boldsymbol{h}^k)$. The vector $\nabla_{\boldsymbol{x}_t}\mathrm{M}(\check{\boldsymbol{x}}^{t,k})$ is defined in a similar way.

3.3. Convergence analysis

The convergence of Algorithm 1 requires the design of a proper strategy to determine the stepsize parameters $\left(\lambda_t^k\right)_{1\leqslant t\leqslant T}$ and μ^k at each iteration $k \in \mathbb{N}$. First, let us state two properties, which are related to the quadratic form of functions Φ and M.

Proposition 1 *Let $\mu \in (0, +\infty)$ and $\boldsymbol{x} \in \mathbb{R}^{TN}$. Let us define the symmetric definite positive matrix*

$$\boldsymbol{A}(\boldsymbol{x}) = \sum_{t=1}^{T} \boldsymbol{X}_t^\top \boldsymbol{X}_t + \mu^{-1}\boldsymbol{I}_P, \tag{12}$$

where $\boldsymbol{I}_P$ is the identity matrix of $\mathbb{R}^P$, and, for every $t \in \{1, \ldots, T\}$, $\boldsymbol{X}_t \in \mathbb{R}^{N\times P}$ is such that $S(\boldsymbol{h})\boldsymbol{x}_t = \boldsymbol{X}_t\boldsymbol{h}$. Then, for every $\boldsymbol{h} \in \mathbb{R}^P$,

$$\mathrm{prox}_{\mu(\Theta+\Phi(\boldsymbol{x},\cdot))}(\boldsymbol{h}) = \mathrm{prox}_{\boldsymbol{A}(\boldsymbol{x}),\Theta}(\boldsymbol{h} - \boldsymbol{A}(\boldsymbol{x})^{-1}\nabla_{\boldsymbol{h}}\Phi(\boldsymbol{x}, \boldsymbol{h})), \tag{13}$$

and, for every $\boldsymbol{x}' \in \mathbb{R}^{TN}$ and $\boldsymbol{h}' \in \mathbb{R}^P$,

$$\Phi(\boldsymbol{x}, \boldsymbol{h}') + \nabla_{\boldsymbol{h}'}\Phi(\boldsymbol{x}', \boldsymbol{h}')^\top(\boldsymbol{h} - \boldsymbol{h}') + \frac{1}{2}\|\boldsymbol{h} - \boldsymbol{h}'\|^2_{\boldsymbol{A}(\boldsymbol{x})} \geqslant \Phi(\boldsymbol{x}, \boldsymbol{h}). \tag{14}$$

Proposition 2 *Let $t \in \{1, \ldots, T\}$ and $\boldsymbol{h} \in \mathbb{R}^P$. The partial gradient function $\boldsymbol{x}_t \mapsto \nabla_{\boldsymbol{x}_t}\Phi(\boldsymbol{x}, \boldsymbol{h})$ is Lipschitzian continuous, with Lipschitz modulus*

$$L_t(\boldsymbol{h}) = \|S(\boldsymbol{h})\|^2 + \sum_{\ell\in\mathcal{V}_t}\left(\beta_{\ell,t} + \beta_{t,\ell}\|\boldsymbol{M}_{t\to\ell}\|^2\right). \tag{15}$$

Combining Propositions 1, 2 and the convergence properties of the BC-VMFB (Block Coordinate Variable Metric Forward Backward) algorithm derived in [26], the convergence of Algorithm 1 can be proved as expressed by the following result:

Theorem 1 *Let $(\boldsymbol{x}^k, \boldsymbol{h}^k)_{k\in\mathbb{N}}$ be a sequence generated by Algorithm 1, where, for all $k \in \mathbb{N}$, $\lambda_t^k = \theta L_t(\boldsymbol{h}^k)^{-1}$ with $\theta \in (0, 2)$ and $L_t(\cdot)$ defined in* (15)*. Then the sequence $(\boldsymbol{x}^k, \boldsymbol{h}^k)_{k\in\mathbb{N}}$ converges to a critical point $(\widehat{\boldsymbol{x}}, \widehat{\boldsymbol{h}})$ of G. Moreover, $(G(\boldsymbol{x}^k, \boldsymbol{h}^k))_{k\in\mathbb{N}}$ is a nonincreasing sequence converging to $G(\widehat{\boldsymbol{x}}, \widehat{\boldsymbol{h}})$.*

3.4. Practical implementation

The proximal steps in Algorithm 1 are not explicit, and subiterations are needed. In practice, the proximity operator of $\Theta + \Phi(\boldsymbol{x}, \cdot)$ is computed with Dykstra algorithm [18], while the Dual Forward-Backward algorithm [27] is used to compute the proximity operators of functions Ψ_t, for $t \in \{1, \ldots, T\}$. The proposed algorithm is thus used under an inexact form for which the conclusions of Theorem 1 still hold.

4. EXPERIMENTAL RESULTS

4.1. Synthetic data

We first demonstrate the practical performance of our method on a synthetic video restoration example. We extract $T = 15$ frames of size $N = 256\times256$ from the video sequence Claire, available at `http://www.cipr.rpi.edu/resource/sequences/cif.html`. To generate the observed video $\boldsymbol{y} = (\boldsymbol{y}_t)_{1\leqslant t\leqslant T}$, we degraded the original one with the horizontal convolution kernel displayed in Figure 2, with size $P = 41$, which corresponds to a realistic model of those affecting old analog television sequences. The video is further corrupted with a white additive zero-mean Gaussian noise. Parameters ρ, η and $(\beta_{\ell,t})_{\ell,t}$ were adjusted to maximize the SNR between the original and reconstructed sequences. Neighboring frames such that $|\ell - t| = 1$ have been taken into account in the regularization term M. The motion matrices $(\boldsymbol{M}_{\ell\to t})_{\ell,t}$ have been estimated from the degraded sequence $\boldsymbol{y}$, using the optical flow estimation algorithm from [17]. Subpixel motions have been ignored, so as to obtain simple expressions for $\left(\boldsymbol{M}_{\ell\to t}^\top\right)_{\ell,t}$ and $\left(\|\boldsymbol{M}_{\ell\to t}\|^2\right)_{\ell,t}$. Algorithm 1 provides estimates $(\widehat{\boldsymbol{x}}, \widehat{\boldsymbol{h}})$ displayed in Figures 1-2, with SNR $= 32.2$ dB, and relative quadratic error on the kernel equal to 0.006. To emphasize the efficiency of the motion-based regularization strategy, note that, when taking $\beta_{\ell,t} \equiv 0$, we observe a decrease of the estimation quality, as the SNR of the restored sequence is equal to 31.2 dB and the kernel estimation error to 0.016.

Fig. 1. 3rd and 12th frames of the Claire sequence: Noisy convolved images, SNR = 25.6 dB (top) and restored images (bottom) with the proposed algorithm, SNR = 32.2 dB.

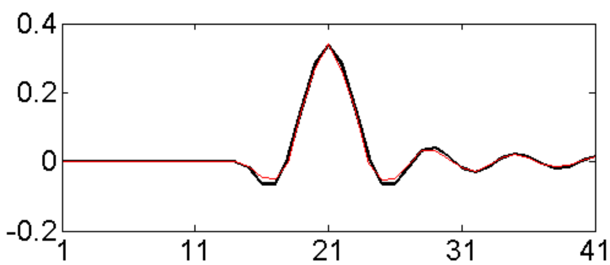

Fig. 2. Original kernel (black) and estimated kernel (red), with relative error 0.006.

4.2. Real data

We now apply our method to a real video sequence (for which no ground truth is available) containing $T = 5$ frames of $N = 720 \times 576$ pixels, extracted from a French broadcast archive programme "Au théâtre ce soir". This sequence, provided by INA, results from a recording taken from a radio frequency analog terrestrial link, affected by multiple paths. Algorithm 1 is employed to restore the luminance component of the YCrCb representation of the sequence, the two chrominance components remaining untouched. To account for the interlacing, the odd and even lines of the video are restored separately, and then gathered at the end of the restoration process. Examples of degraded and restored images, as well as

Fig. 3. 2nd frame of the real INA sequence: Degraded image (top) and restored image (bottom) with the proposed algorithm.

Fig. 4. Zoom on a part of the above images.

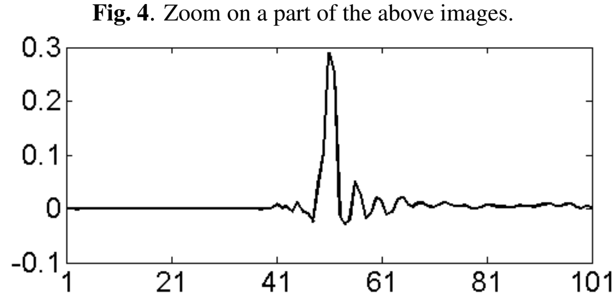

Fig. 5. Estimated kernel

the estimated kernel, are displayed in Figures 3-5. One can observe on Figure 4 the ability of the proposed method to highly improve the sharpness of the image. Moreover, the knowledge of an estimation of the convolution kernel allows us to better analyze the main source of the video degradation process, here mainly due to oscillation effects probably caused by the analog transmission.

5. CONCLUSION

In this work, we have presented a new variational method for blind deconvolution of video sequences. Our approach relies on the minimization, through a proximal alternating algorithm, of a penalized criterion that accounts for the temporal continuity between the video frames. Our experimental results on both synthetic and real data showed that our method leads to well restored images, as well as satisfactory estimations of the blur kernel. It should be emphasized that the versatility of the proposed algorithm makes it possible to handle various forms of blur kernels, and to minimize convex and nonconvex data fidelity and penalty functions. Moreover, using variants of this algorithm could open the gate to faster implementation strategies based on parallel computing architectures.

REFERENCES

[1] J.-H. Chenot, J. O. Drewery, and D. Lyon, "Restoration of archived television programmes for digital broadcasting," Tech. Rep., Sep. 1998, `http://www.bbc.co.uk/rd/publications/rdreport_1998_05`.

[2] V. Naranjo and A. Albiol, "Flicker reduction in old films," in *7th IEEE Int. Conf. on Image Processing (ICIP 2000)*, Vancouver, Canada, 10-13 Sept. 2000, vol. 2, pp. 657–659.

[3] A. Pizurica, V. Zlokolica, and P. Wilfried, "Noise reduction in video sequences using wavelet-domain and temporal filtering," in *Proc. of the SPIE 5266, Wavelet Applications in Industrial Processing*, Providence, RI, 27-30 Oct. 2004, pp. 48–59.

[4] A. C. Kokaram and S. J. Godsill, "MCMC for joint noise reduction and missing data treatment in degraded video," *IEEE Trans. Signal Process.*, vol. 50, no. 2, pp. 189–205, Feb. 2002.

[5] A. C. Kokaram, *Motion Picture Restoration: Digital Algorithms for Artefact Suppression in Degraded Motion Picture Film and Video*, Springer-Verlag, London, UK, 1998.

[6] Y. Tendero and S. Osher, "Blind uniform motion blur deconvolution for image bursts and video sequences based on sensor characteristics," Tech. Rep., 2014, `http://dev.ipol.im/~tendero/icip2014v2.pdf`.

[7] M. S. C Almeida, A. Abelha, and A. B. Almeida, "Multi-resolution approach for blind deblurring of natural images," in *19th IEEE Int. Conf. on Image Processing (ICIP 2012)*, Florida, USA, 30 Sept-3 Oct. 2012, pp. 3041–3044.

[8] E. Thiébaut, "Optimization issues in blind deconvolution algorithms," in *Proc. of SPIE 4847, Astronomical Data Analysis II*, Waikoloa, Hawaii, 27-28 Aug. 2002, pp. 174–183.

[9] J. Bolte, P.L. Combettes, and J.-C. Pesquet, "Alternating proximal algorithm for blind image recovery," in *17th IEEE Int. Conf. on Image Processing (ICIP 2010)*, Hong-Kong, 26-29 Sept. 2010, pp. 26–29.

[10] N. Komodakis and N. Paragios, "MRF-based blind image deconvolution," in *11th Asian Conference on Computer Vision (ACCV 2013)*, Daejeon, Korea, 5-9 Nov. 2013, pp. 361–374.

[11] A. Levin, Y. Weiss, F. Durand, and W. T. Freeman, "Understanding and evaluating blind deconvolution algorithms," in *14th IEEE Conf. on Computer Vision and Pattern Recognition (CVPR 2009)*, Miami, Florida, USA, 20-25 Jun. 2009, pp. 1964–1971.

[12] D. Krishnan, T. Tay, and R. Fergus, "Blind deconvolution using a normalized sparsity measure," in *16th IEEE Conf. on Computer Vision and Pattern Recognition (CVPR 2011)*, Colorado Springs, USA, 20-25 Jun. 2011, pp. 233 – 240.

[13] F. Soulez, E. Thiébaut, Y. Tourneur, A. Gressard, and R. Dauphin, "Blind deconvolution of video sequences," in *15th IEEE Int. Conf. on Image Processing (ICIP 2008)*, San Diego, CA, 12-15 Oct. 2008, pp. 673–676.

[14] F. Sroubek, J. Flusser, and M. Sorel, "Superresolution and blind deconvolution of video," in *19th IEEE Int. Conf. on Pattern Recognition (ICPR 2008)*, Florida, USA, 8-11 Dec. 2008, pp. 1–4.

[15] F. Sroubek and P. Milanfar, "Robust multichannel blind deconvolution via fast alternating minimization," *IEEE Trans. Image Process.*, vol. 21, no. 4, pp. 1687–1700, Apr. 2012.

[16] L. I. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Phys. D.*, vol. 60, pp. 259–268, 1992.

[17] C. Liu, W. T. Freeman, E. H. Adelson, and Y. Weiss, "Human-assisted motion annotation," in *13th IEEE Conf. on Computer Vision and Pattern Recognition (CVPR 2008)*, Anchorage, AK, 23-28 Jun. 2008, pp. 1–8.

[18] P. L. Combettes and J.-C. Pesquet, "Proximal splitting methods in signal processing," in *Fixed-Point Algorithms for Inverse Problems in Science and Engineering*, H. H. Bauschke, R. Burachik, P. L. Combettes, V. Elser, D. R. Luke, and H. Wolkowicz, Eds., pp. 185–212. Springer-Verlag, New York, 2010.

[19] N. Parikh and S. Boyd, "Proximal algorithms," *Foundations and Trends in Optimization*, vol. 1, no. 3, pp. 123–231, 2013.

[20] P. Tseng, "Convergence of a block coordinate descent method for non-differentiable minimization," *J. Optim. Theory Appl.*, vol. 109, no. 3, pp. 475–494, Jun. 2001.

[21] H. Attouch, J. Bolte, and 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.*, vol. 137, pp. 91–129, Feb. 2011.

[22] Y. Xu and W. Yin, "A block coordinate descent method for regularized multi-convex optimization with applications to nonnegative tensor factorization and completion," Tech. Rep., 2012, http://www.optimization-online.org/DB_HTML/2012/08/3565.html.

[23] Z. Q. Luo and P. Tseng, "On the convergence of the coordinate descent method for convex differentiable minimization," *J. Optim. Theory Appl.*, vol. 72, no. 1, pp. 7–35, 1992.

[24] H. Attouch, J. Bolte, P. Redont, and A. Soubeyran, "Proximal alternating minimization and projection methods for nonconvex problems. An approach based on the Kurdyka-Łojasiewicz inequality," *Math. Oper. Res.*, vol. 35, no. 2, pp. 438–457, 2010.

[25] J. Bolte, S. Sabach, and M. Teboulle, "Proximal alternating linearized minimization for nonconvex and nonsmooth problems," *To appear in Math. Program.*, 2013.

[26] E. Chouzenoux, J.-C. Pesquet, and A. Repetti, "A block coordinate variable metric forward-backward algorithm," Tech. Rep., 2013, `http://www.optimization-online.org/DB_HTML/2013/12/4178.html`.

[27] P. L. Combettes, D. Dũng, and B. C. Vũ, "Proximity for sums of composite functions," *J. Math. Anal. Appl.*, vol. 380, no. 2, pp. 680–688, 2011.