Convex Super-Resolution Detection of Lines in Images

Kévin Polisano, Laurent Condat, Marianne Clausel, Valérie Perrier

Convex Super-Resolution Detection of Lines in Images

Kévin Polisano*, Laurent Condat†, Marianne Clausel* and Valérie Perrier*

*Univ. Grenoble Alpes, Laboratoire Jean Kuntzmann, F-38000, Grenoble, France
Email: Kevin.Polisano@imag.fr
†Univ. Grenoble Alpes, GIPSA-lab, F-38000, Grenoble, France
Email: Laurent.Condat@gipsa-lab.grenoble-inp.fr

***Abstract*—In this paper, we present a new convex formulation for the problem of recovering lines in degraded images. Following the recent paradigm of super-resolution, we formulate a dedicated atomic norm penalty and we solve this optimization problem by means of a primal–dual algorithm. This parsimonious model enables the reconstruction of the lines from lowpass measurements, even in presence of a large amount of noise or blur. Furthermore, a Prony method performed on rows and columns of the restored image, provides a spectral estimation of line parameters, with subpixel accuracy.**

I. Introduction

Many restoration or reconstruction imaging problems are ill-posed and must be regularized. So, they can be formulated as convex optimization problems formed by the combination of a data fidelity term with a norm-based regularizer. Typically, given the data $y = Ax^\sharp$, for some unknown image $x^\sharp$ to estimate and known observation operator A, one aims at solving a problem like

$$\text{Find } \hat{x} \in \arg\min_x \|Ax - y\|^2 + \lambda R(x), \tag{1}$$

where lambda controls the tradeoff between data fidelity and regularization and R is a convex functional, which favors some notion of low complexity. We place ourselves in the general framework of *atomic norm* minimization [1]: the sought-after image $x^\sharp$ is supposed to be a sparse positive combination of the elements, called *atoms* and of unit norm, of an infinite dictionary $\mathcal{A}$, indexed by continuously varying parameters. Then, R is chosen as the atomic norm $\|x\|_{\mathcal{A}}$ of the image x, which is simply the sum of the coefficients, when the image is expressed in terms of the atoms. Indeed, by choosing the atoms as the kind of elements we want to promote in images, we can estimate them from degraded measurements in a robust way, even with infinite precision when there is no noise. Methods achieving this goal are qualified as *super-resolution* methods, because they uncover fine scale information, which was lost in the data, beyond the Rayleigh or Nyquist resolution limit of the acquisition system [2], [3]. In this paper, we consider the new setting, where the atoms are *lines*. This approach will provide a very high accuracy for the lines estimation, where the Hough and the Radon transforms fail, due to their discrete nature. Our motivation stems from the frequent presence in biomedical images, e.g. in microscopy, of elongated structures like filaments, neurons, veins, which are deteriorated when reconstructed with classical penalties.

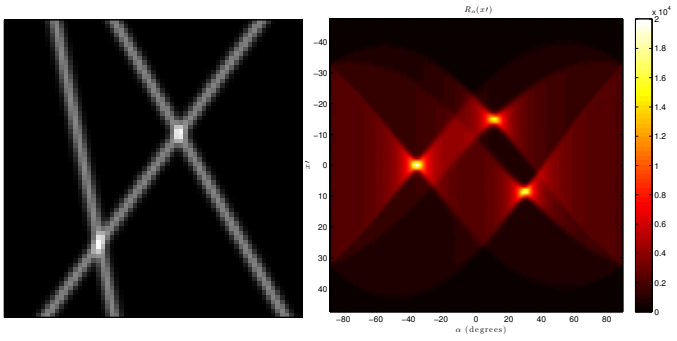


Fig. 1: The image $x^\sharp$ of three blurred lines (on the left) and the Radon transform of $x^\sharp$ (on the right)

II. Problem formulation

Our aim is to restore an image $x^\sharp$ containing lines, and to estimate the parameters— angle, offset, amplitude— of the lines, given degraded data y. In this section, we formulate what we precisely mean by an image containing lines. In short, $x^\sharp$ is a sum of perfect lines, which have been blurred and then sampled.

We place ourselves in the quotient space $\mathbb{P} = \mathbb{R}/(W\mathbb{Z}) \times \mathbb{R}$, corresponding to the 2-D plane with horizontal $W-$periodicity, for some integer $W \geq 1$. To simplify the notations, we suppose that W is odd and we set $M = (W-1)/2$.

The ideal continuous model. A line of infinite length, with angle $\theta \in (-\pi/2, \pi/2]$ with respect to verticality, amplitude $\alpha > 0$, and offset $\gamma \in \mathbb{R}$ from the origin, is defined as the distribution

$$(t_1, t_2) \in \mathbb{P} \mapsto \alpha\delta\big(\cos(\theta)t_1 + \sin(\theta)t_2 + \gamma\big), \tag{2}$$

where δ is the Dirac distribution. We define the distribution $s^\sharp$, which is a sum of K different such perfect lines, for some integer $K \geq 1$, as

$$s^\sharp : (t_1, t_2) \in \mathbb{P} \mapsto \sum_{k=1}^{K} \alpha_k \delta\big(\cos(\theta_k)t_1 + \sin(\theta_k)t_2 + \gamma_k\big). \tag{3}$$

In this paper, to simplify the discussion, we suppose that the lines are rather vertical; that is, $\theta_k \in (-\pi/4, \pi/4]$, for every $k = 1, \dots, K$.

The observed image. The image $x^\sharp$ of size $W \times H$ is obtained by sampling $s^\sharp * \phi$ with unit step:

$$x^\sharp[n_1, n_2] = (s^\sharp * \phi)(n_1, n_2), \quad \forall n_1 = 0, \dots, W-1, \\ n_2 = 0, \dots, H-1, \tag{4}$$

where the point spread function ϕ, which blurs $s^\sharp$ before sampling, is separable: $\phi(t_1, t_2) = \varphi_1(t_1)\varphi_2(t_2)$.

Let us characterize the image $x^\sharp$ more precisely. Since ϕ is separable, the function $s^\sharp * \phi$ can be obtained by a first horizontal convolution with φ_1 and then a second vertical convolution with φ_2. Formally, $s^\sharp * \phi = (s^\sharp * \phi_1) * \phi_2$ with $\phi_1(t_1, t_2) = \varphi_1(t_1)\delta(t_2)$ and $\phi_2(t_1, t_2) = \delta(t_1)\varphi_2(t_2)$. So, after the first horizontal convolution, using the fact that $\delta(at) = \delta(t)/|a|$ for any $a \neq 0$, we obtain the function

$$u^\sharp = s^\sharp * \phi_1 : (t_1, t_2) \in \mathbb{P} \mapsto \sum_{k=1}^{K} \frac{\alpha_k}{\cos(\theta_k)} \varphi_1\Big(t_1 + \\ \tan(\theta_k) t_2 + \frac{\gamma_k}{\cos(\theta_k)}\Big). \tag{5}$$

We can show that, after the second vertical convolution, we get the function

$$s^\sharp * \phi = u^\sharp * \phi_2 : (t_1, t_2) \in \mathbb{P} \mapsto \\ \sum_{k=1}^{K} \alpha_k \psi_k\big(\cos(\theta_k) t_1 + \sin(\theta_k) t_2 + \gamma_k\big), \tag{6}$$

where

$$\psi_k = \left(\frac{1}{\cos(\theta_k)} \varphi_1\Big(\frac{\cdot}{\cos(\theta_k)}\Big)\right) * \left(\frac{1}{\sin(\theta_k)} \varphi_2\Big(\frac{\cdot}{\sin(\theta_k)}\Big)\right) \tag{7}$$

if $\theta_k \neq 0$ and $\psi_k = \varphi_1$ else.

We can notice that eqn. (6) can also be interpreted as follows: every line has undergone a 1-D convolution with ψ_k in the direction transverse to it.

The blur model. We assume that ϕ has the following properties:

- the function $\varphi_1 \in L^1([0, W))$ is W-periodic, bounded, such that $\int_0^W \varphi_1 = 1$, and bandlimited; that is, its Fourier coefficients $(1/W) \int_0^W \varphi_1(t_1) e^{-j2\pi m t_1 / W} dt_1$ are zero for every $m \in \mathbb{Z}$ with $|m| \geq (W+1)/2 = M + 1$. The discrete filter $g[n] = \varphi_1(n)$, with these assumptions on φ_1, has discrete Fourier coefficients which correspond to Fourier coefficients of φ_1.
- $\varphi_2 \in L^1(\mathbb{R})$, with $\int_{\mathbb{R}} \varphi_2 = 1$. In addition, the discrete filter $\big(h[n] = (\varphi_2 * \mathrm{sinc})(n)\big)_{n \in \mathbb{Z}}$, where $\mathrm{sinc}(t_2) = \sin(\pi t_2)/(\pi t_2)$, has compact support of length $2S + 1$, for some $S \in \mathbb{N}$, i.e. $h[n] = 0$ if $|n| \geq S + 1$. Note that this assumption is not restrictive, and that if φ_2 is bandlimited, we simply have $h[n] = \varphi_2(n)$.

We can also notice that if φ_1 and φ_2 are Gaussians and have same variance κ^2, it follows from (7) that ψ_k has variance $\kappa^2\big(\cos(\theta)^2 + \sin(\theta)^2\big) = \kappa^2$ as well.

Now, for every $k = 1, \dots, K$, the assumption $\theta_k \in (-\pi/4, \pi/4]$ yields $|\tan(\theta_k)| \leq 1$. So, the function $u^\sharp$ given in (5), as a function of t_2 at fixed t_1, is bandlimited: for every $t_1 \in [0, W[$, the Fourier transform $\omega_2 \mapsto \int_{\mathbb{R}} u^\sharp(t_1, t_2) e^{-j\omega_2 t_2} dt_2$, which is a distribution (sum of K Dirac combs), is zero for every $|\omega_2| \geq \pi$. Hence, it is equivalent to perform the vertical convolution of $u^\sharp$ with φ_2, with $\varphi_2 * \mathrm{sinc}$, or with the Dirac comb $\gamma : t_2 \mapsto \sum_{n=-S}^{S} h[n]\delta(t_2 - n)$, where $h[n] = (\varphi_2 * \mathrm{sinc})(n)$. So, let us define the image $v^\sharp$ obtained by sampling the function $u^\sharp$ with unit step:

$$v^\sharp[n_1, n_2] = u^\sharp(n_1, n_2), \quad \forall n_1 = 0, \dots, W-1, \\ n_2 = -S, \dots, H - 1 + S. \tag{8}$$

Then, as we have seen, we can express $x^\sharp$ from $v^\sharp$ using a discrete vertical convolution with the filter h:

$$x^\sharp[n_1, n_2] = \sum_{p=-S}^{S} v^\sharp[n_1, n_2 - p] h[p], \quad \forall n_1 = 0, \dots, W-1, \\ n_2 = 0, \dots, H-1. \tag{9}$$

Altogether, we have completely and exactly characterized the sampling process, which involves a continuous blur, using the discrete and finite filters $(g[n])_{n=0}^{W-1}$ and $(h[n])_{n=-S}^{S}$.

The objectives. Since the ideal model $s^\sharp$ is made up of Diracs, the image of the Fourier transform along columns $\hat{w}^\sharp = \mathcal{F}_1 s^\sharp$, is composed of a sum of exponentials. Our goal will be to reconstruct $\hat{w}^\sharp$ from the blurred image $x^\sharp$.

Let us further characterize the image $x^\sharp$ in Fourier domain. We define the image $\hat{v}^\sharp$ obtained by applying the 1-D discrete Fourier transform on every column of $v^\sharp$, defined in (8):

$$\hat{v}^\sharp[m, n_2] = \sum_{n_1=0}^{W-1} v^\sharp[n_1, n_2] e^{-j2\pi m n_1 / W}, \\ \forall m = -M, \dots, M, \ n_2 \in \mathbb{Z}. \tag{10}$$

Then, the $\hat{v}^\sharp[m, n_2]$ are the Fourier coefficients of the function $u^\sharp$:

$$\hat{v}^\sharp[m, n_2] = \frac{1}{W} \int_0^W u^\sharp(t_1, n_2) e^{-j2\pi m t_1 / W} dt_1, \\ \forall m = -M, \dots, M, \ n_2 \in \mathbb{Z}. \tag{11}$$

Consequently, from (5) and (11), we obtain

$$\hat{v}^\sharp[m, n_2] = \hat{g}[m] \hat{w}^\sharp[m, n_2], \forall m = -M, \dots, M, \ n_2 \in \mathbb{Z},$$

$$\hat{w}^\sharp[m, n_2] = \sum_{k=1}^{K} \frac{\alpha_k}{\cos\theta_k} e^{j2\pi(\tan(\theta_k) n_2 + \gamma_k / \cos(\theta_k)) m / W}. \tag{12}$$

Applying a 1-D Discrete Fourier transform on the first component of $x^\sharp[n_1, n_2] = v^\sharp[n_1, :] * h$, leads to the elements $\hat{x}^\sharp[m, n_2] = \hat{v}^\sharp[m, :] * h$. Let $\mathbf{A}$ denote the operator which multiplies each vector $\hat{w}^\sharp[m, :]$ by the corresponding Fourier coefficient $\hat{g}[m]$ and convolves it with the filter h. Thus, we have $\mathbf{A}\hat{w}^\sharp = \hat{x}^\sharp$. The image $x^\sharp$ of the blurred lines is affected by some noise ε, so that we observe the degraded image

$y = x^\sharp + \varepsilon$, with $\varepsilon \sim \mathcal{N}(0, \zeta^2)$ and ζ is the noise level. Our notations are explained in more details, with an illustrating drawing, in a supplementary material available on the webpage of the first author.

III. Super-Resolution Detection of Lines

A. Minimization Problem with Atomic Norm Regularization

We consider the extended image $\hat{w}^\sharp[m, n_2]$ of size $W \times H_S$, with $H_S = H + 2S$, taking into account the addition of S pixels beyond the borders for the convolution by the filter h. We introduce the dictionary $\mathcal{A} = \{a(f, \phi) \in \mathbb{C}^{|I|}, f \in [0, 1], \phi \in [0, 2\pi)\}$, in which the *atoms* are $[a(f, \phi)]_i = e^{j(2\pi f i + \phi)}, i \in I$, and simply $[a(f)]_i = e^{j2\pi f i}, i \in I$, if $\phi = 0$. Then, the rows $\hat{w}_m$ (resp. columns $\hat{w}_{n_2}$) of the matrix $\hat{w}^\sharp$, with $I = \{-M, \dots, M\}$ (resp. $I = \{0, \dots, H_S - 1\}$), can be viewed as a sum of atoms

$$\hat{w}^\sharp_{n_2} = \hat{w}^\sharp[:, n_2] = \sum_{k=1}^{K} c_k a(f_{n_2,k}), \tag{13}$$

$$\left(\text{resp.} \hat{w}^\sharp_m = \hat{w}^\sharp[m, :] = \sum_{k=1}^{K} c_k a(f_{m,k}, \phi_{m,k})^T \right), \tag{14}$$

with

$$c_k = \frac{\alpha_k}{\cos\theta_k}, \quad f_{n_2,k} = \frac{\tan(\theta_k) n_2 + \gamma_k / \cos(\theta_k)}{W},$$
$$\phi_{m,k} = \frac{2\pi\gamma_k m}{\cos(\theta_k) W}, \quad f_{m,k} = \frac{\tan(\theta_k) m}{W}. \tag{15}$$

We define for later use, the horizontal offset $\eta_k = \gamma_k / \cos(\theta_k)$, the frequency $\nu_k = \eta_k / W$ and the coefficients $d_{m,k} = c_k e^{j\phi_{m,k}}$, $e_{m,k} = e^{j\phi_{m,k}}$.

We now define the atomic norm, first introduced in [4], as

$$\|x\|_{\mathcal{A}} = \inf\{t > 0 : x \in t\text{conv}(\mathcal{A})\}$$
$$= \inf_{\substack{c'_k \geqslant 0 \\ f'_k \in [0,1] \\ \phi'_k \in [0, 2\pi)}} \left\{ \sum_k c'_k : x = \sum_k c'_k a(f'_k, \phi'_k) \right\}.$$

The Caratheodory theorem ensures that a vector x of length $N = 2M+1$ is a positive combination ($c_k > 0$) of $K \leqslant M+1$ atoms $a(f_k)$ if and only if $\mathbf{T}(x) \succcurlyeq 0$, where $\mathbf{T}$ is the Toeplitz operator

$$\mathbf{T} : (x_1, \dots, x_N) \mapsto \begin{pmatrix} x_1 & x_2 & \cdots & x_N \\ {x_2}^* & x_1 & \cdots & x_{N-1} \\ \vdots & \vdots & \ddots & \vdots \\ {x_N}^* & {x_{N-1}}^* & \cdots & x_1 \end{pmatrix}, \tag{16}$$

and $\succcurlyeq 0$ denotes positive semi-definiteness, and $*$ is complex conjugation. Also this decomposition is unique, since $K \leqslant M$. Hence,

$$\|\hat{w}^\sharp_{n_2}\|_{\mathcal{A}} = \sum_{k=1}^{K} c_k = \hat{w}^\sharp[0, n_2], \forall n_2 = 0, \dots, H_S - 1, \tag{17}$$

whereas, since the $d_{m,k}$ are complex, we simply have

$$\|\hat{w}^\sharp_m\|_{\mathcal{A}} \leqslant \sum_{k=1}^{K} c_k, \forall m = -M, \dots, M. \tag{18}$$

Moreover, we improved the result of [5, Proposition II.1]:

Proposition 1: The atomic norm $\|\hat{w}^\sharp_{n_2}\|_{\mathcal{A}}$ can be characterized by the following semidefinite program:

$$\|\hat{w}^\sharp_m\|_{\mathcal{A}} = \min_{q_m \in \mathbb{C}^{H_S}} \left\{ q_m[0] : \begin{bmatrix} \mathbf{T}(q_m) & \hat{w}^\sharp_m \\ (\hat{w}^\sharp_m)^* & q_m[0] \end{bmatrix} \succcurlyeq 0 \right\}. \tag{19}$$

The proof is in the supplementary material.

Given $\hat{y} = \hat{x}^\sharp + \hat{\varepsilon}$, we are looking for an image $\hat{w}$ which minimizes $\|\mathbf{A}\hat{w} - \hat{y}\|$, for some norm whose expression is given below, and satisfies properties (17)–(18)–(19). Since $\hat{w}$ and $\hat{y}$ are symmetric in the Fourier domain, we can only deal with the right part $\hat{w}_r$ and $\hat{y}_r$ of the image, i.e. for $m = 0, \dots, M$. We fixed a constant c greater than the oracle $c^\sharp = \sum_{k=1}^{K} c_k$. Consequently, the optimization problem is:

$$\hat{w}^\sharp_r \in \arg\min_{\hat{w}_r, q_r} \frac{1}{2} \|A\hat{w}_r - \hat{y}_r\|^2, \tag{20}$$

$$s.t \quad \begin{cases} \forall n_2 = 0, ..., H_S - 1, \forall m = 0, ..., M, & \\ \hat{w}_r[0, n_2] = \hat{w}_r[0, 0] \leqslant c, & \text{(21a)} \\ q_r[m, 0] \leqslant c, & \text{(21b)} \\ L_1(\hat{w}_r[m, :], q_r[m, :]) \succcurlyeq 0, & \text{(21c)} \\ L_2(\hat{w}_r[:, n_2]) \succcurlyeq 0, & \text{(21d)} \end{cases}$$

where $L_2 = \mathbf{T}$ and the operator L_1 maps $(\hat{w}_m, q_m)$ to the matrix defined in (19). The two subspaces of $\mathcal{M}_{M+1, H_S}(\mathbb{C})$, in which $\hat{w}_r$ and q_r lie, are called $\mathcal{W}$ and $\mathcal{Q}$, characterizing the fact that the entries $\hat{w}_r[0, :]$ and $q_r[:, 0]$ are reals ; they can be respectively endowed with the inner products:

$$\langle \hat{w}_1, \hat{w}_2 \rangle_{\mathcal{W}} = \sum_{n_2=0}^{H_S-1} \hat{w}_1[0, n_2] \hat{w}_2[0, n_2] + 2\Re \left(\sum_{m=1}^{M} \sum_{n_2=0}^{H_S-1} \hat{w}_1[m, n_2] \hat{w}_2[m, n_2]^* \right), \tag{22}$$

$$\langle q_1, q_2 \rangle_{\mathcal{Q}} = 2\Re \left(\sum_{m=0}^{M} \sum_{n_2=0}^{H_S-1} q_1[m, n_2] q_2[m, n_2]^* \right). \tag{23}$$

Let us define $L_1^m : (\hat{w}_r, q_r) \mapsto L_1(\hat{w}_r[m, :], q_r[m, :])$ and $L_2^{n_2} : (\hat{w}_r, q_r) \mapsto L_2(\hat{w}_r[:, n_2])$, $\mathcal{B}$ the subset of the Hilbert space $\mathcal{H} = \mathcal{W} \times \mathcal{Q}$ corresponding to the boundary constraints (21a)–(21b), and $\mathcal{C}$ the cone of positive matrices. Let $X = (\hat{w}_r, q_r) \in \mathcal{H}$ and ι be the indicator function of a set, then the optimization problem (20)–(21) can be rewritten in this way:

$$X^\sharp = \arg\min_{X = (\hat{w}_r, q_r) \in \mathcal{H}} \left\{ \frac{1}{2} \|\mathbf{A}\hat{w}_r - \hat{y}_r\|^2_{\mathcal{W}} + \iota_{\mathcal{B}}(X) + \sum_{m=0}^{M} \iota_{\mathcal{C}}(L_1^m(X)) + \sum_{n_2=0}^{H_S-1} \iota_{\mathcal{C}}(L_2^{n_2}(X)) \right\}. \tag{24}$$

B. Algorithm Design

The optimization problem (24) can be viewed in the framework above, involving Lipschitzian, proximable and linear composite terms [6]:

$$X^\sharp = \arg\min_{X\in\mathcal{H}} \left\{ F(X) + G(X) + \sum_{i=0}^{N-1} H_i(L_i(X)) \right\}, \quad (25)$$

with $F = \frac{1}{2}\|\mathbf{A}\cdot - \hat{y}_r\|^2_{\mathcal{W}}$, with a β–Lipschitz gradient ($\beta = \|\mathbf{A}\|^2 = 1$), $G = \iota_{\mathcal{B}}$, which is proximable, and $N = M + 1 + H_S$ linear composite terms where $H_i = \iota_{\mathcal{C}}$ and $L_i \in \{L_1^m, L_2^{n_2}\}$. We define the real $0 \leqslant \mu \leqslant M + H_S$.
Let $\tau > 0$ and $\sigma > 0$ such that

$$\frac{1}{\tau} - \sigma\mu = \frac{\beta}{1.9}. \quad (26)$$

Then the primal–dual Algorithm 1 converges to a solution $(\hat{x}, \hat{z}_0, ..., \hat{z}_{N-1})$ of the problem (25) [6, Theorem 5.1].

Algorithm 1 Primal–dual splitting algorithm for (25)

Input: $\hat{y}_r$ 1D FFT of the blurred and noisy data image y
Output: $\hat{w}_r^*$ solution of the optimization problem (20)–(21)

```
Initialize all primal and dual variables to zero
for n = 1 to Number of iterations do
    x̃_{n+1} = prox_{τG}(x_n − τ∇F(x_n) − τ Σ_{i=0}^{N−1} L_i^* z_{i,n}),
    for i = 0 to N − 1 do
        z̃_{i,n+1} = prox_{σH_i^*}(z_{i,n} + σL_i(2x̃_{n+1} − x_n)),
    end for
end for
```

The gradient of F is $\nabla F(x_n) = (\mathbf{A}^*(\mathbf{A}\hat{w}_{r,n} - \hat{y}_r), \mathbf{0})^T$. Let be $P_{\mathcal{C}}$ the projection operator onto $\mathcal{C}$, by Moreau identity:

$$\text{prox}_{\sigma H_i^*}(u) = u - \sigma \text{prox}_{\frac{H_i}{\sigma}}\left(\frac{u}{\sigma}\right) = u - \sigma P_{\mathcal{C}}\left(\frac{u}{\sigma}\right). \quad (27)$$

Let be $\hat{w}_{av} = \frac{1}{H_S}\sum_{n_2=1}^{H_S} \hat{w}_r[0, n_2]$, we get $\forall m, n_2$:

$$\text{prox}_{\tau G}(\hat{w}_r, q_r) = \begin{cases} \hat{w}_r[0,n_2] = \hat{w}_{av}, \text{ if } \hat{w}_{av} \leqslant c, \\ w_r[0,n_2] = c, \text{ otherwise,} \\ q_r[m,0] = c, \text{ if } q_r[m,0] > c. \end{cases} \quad (28)$$

Let be $M^{(1)} \in \mathcal{M}_{H_S+1}(\mathbb{C})$ and $M^{(2)} \in \mathcal{M}_{M+1}(\mathbb{C})$. We give the expression of the vectors resulting from these adjoints $L_1^*M_1 = (w_1, q_1) \in (\mathbb{C}^{H_S+1})^2$ and $L_2^*M_2 = w_2 \in \mathbb{C}^{M+1}$:

$$w_1[k] = \frac{1}{2}(m_{k,H_S+1} + {m_{H_S+1,k}}^*),$$

$$q_1[k] = \begin{cases} \Re\left\{\sum_{i=1}^{H_S+1} m_{ii}^{(1)}\right\} \text{ if } k = 1, \\ \frac{1}{2}\sum_{i=1}^{H_S+1-k} (m_{i,k+i-1}^{(1)} + {m_{k+i-1,i}^{(1)}}^*) \text{ if } k > 1, \end{cases}$$

$$w_2[k] = \begin{cases} \frac{1}{2}\Re\left\{\sum_{i=1}^{M+1} m_{ii}^{(2)}\right\} \text{ if } k = 1, \\ \frac{1}{2}\sum_{i=1}^{M+1-k} (m_{i,k+i-1}^{(2)} + {m_{k+i-1,i}^{(2)}}^*) \text{ if } k > 1. \end{cases}$$

Notice that in the previous algorithm, τ must be smaller than $\frac{1.9}{\beta}$, which is a limitation in terms of convergence speed. To overcome this issue, we subsequently developed a second algorithm, similar to Algorithm 1, but with the data fidelity term $\|\mathbf{A}\hat{w}_r - \hat{y}_r\|^2_{\mathcal{W}}$ activated through its proximity operator, instead of its gradient. We use this second algorithm, which is detailed in the supplementary material, in the experiments below, since it turned out to be faster than Algorithm 1.

C. Recovering Line Parameters by Prony Method

The goal is to estimate the parameters $(\tilde{\theta}_k, \tilde{\alpha}_k, \tilde{\eta}_k)$, which characterize the K lines, from the solution of the minimization problem $\hat{w}_r^*$, symmetrized in $\hat{w}^*$ beforehand. Let $x = (x_1, \ldots, x_{|I|})$ be a complex vector, we rearrange the elements x_i in a Toeplitz matrix $\mathbf{T}_K(x)$ of size $(|I| - K) \times (K + 1)$ and rank K as follows

$$\mathbf{T}_K(x) = \begin{pmatrix} x_{K+1} & \cdots & x_1 \\ \vdots & \ddots & \vdots \\ x_{|I|} & \cdots & x_{|I|-K} \end{pmatrix}. \quad (31)$$

We describe the recovering procedure hereafter.
– From $m = 1, \ldots, M$,

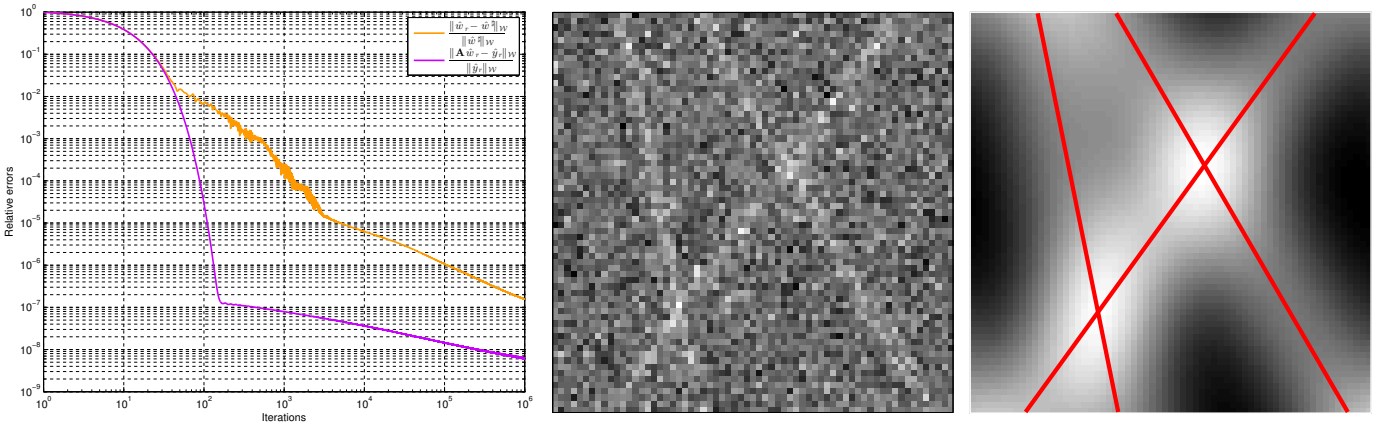
1) Compute $\tilde{f}_{m,k} = \arg(\bar{\xi}_{m,k})/(2\pi)$, where $(\xi_{m,k})_k$ are roots of the polynom $\sum_{k=0}^{K} h_{m,k}z^k$ with $\mathbf{h}_m = [h_{m,0}, \ldots, h_{m,K}]^T$ being the right singular vector of $\mathbf{T}_K(\hat{w}_m^*)$ with $I = \{0, \ldots, H_S - 1\}$. It corresponds to the singular value zero (the smallest value in practise).
2) Compute $\tilde{\theta}_{m,k} = \arctan(W\tilde{f}_{m,k}/m)$ from (15).
3) Form the matrix $\tilde{U}_m = [a(\tilde{f}_{m,1}) \cdots a(\tilde{f}_{m,K})]$, and compute $\tilde{\mathbf{d}}_m = [\tilde{d}_{m,1}, \ldots, \tilde{d}_{m,K}]^T$ by solving the least-squares linear system $U_m^H U_m \mathbf{d}_m = U_m^H \hat{w}_m^*$.
4) Compute $\tilde{c}_{m,k} = |\tilde{d}_{m,k}|$ and $\tilde{\alpha}_{m,k} = \tilde{c}_{m,k}\cos(\tilde{\theta}_{m,k})$.
5) Compute $\tilde{e}_{m,k} = \tilde{d}_{m,k}/|\tilde{d}_{m,k}|$.

– For $k = 1, \ldots, K$

1) Compute the mean of all estimated angles $\tilde{\theta}_k = \frac{1}{M}\sum_{m=1}^{M} \tilde{\theta}_{m,k}$ and amplitudes $\tilde{\alpha}_k = \frac{1}{M}\sum_{m=1}^{M} \tilde{\alpha}_{m,k}$
2) Compute the frequency $\tilde{\nu}_k$ as previously from $\mathbf{T}_K(\tilde{e}_k)$ with $\tilde{e}_k = (\tilde{e}_{m,k})_m$ and $I = \{-M, \ldots, M\}$.
3) Compute the horizontal offset $\tilde{\eta}_k = W\tilde{\nu}_k/(2\pi)$

IV. Experimental Results

The reconstruction procedure described in the previous section, was implemented in Matlab. Codes are available on the webpage of the first author. We consider an image of size $W = H = 65$, containing three lines of parameters $(\theta_1, \eta_1, \alpha_1) = (-\pi/5, 0, 255)$, $(\theta_2, \eta_2, \alpha_2) = (\pi/16, -15, 255)$ and $(\theta_3, \eta_3, \alpha_3) = (\pi/6, 10, 255)$. The first experiment consists in the reconstruction of the lines from $\hat{w}_r^*$ in absence of noise, (1) by applying the operator $\mathbf{A}$ on this solution, possibly with others kernels g and h, and then taking the 1D inverse Fourier transform ; and (2) by applying the Prony method to recover parameters of the lines, in the aim to display these one as vectorial lines. We run the algorithm for 10^6 iterations. Results of relative errors for the solution $\hat{w}_r$ and the estimated parameters are given Fig. 2 (a) and Table



(a) (b) (c)

Fig. 2: (a) Decrease of the relative errors $\frac{\|\hat{w}_r - \hat{w}^\sharp\|_{\mathcal{W}}}{\|\hat{w}^\sharp\|_{\mathcal{W}}}$ and $\frac{\|\mathbf{A}\hat{w}_r - \hat{y}_r\|_{\mathcal{W}}}{\|\hat{y}_r\|_{\mathcal{W}}}$ for the first experiment, (b) Lines affected by a strong noise level ($\zeta = 200$) for the second experiment, (c) Lines degraded by a strong blur ($\kappa = 8$) for the third experiment. In red, the recovered lines by the Prony Method.

TABLE I: Errors on line parameters recovered by the proposed method.

	Experiment 1	**Experiment 2**	**Experiment 3**
Δ_θ/θ	$(10^{-7}, 3.10^{-6}, 7.10^{-7})$	$(10^{-2}, 6.10^{-2}, 9.10^{-2})$	$(6.10^{-7}, 9.10^{-5}, 8.10^{-6})$
Δ_α/α	$(10^{-7}, 10^{-7}, 10^{-7})$	$(10^{-2}, 9.10^{-2}, 2.10^{-1})$	$(4.10^{-5}, 2.10^{-5}, 2.10^{-5})$
Δ_η	$(4.10^{-6}, 7.10^{-6}, 7.10^{-6})$	$(5.10^{-2}, 4.10^{-2}, 3.10^{-2})$	$(5.10^{-5}, 10^{-4}, 3.10^{-4})$

I, where $\Delta_{\theta_i}/\theta_i = |\theta_i - \tilde{\theta}_i|/|\theta_i|$, $\Delta_{\alpha_i}/\alpha_i = |\alpha_i - \tilde{\alpha}_i|/|\alpha_i|$ and $\Delta_{\eta_i} = |\eta_i - \tilde{\eta}_i|$. Although the algorithm is quite slow to achieve high accuracy, we insist on the fact that convergence to the exact solution $x^\sharp$ is guaranteed, when the lines are not too close to each other. The purpose of the second experiment is to highlight the robustness of the method in presence of a strong noise level. With $c = 700$ and only 2.10^3 iterations, we are able to completely remove noise and to estimate the line parameters with an error of 10^{-2}. For both first experiments, we do not depict the estimated images, because it is strictly identical to the one in Fig. 1. Finally, the last experiment for 10^5 iterations, illustrates the efficiency of the method even in presence of a large blur, yielding an error of 10^{-4}. We emphasize that our algorithm has an accuracy which could not be achieved by detecting peaks of the Hough or Radon transform. These methods are relevant for giving a coarse estimation of line parameters. They are robust to strong noise, but completely fail with a strong blur, which prevents peaks detection (see supplementary material).

V. Conclusion

We provided a new formulation for the problem of recovering lines in degraded images using the framework of atomic norm minimization. A primal–dual splitting algorithm has been used to solve the convex optimization problem. We applied it successfully to the deblurring of images, recovering lines parameters by the Prony method, and we showed the robustness of the method for strong blur and strong noise level. We insist on the novelty of our approach, which is to estimate lines with parameters (angle, offset, amplitude) living in a continuum, with perfect reconstruction in absence of noise, without being limited by the discrete nature of the image, nor its finite size. In a future work, we will study the separation conditions under which perfect reconstruction can be guaranteed, we will extend the method with no angle and periodicity restriction, and we will apply it, for instance, to inpainting problems.

Acknowledgment

The authors acknowledge the support of the French Agence Nationale de la Recherche (ANR) under reference ANR-13-BS03-0002-01 (ASTRES).

References

[1] B. N. Bhaskar, G. Tang, and B. Recht, "Atomic norm denoising with applications to line spectral estimation," vol. 61, no. 23, pp. 5987–5999, Dec. 2013.

[2] C. Fernandez-Granda, "Super-resolution and compressed sensing," *SIAM News*, vol. 46, no. 8, Oct. 2013.

[3] E. J. Candès and C. Fernandez-Granda, "Towards a mathematical theory of super-resolution," *Communications on Pure and Applied Mathematics*, vol. 67, no. 6, pp. 906–956, 2014.

[4] V. Chandrasekaran, B. Recht, P. A. Parrilo, and A. S. Willsky, "The convex geometry of linear inverse problems," *Foundations of Computational mathematics*, vol. 12, no. 6, pp. 805–849, 2012.

[5] G. Tang, B. N. Bhaskar, P. Shah, and B. Recht, "Compressed sensing off the grid," *Information Theory, IEEE Transactions on*, vol. 59, no. 11, pp. 7465–7490, 2013.

[6] L. Condat, "A primal–dual splitting method for convex optimization involving lipschitzian, proximable and linear composite terms," *Journal of Optimization Theory and Applications*, vol. 158, no. 2, pp. 460–479, 2013.