EDGE-PRESERVING NONNEGATIVE HYPERSPECTRAL IMAGE RESTORATION

Simon Henrot, Said Moussaoui†, Charles Soussen*, David Brie**

*CRAN UMR 7039, Université de Lorraine, CNRS, Faculté des Sciences, B.P. 70239, F-54506 Vandoeuvre-lès-Nancy, France

† L'UNAM Université, Ecole Centrale Nantes, IRCCyN and UMR-CNRS 6597, 1, rue de la Noé, BP 92101, Nantes, France

ABSTRACT

We consider a hyperspectral image restoration problem in which the solution is known to be nonnegative. The image estimate is obtained as the constrained minimizer of a convex criterion incorporating prior information on its spatial and spectral regularity. We previously proposed a fast algorithm for Tikhonov regularization. Here, we adapt this algorithm to edge-preserving image restoration.

Index Terms— Deconvolution, Hyperspectral imaging, Constrained convex optimization, Half-quadratic criterion

1. INTRODUCTION

In many imaging situations, the data are blurred during the acquisition process (e.g., diffraction-limited blur in fluorescence microscopy, or atmospheric turbulence caused by fluctuations in the refractive index of air in remote sensing). Under the common assumption of linear space-invariant blur, deconvolution methods can be used to restore images with higher resolution. This paper deals with the deconvolution of hyperspectral data, i.e. stacks of two-dimensional (2D) images representing the same scene captured at many overlapping spectral bands or *channels*. Since hyperspectral images may consist of thousands of pixels spanning hundreds of channels, it is crucial to select an efficient restoration strategy. The image estimate is usually obtained as the minimizer of a penalized convex objective criterion, computed as the weighted sum of: 1) a quadratic *data-fidelity* term, measuring the goodness of fit between the observed data and the blurred candidate; 2) a convex *regularization* term which incorporates prior knowledge on the solution, e.g. spatial and spectral regularity. A quadratic (ℓ_2) regularization function favors the reconstruction of smooth images while a convex, non-quadratic, differentiable (e.g. a piecewise $\ell_2-\ell_1$ penalty) function allows to perform edge-preserving restoration. A traditional way to address the latter minimization problem resorts to a majorization-minimization (MM) strategy. The MM approach consists in majorizing the objective at the current solution by a tangent majorant that is easier to minimize. In particular, Half-Quadratic (HQ) methods [1, 2, 3] use a quadratic majoration: the technique amounts to solving a sequence of quadratic minimization problems. In this framework, convergence of HQ methods is ensured [4] and these techniques have been successfully used in the field of image restoration [5, 6]. HQ techniques are commonly used in an unconstrained optimization framework. In many applications such as in computed tomography, the image is known to be nonnegative. In [7], we proposed a fast algorithm to reconstruct smooth hyperspectral data under a non-negativity constraint, with a quadratic penalty approach. In this paper, we propose to combine this method with HQ minimization to perform fast edge-preserving nonnegative restoration of hyperspectral images. This paper is organized as follows. In section 2, the imaging model is introduced and the problem is formulated in the framework of constrained optimization. Section 3 deals with the unconstrained restoration of hyperspectral data using HQ minimization. In section 4, we show how to incorporate the non-negativity constraint into the estimation process and lay out the proposed algorithm. Experimental results are presented in section 5 and we conclude in section 6.

*This work was supported by the French *Agence Nationale de la Recherche*, with reference number ANR-09-BLAN-0336-04 (HAESPRI project).

2. PROBLEM STATEMENT

2.1. Imaging model

Let us consider a hyperspectral image of N pixels acquired in L spectral bands. In each band ℓ, the observed image $\mathbf{y}_\ell$ is obtained from the true image $\mathbf{x}_\ell^0$ according to

$$\mathbf{y}_\ell = \mathbf{H}_\ell \mathbf{x}_\ell^0 + \mathbf{n}_\ell \tag{1}$$

where column vectors $\mathbf{y}_\ell$ and $\mathbf{x}_\ell^0$ are respectively the observed and true image after lexicographical ordering, $\mathbf{H}_\ell$ is the degradation matrix (corresponding to the 2D convolution matrix for the channel point spread function) and $\mathbf{n}_\ell$ is an additive noise term accounting for measurement and model errors, assumed to be white and Gaussian. Denoting the entire hyperspectral image cube by $\mathbf{y} = [\mathbf{y}_1^t \dots \mathbf{y}_L^t]^t$ and the hyperspectral degradation matrix by

$$\mathbf{H} = \begin{bmatrix} \mathbf{H}_1 & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \mathbf{0} \\ \mathbf{0} & \dots & \mathbf{0} & \mathbf{H}_L \end{bmatrix}, \tag{2}$$

(since each image plane is blurred independently), the hyperspectral degradation model simply reads:

$$\mathbf{y} = \mathbf{H}\mathbf{x}^0 + \mathbf{n} \tag{3}$$

The hyperspectral deconvolution problem consists in estimating the true image $\mathbf{x}^0$ given $\mathbf{y}$ and $\mathbf{H}$.

2.2. Problem formulation

The nonnegative restoration problem is formulated as the constrained minimization problem:

$$\min_{\mathbf{x}} \mathcal{J}(\mathbf{x}) \textit{ s.t. } \mathbf{x} \geq \mathbf{0} \tag{4}$$

where $\mathbf{x}$ is a candidate solution and criterion $\mathcal{J}(\mathbf{x})$ writes:

$$\mathcal{J}(\mathbf{x}) = \|\mathbf{y} - \mathbf{H}\mathbf{x}\|^2 + \sum_k \mu_k R_k(\mathbf{x}) \tag{5}$$

where $\| \, . \, \|$ is the Frobenius norm, $\mathcal{J}_0(\mathbf{x}) = \|\mathbf{y} - \mathbf{H}\mathbf{x}\|^2$ is the fidelity-to-data term and $\sum_k \mu_k R_k(\mathbf{x})$ is the convex regularization term (which possibly takes several distinct priors into account). The generic form of $R_k(\mathbf{x})$ is given by

$$R_k(\mathbf{x}) = \sum_{i \sim j} \phi_k(x_i - x_j). \tag{6}$$

where $\sum_{i \sim j}$ denotes summation over i and j that are in the same clique w.r.t. a spatial or spectral neighborhood system. For simplicity and without loss of generality, we rewrite this term as a simple 2D convolution:

$$R_k(\mathbf{x}) = \sum_i \phi_k(\{\mathbf{D}_k \mathbf{x}\}_i) \tag{7}$$

where $\mathbf{D}_k$ is the 2D convolution matrix for the appropriate finite-difference operator. The choice of regularization function ϕ_k dramatically impacts the shape of the solution results: opting for a quadratic (ℓ_2) function yields smooth images while choosing a ($\ell_2 - \ell_1$) function allows to preserve the image edges [1, 3]. A common choice is Huber's function whose piecewise definition is

$$\phi_{\text{Huber}}(t; \eta) = \begin{cases} t^2 & \text{for } |t| < \eta \\ \eta(2|t| - \eta) & \text{otherwise} \end{cases} \tag{8}$$

where parameter η controls the shape of the function: it is quadratic near the origin and linear towards infinity. It is worth reporting here that many works in recent years deal with ℓ_1 regularization (*total variation*) to strictly enforce parsimonious edges: see e.g. [8]. The choice of a ($\ell_2 - \ell_1$) function over a (ℓ_1) function will be justified below.

Hyperspectral images usually present some degree of similarity in spectrally neighboring image planes. Accounting for this prior can be done by adding a regularization term of the form (6) with cliques being considered w.r.t. a first-order neighborhood system in the spectral dimension. Selecting a quadratic regularization function is usually correct if the spectral sampling of the image is fine enough. Our objective function thus finally writes

$$\mathcal{J}(\mathbf{x}) = \|\mathbf{y} - \mathbf{H}\mathbf{x}\|^2 + \mu_1 \sum_i \phi_1(\{\mathbf{D}_1\mathbf{x}\}_i; \eta) + \mu_2 \|\mathbf{D}_2\mathbf{x}\|^2. \tag{9}$$

3. UNCONSTRAINED RESTORATION

3.1. Fast restoration of smooth images

In the case of quadratic regularization, the criterion becomes

$$\mathcal{J}(\mathbf{x}) = \|\mathbf{y} - \mathbf{H}\mathbf{x}\|^2 + \mu_1 \|\mathbf{D}_1\mathbf{x}\|^2 + \mu_2 \|\mathbf{D}_2\mathbf{x}\|^2. \tag{10}$$

The unconstrained minimization of $\mathcal{J}$ is a standard quadratic problem whose solution is given by

$$\mathbf{x} = (\mathbf{H}^t\mathbf{H} + \mu_1 \mathbf{D}_1^t \mathbf{D}_1 + \mu_2 \mathbf{D}_2^t \mathbf{D}_2)^\dagger \mathbf{H}^t \mathbf{y} \tag{11}$$

where $\dagger$ denotes the Moore-Penrose inverse. Because the number of variables involved in hyperspectral imaging is very large, the computational cost of a direct $NL \times NL$ inversion is usually prohibitive. We proposed in [7] a method that takes advantage of the properties of 2D convolution matrices to carry out simpler computations in the Fourier domain using Parseval's theorem. The use of 2D FFTs and element-wise operations yields a method with low complexity, dominated by $N^2/2$ inversions of $L \times L$ matrices.

For some applications (e.g. in image segmentation), the preservation of edges in the various image planes is critical and ℓ_2 regularization will yield over-smooth images. Unfortunately, if one selects Huber's function for ϕ, Parseval's theorem is not applicable anymore. Therefore, performing a fast optimization of (9) is not straightforward. In the next section, we use HQ minimization to combine ($\ell_2 - \ell_1$) regularization with the fast Fourier transform computations mentioned above.

3.2. Edge-preserving restoration

The half-quadratic methodology for image restoration was originally introduced by Geman and Reynolds (GR) [5] and Geman and Yang (GY) [6]; for a thorough understanding of these approaches, see e.g. [1, 2, 4]. In the following, we adopt GY's construction for reasons detailed below. The main idea consists in replacing the original objective $\mathcal{J}(\mathbf{x})$ by an *augmented criterion* $\mathcal{K}(\mathbf{x}, \mathbf{b})$ depending on both the original variables $\mathbf{x}$ and a set of auxiliary variables $\mathbf{b}$:

$$\mathcal{K}(\mathbf{x}, \mathbf{b}) = \|\mathbf{y} - \mathbf{H}\mathbf{x}\|^2 + \mu_1 \sum_i \left(\frac{1}{2}\left[\{\mathbf{D}_1\mathbf{x}\}_i - b_i\right]^2 + \psi(b_i)\right) + \mu_2 \|\mathbf{D}_2\mathbf{x}\|^2. \quad (12)$$

$\mathcal{K}$ is obtained from $\mathcal{J}$ by majorizing the spatial regularization term of (9) by:

$$\sum_i \left(\frac{1}{2}\left[\{\mathbf{D}_1\mathbf{x}\}_i - b_i\right]^2 + \psi_1(b_i)\right) \quad (13)$$

where ϕ_1 now denotes Huber's function defined in (8) and function ψ_1 is related to ϕ_1 by the concept of *convex duality*; see [1] for details (note however that ψ_1 needs not being computed as we will see). Criterion $\mathcal{K}(\mathbf{x}, \mathbf{b})$ is said to be half-quadratic because it is quadratic only w.r.t. $\mathbf{x}$.

In this framework, minimizing $\mathcal{J}$ w.r.t $\mathbf{x}$ is equivalent to minimizing $\mathcal{K}$ w.r.t. $\mathbf{x}$ and $\mathbf{b}$ jointly. The procedure is carried out alternatively:
• the first step consists in minimizing $\mathcal{K}$ w.r.t. $\mathbf{x}$ given $\mathbf{b}$. The solution is given by:

$$\mathbf{x} = (\mathbf{H}^t\mathbf{H} + \frac{\mu_1}{2}\mathbf{D}_1^t\mathbf{D}_1 + \mu_2\mathbf{D}_2^t\mathbf{D}_2)^\dagger(\mathbf{H}^t\mathbf{y} + \frac{\mu_1}{2}\mathbf{D}_1^t\mathbf{b}) \quad (14)$$

Similar to the evaluation of (11), this solution can be efficiently computed in the Fourier domain using the algorithm proposed in [7]. Note that the normal matrix to be inverted does not depend on $\mathbf{b}$ contrarily to GR's approach. This property allows to reduce the computational cost of inversions throughout iterations;
• in the second step, $\mathcal{K}$ is minimized w.r.t. $\mathbf{b}$ given $\mathbf{x}$. Because the problem is separable w.r.t. $\mathbf{b}$, i.e. the b_i's do not interact, it has a closed-form solution:

$$b_i = \{\mathbf{D}_1\mathbf{x}\}_i - \phi'_{\text{Huber}}(\{\mathbf{D}_1\mathbf{x}\}_i; \eta) \quad (15)$$

where ϕ'_{Huber} is the derivative of ϕ_{Huber}.

In order to reduce the computation time in this alternating scheme in large scale problems, recent research on HQ minimization study the influence of computing an approximate minimizer of $\mathcal{K}$ w.r.t. $\mathbf{x}$ given $\mathbf{b}$ (e.g., by using truncated conjugate gradient techniques) on the convergence of the relaxation algorithm. Our approach is different: here, a fast computation of the exact solution of this subproblem is performed in the Fourier domain [7].

4. NONNEGATIVE RESTORATION

4.1. Quadratic penalty method

So far, we have only considered the unconstrained restoration problem. In [7], we proposed to use the quadratic penalty method to enforce the non-negativity constraint. The minimization scheme is actually similar to half-quadratic regularization: we introduce a set of auxiliary variables $\mathbf{s}$ and an augmented criterion

$$\mathcal{L}(\mathbf{x}, \mathbf{s}; \xi) = \mathcal{J}(\mathbf{x}) + \xi\|\mathbf{x} - \mathbf{s}\|^2 \quad (16)$$

in such a way that the constrained minimization of $\mathcal{J}$ is replaced by the minimization of $\mathcal{L}$ with respect to $(\mathbf{x}, \mathbf{s})$. Here, the non-negativity constraint has been transferred to $\mathbf{s}$. The quadratic penalty method [9] consists in alternating between three steps:
• minimization of $\mathcal{L}$ w.r.t. $\mathbf{x}$ given $\mathbf{s}$ and ξ. The solution is given by:

$$\mathbf{x} = (\mathbf{H}^t\mathbf{H} + \mu_1\mathbf{D}_1^t\mathbf{D}_1 + \mu_2\mathbf{D}_2^t\mathbf{D}_2 + \xi\mathbf{I})^\dagger(\mathbf{H}^t\mathbf{y} + \xi\mathbf{s}) \quad (17)$$

where $\mathbf{I}$ is the $NL \times NL$ identity matrix;
• it is easy to see that the minimization of $\mathcal{L}$ w.r.t. $\mathbf{s}$ given $\mathbf{x}$ and ξ simply consists in a thresholding operation:

$$\mathbf{s} = \max(\mathbf{0}, \mathbf{x}); \quad (18)$$

• ξ is multiplied by a constant α greater than one to increasingly force the solution towards the feasible domain.

Within this algorithmic framework, $\mathbf{x}$ is guaranteed to converge towards the constrained minimum [9, Theorem 17.1, page 494].

4.2. Proposed algorithm for handling positivity and edge-preserving restoration

We recall that the proposed method accounts for prior information on cross-spectral regularity, pixel non-negativity and the preservation of spatial edges. After combining half-quadratic regularization and the quadratic penalty method, the minimization of the augmented criterion w.r.t. $\mathbf{x}$ given $\mathbf{b}$ and $\mathbf{s}$ yields

$$\mathbf{x} = (\mathbf{H}^t\mathbf{H} + \frac{\mu_1}{2}\mathbf{D}_1^t\mathbf{D}_1 + \mu_2\mathbf{D}_2^t\mathbf{D}_2 + \xi\mathbf{I})^\dagger(\mathbf{H}^t\mathbf{y} + \frac{\mu_1}{2}\mathbf{D}_1^t\mathbf{b} + \xi\mathbf{s}) \quad (19)$$

The algorithm is summed up in Table 1.

5. EXPERIMENTAL RESULTS

To simulate hyperspectral data, we use the linear mixing model widely used in remote sensing because it allows to account for the similarity between image planes at different wavelengths. We use real bacteria images to create two 100×100 *abundance* maps and synthetic overlapping smooth Gaussian spectra (*endmembers*) sampled on 10 points. Each pixel of the spectral image is then computed as the sum of all sources weighted by their respective abundance to generate a $100 \times 100 \times 10$ data stack. Image degradation is simulated by

— Compute the 2D Fourier Transforms of all observed image planes $\mathbf{y}_\ell$, PSFs $\mathbf{H}_\ell$ and operator $\mathbf{D}_1$. Set $\xi^{(1)} = 1$, $\alpha > 1$, $\tau < 0$, $M > 0$ and initialize $\mathbf{s}$ and $\mathbf{b}$ to zero.
— Repeat:

(a) Repeat for M iterations:
- Compute $\mathbf{x}$ in the Fourier domain using (19).
- Update $\mathbf{b}$ using (15).
- Update $\mathbf{s}$ using (18).

(b) Update $\xi^{k+1} = \alpha\xi^{(k)}$.

until some stopping criterion is met, e.g. the lowest-valued pixel of $\mathbf{x}$ is greater than some tolerance threshold τ.

Table 1. Proposed algorithm

convolving each channel with a 2D Gaussian PSF with an arbitrary fixed kernel size and standard deviation and by adding independent Gaussian noise. The approximation of physical PSFs by 2D Gaussian functions is actually reasonable in numerous cases, including fluorescence microscopy [10] and atmospheric turbulence [11]. Hyperparameters μ_1 and μ_2 are selected empirically in the restoration phase. Operator $\mathbf{D}_1$ is a Laplacian filter and $\mathbf{D}_1$ is the 1D $[-1, 1]$ derivative filter in the spectral dimension. A data set example is provided in figure 1. Our experiments (unsurprisingly) confirm that edges in the image are better preserved by $\ell_2 - \ell_1$ regularization. The running times for our algorithm with $\ell_2 - \ell_1$ and ℓ_2 regularization were respectively 55s and 13s for a MATLAB implementation on a 2.4 Ghz Intel Core 2 Duo processor with a RAM of 4 gigabytes. In our experiments, we empirically set the number of iterations in the inner loop to $M = 5$ to achieve convergence. Future work could include a technical analysis on determining the minimum value of parameter M to minimize running times while preserving convergence of the method.

6. CONCLUSION

In this paper, we proposed an algorithm to restore hyperspectral data while preserving spatial edges and accounting for spectral smoothness and pixel non-negativity. The method was originally based on the fast inversion algorithm derived in [7] for smooth image restoration. We have shown that it can be adapted to non-quadratic regularization at the cost of performing repeated fast inversions. Perspectives include comparisons with other constrained optimization methods, namely large-scale interior point algorithms [12].

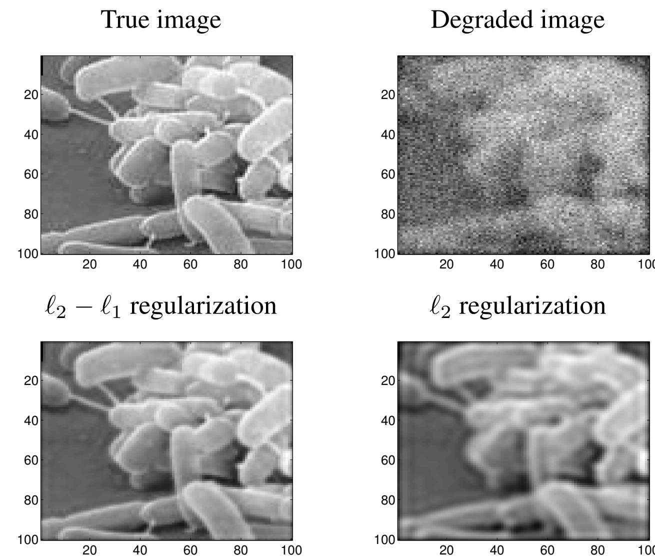

Fig. 1. Example of test images showed at one wavelength.

7. REFERENCES

[1] J. Idier, "Convex half-quadratic criteria and iteracting auxiliary variables for image restoration," *IEEE Trans. Image Processing*, vol. 10, no. 7, pp. 1001–1009, Jul. 2001.

[2] N. Nikolova and M. K. Ng, "Analysis of half-quadratic minimization methods for signal and image recovery," *SIAM J. Sci. Comput.*, vol. 27, no. 3, pp. 937–966, Dec. 2005.

[3] J. Yang, W. Yin, Y. Zhang, and Y. Wang, "A fast algorithm for edge-preserving variational multichannel image restoration," *SIAM J. Imaging Sciences*, vol. 2, no. 2, pp. 569–592, 2009.

[4] M. Allain, J. Idier, and Y. Goussard, "On global and local convergence of half-quadratic algorithms," *IEEE Trans. Image Process*, vol. 15, no. 5, pp. 1130–1142, May 2006.

[5] D. Geman and G. Reynolds, "Constrained restoration and the recovery of discontinuities," *IEEE Trans. Pattern Anal. Mach. Intell.,*, vol. 14, no. 3, pp. 367 –383, Mar. 1992.

[6] D. Geman and C. Yang, "Nonlinear image recovery with half-quadratic regularization," *IEEE Trans. Image Process*, vol. 4, no. 7, pp. 932 –946, Jul. 1995.

[7] S. Henrot, C. Soussen, and D. Brie, "Fast positive deconvolution of hyperspectral images," *IEEE Trans. Image Process*, to be published in 2012.

[8] A. Beck and M. Teboulle, "Fast gradient-based algorithms for constrained total variation image denoising and deblurring problems," *Image Processing, IEEE Transactions on*, vol. 18, no. 11, pp. 2419–2434, Nov.

[9] J. Nocedal and S. Wright, *Numerical Optimization*, Springer, 2006.

[10] B. Zhang, J. Zerubia, and J.-C. Olivo-Marin, "Gaussian approximations of fluorescence microscope point-spread function models," *Applied Optics*, vol. 46, 2007.

[11] X. Zhao, F. Wang, T. Huang, M. Ng, and R. Plemmons, "Deblurring and sparse unmixing for hyperspectral images," *Preprint*, may 2012.

[12] S. Moussaoui, E. Chouzenoux, and J. Idier, "Primal-dual interior point optimization for penalized least squares estimation of abundance maps in hyperspectral imaging," in *Proc. 4th IEEE Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing*, 2012.