ADAPTIVE DECONVOLUTION FOR PUSHBROOM
HYERSPECTRAL IMAGING SYSTEMS

Yingying Song1, El-Hadi Djermoune1, Jie Chen2, Cédric Richard3 and David Brie1

1Université de Lorraine, CNRS, Vandœuvre-lès-Nancy, France
2Northwestern Polytechnical University, Xi’an, China
3Université Côte d’Azur, CNRS, OCA, Nice, France

ABSTRACT
This paper introduces a framework based on the LMS algorithm for sequential deconvolution of images acquired by a pushbroom hyperspectral imaging system. Considering a sequential model of image blurring phenomenon, we derive a zero-attracting LMS (ZA-LMS) algorithm for 2D image deconvolution. Its transient behavior is analyzed in the mean and mean-square sense. For hyperspectral images, a spectral regularizer is added to promote a coupling between spectral bands. Simulation results are presented to evaluate the accuracy of the transient performance analysis and to illustrate the effectiveness of the deconvolution algorithm.

Index Terms—Hyperspectral image, adaptive deconvolution, LMS, ZA-LMS

1. INTRODUCTION

Hyperspectral image deconvolution consists of restoring an original image from blurred and generally noisy observations. Multichannel images restoration by Wiener methods are considered in [1,2]. Other strategies are discussed in [3–6] but only for offline processing. The aim of this paper is to address the problem of online (sequential) deconvolution of hyperspectral images provided by a pushbroom imaging system such as those described in [7,8]. These imaging systems have been used in many areas such as food safety [9,10], georeferencing [11] and material sorting [12,13].

Consider an hyperspectral image $Y \in \mathbb{R}^{N \times P \times K}$ collected by an industrial hyperspectral imaging system (pushbroom), where $N, P, K$ denote the number of spatial, spectral and time samples, respectively. The samples to be imaged are carried by a conveyor moving at constant speed. The hyperspectral image is then acquired slice by slice, each of which is denoted by $Y_k \in \mathbb{R}^{N \times P}$ (see Figure 1). Thus, the size of $Y$ increases with $k$ which can possibly grow to infinity. In a pushbroom imaging system, pixels (represented by the dots in Figure 2) are collected sequentially. The acquisition parameters are the spatial sampling $\Delta$ and integration time $T$. Assuming a constant instrument resolution (represented by the circles in Figure 2), it appears that increasing the spatial resolution, i.e., decreasing $\Delta$, results in a blurring of the image. The integration time $T$ mainly acts on the noise level: a low value for $T$ results in a low signal-to-noise ratio (SNR). Consequently, assuming a constant acquisition velocity $\Delta/T$, any resolution increase results in an increase of both blurring and noise level. This motivates the development of sequential deconvolution algorithms that are able of restoring, in an online way, an hyperspectral image $X$ from a noisy and blurred hyperspectral image $Y$.

Fig. 1: Data acquisition by a pushbroom imaging system

Fig. 2: Data acquisition with (a) low and (b) high spatial resolution. High resolution results in a blurring of the image.

The main contribution of this work is to propose an LMS framework for sequential deconvolution of images. Our algorithm operates in the spirit of LMS algorithms used for adaptive signal processing [14–16]. However, the analysis of the transient behavior differs from the canonical case (independent input sequence) usually considered for adaptive filters. In particular, to handle the correlation introduced by the convolution filter, it is necessary to introduce an appropriate extended state vector. In [17], an LMS algorithm is derived for the super-resolution restoration of image sequences. Never-
theless, it does not promote sparsity and no transient analysis is provided.

2. SEQUENTIAL IMAGE BLURRING MODEL

Following [5], hyperspectral image blurring can be seen as $P$ simultaneous spatial convolutions. For each wavelength $\lambda_p$, $p = 1, \ldots, P$, the blurred spatial image $Y^p \in \mathbb{R}^{N \times K}$ corresponds to a 2D convolution:

$$Y^p = H^p * X^p + Z^p$$

where $*$ denotes the 2D convolution operator, $X^p \in \mathbb{R}^{N \times K}$ is the image to restore, $H^p \in \mathbb{R}^{M \times L}$ is a convolution filter and $Z^p$ is an additive i.i.d. noise. In what follows, we derive a sequential formulation of model (1). Without loss of generality, we will firstly state the sequential model for 2D images, by omitting the dependence with respect to $p$. The image $X$ collected in an online way, can be represented as a sequence of vectors $x_k := [x_{1,k}, \ldots, x_{N,K}]^\top$, $k = 1, \ldots, K$, where $\top$ denotes the transpose of a matrix. We use the same notation for $Y$. The convolution filter is written as: $H = [h_1, \ldots, h_1]$ where $h_k = [h_{M,\ell}, h_{M-1,\ell}, \ldots, h_{1,\ell}]^\top$ with $\ell = 1, \ldots, L$. Model (1) can be written as follows:

$$y_k = \sum_{\ell=1}^L H_\ell x_{k-\ell+1} + z_k$$

with $z_k$ a zero-mean measurement noise of covariance $\sigma^2 I_N$, and statistically independent of other signals. $I_N$ denotes the identity matrix of size $N \times N$, and $H_\ell$ is the $N \times N$ Toeplitz matrix with first column given by $[h_{1,\ell}, \ldots, h_{M,\ell}, 0, \ldots, 0]^\top$ and first row given by $[h_{1,\ell}, 0, \ldots, 0]^\top$. Relation (2) introduces a delay in both dimensions $n$ and $k$, allowing the filter to be causal along these dimensions. To avoid the storage of $H_\ell$, convolution can be written in the frequency domain.

3. ADAPTIVE IMAGE DECONVOLUTION

We shall now derive an efficient approach for sequentially estimating vector $x_k$. Consider the following criterion which depends on past estimates $\hat{x}_{k-1}, \ldots, \hat{x}_{k-L+1}$:

$$J(\hat{x}_k | \hat{x}_{k-1}, \ldots, \hat{x}_{k-L+1}) = E\|y_k - H_\ell x_k - \sum_{\ell=2}^L H_\ell \hat{x}_{k-\ell+1}\|^2 + \eta \|x_k\|_1$$

where $\|x_k\|_1 = \sum_{\ell=1}^L |x_{\ell,k}|$ is an $\ell_1$-norm spatial regularizer promoting sparsity along the columns of the image to restore. This is motivated by the targeted application. It concerns the inspection of wood pieces on a conveyor belt which, after background calibration and correction, should ideally be zero. The strength of this regularizer is controlled by $\eta \geq 0$.

3.1. Zero-Attracting LMS algorithm

The Zero-Attracting LMS (ZA-LMS) was proposed in [15] for sparse system identification. Its application to image deconvolution was never reported. Approximating the gradient of (3) by its instantaneous value yields the ZA-LMS deconvolution algorithm defined as:

$$x_{k+1} = x_k + \mu H_\ell^\top \left(y_k - \sum_{\ell=1}^L H_\ell x_{k-\ell+1} - \rho \text{sign}(x_k)\right)$$

where $\rho = \eta \mu / 2$. Parameter $\mu$ is the step size that controls the convergence rate and stability. The sign function is defined as $\text{sign}(x) = 0$ for $x = 0$, and $\text{sign}(x) = x/|x|$ otherwise. Stability of the ZA-LMS algorithm was analyzed in [15], [18]. Despite its significant interest, few works analyzed its transient performance. We shall now characterize its stochastic behavior in the mean and mean-square sense for a time-invariant solution given by:

$$x_k = x^o$$

for all $k$. The analysis using more elaborate models such as in [19] will be considered in future works.

3.1.1. Mean behavior analysis

Let $v_k = \hat{x}_k - x^o$ be the error vector. From (4)–(5), $v_{k+1}$ can be expressed as:

$$v_{k+1} = v_k + \mu H_\ell^\top \epsilon_k - \rho \text{sign}(v_k + x^o)$$

where $\epsilon_k = y_k - \sum_{\ell=1}^L H_\ell \hat{x}_{k-\ell+1}$. Using (2) and taking the expectation of both sides, we obtain that the mean error vector evolves according to the following recursion:

$$E\{v_{k+1}\} = E\{v_k\} - \mu H_\ell^\top \sum_{\ell=1}^L H_\ell E\{\epsilon_{k-\ell+1}\} - \rho E\{\text{sign}(v_k + x^o)\}$$

The main difficulty in (7) lies in evaluating the expectation of the sign function. This point will be discussed in Section 3.1.2. Let us examine now how this expression differs from the classical system identification problem. First, observe that no approximation is required to get the mean error vector because matrices $H_\ell$ are deterministic. Because the convolution kernel is known, higher-order correlations have to be taken into account. To handle them properly, it is necessary to introduce an extended weight error vector as explained in the next section.

3.1.2. Second-order moment analysis

Consider the extended error vector $w_k \triangleq \text{col}\{v_{k-\ell+1}\}_{\ell=1}^L$, where $\text{col}\{\cdot\}$ stacks the column vectors entries on top of each
other. Let \( W_k \triangleq \mathbb{E}\{w_k w_k^\top\} \). As \( z_k \) is assumed independent of other variables, the mean-square error can be expressed as:
\[
\mathbb{E}\{\|e_k\|^2\} = N\sigma_e^2 + \text{trace}(RW_k)
\]
where \( R \triangleq [H_1, \ldots, H_L]^\top [H_1, \ldots, H_L] \). Using (6), the extended mean error vector \( w_k \) in \( W_k \) can be updated as:
\[
w_{k+1} = \text{AW}_k + \mu_B z_k - \rho C \text{sign}(w_k + u^o)
\]
with \( u^o \equiv [x^o, 0_{N \times N(L-1)}]^\top \), \( A \) defined above and:
\[
B \triangleq [H_1, 0_{N \times N(L-1)}]^\top,
\%
C \triangleq \begin{bmatrix} I_N & 0_{N \times N(L-1)} \\
0_{N(L-1) \times N} & 0_{N(L-1) \times N(L-1)}
\end{bmatrix}
\]
It follows that matrix \( W_k \) can be updated as follows:
\[
W_{k+1} = \text{AW}_k A^\top + \mu^2 \sigma_e^2 BB^\top
\]
\[-\rho A \mathbb{E}\{w_k \text{sign}(w_k + u^o)^\top\} C^\top
\]
\[-\rho C \mathbb{E}\{\text{sign}(w_k + u^o) w_k\} A^\top
\]
\[+ \rho^2 C \mathbb{E}\{\text{sign}(w_k + u^o) \text{sign}(w_k + u^o)^\top\} C^\top
\]
Again, the main difficulty lies in the evaluation of the expectations \( \mathbb{E}\{\text{sign}(u) \text{sign}(v)\} \) and \( \mathbb{E}\{u \text{sign}(v)\} \). We consider the first-order approximations:
\[
\mathbb{E}\{\text{sign}(w_k + u^o)\} \approx \text{sign}(\mathbb{E}\{w_k\} + u^o)
\]
\[\mathbb{E}\{w_k \text{sign}(w_k + u^o)^\top\} \approx \mathbb{E}\{w_k\} \text{sign}(\mathbb{E}\{w_k\} + u^o)^\top
\]
\[\mathbb{E}\{\text{sign}(w_k + u^o) \text{sign}(w_k + u^o)^\top\} \approx \text{sign}(\mathbb{E}\{w_k\} + u^o) \text{sign}(\mathbb{E}\{w_k\} + u^o)^\top
\]
which lead to accurate-enough models when the noise level is low. An analysis of the stochastic behavior of ZA-LMS is carried out in [20]. Exact expressions of these expectations are proposed under a Gaussian assumption on the error vectors.

### 3.2. Adaptive hyperspectral image deconvolution

Consider now the problem of 3D hyperspectral image deconvolution which aims at restoring slices \( X_k \in \mathbb{R}^{N \times P} \) in a sequential way. In an equivalent way, we shall consider vectorized data \( x_k \triangleq \text{col}\{x_k\}_{p=1}^P \) and \( y_k \triangleq \text{col}\{y_k\}_{p=1}^P \) where superscript \( p \) refers to the spectral band. Adding a spectral regularization term to promote spectral smoothness of the image (first-order derivative filter along the spectral dimension) leads to the criterion:
\[
\mathcal{J}(x_k | x_{k-1}, \ldots, x_{k-L+1}) = \mathbb{E}\|y_k - G_1 x_k - \sum_{\ell=2}^L G_{\ell} \hat{x}_{k-\ell+1}\|^2 
\]
\[+ \eta \|x_k\|_1 + \alpha \|G x_k\|^2
\]
where \( G_\ell \triangleq \text{blkdiag}(H^\ell)^P \) is a block-diagonal matrix, and \( \Gamma \) defined above promotes spectral smoothness via coefficients \( \{c_p\}_{p=1}^P \). Parameters \( \eta \) and \( \alpha \) define the trade-off between fitting the data, model smoothness across frequency, and sparsity across time and space.

The ZA-LMS algorithm for hyperspectral image deconvolution can be expressed as:
\[
\hat{x}_{k+1} = \hat{x}_k + \mu G_1^\top y_k - \mu \sum_{\ell=1}^L G_{\ell}^\top G_{\ell} \hat{x}_{k-\ell+1}
\]
\[-\rho \text{sign}(\hat{x}_k) - \mu \alpha \Gamma \hat{x}_k
\]
Let us mention that the hyperspectral image deconvolution shares some similarities in the problem formulation with clustered multitask networks as introduced in [21]. This will be used as a starting point to analyze the transient behavior of the algorithm (13).

### 4. EXPERIMENTAL RESULTS

The first experiment aims at validating the transient behavior analysis and comparing the algorithm with the LMS algorithm (\( \rho = 0 \)). The simulated image was constant over time \( k \). Its columns were set to: \( x^o = [0_{1 \times 3}, 1, 0.9, \ldots, 0.1, 0_{1 \times 3}]^\top \). The initial value \( x_0 \) was set to zero. The additive noise \( z_k \) was a zero-mean i.i.d. Gaussian sequence with variance 0.1. The convolution filter was a Gaussian filter of size \( 3 \times 3 \). Simulation results were obtained by averaging over 50 runs. The convergence behavior in the mean and mean-square sense of both the standard LMS and ZA-LMS are presented in Figure 3. The simulated curves (in blue) and the theoretical curves (7)–(8) (in red) are superimposed. The zero-attracting property of the ZA-LMS results in a faster convergence to
zero than the standard LMS. However, it also introduces a bias which depends on $\eta$.

Experimental results obtained on 2D images are reported in Figure 4. The original image is shown in Figure 4(a). The convolution filter $H$ was a Gaussian filter of size $30 \times 30$ with full width at half-maximum set to 15 pixels in both dimensions. Noise $z_k$ was Gaussian with SNR $= -15$ dB. The blurred noisy image is shown in Figure 4(b). Figure 4(c) and 4(d) show the results obtained with LMS and ZA-LMS, respectively. The delay due to the causal filter can be observed on the noisy and estimated images. The zero-attracting property of the $\ell_1$-norm regularizer yields a more interpretable result. However, low amplitude objects are partially eroded, which is a consequence of the bias induced by ZA-LMS.

The third experiment was designed to mimic objects carried by the conveyor of an imaging system. The resulting hyperspectral image of size $100 \times 800 \times 8$ is shown in Figure 5(a) (only 3 wavelengths are shown due to space limitation). The blurred noisy image is shown in Figure 5(b). The convolution filter was a Gaussian filter of size $20 \times 20$. Its full width at half-maximum was set to 10 pixels. The additive noise was Gaussian with SNR $= 2$ dB. The coefficients $c_p$ were all set to 1. The image restored with ZA-LMS ($\mu = 0.06, \rho = 1.2, \alpha = 0.005$) is shown in Figure 5(c). Noise removal is quite satisfying but image deblurring is limited. Better results are obtained when the size of the convolution kernel decreases.

![Image](image1)

**Fig. 3:** Transient behavior model validation

![Image](image2)

**Fig. 4:** Estimation results

![Image](image3)

**Fig. 5:** Hyperspectral image restoration

5. CONCLUSION

In this work, we proposed a ZA-LMS algorithm for adaptive deconvolution of images collected by a pushbroom imaging system. We analyzed the transient behavior of the algorithm taking into account high order correlations introduced by convolution kernel. The ZA-LMS algorithm turns out to have good noise removal property but has an limited deblurring capability. Future works will be focused on the development of a more general sequential deconvolution framework that should allow to reach the performance of batch algorithms.

6. REFERENCES


