# ESTIMATING FIRST-ORDER FINITE-DIFFERENCE INFORMATION IN IMAGE RESTORATION PROBLEMS

*Patrick L. Combettes*[1] *and Jean-Christophe Pesquet*[2]

[1]Laboratoire Jacques-Louis Lions, Université Pierre et Marie Curie – Paris 6, 75005 Paris, France
[2]Institut Gaspard Monge and UMR-CNRS 8049, Université de Marne la Vallée, 77454 Marne la Vallée Cedex 2, France

## ABSTRACT

First-order finite-difference information has been exploited in a variety of image and signal restoration settings. These approaches typically require – implicitly or explicitly – that certain attributes of the finite difference images be known *a priori*. In this paper, we propose a new statistical framework in which such attributes are estimated *a posteriori* from the observed data under the assumption that the noise is additive and Gaussian. Our analysis can be directly applied to the construction of property sets in set theoretic estimation methods. The proposed framework is illustrated through an application to image denoising.

## 1. INTRODUCTION

The standard discrete linear restoration problem is to estimate an image $\overline{x}$ in $\mathcal{H} = \mathbb{R}^{N\times M}$ from the observation of an image

$$y = L\overline{x} + u \tag{1}$$

in $\mathcal{H}$, where $L\colon \mathcal{H} \to \mathcal{H}$ is a linear operator modeling the image formation process and $u \in \mathcal{H}$ is a noise component. Typically, such a problem can be approached through a constrained minimization formulation of the form

$$\text{Find} \;\; x \in S = \bigcap_{i=1}^{m} S_i \quad \text{such that} \;\; J(x) = \inf J(S), \tag{2}$$

where $(S_i)_{1\leq i\leq m}$ are constraint sets in $\mathcal{H}$ confining the candidate solutions to the feasibility set $S$ and $J\colon \mathcal{H} \to \left]-\infty, +\infty\right]$ an objective function. In this approach, the sets $(S_i)_{1\leq i\leq m}$ are constructed from *a priori* knowledge about the original image $\overline{x}$ and the image formation model (1), e.g., [2, 5, 11, 15] and the references therein. Ultimately, one aims at incorporating as much information as possible in the formulation so as to achieve a small feasibility set $S$ and thereby obtain a reliable restoration.

Of special interest in restoration problems is information pertaining to first-order finite-differences (discrete partial derivatives) [1, 3, 5, 7, 8, 9, 10, 13, 14]. Recall that the (normalized) vertical and horizontal first-order finite-difference images associated with an image $z = (z^{k,\ell})_{1\leq k\leq N, 1\leq \ell\leq M}$ in $\mathcal{H}$ are the images $D_1 z$ and $D_2 z$ defined pixel-wise as

$$(D_1 z)^{k,\ell} = \frac{z^{k+1,\ell} - z^{k,\ell}}{\sqrt{2}}, \begin{cases} 1 \leq k \leq N_1 = N-1 \\ 1 \leq \ell \leq M_1 = M, \end{cases} \tag{3}$$

and

$$(D_2 z)^{k,\ell} = \frac{z^{k,\ell+1} - z^{k,\ell}}{\sqrt{2}}, \begin{cases} 1 \leq k \leq N_2 = N \\ 1 \leq \ell \leq M_2 = M-1 \end{cases} \tag{4}$$

($k$ and $l$ denote the vertical and horizontal coordinates, respectively). It will be convenient to define the discrete gradient of $z$ as

$$Dz = (D_1 z, D_2 z). \tag{5}$$

In (2), $Dx$ can be controlled either through the objective or through the constraints. In the former case, undesirable effects ranging from oversmoothing to staircasing can be observed. In the latter case, a finer control of the gradient is possible but it requires – implicitly or explicitly – that information on certain attributes of the original image (upper bounds on gradient energy, on total variation, on Fisher information, etc) be known *a priori*. In this paper, we propose a new statistical framework in which attributes of the discrete gradient of the degraded image $L\overline{x}$ are estimated *a posteriori* from the observed data $y$ under the assumption that the noise $u$ in (1) is Gaussian. In turn, we shall obtain closed and convex constraint sets of the form

$$S_i = \big\{x \in \mathcal{H} \mid \varphi_i(DLx) \leq \delta_i\big\}, \tag{6}$$

where $\varphi_i$ is a real-valued convex function and $\delta_i \in \mathbb{R}$. The addition of these sets in the set theoretic formulation lead to improved estimates. Numerically, the sets thus obtained can be handled via convex projection techniques such as those discussed in [5].

In the next section, we provide some background and describe the proposed set construction technique. An asymptotic statistical analysis is performed in Section 3. Finally, in Section 4, the proposed framework is illustrated through an application to set theoretic image denoising.

0-7803-8554-3/04/$20.00 ©2004 IEEE.

## 2. CONSTRAINTS BASED ON FIRST-ORDER FINITE-DIFFERENCE INFORMATION

Throughout, $(\psi_i)_{1\le i\le q}$ are real-valued function defined on $\mathbb{R}$ and

$$\begin{aligned}\psi\ &: \mathbb{R}\to\mathbb{R}^q\\ &u\mapsto\ [\psi_1(u),\dots,\psi_q(u)]^\top.\end{aligned} \tag{7}$$

A key tool in our framework is the so-called Stein's principle.

**Lemma 1** [12] *Suppose that $A$ and $B$ are real-valued random variables such that*

i) $\mathsf{E}|A|^2<+\infty$*;*

ii) *$B-A$ is a zero-mean Gaussian random variable with variance $\sigma^2$;*

iii) *$A$ and $B-A$ are independent;*

iv) *$\psi_i$ is continuous, piecewise differentiable, and*

$$(\forall\theta\in\mathbb{R})\ \lim_{|v|\to+\infty}\psi_i(v)\exp\Big(-\frac{(v-\theta)^2}{2\sigma^2}\Big)=0;$$

v) $0<\mathsf{E}|\psi_i(B)|^2<+\infty$ *and* $\mathsf{E}|\psi_i'(B)|<+\infty$.

*Then* $\mathsf{E}(A\psi_i(B))=\mathsf{E}(B\psi_i(B))-\sigma^2\mathsf{E}\psi_i'(B)$.

For $d\in\{1,2\}$, we can write $\beta_d=\overline{\alpha}_d+\gamma_d$, where

$$\overline{\alpha}_d=D_d(L\overline{x}),\ \beta_d=D_dy,\ \text{and}\ \gamma_d=D_du. \tag{8}$$

If we assume that $L\overline{x}$ and $u$ are realizations of some random fields, then $\overline{\alpha}_d^{k,\ell}$, $\beta_d^{k,\ell}$, and $\gamma_d^{k,\ell}$ are realizations of some random variables $\overline{A}_d^{k,\ell}$, $B_d^{k,\ell}$, and $C_d^{k,\ell}$, respectively. Now suppose that $\psi_i$, $\overline{A}_d^{k,\ell}$, and $B_d^{k,\ell}$ satisfy the assumptions stated in Lemma 1 (in particular, $C_d^{k,\ell}$ is Gaussian with variance $\sigma^2$). Then we derive from Lemma 1 the following vector identity

$$\mathsf{E}\big(\overline{A}_d^{k,\ell}\psi(B_d^{k,\ell})\big)=\mathsf{E}\big(B_d^{k,\ell}\psi(B_d^{k,\ell})\big)-\sigma^2\mathsf{E}\psi'(B_d^{k,\ell}), \tag{9}$$

where $\psi'$ is the gradient of $\psi$. In practice, empirical means will be used instead of the above expectations. Thus, under the assumptions to be explicitly stated in Assumption 2, these expectations will be estimated by the consistent statistics

$$\mathsf{E}\big(\overline{A}_d^{k,\ell}\psi(B_d^{k,\ell})\big)\approx\frac{1}{N_dM_d}\sum_{k=1}^{N_d}\sum_{\ell=1}^{M_d}\overline{\alpha}_d^{k,\ell}\psi(\beta_d^{k,\ell}) \tag{10}$$

and

$$\mathsf{E}\big(B_d^{k,\ell}\psi(B_d^{k,\ell})\big)-\sigma^2\mathsf{E}\psi'(B_d^{k,\ell})\approx\frac{\eta_d^{N,M}}{N_dM_d}, \tag{11}$$

where

$$\eta_d^{N,M}=\sum_{k=1}^{N_d}\sum_{\ell=1}^{M_d}\beta_d^{k,\ell}\psi(\beta_d^{k,\ell})-\sigma^2\sum_{k=1}^{N_d}\sum_{\ell=1}^{M_d}\psi'(\beta_d^{k,\ell}). \tag{12}$$

Naturally, due to the error incurred by this empirical estimation process, the equality in (9) must be reformulated as a statistically relevant inequality. In other words, (9) must be replaced by

$$\left\|(V_d^{N,M})^{-1/2}\Big(\sum_{k=1}^{N_d}\sum_{\ell=1}^{M_d}\overline{\alpha}_d^{k,\ell}\psi(\beta_d^{k,\ell})-\eta_d^{N,M}\Big)\right\|\le\zeta_d^{N,M}, \tag{13}$$

where $\|\cdot\|$ denotes the Euclidean norm on $\mathbb{R}^q$, and where the $q\times q$ symmetric positive definite matrix $V_d^{N,M}$ and the upper bound $\zeta_d^{N,M}$ will be derived from the asymptotic distribution of the statistics and some confidence level. To sum up, (13) confines a solution $x$ to the recovery problem to the set

$$S_d=\Big\{x\in\mathcal{H}\mid\ \Big\|(V_d^{N,M})^{-1/2}\Big(\sum_{k=1}^{N_d}\sum_{\ell=1}^{M_d}\alpha_d^{k,\ell}\psi(\beta_d^{k,\ell})-\eta_{N,M}^d\Big)\Big\|\le\zeta_d^{N,M}\Big\}, \tag{14}$$

where $DLx=(\alpha_1,\alpha_2)$.

## 3. ASYMPTOTIC ANALYSIS

The objective of this section is to determine the matrix $V_d^{N,M}$ and the confidence bound $\zeta_d^{N,M}$ arising in the construction of the property set (14).

We assume that the images $((L\overline{x})_{k,\ell})_{1\le k\le N,1\le\ell\le M}$, $(y_{k,\ell})_{1\le k\le N,1\le\ell\le M}$ and $(u_{k,\ell})_{1\le k\le N,1\le\ell\le M}$ are limited-support realizations of random fields $((L\overline{X})_{k,\ell})_{(k,\ell)\in\mathbb{Z}^2}$, $(Y_{k,\ell})_{(k,\ell)\in\mathbb{Z}^2}$, and $(U_{k,\ell})_{(k,\ell)\in\mathbb{Z}^2}$ satisfying the assumptions listed below. The gradients of these fields will be denoted as in the previous sections.

**Assumption 2**

i) $((L\overline{X})_{k,\ell})_{(k,\ell)\in\mathbb{Z}^2}$ *has finite variance and is a stationary $K$-dependent field, i.e., for all subsets $\Lambda_1$ and $\Lambda_2$ of $\mathbb{Z}^2$ such that*

$$\inf_{\substack{(k_1,\ell_1)\in\Lambda_1\\(k_2,\ell_2)\in\Lambda_2}}\max\{|k_1-k_2|,|\ell_1-\ell_2|\}>K, \tag{15}$$

*the families $((L\overline{X})_{k,\ell})_{(k,\ell)\in\Lambda_1}$ and $((L\overline{X})_{k,\ell})_{(k,\ell)\in\Lambda_2}$ are independent;*

ii) *$(U_{k,\ell})_{(k,\ell)\in\mathbb{Z}^2}$ is a zero-mean Gaussian i.i.d. field with variance $\sigma^2$;*

iii) $((L\overline{X})_{k,\ell})_{(k,\ell)\in\mathbb{Z}^2}$ *and* $(U_{k,\ell})_{(k,\ell)\in\mathbb{Z}^2}$ *are independent;*

iv) $\psi_i$ *is continuous, piecewise differentiable, and*

$$(\forall\theta\in\mathbb{R})\ \lim_{|v|\to+\infty} v\,\psi_i(v)^2\exp\Big(-\frac{(v-\theta)^2}{2\sigma^2}\Big)=0;$$

v) *for* $d\in\{1,2\}$, $0<\mathsf{E}\|\psi(B_d^{0,0})\|^2<+\infty$ *and* $\mathsf{E}\|\psi'(B_d^{0,0})\|^2<+\infty$.

In view of (13), to obtain the desired asymptotic results, it is useful to study the properties of the random vector

$$\begin{aligned} Z_d^{k,\ell} &= \overline{A}_d^{k,\ell}\psi(B_d^{k,\ell}) - B_d^{k,\ell}\psi(B_d^{k,\ell}) + \sigma^2\psi'(B_d^{k,\ell}) \\ &= -C_d^{k,\ell}\psi(B_d^{k,\ell}) + \sigma^2\psi'(B_d^{k,\ell}). \end{aligned} \tag{16}$$

It must be noted that Assumption 2.ii) implies that $(C_d^{k,\ell})_{(k,\ell)\in\mathbb{Z}^2}$ is a Gaussian moving average (MA) field of order 1 with variance $\sigma^2$. It therefore follows from Assumptions 2.i) and 2.iii) that $(B_d^{k,\ell})_{(k,\ell)\in\mathbb{Z}^2}$ is a $(K+1)$-dependent field. After some calculations, we arrive at the following facts [6].

**Lemma 3** *The random vector field* $(Z_1^{k,\ell})_{(k,\ell)\in\mathbb{Z}^2}$ *is a zero-mean,* $(K+1)$*-dependent stationary field with covariance matrix function given, for all* $(p,q)\in\mathbb{Z}^2$, *by* $\Gamma_1^{p,q}=\mathsf{E}(Z_1^{k+p,\ell+q}(Z_1^{k,\ell})^\top)$ *where*

$$\begin{aligned} \Gamma_1^{0,0} &= \sigma^2\mathsf{E}(\psi(B_1^{k,\ell})\psi(B_1^{k,\ell})^\top) + \\ &\qquad \sigma^4\mathsf{E}(\psi'(B_1^{k,\ell})\psi'(B_1^{k,\ell})^\top), \\ \Gamma_1^{1,0} &= (\Gamma_1^{-1,0})^\top = -\frac{\sigma^2}{2}\mathsf{E}(\psi(B_1^{k+1,\ell})\psi(B_1^{k,\ell})^\top) + \\ &\qquad \frac{\sigma^4}{4}\mathsf{E}(\psi'(B_1^{k+1,\ell})\psi'(B_1^{k,\ell})^\top), \\ \Gamma_1^{p,q} &= 0, \quad \text{if } |p|>1 \text{ or } q\neq 0. \end{aligned}$$

**Theorem 4** [6] *Let* $\mathcal{N}$ *be a standard normal random vector of dimension q, suppose that Assumption 2 is satisfied, and set*

$$\begin{aligned} D_{0,1}^{N,M} &= \sigma^2\sum_{k=1}^{N-1}\sum_{\ell=1}^{M}\psi(B_1^{k,\ell})\psi(B_1^{k,\ell})^\top + \\ &\qquad \sigma^4\sum_{k=1}^{N-1}\sum_{\ell=1}^{M}\psi'(B_1^{k,\ell})\psi'(B_1^{k,\ell})^\top, \\ D_{1,1}^{N,M} &= -\frac{\sigma^2}{2}\sum_{k=1}^{N-2}\sum_{\ell=1}^{M}\psi(B_1^{k+1,\ell})\psi(B_1^{k,\ell})^\top + \\ &\qquad \frac{\sigma^4}{4}\sum_{k=1}^{N-2}\sum_{\ell=1}^{M}\psi'(B_1^{k+1,\ell})\psi'(B_1^{k,\ell})^\top \\ V_1^{N,M} &= D_{0,1}^{N,M} + D_{1,1}^{N,M} + (D_{1,1}^{N,M})^\top, \end{aligned}$$

*and*

$$E_1^{N,M} = \sum_{k=1}^{N-1}\sum_{\ell=1}^{M}\overline{A}_1^{k,\ell}\psi(B_1^{k,\ell}) - B_1^{k,\ell}\psi(B_1^{k,\ell}) + \sigma^2\psi'(B_1^{k,\ell}).$$

*Then, as* $N,M\to+\infty$, $\mathsf{Cov}(E_1^{N,M})^{-1}V_1^{N,M}\overset{\text{a.s.}}{\to} I$ *and* $(V_1^{N,M})^{-1/2}E_1^{N,M}\overset{\text{d}}{\to}\mathcal{N}$.

Note that $V_1^{N,M}$ can be computed from the observed data. Besides, for $N$ and $M$ large enough, the bound $\zeta_1^{N,M}$ in (14) can be determined for a preset confidence level $p_1\in]0,1[$ as the solution to the equation

$$\int_{\substack{u\in\mathbb{R}^q\\ \|u\|\leq\zeta_1^{N,M}}}\exp\Big(-\frac{\|u\|^2}{2}\Big)du = (2\pi)^{q/2}p_1. \tag{17}$$

Similar results can be obtained for the constraints on the horizontal finite differences ($d=2$) by interchanging the roles played by the spatial indices $k$ and $\ell$.

It should be noted that, at the expense of a finer analysis, assumption 2.i) above can be replaced by weaker mixing assumptions [6].

## 4. SIMULATION EXAMPLE

The original image is the $256\times256$ 8-bit image $\overline{x}$ shown in Fig. 1. The degraded image $y$ is obtained by adding zero-mean Gaussian white noise to $\overline{x}$. The MSE is 897. Shown in Fig. 2 is a $88\times88$ portion of $y$. The constraint sets $S_1$ and $S_2$ defined in (14) are used with an overall 0.8 confidence level. The functions involved in this example are

- $\psi_1\colon v\mapsto\tanh(v/a)$,
- $\psi_2\colon v\mapsto v\big(\tanh((v+\chi)/a)-\tanh((v-\chi)/a)\big)$,

where $a\in\mathbb{R}_+^*$ and $\chi\in\mathbb{R}_+^*$. Considerations leading to the choice of these functions can be found in [6].

We assume that the variance $\sigma^2$ the noise is known, as well as the range of the pixel values of $\overline{x}$ and their mean $\mu$. This leads to the constraint sets $S_3=[0,255]^{N\times M}$ and

$$S_4=\Big\{x\in\mathbb{R}^{N\times M}\mid\sum_{k=1}^{N}\sum_{\ell=1}^{M}x^{k,\ell}=NM\mu\Big\}. \tag{18}$$

The denoised image is obtained via the variational formulation (2) in which

$$J\colon x\mapsto\|x-y\|^2. \tag{19}$$

The resulting constrained optimization problem was solved numerically with the algorithm proposed in [4]. The MSE is 188. The $88\times88$ portion of this image corresponding to Fig. 2 is shown in Fig. 3. The denoising performance of the proposed approach based on gradient constraints can also be assessed qualitatively by comparing Figs. 2 and 3.

## 5. REFERENCES


[1] J. M. Borwein, A. S. Lewis, M. N. Limber, and D. Noll, “Maximum entropy reconstruction using derivative information part 2: Computational results,” *Numer. Math.,* vol. 69, pp. 243–256, 1995.

[2] P. L. Combettes, “The convex feasibility problem in image recovery,” in *Advances in Imaging and Electron Physics*, vol. 95, pp. 155-270. New York: Academic, 1996.

[3] P. L. Combettes and J. Luo, “An adaptive level set method for nondifferentiable constrained image recovery,” *IEEE Trans. Image Processing,* vol. 11, pp. 1295–1304, Nov. 2002.

[4] P. L. Combettes, “A block-iterative surrogate constraint splitting method for quadratic signal recovery,” *IEEE Trans. Signal Processing,* vol. 51, pp. 1771-1782, July 2003.

[5] P. L. Combettes and J.-C. Pesquet, “Image restoration subject to a total variation constraint,” *IEEE Trans. Image Processing,* vol. 13, 2004.

[6] P. L. Combettes and J.-C. Pesquet, “Stein’s principle in set theoretic estimation,” Technical report, 2004.

[7] T. Kärkkäinen, K. Majava, and M. M. Mäkelä, “Comparison of formulations and solution methods for image restoration problems,” *Inverse Problems,* vol. 17, pp. 1977–1995, 2001.

[8] S. L. Keeling, “Total variation based convex filters for medical imaging,” *Appl. Math. Comput.,* vol. 139, pp. 101–119, 2003.

[9] M. Nikolova, “Local strong homogeneity of a regularized estimator,” *SIAM J. Appl. Math.,* vol. 61, pp. 633–658, 2000.

[10] S. Osher and N. Paragios (Eds.), *Geometric Level Set Methods in Imaging, Vision, and Graphics,* New York: Springer, 2003.

[11] H. Stark and Y. Yang, *Vector Space Projections : A Numerical Approach to Signal and Image Processing, Neural Nets, and Optics.* New York: Wiley, 1998.

[12] C. M. Stein, “Estimation of the mean of a multivariate normal distribution,” *Ann. Stat.*, vol. 9, pp. 1135-1151, 1981.

[13] S. Teboul, L. Blanc-Féraud, G. Aubert, and M. Barlaud, “Variational approach for edge-preserving regularization using coupled PDE’s,” *IEEE Trans. Image Processing,* vol. 7, pp. 387–397, Mar. 1998.

[14] D. M. Titterington, “General structure of regularization procedures in image reconstruction,” *Astron. Astrophys.,* vol. 144, pp. 381–387, Mar. 1985.

[15] D. C. Youla and H. Webb, “Image restoration by the method of convex projections: Part 1 – theory,” *IEEE Trans. Medical Imaging,* vol. 1, pp. 81-94, Oct. 1982.


**Fig. 1**. Original image.

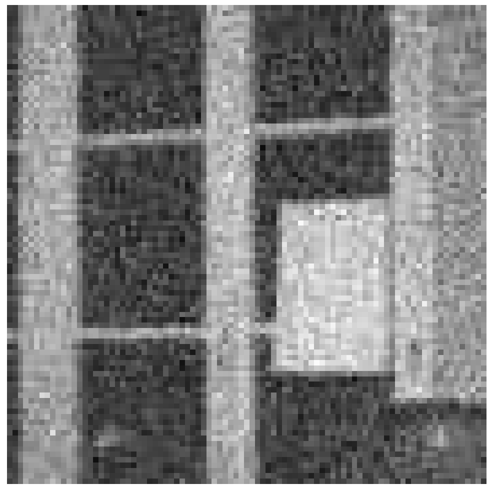

**Fig. 2**. $88 \times 88$ portion of the noisy image.

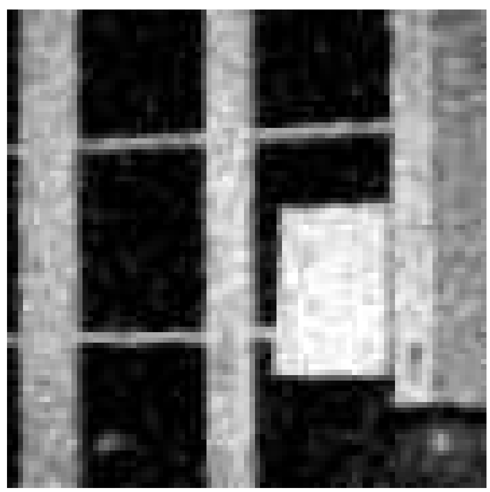

**Fig. 3**. $88 \times 88$ portion of the denoised image.