Available online at www.sciencedirect.com
ScienceDirect
Journal of Computational and Applied Mathematics 210 (2007) 47–55
JOURNAL OF COMPUTATIONAL AND APPLIED MATHEMATICS
www.elsevier.com/locate/cam

Statistically based multiwavelet denoising☆

S. Bacchelli, S. Papi*

Department of Mathematics, University of Bologna, Italy

Received 19 July 2005; received in revised form 15 June 2006

Abstract

In this work, we consider a statistically based multiwavelet thresholding method which acts on the empirical wavelet coefficients in groups, rather than individually, in order to obtain an edge-preserving image denoising technique. Our strategy allows us to exploit the dependencies between neighboring coefficients to make a simultaneous thresholding decision, so that estimation accuracy is increased.

By interpreting the multiwavelet analysis in a statistical context, we propose a new weighted multiwavelet matrix thresholding rule, based on the statistical modeling of empirical coefficients. This allows the thresholding decision to be adapted to the local structure of the underlying image, hence producing edge-preserving denoising. Extensive numerical results are presented showing the performance of our denoising procedure.

MSC: 65D; 65Y20; 65F

Keywords: Multiwavelets thresholding; Image denoising; Weighted thresholding rule

1. Introduction

In this work, we address the classical problem of removing additive Gaussian noise from a corrupted image, namely, the denoising problem. The goal of any denoising method is to eliminate the noise parts while retaining as much as possible of the important signal characteristics, such as, for example, the edges and sharp features. More precisely, given a noisy image represented by a function $\bar{f}(x, y)$, defined on a square domain I, it can be interpreted as the following sum:

$$\bar{f}(x, y) = f(x, y) + \varepsilon(x, y),$$

where $f(x, y)$ is the original image, and $\varepsilon(x, y)$ is a Gaussian noise component. Our goal is to find an approximation of $\bar{f}$ which is as close as possible to the function f, corresponding to the original non-perturbed image. This can be reformulated as the following variational problem. Let λ be a positive parameter; we wish to find a function g_λ^* that minimizes, over all possible functions g in a smoothness space Y, the functional

$$\|\bar{f} - g\|^2_{L_2(I)} + 2\lambda \|g\|_Y, \tag{1}$$

☆ This research was supported by MIUR, Cofin2004 and R.F.O. projects.
* Corresponding author.
E-mail addresses: silvia@csr.unibo.it (S. Bacchelli), spapi@csr.unibo.it (S. Papi).

0377-0427/$ - see front matter © 2006 Published by Elsevier B.V.
doi:10.1016/j.cam.2006.10.091

where the L_2-norm $\|\bar{f} - g\|_{L_2(I)}$ measures the difference between $\bar{f}$ and g, and $\|g\|_Y$ is the norm in the space Y. The positive parameter λ balances the smoothness of g with the goodness of the fit, determining the amount of noise that must be removed to obtain a good approximation. In [4], it has been shown that, if Y is the Besov space $Y = B_1^1(L_1(I))$, when the wavelet expansion of $\bar{f}$ and g are considered, the exact minimizer of (1) is obtained by means of the so-called *soft thresholding rule*, which was previously introduced by Donoho and Johnstone in [7]. This procedure estimates wavelet coefficients term by term, on the basis of their individual magnitudes. Other coefficients have no influence on the treatment of any singular coefficient. On the other hand, in [1–3,10,11,13,5,9], soft thresholding is performed by considering empirical wavelet coefficients in groups, rather than individually. Since the proper setting for working with blocks of wavelet coefficients is the multiwavelet framework, the main idea of this paper is to consider a multiwavelet decomposition of the original noisy image, in order to naturally take into account the local dependencies of neighboring coefficients. In this context, we propose a *weighted matrix shrinkage rule* which yields the exact minimizer of a new functional involving a roughness penalty term. This consists in a sum of a weighted penalty for each multiwavelet coefficient. The crucial point of our approach is to find suitable weights, depending only on the starting data, that allow us to eliminate the noise, while retaining the important signal features. Since it is well known that the multiwavelet transform coefficients can be modeled, within each subband, as independent identically distributed (i.i.d.) random variables with Generalized Gaussian distribution [13], we use the entries of each block multiwavelet coefficient to statistically estimate the standard deviation of the original non-corrupted image. This allows us to choose, for the entries of each block, weights that are inversely proportional to the ratio between the energy of the original signal and the energy of the noise, and to the magnitude of the entries themselves. In addition, we perform a data-driven selection of the parameter λ, by suitably adapting the sure estimation procedure proposed by Donoho and Johnstone [8].

Our paper is organized as follows. In Section 2, some basics on multiwavelets are given. Section 3 is devoted to our proposed *weighted matrix shrinkage*. The choice of the weight matrices is considered in Section 4, while the choice of the thresholding parameter and our encouraging numerical results are discussed in Section 5.

2. Basics on multiwavelets

Given a function $f \in L_2(\mathbb{R})$, and an orthogonal multiwavelet basis $\{\boldsymbol{\psi}_{j,k}\}_{j,k\in\mathbb{Z}}$ of $L_2(\mathbb{R})^r$, obtained by dilating and translating a vector-valued "*mother wavelet*" $\boldsymbol{\psi}(\cdot) = [\psi_1(\cdot), \psi_2(\cdot), \ldots, \psi_r(\cdot)]^{\mathrm{T}}$, i.e., $\boldsymbol{\psi}_{j,k}(\cdot) := 2^{j/2}\boldsymbol{\psi}(2^j \cdot - k)$, the multiwavelet expansion of f is given by

$$f = \sum_{j,k\in\mathbb{Z}} \mathbf{d}_{j,k}^{\mathrm{T}} \boldsymbol{\psi}_{j,k} = \sum_{j,k\in\mathbb{Z}} \langle f, \boldsymbol{\psi}_{j,k}\rangle^{\mathrm{T}} \boldsymbol{\psi}_{j,k}, \tag{2}$$

where $\mathbf{d}_{j,k} = \langle f, \boldsymbol{\psi}_{j,k}\rangle = \int_{\mathbb{R}} f(x)\boldsymbol{\psi}_{j,k}(x)\,\mathrm{d}x$ are r-vectors. A simple construction of the *mother multiwavelet* $\boldsymbol{\psi}$ can be realized by introducing the concept of multiresolution analysis (MRA) of multiplicity r, namely, a nested sequence of subspaces of $L_2(\mathbb{R})^r$, $\{V_j\}_{j\in\mathbb{Z}}$, satisfying $\ldots V_{j-1} \subset V_j \subset V_{j+1}, \ldots$, and a vector-valued *scaling function* $\boldsymbol{\phi}$, such that, for each $j \in \mathbb{Z}$, the integer translates of the jth diadic dilates of $\boldsymbol{\phi}$, $\{\boldsymbol{\phi}_{j,k}(x) := 2^{j/2}\boldsymbol{\phi}(2^j x - k), k \in \mathbb{Z}\}$ form an *orthonormal basis* for V_j (see [6] for details). More precisely, the space V_j is defined as

$$V_j := \overline{\operatorname{span}\{2^{j/2}\phi_i(2^j - k), 1 \leqslant i \leqslant r, k \in \mathbb{Z}\}}.$$

For an assigned MRA of multiplicity r $\{V_j\}_{j\in\mathbb{Z}}$ we can define the complementary space W_j of V_j, for every $j \in \mathbb{Z}$, such that $V_{j+1} = V_j \oplus W_j$. In this context, for each $j \in \mathbb{Z}$, the integer translates of the jth diadic dilates of the mother multiwavelet $\boldsymbol{\psi}$, $\{\boldsymbol{\psi}_{j,k}(x) := 2^{j/2}\boldsymbol{\psi}(2^j x - k), k \in \mathbb{Z}\}$ form an *orthonormal basis* for W_j. More precisely, the space W_j is defined as

$$W_j := \overline{\operatorname{span}\{2^{j/2}\psi_i(2^j - k), 1 \leqslant i \leqslant r, k \in \mathbb{Z}\}}.$$