# FAST WAVELET TECHNIQUES FOR NEAR-OPTIMAL IMAGE PROCESSING

RONALD A. DeVORE
Department of Mathematics
University of South Carolina
Columbia, South Carolina 29208
devore@math.scarolina.edu

BRADLEY J. LUCIER
Department of Mathematics
Purdue University
West Lafayette, Indiana 47907
lucier@math.purdue.edu

## 1. Introduction

In recent years many authors have introduced certain nonlinear algorithms for data compression, noise removal, and image reconstruction. These methods include wavelet and quadrature-mirror filters combined with quantization of filter coefficients for image compression, wavelet shrinkage estimation for Gaussian noise removal, and certain minimization techniques for image reconstruction. In many cases these algorithms have out-performed linear algorithms (such as convolution filters, or projections onto fixed linear subspaces) according to both objective and subjective measures. Very recently, several authors, e.g., [5], [7], [14], have developed mathematical theories that explain the improved performance in terms of how one measures the smoothness of images or signals. In this paper we present a unified mathematical approach that allows one to formulate both linear and nonlinear algorithms in terms of minimization problems related to the so-called $K$-functionals of harmonic analysis. We then summarize the previously developed mathematics that analyzes the image compression and Gaussian noise removal algorithms. We do not know of a specific formulation of the image reconstruction problem that supports an analysis or even definition of optimal solution. Although our framework and analysis can be applied to any $d$-dimensional signals ($d = 2$ for images, $d = 1$ for audio signals, etc.), we restrict our discussion in this paper to images.

The outline of our paper is as follows. We want to find an approximation $\tilde{f}$ to a given image $f$ on a square domain $I$, either to compress the image, remove noise from the image, etc. The size of the difference between $f$ and $\tilde{f}$ is measured by a norm, which we shall take in this paper to be the $L_2(I)$ (mean-square) norm. (We emphasize that we do *not* believe that the $L_2(I)$ norm matches the spatial-frequency–contrast response of the human visual system—we use the $L_2(I)$ norm here only because the presentation is simpler; see, e.g., [5], where we develop a theory of image compression in $L_p(I)$, with $0 < p < \infty$.) We wish to balance the smoothness of $\tilde{f}$ with the goodness of fit $\|f - \tilde{f}\|_{L_2(I)}$; to this end we consider the problem of minimizing

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

where $Y$ is a space that measures the *smoothness* of the approximations $g$. We take $\tilde{f}$ to be a function that minimizes (1). If the positive parameter $\lambda$ is large, then the smoothness of $g$ is important; if it is small, the approximation error between $f$ and $g$ is important. We consider two families of spaces in which to measure the smoothness of $\tilde{f}$, the Sobolev spaces $W^\alpha(L_2(I))$, whose functions have $\alpha$ "derivatives" in $L_2(I)$, and the Besov spaces $B^\alpha_\sigma(L_\sigma(I))$ with $\sigma = 2/(1+\alpha)$, which contain functions with $\alpha$ "derivatives" in $L_\sigma(I)$. The first family of spaces will likely be more familiar to the reader than the second, but the second family is also quite natural to consider, as it is precisely the scale of spaces with minimal smoothness to be embedded in $L_2(I)$. By using wavelet decompositions of the functions $f$ and $\tilde{f}$, we show that one can find near-optimal solutions to the variational problem (1) (or one close to it). This is possible because the norms $\|f\|_{L_2(I)}$, $\|f\|_{W^\alpha(L_2(I))}$, and $\|f\|_{B^\alpha_\sigma(L_\sigma(I))}$ can be determined *simultaneously* be examining the wavelet coefficients of $f$. The approximate minimizers $\tilde{f}$ of (1) that arise from these algorithms have surprisingly good approximation properties to $f$. In fact, when considering the problem of image compression, by setting $Y = W^\alpha(L_2(I))$ the procedure above defines a mapping $f \to \tilde{f}$ that is a near-optimal approximation from all *linear* approximation procedures to functions in $W^\alpha(L_2(I))$; and by setting $Y = B^\alpha_\sigma(L_\sigma(I))$, one finds a mapping $f \to \tilde{f}$ that is near-optimal among all *nonlinear* approximations to functions in $B^\alpha_\sigma(L_\sigma(I))$ that satisfy a certain continuous selection property. Thus, one derives both linear and nonlinear near-optimal algorithms by considering the same principle, only with different families of smoothness spaces. The analysis for noise removal methods is not as complete as for compression methods, but we report here several results of Donoho and Johnstone [13] [14], and, to a lesser extent, the present authors [8], about nonlinear noise reduction techniques. We finish with sample images that illustrate the theory in the text.

## 2. Images, smoothness spaces, and minimization problems

We begin with a function $f$ that represents a grey-scale image. To be specific, we assume that as $x := (x_1, x_2)$, $0 \le x_1 \le 1$, $0 \le x_2 \le 1$, ranges over the unit square $I := [0,1]^2 := [0,1] \times [0,1]$, $f(x)$ denotes the intensity of

MILCOM '92, IEEE Military Communications Conference Record, San Diego, California, October 11–14, 1992, IEEE, Piscataway, NJ, 1992, pages 1129–1135.

the image at location $x$. (We use $:=$ to mean "defined by.") We want to "approximate" $f$ by a "smoother" function $g$. To be precise, we must define what we mean by "approximate" and "smoother," and how we wish to trade off the size of the approximation error and the smoothness of $g$. To this end, we introduce the "error" function space $X$ with norm $\|\cdot\|_X$ and the "smoothness" function space $Y$ with corresponding norm $\|\cdot\|_Y$. (The dot in the previous expressions indicates where the missing arguments are to be placed. Thus, we could talk of a function $f$ or $f(\cdot)$.) The smoothness of the function $g$ will be measured by $\|g\|_Y$ and the size of the error between $f$ and $g$ will be given by $\|f-g\|_X$.

We shall require that any smooth function have a finite $X$-norm, so that there exists a constant $C$ such that for all $g$ in $Y$, $\|g\|_X \le C\|g\|_Y$. We shall assume also that the space $Y$ is dense in $X$, so that given any function $f$ and any positive $\epsilon$, there is a $g \in Y$ with $\|f-g\|_X < \epsilon$.

Finally, we set up our minimization problem as: fix a positive parameter $\lambda$ and find a function $\tilde{f}$ that minimizes the right side of

$$K(f,\lambda,X,Y) := \inf_{g\in Y}\{\|f-g\|_X + \lambda\|g\|_Y\}. \tag{2}$$

The function $K(f,\lambda,X,Y)$ is called the $K$-*functional* between $X$ and $Y$. The parameter $\lambda$ determines the relative importance of the smoothness of $g$ and the approximation error of $f$.

We note several properties of $K(f,\lambda,X,Y)$. First, $K(f,\lambda,X,Y) \le \|f\|_X$, since we can take $g=0$. Second, $K(f,\lambda,X,Y) \to 0$ as $\lambda \to 0$, since $Y$ is dense in $X$. In fact, one can measure the smoothness of any $f$ in $X$ in terms of how quickly $K(f,\lambda,X,Y)\to 0$ as $\lambda \to 0$. We say that $f$ is in the *interpolation space* $(X,Y)_{\theta,q}$, $0<\theta<1$, $0<q\le\infty$, if $K(f,\lambda,X,Y)$ decays fast enough as $\lambda\to 0$ that

$$\|f\|_{\theta,q} := \left(\int_0^\infty [\lambda^{-\theta}K(f,\lambda,X,Y)]^q \frac{d\lambda}{\lambda}\right)^{1/q} < \infty.$$

The spaces $(X,Y)_{\theta,q}$ are somehow intermediate in smoothness to $X$ and $Y$.

We must choose spaces $X$ and $Y$ useful for image processing. To simplify our presentation, we shall assume that $X$ is the space $L_2(I)$, the space of square-integrable functions. (Other spaces are of both theoretical and practical importance; see, e.g., image compression algorithms in $L_p(I)$, $p\neq 2$, as described in [5].)

A function $f$ is in $L_2(I)$ if

$$\|f\|^2_{L_2(I)} := \int_I |f(x)|^2\,dx < \infty.$$

Two families of spaces are useful for measuring the smoothness of $g$. The more classical spaces are the Sobolev spaces $W^\beta(L_2(I))$, which consist of all functions with $\beta$ derivatives in $L_2(I)$. When $\beta$ is an integer, we can define

$$\|g\|^2_{W^\beta(L_2(I))} := \sum_{|m|\le\beta}\|D^m f\|^2_{L_2(I)};$$

here $m:=(m_1,m_2)$, $m_i$ is a nonnegative integer, $|m| := m_1+m_2$, $D^m f := D_1^{m_1}D_2^{m_2}f$, and $D_i f := \partial f/\partial x_i$.

When $\beta$ is not an integer, the Sobolev spaces $W^\beta(L_2(I))$ can be defined in terms of the Besov spaces $B^\beta_q(L_p(I))$. Functions in these spaces have, roughly, $\beta$ "derivatives" in $L_p(I)$; the parameter $q$, which measures more subtle gradations in smoothness, is necessary for certain theorems. A precise definition of the Besov spaces of interest to us can be found in [5]. The spaces $W^\beta(L_2(I))$ can be defined by $W^\beta(L_2(I)) = B^\beta_2(L_2(I))$. In computations, it is more convenient to consider the functional

$$\inf_{g\in W^\beta(L_2(I))}\{\|f-g\|^2_{L_2(I)} + \lambda\|g\|^2_{W^\beta(L_2(I))}\}, \tag{3}$$

which, by a certain abuse of notation, we shall also call $K(f,\lambda,L_2(I),W^\beta(L_2(I)))$.

We consider a second class of smoothness spaces, the Besov spaces $B^\beta_\tau(L_\tau(I))$ with $1/\tau = \beta/2+1/2$. These spaces have $\beta$ "derivatives" in $L_\tau(I)$, where $\tau = 2/(\beta+1)$. Since $\tau<2$, and, indeed, if $\beta>1$ then $\tau<1$, we see that functions need less smoothness to be in $B^\beta_\tau(L_\tau(I))$ than to be in $W^\beta(L_2(I)) = B^\beta_2(L_2(I))$. Stated another way, there are more functions (images) with $\beta$ derivatives in $L_\tau(I)$ than there are functions with $\beta$ derivatives in $L_2(I)$.

For the pair $(X,Y) = (L_2(I), B^\beta_\tau(L_\tau(I)))$ we shall consider the functional

$$\inf_{g\in B^\beta_\tau(L_\tau(I))}\{\|f-g\|^2_{L_2(I)} + \lambda\|g\|^\tau_{B^\beta_\tau(L_\tau(I))}\}, \tag{4}$$

which we, yet again, denote by $K(f,\lambda,L_2(I),B^\beta_\tau(L_\tau(I)))$. One can show that all these choices of $K(f,\lambda,X,Y)$ yield the same family of interpolation spaces $(X,Y)_{\theta,q}$ with the natural transformation of parameters; see Bergh and Löfström [1], page 68.

The spaces $B^\beta_\tau(L_\tau(I))$ are quite natural when considering approximation in $L_2(I)$, because this family of spaces has the minimal smoothness to be embedded in $L_2(I)$. This means that given a pair $\beta$ and $\tau = 2/(1+\beta)$, there is no $\beta'<\beta$ or $\tau'<\tau$ such that $B^{\beta'}_\tau(L_\tau(I))$ or $B^\beta_{\tau'}(L_{\tau'}(I))$ is contained in $L_2(I)$.

One can ask two questions: which of these spaces can contain common images and what are the interpolation spaces between $L_2(I)$ and $W^\beta(L_2(I))$ or $B^\beta_\tau(L_\tau(I))$? First, one should note that the intensity of images often is discontinuous across lines or one-dimensional curves, and one finds that, necessarily, images with such a property (hereinafter called "images with edges") cannot be in $W^\beta(L_2(I))$ if $\beta\ge 1/2$, and cannot be in $B^\beta_\tau(L_\tau(I))$ if $\beta \ge 1$; see [5] for the calculation. So images can be in smoother classes $B^\beta_\tau(L_\tau(I))$ than $W^\beta(L_2(I))$. Second, the interpolation spaces between $L_2(I)$ and our smoothness spaces are again spaces in the same class; specifically, for $\alpha = \theta\beta$, $0<\theta<1$,

$$(L_2(I), W^\beta(L_2(I)))_{\theta,2} = W^\alpha(L_2(I))$$

(see Bergh and Löfström [1]) and

$$(L_2(I), B^\beta_\tau(L_\tau(I)))_{\theta,\sigma} = B^\alpha_\sigma(L_\sigma(I)) \quad \sigma = 2/(1+\alpha),$$

(see DeVore and Popov [11]).

## 3. Wavelet transforms

Even though our methods and analysis apply to any wavelet transform, we shall consider here only the well-known Haar transform of images; this avoids certain technicalities near the boundary of the square $I$. Some readers may object that the Haar transform lacks smoothness. In reply, we point out that if one works with images with edges, then the Haar transform has enough smoothness to prove optimal convergence results in $L_2(I)$. For example, no smoother wavelet transforms can achieve higher global asymptotic rates of convergence in image compression applications—using smoother wavelets may decrease the error by a constant factor, because of their better performance in smooth regions of an image, but they will not achieve a better *rate* of convergence.

To briefly recall the Haar transform, we introduce the functions

$$\psi^{(1)}(x) := \begin{cases} 1, & x_1 \le \frac{1}{2}, \\ -1, & x_1 > \frac{1}{2}, \end{cases} \quad \psi^{(2)}(x) := \begin{cases} 1, & x_2 \le \frac{1}{2}, \\ -1, & x_2 > \frac{1}{2}, \end{cases}$$

$\psi^{(3)}(x) = \psi^{(1)}(x)\psi^{(2)}(x)$, and $\psi^{(4)}(x) \equiv 1$ for $x \in I$. If we define

$$\Psi_k := \begin{cases} \{\psi^{(1)}, \psi^{(2)}, \psi^{(3)}, \psi^{(4)}\}, & k = 0, \\ \{\psi^{(1)}, \psi^{(2)}, \psi^{(3)}\}, & k > 0, \end{cases}$$

and $\mathbb{Z}_k^2 := \{j := (j_1, j_2) \mid 0 \le j_1 < 2^k,\ 0 \le j_2 < 2^k\}$, then the set

$$\{\psi_{j,k}(x) := 2^k \psi(2^k x - j),\ x \in I \mid \psi \in \Psi_k,\ j \in \mathbb{Z}_k^2,\ k \ge 0\}$$

forms an orthonormal basis for $L_2(I)$. This means that any function $f \in L_2(I)$ can be written as

$$f = \sum_{0 \le k} \sum_{j \in \mathbb{Z}_k^2} \sum_{\psi \in \Psi_k} c_{j,k,\psi} \psi_{j,k},\ c_{j,k,\psi} = \int_I f(x) \psi_{j,k}(x)\, dx.$$

More importantly, we have the following equivalences (see, e.g., [15]):

$$\begin{aligned} \|f\|_{L_2(I)}^2 &= \sum_{0 \le k} \sum_{j \in \mathbb{Z}_k^2} \sum_{\psi \in \Psi_k} |c_{j,k,\psi}|^2; \\ \|f\|_{W^\beta(L_2(I))}^2 &\asymp \sum_{0 \le k} \sum_{j \in \mathbb{Z}_k^2} \sum_{\psi \in \Psi_k} 2^{2\beta k} |c_{j,k,\psi}|^2,\ \beta < 1/2; \\ \|f\|_{B_\tau^\beta(L_\tau(I))}^\tau &\asymp \sum_{0 \le k} \sum_{j \in \mathbb{Z}_k^2} \sum_{\psi \in \Psi_k} |c_{j,k,\psi}|^\tau \end{aligned} \tag{5}$$

for $\beta < 1$ and $\tau = 2/(1+\beta)$. (Two notes: We have $1/\tau = \beta/d + 1/2$ for $d$-dimensional signals, and $A(f) \asymp B(f)$ means that $A(f)/B(f)$ is bounded above and below by positive constants that are independent of $f$.) Thus, the norms $\|f\|_{L_2(I)}$, $\|f\|_{W^\beta(L_2(I))}$, and $\|f\|_{B_\tau^\beta(L_\tau(I))}$ are *simultaneously* equivalent to the corresponding sequence norms on the right hand side of (5). We haven't yet bothered to define precisely the norms $\|f\|_{W^\beta(L_2(I))}$ and $\|f\|_{B_\tau^\beta(L_\tau(I))}$; we now use the sequence norms on the right side of (5) as the *definition* of these quantities.

## 4. Solving minimization problems

We consider how to minimize (3). We decompose $f$ as above, and write $g = \sum_{0 \le k} \sum_{j \in \mathbb{Z}_k^2} \sum_{\psi \in \Psi_k} d_{j,k,\psi} \psi_{j,k}$. Using the equivalent sequence norms, we wish to minimize

$$\sum_{0 \le k} \sum_{j \in \mathbb{Z}_k^2} \sum_{\psi \in \Psi_k} [(c_{j,k,\psi} - d_{j,k,\psi})^2 + \lambda 2^{2\beta k} |d_{j,k,\psi}|^2].$$

Clearly, one minimizes the entire expression by minimizing each term separately. So we consider the problem: given $t$ and $\mu > 0$, find $s$ that minimizes

$$E := (t - s)^2 + \mu s^2.$$

It is easy to see that this expression attains its minimum of $t^2\mu/(1+\mu)$ when $s = t/(1+\mu)$. By considering separately the cases $\mu \le 1$ and $\mu \ge 1$, one sees that for all $\mu > 0$,

$$\frac{1}{2} \min(\mu, 1) \le \frac{\mu}{1+\mu} \le \min(\mu, 1),$$

so that the following algorithm, which gives $E$ a value of $\min(\mu, 1)t^2$, minimizes $E$ to within a factor of 2: If $\mu < 1$ then set $s = t$, otherwise set $s = 0$. Thus, an approximate minimizer of (3) is

$$\tilde{f} = \sum_{\lambda 2^{2\beta k} < 1} \sum_{j \in \mathbb{Z}_k^2} \sum_{\psi \in \Psi_k} c_{j,k,\psi} \psi_{j,k}. \tag{6}$$

This sum consists of all $c_{j,k,\psi}\psi_{j,k}$ with $k < K$, where $K$ is the smallest integer for which $\lambda 2^{2\beta k} \ge 1$; this is a linear algorithm for finding $\tilde{f}$.

When $Y = B_\tau^\beta(L_\tau(I))$, we wish to minimize (4) or

$$\sum_{0 \le k} \sum_{j \in \mathbb{Z}_k^2} \sum_{\psi \in \Psi_k} (c_{j,k,\psi} - d_{j,k,\psi})^2 + \lambda \sum_{0 \le k} \sum_{j \in \mathbb{Z}_k^2} \sum_{\psi \in \Psi_k} |d_{j,k,\psi}|^\tau.$$

Again, one should minimize each term separately, so we consider the problem of minimizing

$$E := (t - s)^2 + \lambda |s|^\tau.$$

We can assume without loss of generality that $t \ge 0$, and, trivially, $0 \le s \le t$. Note that if $s \le t/2$ then $E \ge t^2/4$, and if $s \ge t/2$ then $E \ge \lambda t^\tau / 2^\tau$. Since $\tau < 2$, $2^\tau < 4$ and $E$ is bounded below by

$$\frac{1}{4} \min(t^2, \lambda t^\tau).$$

Thus, the following algorithm, which gives a value of $\min(t^2, \lambda t^\tau)$ to $E$, minimizes $E$ to within a factor of 4: If $\lambda t^\tau \le t^2$, set $s = t$, otherwise, set $s = 0$. Thus, an approximate minimizer of (4) is

$$\tilde{f} = \sum_{\lambda |c_{j,k,\psi}|^\tau \le |c_{j,k,\psi}|^2} c_{j,k,\psi} \psi_{j,k}. \tag{7}$$

Note that we choose all $c_{j,k,\psi}$ greater than a certain threshold, namely $|c_{j,k,\psi}| \ge \lambda^{1/(2-\tau)}$. Thus, this is a nonlinear method for choosing $\tilde{f}$.

## 5. Image compression

We assume the support of our image is the square $I$, and that a function $F(x)$ measures the intensity distribution of light at each point $x$ in $I$. We shall assume that $I$ is covered by $2^{2m}$ pixels, in $2^m$ columns and $2^m$ rows, and that each pixel $p_j$, $j \in \mathbb{Z}^2_m$, is the average of the intensity $F(x)$ on $I_{j,m}$, which we define to be the square of side-length $2^{-m}$ and lower left corner located at $j/2^m$. (This is a fair approximation of what happens with CCD cameras.) It is not hard to see that if we set $f(x) = p_j$ for $x \in I_{j,m}$ and

$$F = \sum_{0\le k}\sum_{j\in\mathbb{Z}^2_k}\sum_{\psi\in\Psi_k} c_{j,k,\psi}\psi_{j,k}, \text{ then}$$

$$f = \sum_{0\le k<m}\sum_{j\in\mathbb{Z}^2_k}\sum_{\psi\in\Psi_k} c_{j,k,\psi}\psi_{j,k}.$$

In what follows, we consider smoothness of any order $0 < \alpha < \beta$, because we want to make clear that our algorithms depend only on what family of smoothness spaces we use, not on the precise value of $\alpha$. Clearly, the $L_2(I)$, $W^\alpha(L_2(I))$, $0 < \alpha < 1/2$, and $B^\alpha_\sigma(L_\sigma(I))$, $0 < \alpha < 1$ and $\sigma = 2/(1+\alpha)$, norms of $f$ are bounded by the corresponding norms of $F$.

The linear approximation $\tilde f$ to $f$ is given by

$$\tilde f = \sum_{0\le k<K}\sum_{j\in\mathbb{Z}^2_k}\sum_{\psi\in\Psi_k} c_{j,k,\psi}\psi_{j,k}$$

where $K$ is the smallest integer such that $\lambda 2^{2\beta K} \ge 1$; see (6). There are $N := 2^{2K}$ terms in this sum, and the error between $f$ and $\tilde f$ is bounded by

$$\begin{aligned}\|f-\tilde f\|^2_{L_2(I)} &= \sum_{K\le k<m}\sum_{j\in\mathbb{Z}^2_k}\sum_{\psi\in\Psi_k} |c_{j,k,\psi}|^2\\ &\le 2^{-2\alpha K}\sum_{K\le k<m}\sum_{j\in\mathbb{Z}^2_k}\sum_{\psi\in\Psi_k} 2^{2\alpha k}|c_{j,k,\psi}|^2\\ &\le 2^{-2\alpha K}\|f\|^2_{W^\alpha(L_2(I))}.\end{aligned} \tag{8}$$

Thus, we achieve a rate of approximation

$$\|f-\tilde f\|_{L_2(I)} \le N^{-\alpha/2}\|f\|_{W^\alpha(L_2(I))}.$$

This is an optimal rate of *linear* approximation for functions in $W^\alpha(L_2(I))$; see [17], pp. 3 ff. Recall that $\alpha < 1/2$ for images with edges.

The analysis of the nonlinear algorithm is only slightly more difficult. Here, our approximation has the form

$$\tilde f = \sum_{|c_{j,k,\psi}|\ge\epsilon} c_{j,k,\psi}\psi_{j,k},$$

where $\epsilon = \lambda^{1/(2-\tau)}$. Again, we wish to bound (1) the number of terms $N$ in this sum and (2) the error $\|f - \tilde f\|_{L_2(I)}$ in terms of $N$ and $\|f\|_{B^\alpha_\sigma(L_\sigma(I))}$. We note that this is an *optimal* way to partition the coefficients: If $\Lambda$ is *any* set of $N$ coefficients, then, because we include the $N$ *largest* coefficients in $\tilde f$,

$$\|f-\tilde f\|_{L_2(I)} \le \|f - \sum_{c_{j,k,\psi}\in\Lambda} c_{j,k,\psi}\psi_{j,k}\|_{L_2(I)}.$$

Because

$$\sum_{0\le k<m}\sum_{j\in\mathbb{Z}^2_k}\sum_{\psi\in\Psi_k} |c_{j,k,\psi}|^\sigma = \|f\|^\sigma_{B^\alpha_\sigma(L_\sigma(I))},$$

it is clearly necessary that $N\epsilon^\sigma \le \|f\|^\sigma_{B^\alpha_\sigma(L_\sigma(I))}$. Consequently, $N \le \epsilon^{-\sigma}\|f\|^\sigma_{B^\alpha_\sigma(L_\sigma(I))}$ and, if $N > 0$,

$$\epsilon^{2-\sigma} \le N^{1-2/\sigma}\|f\|^{2-\sigma}_{B^\alpha_\sigma(L_\sigma(I))}.$$

Therefore, we have

$$\begin{aligned}\|f-\tilde f\|^2_{L_2(I)} &= \sum_{|c_{j,k,\psi}|<\epsilon} |c_{j,k,\psi}|^2\\ &\le \epsilon^{2-\sigma}\sum_{|c_{j,k,\psi}|<\epsilon} |c_{j,k,\psi}|^\sigma\\ &\le N^{1-2/\sigma}\|f\|^{2-\sigma}_{B^\alpha_\sigma(L_\sigma(I))}\|f\|^\sigma_{B^\alpha_\sigma(L_\sigma(I))}\\ &= N^{-\alpha}\|f\|^2_{B^\alpha_\sigma(L_\sigma(I))},\end{aligned} \tag{9}$$

since $1 - 2/\sigma = -\alpha$. Thus, for the nonlinear algorithm,

$$\|f-\tilde f\|_{L_2(I)} \le N^{-\alpha/2}\|f\|_{B^\alpha_\sigma(L_\sigma(I))}.$$

This is an optimal rate of *nonlinear* approximation for functions in $B^\alpha_\sigma(L_\sigma(I))$ among all methods that satisfy a so-called continuous selection property; see DeVore and Yu [12]. Even though the linear and nonlinear methods have the same convergence rate of $O(N^{-\alpha/2})$, the nonlinear method is better for two reasons: For a given $\alpha$, the nonlinear method achieves this convergence rate for more functions, since $B^\alpha_\sigma(L_\sigma(I))$ is much larger (contains many more images) than $W^\alpha(L_2(I))$; and, while $\alpha$ must be less than $1/2$ for the linear method applied to images with edges, the nonlinear method allows $1/2 \le \alpha < 1$.

Note that neither the linear nor nonlinear algorithms depend on the precise value of $\alpha$. In the linear case, we choose all coefficients $c_{j,k,\psi}$ with $k$ less than a predetermined value $K$. This achieves results similar to linear filtering or projections onto any other finite-dimensional subspace of dimension $2^{2K}$ of $L_2(I)$. In the nonlinear case, we choose all coefficients $c_{j,k,\psi}$ with $|c_{j,k,\psi}|$ greater than a predetermined parameter $\epsilon$. In the engineering literature, this is known as *threshold coding.* In each case, only the resulting rates of convergence depend on $\alpha$.

## 6. Removing Gaussian noise from images

In the previous section, we assumed that the measured pixel values $p_j$, $j \in \mathbb{Z}^2_m$, were the exact averages of the intensity $F(x)$ on the pixel squares $I_{j,m}$. In this section, we assume that our measurements are corrupted by Gaussian noise, that is, that we measure not $p_j$, but $\bar{p}_j := p_j + \epsilon_j$, where $\epsilon_j$ are i.i.d. (independent, identically distributed) normal random variables with mean 0 and variance $\sigma_0^2$ (denoted by $N(0,\sigma_0^2)$). (This $\sigma_0$ should not be confused with the $\sigma$ in $B^\alpha_\sigma(L_\sigma(I))$.) Because the mapping $\{2^{-m}p_j\} \to \{c_{j,k,\psi}\}$, which takes scaled pixel values to wavelet coefficients, is an orthonormal transformation, the noisy image can be represented by

$$\bar{f} = \sum_{0\le k<m} \sum_{j\in\mathbb{Z}^2_k} \sum_{\psi\in\Psi_k} \bar{c}_{j,k,\psi}\psi_{j,k},$$

where $\bar{c}_{j,k,\psi} := c_{j,k,\psi} + \delta_{j,k,\psi}$ and the $\delta_{j,k,\psi}$ are i.i.d. $N(0, 2^{-2m}\sigma_0^2)$ random variables. This model assumes that the expected value of the noise,

$$E(\|f-\bar{f}\|^2_{L_2(I)}) = \sigma_0^2,$$

is independent of the number of pixels. We now examine how the linear and nonlinear algorithms for calculating $\tilde{f}$ can be adapted for noise removal.

Starting with $\bar{f}$, the linear algorithm (6) calculates

$$\tilde{f} = \sum_{0\le k<K} \sum_{j\in\mathbb{Z}^2_k} \sum_{\psi\in\Psi_k} \bar{c}_{j,k,\psi}\psi_{j,k};$$

we will choose $K$ (hence $\lambda$) to minimize $E(\|F-\tilde{f}\|^2_{L_2(I)})$. Using the wavelet decompositions of $\tilde{f}$ and $F$, we calculate

$$\begin{aligned} E(\|F-\tilde{f}\|^2_{L_2(I)}) &= \sum_{0\le k<K} \sum_{j\in\mathbb{Z}^2_k} \sum_{\psi\in\Psi_k} E([c_{j,k,\psi}-\bar{c}_{j,k,\psi}]^2) \\ &\quad + \sum_{K\le k} \sum_{j\in\mathbb{Z}^2_k} \sum_{\psi\in\Psi_k} |c_{j,k,\psi}|^2 \\ &\le \sum_{0\le k<K} \sum_{j\in\mathbb{Z}^2_k} \sum_{\psi\in\Psi_k} E(\delta^2_{j,k,\psi}) \\ &\quad + 2^{-2\alpha K}\|F\|^2_{W^\alpha(L_2(I))} \\ &= 2^{2K-2m}\sigma_0^2 + 2^{-2\alpha K}\|F\|^2_{W^\alpha(L_2(I))}. \end{aligned}$$

The inequality follows from our estimate (8), and the second equality holds because the $2^{2K}$ random variables $\delta_{j,k,\psi}$ each have variance $2^{-2m}\sigma_0^2$.

Again we set $N := 2^{2K}$ and we minimize $E(\|F-\tilde{f}\|^2_{L_2(I)})$ with respect to $N$. Calculus shows that we overestimate the error by at most a factor of 2 if we set the two terms in our bound equal to each other, i.e., $N\sigma_0^2 2^{-2m} = N^{-\alpha}\|F\|^2_{W^\alpha(L_2(I))}$. This yields

$$N = \left(\frac{\|F\|^2_{W^\alpha(L_2(I))}2^{2m}}{\sigma_0^2}\right)^{1/(\alpha+1)},$$

and

$$E(\|F-\tilde{f}\|^2_{L_2(I)}) \le 2(2^{-2m}\sigma_0^2)^{\alpha/(\alpha+1)}\|F\|^{2/(\alpha+1)}_{W^\alpha(L_2(I))}.$$

This linear algorithm removes all terms $c_{j,k,\psi}\psi_{j,k}$ with $k$ greater than a threshold $K$; these terms can be considered to have frequency at least $2^K$. Thus, the linear method considers any low-frequency structure to be signal, and any high-frequency structure to be noise, no matter how large $c_{j,k,\psi}$, the scaled amplitude of the signal, might be. This is not acceptable to people, such as astronomers, who deal with high amplitude, small extent (and hence high frequency) signals. The nonlinear algorithm presented next recognizes high-amplitude, high-frequency structures as signals. It has been used in astronomical calculations, e.g., by White [20].

The bound for the nonlinear algorithm is much more complicated, and will not be derived here. We will, however, give the following lower bound, derived independently by Donoho and Johnstone [13], which is based on the assumption that we have extra information about which of the true coefficients $c_{j,k,\psi}$ are large.

We limit our estimator to the form

$$\tilde{f} = \sum_{0\le k} \sum_{j\in\mathbb{Z}^2_k} \sum_{\psi\in\Psi_k} \tilde{c}_{j,k,\psi}\psi_{j,k}$$

where either $\tilde{c}_{j,k,\psi} = \bar{c}_{j,k,\psi}$ or $\tilde{c}_{j,k,\psi} = 0$. In the first case, $E([c_{j,k,\psi}-\tilde{c}_{j,k,\psi}]^2) = E(\delta^2_{j,k,\psi}) = 2^{-2m}\sigma_0^2$, and in the second case, $E([c_{j,k,\psi}-\tilde{c}_{j,k,\psi}]^2) = |c_{j,k,\psi}|^2$. Thus, if we knew which coefficients $c_{j,k,\psi}$ satisfy $|c_{j,k,\psi}|^2 < 2^{-2m}\sigma_0^2$, we would obtain an optimal estimator

$$\tilde{f} = \sum_{|c_{j,k,\psi}|\ge 2^{-m}\sigma_0} \bar{c}_{j,k,\psi}\psi_{j,k}.$$

From Section 5, coefficients satisfying $|c_{j,k,\psi}| \ge 2^{-m}\sigma_0$ number at most $(2^{-m}\sigma_0)^{-\sigma}\|F\|^\sigma_{B^\alpha_\sigma(L_\sigma(I))}$, while

$$\sum_{|c_{j,k,\psi}|<2^{-m}\sigma_0} |c_{j,k,\psi}|^2 \le (2^{-m}\sigma_0)^{2-\sigma}\|F\|^\sigma_{B^\alpha_\sigma(L_\sigma(I))}.$$

Thus, $E(\|F-\tilde{f}\|^2_{L_2(I)})$ equals

$$\begin{aligned} \sum_{|c_{j,k,\psi}|\ge 2^{-m}\sigma_0} 2^{-2m}\sigma_0^2 &+ \sum_{|c_{j,k,\psi}|<2^{-m}\sigma_0} |c_{j,k,\psi}|^2 \\ &\le 2(2^{-m}\sigma_0)^{2-\sigma}\|F\|^\sigma_{B^\alpha_\sigma(L_\sigma(I))} \\ &= 2(2^{-2m}\sigma_0^2)^{\alpha/(\alpha+1)}\|F\|^{2/(\alpha+1)}_{B^\alpha_\sigma(L_\sigma(I))}, \end{aligned}$$

since $2-\sigma = 2\alpha/(\alpha+1)$ and $\sigma = 2/(\alpha+1)$. By considering functions $F$ all of whose nonzero coefficients $c_{j,k,\psi}$ satisfy either $|c_{j,k,\psi}| = 2^{-m}\sigma_0$ or $|c_{j,k,\psi}| = 2^{-m}\sigma_0 - \delta$ for some small $\delta > 0$, we see that half this value is the promised lower bound. Note that we get the same rate of approximation as in the linear case, but we are using a much weaker norm to measure the smoothness of the image $F$.

In practice, we must guess from the noisy coefficients $\bar{c}_{j,k,\psi}$ which of the $c_{j,k,\psi}$ are large. Our nonlinear approximation algorithm, when applied to the noisy data $\bar{c}_{j,k,\psi}$, is of the form

$$\tilde{f} = \sum_{|\bar{c}_{j,k,\psi}|>\epsilon} \bar{c}_{j,k,\psi}\psi_{j,k}, \tag{10}$$

and the problem is how to choose $\epsilon$ ($= \lambda^{1/(2-\tau)}$) to minimize the maximum expected error. If one chooses $\epsilon = a2^{-m}\sigma_0$ with $a$ fixed, then the expected error for $F \equiv 0$ is

$$\sigma_0^2 \frac{1}{\sqrt{2\pi}} \int_{|x|>a} x^2 e^{-x^2/2}\, dx,$$

which does not tend to zero as the number of data points $2^{2m}$ increases. Thus, $a$ must increase without bound as the number of data points tends to infinity.

The problem of optimal estimation was studied in several papers by Donoho and Johnstone culminating in [14], and in [8] by the present authors. In various settings, Donoho and Johnstone show that, asymptotically, the optimal choice of $a$ is $C\sqrt{\log(2^{2m})}$. Instead of keeping all coefficients with $|\bar{c}_{j,k,\psi}| \geq a\sigma_0 2^{-m}$, they take

$$\tilde{c}_{j,k,\psi} = \begin{cases} \bar{c}_{j,k,\psi} - a\sigma_0 2^{-m}, & \bar{c}_{j,k,\psi} > a\sigma_0 2^{-m}, \\ 0, & |\bar{c}_{j,k,\psi}| \leq a\sigma_0 2^{-m}, \\ \bar{c}_{j,k,\psi} + a\sigma_0 2^{-m}, & \bar{c}_{j,k,\psi} < -a\sigma_0 2^{-m}, \end{cases}$$

a technique they call “wavelet shrinkage.” In [8] by the present authors, it is shown for $F$ in $B^\alpha_\sigma(L_\sigma(I))$ that $a = \sqrt{C\log(2^{2m})}$ for some $C$ yields an asymptotically near-optimal algorithm of the form (10). With this choice, the error is proportional to

$$(\log 2^{2m})^{\alpha/(\alpha+1)}(2^{-2m})^{\alpha/(\alpha+1)}.$$

Thus, we incur an extra factor of $(\log 2^{2m})^{\alpha/(\alpha+1)}$ because we don’t know which coefficients $c_{j,k,\psi}$ are large.

## 7. Image reconstruction

Several authors (e.g., [2], [16]) use functionals

$$\min_{\|g\|_Y<\infty} \{\|f-f\|^2_{L_2(I)} + \lambda\|g\|_Y\}$$

for purposes of image reconstruction, where $Y$ is close to the norms considered here. For example, Bouman and Sauer [2] investigate minimizing for $p \geq 1$

$$\sum_{j\in\mathbb{Z}^2_m} |f_j - g_j|^2 + \mu \sum_{j\in\mathbb{Z}^2_m} [|g_j - g_{j+(1,0)}|^p + |g_j - g_{j+(0,1)}|^p],$$

where $f_j$ are the pixels of $f$ and $g_j$ are the pixels of $g$. In the continuous limit, this is similar to minimizing

$$\|f-g\|^2_{L_2(I)} + \lambda[\|D_1 g\|^p_{L_p(I)} + \|D_2 g\|^p_{L_p(I)}], \quad p \geq 1,$$

where $\|g\|^p_{L_p(I)} := \int_I |g(x)|^p\, dx$. They found that for images with edges, $p = 1$ was most effective in reconstructing the original image after bluring. In this case, their functional is precisely

$$\|f-g\|^2_{L_2(I)} + \lambda\|g\|_{\mathrm{BV}}, \tag{11}$$

where BV is the space of functions of bounded variation on $I$. We have $B^1_1(L_1(I)) \subset \mathrm{BV} \subset B^1_\infty(L_1(I))$, and $B^1_1(L_1(I))$ is one of our nonlinear spaces. Thus, minimizing (11) should yield a reconstructed $\tilde{f}$ for $f$ that is close to the usual $\tilde{f}$ we get from our nonlinear algorithm.

**Acknowledgments**

This work was supported in part by the National Science Foundation (grants DMS-8922154 and DMS-9006219), the Air Force Office of Scientific Research (contract 89-0455-DEF), the Office of Naval Research (contracts N00014-91-J-1152 and N00014-91-J-1076), the Defense Advanced Research Projects Agency (AFOSR contract 90-0323), and the Army High Performance Computing Research Center (Army Research Office contract DAAL03-89-C-0038).

## References

$K$-functionals are discussed in [1]. Aspects of wavelet theory and practice are surveyed in, e.g., [3], [9], [15], and [18]. The theory of nonlinear data compression using wavelets is developed in [7], [4], [5], [6], and [10]. Nonparametric estimation using functionals similar to (3) is surveyed in [19].

1. J. Bergh and J. Löfström, *Interpolation Spaces: An Introduction*, Springer-Verlag, New York, 1976.
2. C. Bouman and K. Sauer, *Bayesian estimation of transmission tomograms using segmentation based optimization*, IEEE Trans. Nuclear Science **39** (1992), no. 4 (to appear).
3. I. Daubechies, *Ten Lectures on Wavelets*, CBMS-NSF Regional Conference Series in Applied Mathematics, vol. 91, SIAM, Philadelphia, 1992.
4. R. DeVore, B. Jawerth, and B. Lucier, *Data compression using wavelets: Error, smoothness, and quantization*, DCC 91: Data Compression Conference (J. Storer and J. Reif, eds.), IEEE Computer Society Press, Los Alamitos, CA, 1991, pp. 186–195.
5. ———, *Image compression through wavelet transform coding*, IEEE Trans. Information Theory **38** (1992), no. 2, 719–746, Special issue on Wavelet Transforms and Multiresolution Analysis.
6. ———, *Surface compression*, Computer Aided Geom. Design (to appear).
7. R. DeVore, B. Jawerth, and V. Popov, *Compression of wavelet decompositions*, Amer. J. Math. (to appear).
8. R. DeVore and B. Lucier, *Nonlinear wavelet algorithms for removing Gaussian noise*, in preparation.
9. ———, *Wavelets*, Acta Numerica 1992 (A. Iserles, ed.), Cambridge University Press, New York, 1992, pp. 1–56.
10. R. A. DeVore, P. Petrushev, and X. M. Yu, *Nonlinear wavelet approximations in the space $C(\mathbb{R}^d)$*, Proceedings of the US/USSR Conference on Approximation, Tampa, Springer-Verlag, New York, 1991 (to appear).
11. R. DeVore and V. Popov, *Interpolation of Besov spaces*, Trans. Amer. Math. Soc. **305** (1988), 397–414.
12. R. DeVore and X. Yu, *Nonlinear n-widths in Besov spaces*, Approximation Theory VI: Vol. 1, C. K. Chui, L. L. Schumaker, and J. D. Ward, eds., Academic Press, New York, 1989, pp. 203–206.
13. D. Donoho and I. Johnstone, *Ideal spatial adaptation by wavelet shrinkage*, preprint.
14. D. Donoho and I. Johnstone, *Minimax estimation via wavelet shrinkage*, preprint.
15. Y. Meyer, *Ondelettes et Opérateurs I: Ondelettes*, Hermann, Paris, 1990.
16. D. Mumford and J. Shah, *Optimal approximations by piecewise smooth functions and associated variational problems*, Comm. Pure Appl. Math. **XLII** (1989), 577–685.
17. A. Pinkus, *N-widths and Approximation Theory*, Springer-Verlag, New York, 1985.
18. O. Rioul and M. Vetterli, *Wavelets and signal processing*, IEEE Signal Processing Magazine **8** (1991), no. 4, 14–38.
19. G. Wahba, *Spline Models for Observational Data*, CBMS-NSF Regional Conference Series in Applied Mathematics, vol. 59, SIAM, Philadelphia, 1990.
20. R. White, *High-performance compression of astronomical images* (abstract only), DCC 92: Data Compression Conference (J. Storer and M. Cohn, eds.), IEEE Computer Society Press, Los Alamitos, CA, 1992, p. 403.

Fig. 1. $512 \times 512$ lena, green component of color image.

Fig. 2. Compressed with the linear method, $N = 16384$.

Fig. 3. Compressed with the nonlinear method, $N = 16384$.

Fig. 4. With 32 grey scales RMS noise added.

Fig. 5. Noise removed with linear method, $N = 16384$.

Fig. 6. Noise removed with wavelet shrinkage, $a = 1$.