Wavelet-based minimal-energy approach to image restoration

Charles K. Chui\textsuperscript{a,\ast,1}, Jianzhong Wang\textsuperscript{b,2}

\textsuperscript{a} Department of Mathematics and Computer Science, University of Missouri–St. Louis, St. Louis, MO 63121, USA
\textsuperscript{b} Department of Mathematics and Statistics, Sam Houston State University, Huntsville, TX 77341, USA

Received 14 February 2006; revised 13 December 2006; accepted 12 January 2007

Communicated by Zuowei Shen

Abstract

The popular mathematical models for digital image restoration, formulated as a minimization problem of certain total energy functionals, give rise to the variational/PDE-based approach to process the input image in the spatial (or physical) domain. In general, the total energy in these models consists of two additive terms, namely: the internal energy for dictating the image quality in terms of image smoothness and image feature preservation, and the external energy for ensuring the output image not to deviate too far from the input image. In formulating the internal energy, a specific energy density function is chosen and applied to the magnitude of the gradient operation for extracting image features to be preserved. In this paper, we replace the gradient operation by some wavelet transform that has better performance in feature extraction and eliminates the need of iterative steps. Thus, the problem of image processing is performed in the wavelet domain instead. By taking advantage of the multi-scale structure of wavelets and the corresponding multi-level singularity detection capability, the proposed approach should facilitate further development of fast and effective algorithms for the variational approach, perhaps even with significant reduction in computational complexity as compared with the traditional approach to digital image restoration.

© 2007 Elsevier Inc. All rights reserved.

1. Introduction

This paper is concerned with digital image noise reduction/removal. Although the wavelet approach introduced in this paper can be modified to solving other image restoration problems, we only focus on the basic problem of removing additive noise in order to demonstrate the power and flexibility of the wavelet-based optimization consideration. Precisely, we will only study the problem of “recovering” an unknown deterministic (target) image, represented by the function \( v \) from one single observation of a noisy image sample:

\[
u_0 = v + \xi, \quad v \in L^2(\Omega).
\]

\textsuperscript{\ast} Corresponding author.
\textsuperscript{1} Research supported by ARO Grant #W911NF-04-1-0298. This author is also with the Department of Statistics, Stanford University, Stanford, CA 94305, USA.
\textsuperscript{2} Research supported by SHSU EGR-2006 Grant.

1063-5203/$ – see front matter © 2007 Elsevier Inc. All rights reserved.
doi:10.1016/j.acha.2007.01.006

where $\xi$ is a sample of some random function defined on $\Omega$ and is often assumed to be ergodic. The variational method to be considered assumes that the target image to be recovered has minimum total energy, defined as the sum of two energy functionals, namely: the internal energy functional for dictating the image quality in terms of image smoothness and image feature preservation, and the external energy functional for ensuring the output image not to deviate too far from the input sample.

Let $H(\Omega) \subset L^2(\Omega)$ denote the physical space of images (represented by functions) with domain $\Omega \subset \mathbb{R}^2$. The ideal target image $u^\ast$ could be considered as the solution of the minimization problem:

$$ u^\ast = \arg \min_{u \in H(\Omega)} \| u - v \|_2. $$

However, since $v$ is an unknown quantity, this problem does not make any sense, and the minimization problem must be posed in terms of a known quantity, such as the observed image $u_0$. The most popular approach is to change the above minimization problem to the problem of minimizing the total energy functional

$$ E(\rho, \lambda, u) = E_i + \lambda E_e = \int_{\Omega} \rho(|\nabla u|) \, dx + \frac{\lambda}{2} \int_{\Omega} (u - u_0)^2 \, dx, \quad u \in H(\Omega), $$

with internal energy $E_i = E_i(\rho, u)$ of $u$, governed by some energy density function $\rho$; external energy $E_e = E_e(u_0, u)$ of $u$, in terms of the observed image $u_0$; and a positive parameter $\lambda$ that plays the role of balancing between the internal and external energies. Here, $\rho$ is a non-negative function defined on $\mathbb{R}^+$ and is assumed to satisfy $\rho(0) = 0$. The restored image is then represented by the function

$$ u_\lambda^\ast = \arg \min_{u \in B_\sigma} E(\rho, \lambda, u). $$

The gradient operation in the formulation of the internal energy is introduced to extract image features, with increase in the amount of image features for larger magnitude of the gradient and with absolutely no image feature at all when the gradient value is zero. Hence, the assumption of $\rho(0) = 0$ is natural; and for the minimization problem to capture the essential ingredient of image feature preservation, the increase in the amount of image features must contribute to stronger internal energy. For this reason, we will assume, in addition, that the energy density function $\rho$ is a non-decreasing function on $\mathbb{R}^+$.

Observe that the minimization problem (2) does not take the knowledge, if any, of the noise process $\xi$ into consideration. If, for example, $\xi$ is a sample of some zero-mean Gaussian noise $\Sigma$ with known variance $\sigma^2$, then by using the standard notation $\mathbb{E}(X)$ for the expectation of a random variable $X$ and the notation

$$ \mathcal{R}(U, v) := \frac{1}{|\Omega|} \mathbb{E}\| v - U \|_2^2 = \sigma^2 $$

for the risk of the random image $U = v + \Sigma$, the restored image can be considered as solution of the constrained minimization problem

$$ u^{**} = \arg \min_{u \in B_\sigma} \left\{ E_i(\rho, u); \ B_\sigma = \left\{ u: \frac{1}{|\Omega|} \| u - u_0 \|_2^2 = \sigma^2 \right\} \right\}. $$

Let us first return to the discussion of the minimization problem (2). By introducing the function

$$ c(p) = \frac{\rho'(p)}{p}, $$

the Euler–Lagrange equation corresponding to this minimization problem is given by the steady-state diffusion-reaction equation:

$$ - \text{div}(c(|\nabla u_\lambda^\ast|) \nabla u_\lambda^\ast) + \lambda (u_\lambda^\ast - u_0) = 0, $$

with (heat) conductivity $c(p)$. Of course, the quality of the restored image $u_\lambda^\ast$ depends on the choice of the parameter $\lambda$. In general, smaller values of $\lambda$ result in a smoother image $u_\lambda^\ast$, and larger values of $\lambda$ result in a sharper $u_\lambda^\ast$.

However, without any knowledge of the noise random variable, it is not possible to determine a precise value of $\lambda$ for which the quantity $\| u_\lambda^\ast - v \|_2$ is minimized. On the other hand, if the noise variance is known, then by applying
the Lagrange multiplier method, the Euler–Lagrange equation for the minimization problem (3) becomes the same
steady-state diffusion–reaction equation:
\[- \nabla (c(|\nabla u^*|) \nabla u^*) + \lambda (u^* - u_0) = 0,\]
but with the side condition:
\[\frac{1}{|\Omega|} \|u^* - u_0\|_2^2 = \sigma^2,\]  
which can be used to determine the optimal value of the parameter $\lambda$.

The fast numerical algorithms have been developed to solve the Euler–Lagrange equations. They include the
nonlinear primal-dual method [7,9], the lagged diffusivity fixed point iteration method [40], the second-order cone
programming method [26], and the multilevel algorithm [10]. The method of steepest descent also applies to compute
(usually) satisfactory numerical solutions of the unconstrained minimization problem, though the task becomes more
difficult for the constrained one. More precisely, by introducing the time variable $t$ in the formulation of the equation,
the minimization problem (3) could be solved numerically by solving the (anisotropic) diffusion–reaction PDE
(partial differential equation):
\[\frac{\partial u}{\partial t} = \nabla (c(|\nabla u|) \nabla u) - \lambda (u - u_0),\]
\[u(0, x, y) = u_0(x, y), \ (x, y) \in \Omega, \quad \frac{\partial u}{\partial \vec{n}} = 0, \ (x, y) \in \partial \Omega,\]  
with the steady-state solution (i.e., sufficiently large values of $t$) to be considered as the solution $u^*_\lambda$ in (4).

To apply the method of steepest descent to the corresponding constrained minimization problem, we may abandon
the task of finding the steady-state solution by considering, instead, the Perona–Malik anisotropic diffusion equation [37]:
\[\frac{\partial u}{\partial t} = \nabla (c(|\nabla u|) \nabla u), \quad u(0, x, y) = u_0(x, y), \ (x, y) \in \Omega,\]
\[\frac{\partial u}{\partial \vec{n}} = 0, \quad (x, y) \in \partial \Omega,\]  
with stop-time $t_0$ to be determined by the constraint
\[\frac{1}{|\Omega|} \|u(t_0) - u_0\|_2^2 = \sigma^2.\]  
Observe that the condition (5) of the constrained minimization problem is replaced by the stop-time criterion (8).
Hence, by some appropriate discretization of the time variable $t$, an iterative scheme can be designed to solve (7)
numerically, with termination of the iteration process to be determined by a discrete version of (8).

To complete our discussion of the variational/PDE approach, we give several examples of the conductivity func-
tions $c(p)$, or equivalently, the energy density functions $\rho(p)$ in the corresponding energy minimization problem.
The following examples that can be found in [4,37–39,41], are among the commonly used ones in the literature.

1. Linear conductivity: $c(p) = c$, where $c$ is a positive constant, corresponds to the energy density function $\rho(p) = \frac{c}{2} p^2$ and gives rise to the isotropic diffusion
   \[\frac{\partial u}{\partial t} = \nabla c(|\nabla u|) \nabla u = c \Delta u.\]

2. Inverse proportional conductivity: $c(p) = \frac{1}{p}$, which corresponds to the energy density function $\rho(p) = p$, gives rise to the mean curvature motion model
   \[\frac{\partial u}{\partial t} = \nabla \left( \frac{\nabla u}{|\nabla u|} \right) \]
   (see [24,27]). This is also the origin of the popular total variation (TV) model (see, for example, [7,8,11]).
3. Gaussian conductivity: \( c(p) = e^{-\frac{p^2}{K^2}} \), which corresponds to the energy density function

\[
\rho(p) = \frac{K^2}{2} \left( 1 - e^{-\frac{p^2}{K^2}} \right)
\]

and first appears in the Perona–Malik paper [37], gives rise to the anisotropic diffusion PDE

\[
\frac{\partial u}{\partial t} = \text{div} \left( e^{-\frac{\|\nabla u\|^2}{K^2}} \nabla u \right).
\]

4. Exponential decay conductivity: \( c(p) = e^{-\frac{p}{K}} \), which corresponds to the energy density function

\[
\rho(p) = \int_0^p xe^{-\frac{x}{K}} \, dx
\]

gives rise to the diffusion PDE model [4]

\[
\frac{\partial u}{\partial t} = \text{div} \left( e^{-\frac{\|\nabla u\|}{K^2}} \nabla u \right).
\]

5. Lorentz conductivity: \( c(p) = \frac{1}{1 + \frac{p^2}{K^2}} \), which corresponds to the energy density function

\[
\rho(p) = \frac{K^2}{2} \ln \left( 1 + \left( \frac{p^2}{K^2} \right) \right)
\]

gives rise to the diffusion PDE model [38]

\[
\frac{\partial u}{\partial t} = \text{div} \left( \frac{K^2}{K^2 + \|\nabla u\|^2} \nabla u \right).
\]

The positive parameter \( K \) in the above anisotropic diffusion PDE models dictates the trade-off between image sharpness and image smoothness in terms of the degree of image edge preserving. For instance, smaller values of \( K \) tend to better preserve less visible image edges at the expense of over-smoothing the image. From another point of view, the relation between the amount of diffusion and the magnitude of the image gradient for each pixel is a function of the conductivity \( c(p) \), namely: stronger diffusion of the pixel value corresponds to smaller magnitude of the image gradient at this pixel. The graphs of various commonly used conductivity functions are shown in Fig. 1.

The significance of the gradient operation in the formulation of the internal energy functional is its capability to extract image features for better image quality preservation. Unfortunately, numerical solution of the nonlinear PDE requires numerous iterative steps. To improve the capability in image feature extraction and to eliminate the need of the iterative process, we propose to replace the gradient operation by some wavelet transform and consider solution of the optimization problem more directly. In this regard, since numerical solution for the variational approach

![Conductivity functions](image-url)
2. Minimum-energy image processing in the wavelet domain

To perform image processing in the wavelet domain, it is necessary to convert the image in the function space \( H(\Omega) \) to its representation in some wavelet space. The theory of wavelet analysis can be found in many books such as the older ones [17,19,33].

2.1. Two-dimensional wavelet transforms

For computational efficiency and convenience of our discussion, we only consider tensor-product orthonormal wavelet bases of \( L^2(\Omega) \) derived from some orthonormal multiresolution analysis (MRA) on a bounded interval \( I \subset \mathbb{R} \) (see, for example, [18] and [30]), where \( \Omega \) is a rectangular domain. Let

\[ V_0 \subset V_1 \subset \cdots \]

be an MRA of \( L^2(I) \) generated by some orthonormal scaling function \( \phi \) with appropriate boundary corrections and \( \psi \) be a corresponding orthonormal wavelet (again with boundary corrections). As usual, let \( W_j \) be the wavelet subspace defined by

\[ V_j \oplus W_j = V_{j+1}, \quad V_j \perp W_j, \quad j = 0, 1, \ldots. \]

Then if \( \dim V_0 = n_0 \), we have \( \dim V_j \simeq \dim W_j \simeq 2^j n_0 \), where the sign \( \simeq \) denotes “essential equality, or equality with some minor correction which is independent of \( j \), due to the boundary corrections”. Set \( n_j = 2^j n_0 \) and consider

\[ \phi_{jk} \simeq 2^{j/2} \phi(2^j \cdot -k), \quad \psi_{jk} \simeq 2^{j/2} \psi(2^j \cdot -k), \quad j = 0, 1, \ldots. \]

Then \( \{\phi_{jk}\}_{k=1}^{n_j} \) and \( \{\psi_{jk}\}_{k=1}^{n_j} \) are the orthonormal bases of \( V_j \) and \( W_j \) respectively. Without loss of generality, we may assume that \( \Omega = I \times I \). By setting

\[ W_j^h = W_j \otimes V_j, \quad W_j^v = V_j \otimes W_j, \quad W_j^d = W_j \otimes W_j; \]

\[ W_j = W_j^h \oplus W_j^v \oplus W_j^d, \quad V_j = V_j \oplus V_j, \]

we again have a nested sequence of subspaces

\[ V_0 \subset V_1 \subset \cdots \]

which constitutes an MRA of \( L^2(\Omega) \), and that \( V_{j+1} = V_j \oplus W_j \). In addition, by setting \( s_j = n_j^2 \), it is clear that

\[ \dim V_j \simeq \dim W_j^h \simeq \dim W_j^v \simeq \dim W_j^d \simeq s_j. \]

Let \( N_s = \{1, 2, \ldots, s\} \) and consider the mapping \( \tau \) from \( N_s \times N_s \) onto \( N_s^2 \) defined by \( \tau(n, m) = ns + m \). For convenience, set \( k = \tau(n, m) \) and denote the tensor-product basis functions by

\[ \varphi_{jk}(x, y) = \varphi_{jn}(x)\varphi_{jm}(y), \quad \psi_{jk}(x, y) = \psi_{jn}(x)\psi_{jm}(y), \]

\[ \psi_{jk}^v(x, y) = \varphi_{jn}(x)\psi_{jm}(y), \quad \psi_{jk}^d(x, y) = \psi_{jn}(x)\psi_{jm}(y). \]

It is clear that \( \{\varphi_{jk}\}_{k=1}^{s_j}, \{\psi_{jk}^v\}_{k=1}^{s_j}, \{\psi_{jk}^d\}_{k=1}^{s_j} \), and \( \{\psi_{jk}^h\}_{k=1}^{s_j} \) are orthonormal bases of \( V_j, W_j^h, W_j^v \), and \( W_j^d \), respectively.

In the following, we will also employ the commonly used notation \( \psi_{jk} \) for \( \psi_{jk}^h \), \( \psi_{jk}^v \), or \( \psi_{jk}^d \), with sub-indices arranged in an appropriate order, resulting in

\[ \{\psi_{jk}\}_{k=1}^{3s_j} = \{\psi_{jk}^h\}_{k=1}^{s_j} \cup \{\psi_{jk}^v\}_{k=1}^{s_j} \cup \{\psi_{jk}^d\}_{k=1}^{s_j}, \]

so that \( \{\psi_{jk}\}_{k=1}^{3s_j} \) is an orthonormal basis of \( W_j \).
The wavelet decomposition of any function \( f \in L^2(\Omega) \) is then described by

\[
  f_j = P_j f = \sum_{k=1}^{s_j} a_{jk} \phi_{jk}, \quad g_j = Q_j f = \sum_{k=1}^{3s_j} d_{jk} \psi_{jk},
\]

where \( P_j \) and \( Q_j := [Q^h_j, Q^v_j, Q^d_j] \) are the orthogonal projection operators from \( L^2(\Omega) \) to \( V_j \) and \( W_j := [W^h_j, W^v_j, W^d_j] \) respectively. Hence, we have

\[
  f_j = \sum_{k=1}^{s_j} a_{0jk} \phi_{0jk} + \sum_{l=0}^{j-1} \sum_{k=1}^{3s_j} d_{lk} \psi_{lk}.
\]

In what follows, we will use the notations: \( a_j = (a_{jk})_{k=1}^{s_j}, d_j = (d_{jk})_{k=1}^{3s_j} \) and \( b_j = (a_0, d_0, \ldots, d_{j-1}) \). Then we have the following well-known result.

**Theorem 1.** The linear transformation \( T_j : \mathbb{R}^{s_j} \to \mathbb{R}^{s_j} \), defined by \( T_j(a_j) = b_j \), is unitary.

The unitary (or orthogonal) transformation \( T_j \) in the above theorem is also called an orthonormal discrete wavelet transform (DWT). Later, we may omit the index \( j \) in some occasions when \( j \) is fixed and denotes the highest wavelet level. We will also employ the notations:

\[
  a_l = T_l^a(a), \quad d_l = T_l^d(a), \quad d^h_l = T_l^h(a), \quad d^v_l = T_l^v(a), \quad d^d_l = T_l^d(a),
\]

where \( d^h_l, d^v_l, \) and \( d^d_l \), denote the horizontal, vertical, and diagonal components of \( d_l \), respectively. Of course, all of these transformations are orthogonal projections. For convenience, we will also abuse notations by directly writing \( a_l \in V_l, d_l \in W_l \), and so forth.

### 2.2. Wavelet representations of digital images

A digital image may be considered as an array of sampling data called pixels (or pixel values) defined on a rectangular grid of uniformly spaced pixel locations. For the sake of discussion only, let us arrange the pixel locations as a one-dimensional sequence \( \{p_i\}_{i=0}^{s-1} \), where the length \( s \) of the sequence is the number of pixels, called the image resolution. Hence, the sequence of ordered pairs \( \{(p_i, x_i)\}_{i=0}^{s-1} \), where \( x_i \) denotes the image pixel at the pixel location \( p_i \), completely determines the digital image. Let the (observed) digital image be denoted by \( u_0 \). Then assuming that the image resolution \( s \) matches the dimension of some subspace \( V_j \) of an MRA (i.e., \( \dim V_j = s \)), we can embed \( u_0 \) into the space \( V_j \) in the form of

\[
  x_i := u_0(p_i) = \sum_{i=1}^{s} y_i \varphi_{ji}(p_i).
\]

Set \( x := (x_i)_{i=1}^{s} \) and \( y = (y_i)_{i=1}^{s} \). Since the mapping from \( x \) to \( y \) is one-to-one, we may consider \( y \) as \( x \), and will call it the observed image, to be written as

\[
  y = a + \xi,
\]

where the symbol \( a \) now represents the target image to be recovered, and \( \xi \) is the image noise. The following well known result (see, for example, [20]) will be useful for our discussion.

**Lemma 2.** If \( \xi \) is an independent identically-distributed (i.i.d.) random sequence \( N(0, \sigma^2) \) and \( P \) is an orthogonal projection, then \( P \xi \) is also an i.i.d. random sequence \( N(0, \sigma^2) \).

Now, let \( T \) be some wavelet transform from \( V_j \) to \( V_0 \oplus W \), where

\[
  W = \bigoplus_{j=0}^{J-1} W_j.
\]
and let \( u = Ty \). Then we have

\[
u = T_0^a y + T_{J-1}^g y = u^a + w, \quad u^a \in V_0, \quad w \in W,
\]

where \( u^a \) and \( w \) are called the blur (or smooth, low-frequency) and wavelet (or high-frequency) components of \( u \), respectively.

Let \( \eta = T \xi \). Then it follows from the above lemma that if \( \xi \) is an i.i.d. noise random variable \( \sim N(0, \sigma^2) \), and so is \( \eta \). In addition, when \( \eta \) is decomposed into

\[
\eta = T_0^a \xi + T_{J-1}^g \xi = \eta^a + \eta^g
\]

and \( T a \) into

\[
b = T_0^a a + T_{J-1}^g a = b^a + b^g,
\]

then we have

\[
u^d = T_0^a y = T_0^a u^a + T_0^g \xi = b^a + \eta^a, \quad u^a \in V_0.
\]

Hence, the image recovered from the smooth version \( u^a \) is represented by

\[
y^b = (T_0^a)^T u^a = (T_0^a)^T b^a + (T_0^a)^T \eta^a = a^b + \xi^b, \quad y^b \in V_J,
\]

where \( \xi^a \) is the noise component of \( y^a \). Summarizing the above discussion, we have determined the variance of the noise component \( \xi^a \), as follows.

**Theorem 3.** Let \( y \) and \( a \) denote the observed and target images in \( V_J \), respectively. Let \( T_0^a \) be the orthogonal projection from \( V_J \) to \( V_0 \). If \( \xi \) is an i.i.d. noise carried by \( y \), \( y^b = (T_0^a)^T (T_0^a y) \in V_J \), and \( a^b = (T_0^a)^T (T_0^a a) \in V_J \), then the risk is given by \( \mathcal{R}(y^b, a^b) = \sigma^2 / J \).

**Proof.** Indeed, by setting \( \eta^b = T_0^a \xi \) and \( \xi^b = (T_0^a)^T (T_0^a \xi) \), we see that \( \mathcal{E}(\xi^b) = 0 \) and \( T_0^a \) is an orthogonal projection from \( V_J \) to \( V_0 \). It follows from Lemma 2 that \( \eta^b \) is an i.i.d. noise \( \sim N(0, \sigma^2) \). Since, in addition, \( \xi^b \in V_J \), and \( \eta^a \in V_0 \), by the definition (3), we have

\[
\mathcal{R}(y^b, a^b) = \frac{1}{\dim V_J} \mathcal{E}((y^b - a^b)^T (y^b - a^b)) = \frac{1}{\dim V_J} \mathcal{E}((\xi^b)^T \xi^b)
\]

\[
= \frac{1}{\dim V_J} \mathcal{E}((\eta^a)^T T_0^a (T_0^a)^T \eta^a)
\]

\[
= \frac{1}{\dim V_J} \mathcal{E}((\eta^a)^T \eta^a) = \frac{\dim V_0}{\dim V_J} \sigma^2 = \frac{\sigma^2}{4J}.
\]

This completes the proof of the theorem. \( \Box \)

In view of Theorem 3 and that \( \mathcal{R}(y^b, a^b) \) is near zero for sufficiently large \( J \), it is usually advisable not to touch the blur component \( u^a \) in (10), when image de-noising is performed in the wavelet space.

### 2.3. Energy functionals in the wavelet domain

Many researches have been done on the study of image denoising in the wavelet domain. Donoho and Johnstone [20–22] first created the wavelet shrinkage method. Thereafter, integrating with maximum a posteriori (MAP) estimate and other estimates from Bayesian decision theory, various Bayesian wavelet shrinkage methods were proposed (see [1,16.25,29]). The variational wavelet shrinkage methods were also proposed based on the function space decomposition [5,6,28,33]. Recently, the research on combining wavelet with variational PDE approaches and on TV wavelet processing have provoked wide interest (see [12–15,23,31]).

The objective of this paper is to study an analogous minimization problem as discussed in the Introduction section, when the total energy functional is formulated in the wavelet domain instead of the physical domain. Since the motivation is to replace the gradient operation in the formulation of the internal energy by some suitable wavelet transform
to gain better feature extraction capability, an ultimate goal of this wavelet de-noising approach is to search for the most effective internal energy density functions.

Let \( y_0 \) be an observed image with discrete wavelet transform denoted by \( u_0 \), which is decomposed into \( u_0^b + w_0 \) as in (10) for some relatively large \( J \). As discussed above, we will not touch the blur component \( u_0^b \), but only consider the wavelet component \( w_0 \), given by

\[
w_0 = b^g + \eta^g, \tag{11}
\]

where \( b^g \) is the wavelet component of the target image to be recovered, and \( \eta^g \) an independent random noise in the space \( W \). Hence, the total energy functional under consideration can be formulated as

\[
E_\lambda(\rho, w) = \lambda E_i(\rho, w) + E_e; \quad \text{with } E_e = \frac{1}{2} (w - w_0)^T (w - w_0), \tag{12}
\]

where \( \rho \) is the (internal) energy density function under investigation, and the positive parameter \( \lambda \) is used for an appropriate balance between the internal energy \( E_i(\rho, w) \) and external energy \( E_e \). However, different from (1), we prefer to put the balancing parameter \( \lambda \) with the internal rather than the external energy functionals. The reason for this preference will be clear from our discussion in the next subsection.

The wavelet component of the (wavelet transform of the) target image is the solution of the unconstrained minimization problem:

\[
w_\lambda^* = \arg \min_{w \in W} E_\lambda(\rho, w), \quad w \in W. \tag{13}
\]

If \( \eta^g \) in (9) is i.i.d. \( \sim N(0, \sigma^2) \), then \( R(w_0, b^g) = \sigma^2 \). Motivated by the above argument, we are therefore led to consider the constrained minimization problem

\[
w^{**} = \arg \min_{w \in B} E_e(\rho, w), \tag{14}
\]

with

\[B = \left\{ w : \frac{1}{n} (w - w_0)^T (w - w_0) = \sigma^2 \right\}.
\]

Set

\[
y^*_g = (T_0^g)^{-1} u_0^a + (T_{J-1}^g) w^*_\lambda \quad \text{and} \quad y^{**} = (T_0^g)^{-1} u_0^a + (T_{J-1}^g) w^{**},
\]

which we will call the minimal-energy wavelet restoration of the (image) sample \( y_0 \), for the unconstrained and constrained considerations, respectively.

In the following, we will only consider internal energy functionals that can be represented as a positive definite matrix \( G_w \), which is allowed to depend on \( w \), as follows:

\[
E_i(\rho, w) = w^T G_w w, \quad w = (w_1, \ldots, w_n)^T, \tag{15}
\]

where \( n = \dim W \) is given by \( s_f - s_0 \). Under the assumption that \( \xi \) is an independently distributed random sequence, we may assume that \( G_w \) is a diagonal matrix, namely:

\[G_w = \text{diag}(d(|w_1|), d(|w_2|), \ldots, d(|w_n|))
\]

for some positive function \( d \) defined on \( \mathbb{R}^+ \), called the energy weight function. Then the (internal) energy density function under consideration is given by \( \rho(x) = x^2 d(x), x \in \mathbb{R}^+ \), so that

\[
E_i(\rho, w) = \sum_{i=1}^n w_i^2 d(|w_i|) = \sum_{i=1}^n \rho(|w_i|),
\]

where the integral in the definition of the internal energy in (1) is now replaced by summation in the discrete setting. Since small values of the wavelet coefficients often contribute to noise, while large values tend to represent image feature (see [20]), it is advisable to assign larger values of weights to smaller wavelet coefficients and smaller values of weights to larger coefficients. Therefore, in what follows, the positive function \( d \), defined on \( \mathbb{R}^+ \), is assumed to decrease to 0 at \( \infty \); while being an energy density function, \( \rho(x) \) is assumed to be non-decreasing on \([0, \infty)\), with \( \rho(0) = 0 \). In the following discussion, we will assume that \( \rho \) is piecewise differentiable on \((0, \infty)\).

Then by setting $g(w) = (\rho(|w_i|), \ldots, \rho(|w_n|))$, the Euler–Lagrange equation for the minimization problem (13) can be formulated as

$$w + \lambda g'(w) = w^0.$$  \hspace{1cm} (16)

Hence, for each $i$, the component $(w^*_i)_i$ of the solution $w^*_i$ either satisfies

$$(w^*_i)_i + \lambda \rho'(\sum (w^*_i)_i) \text{sgn}(w^*_i)_i = w_i^0, \quad i = 1, \ldots, n,$$

or $(w^*_i)_i = 0$, where $\rho'(\sum w_i)$ does not exist.

Consequently, applying the Lagrange multiplier method to the constrained minimization problem (14) and re-scaling the multiplier, we may conclude that for each $i$, either $w_i^*$ satisfies Eq. (17) or $w_i^* = 0$, and all of them have to satisfy

$$\frac{1}{n} \sum_{i=0}^{n} (w_i^* - w_i^0)^2 = \sigma^2,$$

which can be used to formulate an equation that governs the parameter $\lambda$.

### 2.4. Examples of internal energy density functions

In this section, we consider several examples of energy density functions for the discrete wavelet transform consideration. Let us first study the unconstrained minimization problem (13).

**Example 1 (Constant weight).** If $d(s) = \frac{1}{2} c$ is a constant, then $\rho(s) = \frac{1}{2} cs^2$, and Eq. (16) becomes

$$w + \lambda c w = w_0,$$

or

$$w = \frac{1}{1 + \lambda c} w_0.$$

Therefore, this choice of the energy density yields the wavelet shrinkage model with uniform shrinkage rate.

**Example 2 (Soft wavelet thresholding).** Based on the assumption that noise on each component of $w_0$ is a random variable $\sim N(0, \sigma^2)$, we adopt the energy density function

$$\rho(s) = |s|.$$  \hspace{1cm} (19)

For this second example, we have the following result.

**Theorem 4.** Let the energy density function $\rho(s)$ be given by the function defined in (19). Then the minimum energy of the unconstrained minimization (13) is given by

$$E_\lambda = \frac{1}{2} \left( \lambda \sum_{|w_i^0| > \lambda} (2|w_i^0| - \lambda) + \sum_{|w_i^0| \leq \lambda} (w_i^0)^2 \right)$$

and this minimum energy is attained at

$$(w^*_i)_i = \begin{cases} 0, & |w_i^0| \leq \lambda, \\ w_i^0 - \lambda \text{sgn}(w_i^0), & |w_i^0| > \lambda. \end{cases}$$

**Proof.** Let

$$\epsilon(s) = \lambda \rho(s) + \frac{1}{2} (s - w_i^0)^2.$$

First observe that since $\lim_{s \to \pm \infty} \epsilon_i(s) = \infty$, the problem (13) has a global minimum. In addition, since

$$\epsilon'(s) = \lambda \rho'(s) + (s - w_i^0) = \lambda \text{sgn}(s) + (s - w_i^0),$$

so that $e'(0)$ does not exist, we see that 0 is the only critical point of $e(s)$ if $|w_i^0| \leq \lambda$. In this case, it is clear that $e(s)$ has its global minimum at 0. On the other hand, if $|w_i^0| > \lambda$, then $e(s)$ has two critical points, namely: 0 and $s^* := w_i^0 - \lambda \text{sgn}(w_i^0)$; and in this case, we have

$$
\lambda \rho(s^*) + \frac{1}{2} (s^* - w_i^0)^2 = \lambda (|w_i^0| - \lambda) + \frac{\lambda^2}{2} - \frac{\lambda^2}{2} < \frac{1}{2}(w_i^0)^2,
$$

$$
\lambda \rho(0) + \frac{1}{2} (0 - w_i^0)^2 = \frac{1}{2}(w_i^0)^2;
$$

so that $e(s)$ attains its minimum at $s^*$. Hence, the minimum of the problem (13) is attained at

$$
(w_{\lambda}^*)_i = \begin{cases} 0, & |w_i^0| \leq \lambda, \\ w_i^0 - \lambda \text{sgn}(w_i^0), & |w_i^0| > \lambda, \end{cases}
$$

with minimum value given by

$$
E_\lambda = \lambda \sum_{|w_i^0| > \lambda} (|w_i^0| - \lambda) + \frac{1}{2} \sum_{|w_i^0| \leq \lambda} (w_i^0)^2 + \frac{1}{2} \sum_{|w_i^0| > \lambda} \lambda^2 = \frac{1}{2} (\lambda \sum_{|w_i^0| > \lambda} (2|w_i^0| - \lambda) + \sum_{|w_i^0| \leq \lambda} (w_i^0)^2).
$$

This completes the proof of the theorem. \(\Box\)

We remark that the formulation in (20) agrees with the notion of soft thresholding in the study of wavelet shrinkage (see [20]).

**Example 3 (Hard thresholding).** By considering the energy density function

$$
\rho(s) = \begin{cases} |s| - \frac{1}{2} \lambda^2 s^2, & |s| \leq \lambda, \\ \frac{1}{2} \lambda, & |s| > \lambda, \end{cases}
$$

we will see that the corresponding minimization problem agrees with the notion of hard thresholding in the study of wavelet shrinkage (see [20]).

**Theorem 5.** Let the energy density function $\rho(s)$ be given by (21). Then the minimum of the unconstrained minimization problem (13) is attained at

$$
(w_{\lambda}^*)_i = \begin{cases} 0, & |w_i^0| \leq \lambda, \\ w_i^0, & |w_i^0| > \lambda, \end{cases}
$$

with minimum value given by

$$
E_\lambda = \frac{1}{2} \sum_{|w_i^0| > \lambda} \lambda^2 + \frac{1}{2} \sum_{|w_i^0| \leq \lambda} (w_i^0)^2,
$$

which is the minimum energy of the minimization problem.

**Proof.** Let $e(s) = \lambda \rho(s) + \frac{1}{2} (s - w_i^0)^2$. Then we have $\lim_{s \to \pm \infty} e_1(s) = \infty$, so that the problem (13) has a global minimum. In addition, we have

$$
e'(s) = \begin{cases} \lambda \text{sgn}(s) - w_i^0, & 0 < |s| < \lambda, \\ \lambda(s - w_i^0), & |s| > \lambda. \end{cases}
$$

Hence, the same argument as given in the proof of the previous theorem applies to complete the proof of this theorem. \(\Box\)

We will now turn to discuss the solution of the constrained minimization problem (14) for Examples 2 and 3, but first remark that as a result of the discussion in the previous subsection, the solutions of these two constrained minimization problems are given by the formulas (20) and (22) respectively, with the exception that the value of $\lambda$ is determined by the constraint (18). We have the following result.
Theorem 6. Let \( \eta_i \) be an i.i.d. \( \sim N(0, \sigma^2) \).
If \( \rho(s) = |s| \), then the solution of (14) is given by (20), with the value of \( \lambda \) determined by

\[
\sum_{|w_i|^\leq \lambda} (w_i^0)^2 + \sum_{|w_i|>\lambda} \lambda^2 = n\sigma^2.
\] (23)

On the other hand, if

\[
\rho(s) = \begin{cases} 
|s| - \frac{1}{2} \lambda^2, & |s| \leq \lambda, \\
\frac{1}{2} \lambda, & |s| > \lambda,
\end{cases}
\]

then the solution of (14) is given by (22), with the value of \( \lambda \) determined by

\[
\sum_{|w_i|^\leq \lambda} (w_i^0)^2 = n\sigma^2.
\] (24)

Proof. If \( \rho(s) = |s| \), then the solution of (14), denoted by \( w^{**} = (w_i)^n \), satisfies (20), with the value of \( \lambda \) determined by

\[
\sigma^2 = \sum_{i=1}^n (w_i^{**} - w_i^0)^2 = \sum_{|w_i|^\leq \lambda} (w_i^0)^2 + \sum_{|w_i|>\lambda} \lambda^2,
\]

which yields (23). The proof for (24) is similar. \( \square \)

It is well known that formulation of the threshold selection rule is a central problem in wavelet shrinkage. As mentioned in the Introduction, the threshold selection rules in Statistics can be formulated in various ways. A popular rule being used in wavelet shrinkage is the Square Root of Logarithm Rule, based on ideal spatial adaption [20], which is denoted by ‘sqtwolog’ (see [34]). In the following, we will denote the rule derived from the (simple) wavelet energy functionals in this subsection by ‘smplengy’.

A comparison of these two rules in removing pure noise is shown in Table 1, where the Gaussian white noise with maximal absolute-value 2.7316 is used. Note that the ‘sqtwolog’ rule is independent of soft or hard thresholding, while the ‘smplengy’ selection rule does (see Theorem 6).

Since the signal is a pure white noise, the proper threshold value close to 2.7316 is considered to be accurate. Observe that the results in Table 1 show that the ‘smplengy’ selection rule (for both soft and hard thresholding) achieves the best threshold value, while the ‘sqtwolog’ selection rule gives a threshold value that over-estimates the proper threshold value.

For image processing, the image quality is usually measured in terms of peak-signal to noise ratios (PSNR), which is defined by \( PSNR = 20\log_{10}(\|I\|/\|n\|) \), where \( \|I\| \) is the intensity of the image, and \( \|n\| \) is the energy of the noise. Here, \( n \) is so normalized that its standard deviation is in the [0, 1] range.

Table 2 is used to compare the ‘smplengy’ and ‘sqtwolog’ threshold selection rules again, when they are both applied to de-noise four noisy ‘Lena’ images, decomposed into 2 wavelet levels by applying the Daubechies wavelet

<table>
<thead>
<tr>
<th>Table 1</th>
<th>Threshold selection comparison for a pure white noise signal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Selection rule</td>
<td>sqtwolog</td>
</tr>
<tr>
<td>Threshold value</td>
<td>3.7169</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 2</th>
<th>Threshold selection comparison for 4 noisy ‘Lena’ images</th>
</tr>
</thead>
<tbody>
<tr>
<td>Noise level</td>
<td>1</td>
</tr>
<tr>
<td>Noisy image</td>
<td>26.11</td>
</tr>
<tr>
<td>sqtwolog/soft</td>
<td>26.25</td>
</tr>
<tr>
<td>sqtwolog/hard</td>
<td>27.65</td>
</tr>
<tr>
<td>smplengy/soft</td>
<td>29.56</td>
</tr>
</tbody>
</table>
‘DB3’ [34] for applying wavelet shrinkage. In Table 2, the second row shows the PSNR values of four noisy ‘Lena’
at different noise levels; the third row gives the PSNR values of the images de-noised by using ‘sqtwolog’ soft thresholding; the fourth row gives the PSNR values of the image de-noised by using ‘sqtwolog’ hard thresholding, and the fifth row gives the PSNR values of the image de-noised by using ‘smplengy’ soft thresholding. The results show that the ‘smplengy’ selection rule has significantly better performance than ‘sqtwolog’.

2.5. Blended wavelet energy functionals of image data

In the above discussion, the wavelet component \( \mathbf{w} \) is considered as a general data sequence. However, since the sequence \( \mathbf{w} \) consists of multi-level wavelet components of the 2-dimensional image data, that represent somewhat unique features, particularly corners and edges, as contributed by the wavelet transform in the diagonal and horizontal/vertical directions, respectively, the noise removal capability can be improved by taking level-dependence as well as spatial-dependence into consideration in the design of the energy density functions of the blended wavelet energy functionals. An illustration of the noise behavior in two wavelet levels with spatial wavelet features is given in Appendix A.

To formulate the blended wavelet energy functionals, let the \( j \)-level wavelet component of \( \mathbf{w} \) be denoted by \( \mathbf{w}_j \), which is further separated into the horizontal, vertical, and diagonal components, to be denoted by

\[
\mathbf{w}_j^h = (w_{j,1}^h, \ldots, w_{j,J}^h), \quad \mathbf{w}_j^v = (w_{j,1}^v, \ldots, w_{j,J}^v), \quad \mathbf{w}_j^d = (w_{j,1}^d, \ldots, w_{j,J}^d),
\]

associated with the wavelet spaces \( W_{j-1}^h \), \( W_{j-1}^v \), and \( W_{j-1}^d \), respectively. Hence, each sequence \( \mathbf{w} \in \mathbf{W} \) can be written as

\[
\mathbf{w} = (\mathbf{w}_j^h, \mathbf{w}_j^v, \mathbf{w}_j^d), \quad j = 1, \ldots, J.
\]

Then the energy density functions of the blended wavelet energy functionals could be designed under the following assumption: the diagonal component \( \mathbf{w}_j^d \) carries more noise than the horizontal and vertical components \( \mathbf{w}_j^h \) and \( \mathbf{w}_j^v \). Under this assumption, we could introduce, only as an example, the following blended wavelet energy functional \( E_\lambda(\mathbf{w}, \mu) \) by isolating the diagonal components from the horizontal and vertical components, namely:

\[
E_\lambda(\mathbf{w}, \mu) = \sum_{j=1}^{J} \sum_{i=1}^{s_j} \lambda_i \left( \rho(m_{ji}(\mu)) + \rho(w_{ji}^d) \right) + \frac{1}{2} (\mathbf{w} - \mathbf{w}_0)^T (\mathbf{w} - \mathbf{w}_0),
\]

\[
m_{ji}(\mu) = (|w_{ji}^h|^\mu + |w_{ji}^v|^\mu)^{1/\mu}, \quad 1 \leq \mu < \infty.
\]

Here, as an illustration, we simply use the same energy density function \( \rho \) for all wavelet levels and spatial components, without tacking on certain weights for convenience. Note that when \( \mu = 1 \), the energy function (25) is the same as the energy function (12). In the following, we only consider the case of \( \mu = 2 \). It is clear that the discussion for other values of \( \mu \) is similar.

**Theorem 7.** For \( \rho(s) = |s| \), the minimum value of the energy functional \( E_\lambda(\mathbf{w}, 2) \) is attained at

\[
(w_{ji}^h)^{\ast} = (w_{ji}^v)^{\ast} = (w_{ji}^d)^{\ast} = 0, \quad j = 0, \ldots, J - 1, \quad i = 1, \ldots, s_j,
\]

\[
(w_{ji}^d)^{\ast} = \text{sgn}(w_{ji}^d)(|w_{ji}^d| - \lambda)_+^+, \quad j = 0, \ldots, J - 1, \quad i = 1, \ldots, s_j.
\]

Furthermore, for

\[
\rho(s) = \begin{cases}
|s| - \frac{s^2}{2}, & |s| \leq \lambda, \\
\frac{1}{2} \lambda, & |s| > \lambda,
\end{cases}
\]

the minimal value of the energy functional \( E_\lambda(\mathbf{w}, 2) \) is attained at

\[
(w_{ji}^h)^{\ast} = \begin{cases}
0, & m_{ji}^h \leq \lambda, \\
(w_{ji}^h)^{\ast}, & m_{ji}^h > \lambda,
\end{cases} \quad j = 0, \ldots, J - 1,
\]

\( (w^d_{ji})^j = \begin{cases} 0, & |w^0_{ji}| \leq \lambda, \\
 w^0_{ji}, & |w^0_{ji}| > \lambda, \end{cases} \quad j = 0, \ldots, J - 1. \) \hfill (29)

**Proof.** Since the proof of (28) and (29) is similar to that of (26) and (27), we only elaborate on the proof of (26) and (27). Let

\[
e(t^h, t^v, t^d) = \lambda \sqrt{(t^h)^2 + (t^v)^2 + |t^d|^2} + \frac{1}{2} ((t^h - w^{0,h}_{ji})^2 + (t^v - w^{0,v}_{ji})^2 + (t^d - w^{0,d}_{ji})^2)
\]

and consider

\[
\frac{\partial e}{\partial t^h} = \frac{\lambda t^h}{\sqrt{(t^h)^2 + (t^v)^2}} + t^h - w^{0,h}_{ji},
\]
\[
\frac{\partial e}{\partial t^v} = \frac{\lambda t^v}{\sqrt{(t^h)^2 + (t^v)^2}} + t^v - w^{0,h}_{ji},
\]
\[
\frac{\partial e}{\partial t^d} = \lambda \text{sgn}(t^d) + t^d - w^{0,h}_{ji}.
\]

Then the (non-linear) system of Eqs. (30)–(32) has a solution only when \( m^0_{ji} > \lambda \) and \( |w^{0,d}_{ji}| > \lambda \), and in this case the solution is given by

\[
s^h = (w^{0,h}_{ji}) \left(1 - \frac{\lambda}{m^0_{ji}} \right),
\]
\[
s^v = (w^{0,v}_{ji}) \left(1 - \frac{\lambda}{m^0_{ji}} \right),
\]
\[
s^d = (w^{0,d}_{ji} - \lambda \text{sgn}(w^{0,d}_{ji})).
\]

For \( m^0_{ji} \leq \lambda \), we note that \((0, 0)\) is a critical point for \((s^h, s^v)\); while for \( |w^{0,d}_{ji}| < \lambda, 0 \) is a critical point for \( s^d \). Hence, by computing the values at the critical points and comparing values of the energy functionals, we arrive at (26) and (27). This completes the proof of the theorem. \( \square \)

**Remark 1.** We can also derive the threshold selection rule for image de-noising, by considering pure white noise with i.i.d. \( \sim N(0, \sigma^2) \) in the same fashion as the ‘smpkengy’ rule. The derivation is similar to that of Theorem 6. We denote the selection rule derived this way by ‘blndengy’, which is dependent on soft or hard thresholding. Similar arguments apply to deriving other selections for other energy density functions \( \rho \), even with different multiplicative weights for the diagonal components and/or different levels.

Comparison of our method of the blended wavelet energy functionals as discussed above with the Square Root of Logarithm Rule is included in Appendix B. The following conclusion is based on various experimental results.

**Conclusion 1.** Experimental results show that thresholds based on the blended energy minimization out-perform the Square Root of Logarithm Rule.

**Uncited references**

[2] [3] [32] [35] [36]

**Acknowledgment**

The authors would like to thank the editor and, especially, the reviewers of this paper for their comments.
Appendix A

In this appendix, we illustrate, with an example of the ‘Lena’ image, the noise behavior in consecutive levels, with features shown along the three different wavelet spatial directions. The image without noise, as shown in Fig. 2, is decomposed into two wavelet levels, as shown in Fig. 5, using the Daubechies wavelet ‘BD3’, where the finer level is arranged in the first row and the coarser level in the second row, both in the order from left to right, of the blur (or low-frequency) component in $V$, the horizontal wavelet component in $W_h$, the vertical wavelet component in $W_v$, and the diagonal wavelet component in $W_d$, respectively.

Gaussian noise is added to the image in Fig. 2 to give the noisy image in Fig. 3, which is then decomposed and arranged, in the same manner, as shown in Fig. 6. Observe that most of image features are inundated in the noise in the finer wavelet level (i.e., the second, third, and fourth pictures on the first row), particularly in the diagonal component (i.e., the fourth picture).

Appendix B

In this appendix, we study the performance in image noise reduction by wavelet energy minimization by comparing thresholding based on minimum blended-energy functional (denoted by ‘blndengy’) with the Square Root of Logarithm Rule (‘sqtwolog’) [20]. The de-noised quality is measured in terms of PSNR values. We compare the noise reduction results for the four images: Lena, Cameraman, Peppers, and Testpat as shown in Fig. 4.

The wavelet ‘DB3’ is applied for wavelet decomposition into 2 wavelet levels. Four different noise levels are used for this comparison. The results are shown in Tables 4–7. In these four tables, the first row gives the PSNR values of the four noisy images, the second and the third rows give the PSNR values of de-noised results by using ‘sqtwolog’
applied to soft and hard thresholding, respectively, and the fifth row gives the PSNR values of the de-noised results by using ‘blindengy’ applied to soft thresholding. Here, NSD stands for Noise Standard Deviation. These results show that ‘blindengy’ applied to soft thresholding out-performs ‘sqtwolog’ applied to both soft and hard thresholding.
Table 5

PSNR Comparison for noisy images with NSD 0.100

<table>
<thead>
<tr>
<th>Image name</th>
<th>Lena</th>
<th>Cameraman</th>
<th>Peppers</th>
<th>Testpat</th>
</tr>
</thead>
<tbody>
<tr>
<td>Noisy image</td>
<td>20.21</td>
<td>20.40</td>
<td>20.14</td>
<td>20.37</td>
</tr>
<tr>
<td>sqtwolog/soft</td>
<td>24.37</td>
<td>23.14</td>
<td>27.12</td>
<td>22.19</td>
</tr>
<tr>
<td>sqtwolog/hard</td>
<td>24.88</td>
<td>24.07</td>
<td>27.38</td>
<td>23.11</td>
</tr>
<tr>
<td>blindengy/soft</td>
<td>25.91</td>
<td>25.14</td>
<td>27.84</td>
<td>24.62</td>
</tr>
</tbody>
</table>

Table 6

PSNR Comparison for noisy images with NSD 0.125

<table>
<thead>
<tr>
<th>Image name</th>
<th>Lena</th>
<th>Cameraman</th>
<th>Peppers</th>
<th>Testpat</th>
</tr>
</thead>
<tbody>
<tr>
<td>Noisy image</td>
<td>18.40</td>
<td>18.51</td>
<td>18.31</td>
<td>18.53</td>
</tr>
<tr>
<td>sqtwolog/soft</td>
<td>23.87</td>
<td>22.44</td>
<td>26.33</td>
<td>21.68</td>
</tr>
<tr>
<td>sqtwolog/hard</td>
<td>24.13</td>
<td>23.00</td>
<td>26.47</td>
<td>22.17</td>
</tr>
<tr>
<td>blindengy/soft</td>
<td>24.81</td>
<td>23.79</td>
<td>26.65</td>
<td>23.37</td>
</tr>
</tbody>
</table>

References


