UCLA
COMPUTATIONAL AND APPLIED MATHEMATICS

Optimal Constructions of Wavelet Coefficients Using Total Variation Regularization in Image Compression

Tony F. Chan
H.M. Zhou

July 2000

CAM Report 00-27

Department of Mathematics
University of California, Los Angeles
Los Angeles, CA. 90095-1555

http://www.math.ucla.edu/applied/cam/index.html

OPTIMAL CONSTRUCTIONS OF WAVELET COEFFICIENTS USING TOTAL VARIATION REGULARIZATION IN IMAGE COMPRESSION *

TONY F. CHAN † AND H.M. ZHOU †

Abstract. In this paper, we propose using Partial Differential Equation (PDE) techniques in wavelet based image processing to reduce edge artifacts generated by wavelet thresholding. We employ a variational framework, in particular the minimization of total variation (TV), to select and modify the retained standard wavelet coefficients so that the reconstructed images have fewer oscillations near edges. Numerical experiments show that this approach improves the reconstructed image quality in wavelet compression and in denoising.

1. Introduction. In this paper, we are concerned with the suppression of edge artifacts caused by wavelet thresholding in digital image denoising and compression. It is well know that wavelet thresholding, including linear (i.e. truncating the high frequencies) and nonlinear thresholding (i.e. retaining large coefficients,) may generate oscillations near discontinuities, especially when the functions or the images contain high level noise. This Gibbs' phenomenon is the primary reason for edge artifacts in digital image processing.

Many methods have been proposed to overcome this problem. Donoho's soft thresholding truncates wavelet coefficients on different scale levels subject to different thresholds [14]. Another type of approach also due to Donoho is to a construct special basis for discontinuities; such as wedgelets [15], ridgelets [16], and curvelets [5]. A different approach is to modify the wavelet transforms so that fewer large high frequency coefficients are generated near discontinuities, resulting in fewer large coefficients truncated in the thresholding process. Along this direction, Claypoole, Davis, Sweldens and Baraniuk [13] proposed an adaptive lifting scheme which lowers the order of approximation near jumps, thus minimizing the Gibbs' effects. We have proposed ENO-wavelet transforms which apply the one-side approximation idea of constructing Essentially Non-Oscillatory (ENO) schemes in numerical shock capturing to design adaptive wavelet transforms such that no large high frequency coefficients are generated through differencing across discontinuities [11], essentially eliminating oscillations in the reconstructed images.

In this paper, we propose an alternative method, which uses other Partial Differential Equation (PDE) techniques, especially PDE's from variational principles, to reduce the oscillations in wavelet thresholding approximations. In fact, variational PDE models have been commonly used in image processing since the end of the 1980's, for example, Mumford-Shah's functional [21] in segmentation, Rudin-Osher-Fatemi's Total Variation in restoration

[1] Research supported in part by grants ONR-N00017-96-1-0277 and NSF DMS-96-26755.

[2] Department of Mathematics, the University of California, Los Angeles, CA90095-1555. email:{chan,hmzhou}@math.ucla.edu

[25], Alveraz-Morel's formalization in image analysis [2], Perona-Malik's diffusion model [24], Sapiro-Tannenbaum's affine scale space [26], Alveraz-Guichard-Lions-Morel's fundamental equations for image processing [1], Morel-Solimini's book on variational method for segmentation [20], and some more recent works such as Chan-Vese's active contour [9], Chambolle-Lions's total variation analysis [7], Bertalmio-Sapiro-Caselles-Ballester's image inpainting [3] and the papers collected in the IEEE special issue on PDE and geometry-driven diffusion image processing [18]. The crucial observation which makes these methods successful is in viewing images as piecewise smooth functions connected by large jumps (edges) and realizing the similarity between images and piecewise smooth solutions of certain kinds of PDE's, then employing well-developed PDE's techniques to handle the edges. Based on this observation, one can also apply PDE techniques to wavelet image processing to reduce the edge oscillations. Our goal is to use a variational framework, in particular, the techniques for minimizing total variation (TV), to select and modify the retained standard wavelet coefficients so that the reconstructed images have fewer oscillations near the edges.

It has been shown through many simulations in the literature that the TV model can effectively suppress noise while retaining sharp edges in images ([25], [10], and [4]). Chambolle, DeVore, Lee and Lucier [6] attempted to use wavelet based variational forms to accomplish compression and denoising. Using wavelet coefficients, they compute the best fitting of the observed images subject to minimizing certain norms in Besov spaces, which are close to the Bounded Variation (BV) space corresponding to the TV norm. An essential difference between the Besov spaces and the BV space is that Besov spaces do not admit the discontinuous functions. Therefore, sharp edges are unaviodably smoothed out in the reconstructed images. In [12], we demonstrated that compressing TV denoised images produces higher ratio compression and better quality than denoising and compressing the images by directly using wavelets. On the other hand, edge oscillations caused by standard wavelet thresholding significantly increase the TV norm of the reconstructed images. All this motivates us to select and modify the nonzero wavelet coefficients in the thresholding procedure subject to minimizing the TV norm of the reconstructed images so that they can produce fewer edge artifacts while retaining sharp edges.

In general, minimizers of such variational problems can be found by solving their associated Euler-Lagrangian equations, which are PDE's. In particular, the PDE produced by the TV minimization problem is highly nonlinear and usually degenerate at flat regions. Many works have been advocated to speed up the solvers in physical space, for instance, see [28], [8] and [23]. In the present work, we deduce the corresponding PDE's in wavelet space and solve them in analogous ways. We will discuss some aspects of the numerics in this paper as well.

The above described method can be easily embedded into a image compression framework by simply replacing the standard wavelet thresholding step by TV regularized wavelet

thresholding . The produced non-zero wavelet coefficients can then be forwarded for quantizing and coding in the standard ways. In this situation, at the reconstruction end, the standard wavelet procedure will automatically restore the images with fewer edge artifacts. We will concentrate on selecting and modifying the non-zero wavelet coefficients subject to minimizing the TV norm of the reconstructed images, and we will not consider the quantization and coding steps. In addition, the ideas introduced here can also be used as a post-processing technique for the reconstructed images so that it can suppress the edge oscillations generated in the compression process.

This paper is arranged in the following way. In section 2, we give the TV regularized wavelet compression model for wavelet thresholding. In section 3, we study some relaxations of the introduced model and their associated PDE's. In section 4, we discuss some numerical aspects of solving these PDE's. And in section 5, we show some examples to illustrate the results of the models.

2. TV Regularized Wavelet Compression Model for Thresholding. In this section, we give our TV regularized Wavelet Compression model for suppressing the oscillations generated by wavelet thresholding.

Suppose we are given an observed image $z(x) = u_0(x) + n(x)$, where $u_0(x)$ is the original noise free image and $n(x)$ the Gaussian white noise with $\|n(x)\|_2 = \sigma$. Let us denote the standard orthonormal wavelet transform of $z(x)$ by:

$$z(\vec{\alpha}, x) = \sum_{j,k} \alpha_{j,k} \phi_{j,k}(x), \tag{1}$$

where $\phi_{j,k}(x)$ are wavelet basis functions and $\vec{\alpha} = \{\alpha_{j,k}\}$ the corresponding coefficients defined by

$$\alpha_{j,k} = \int z(x) \phi_{j,k}(x) dx. \tag{2}$$

One way to describe the wavelet thresholding technique is to prescribe a wavelet coefficient index set I, then retain all coefficients with indices belong to I and truncate the other coefficients to zero:

$$\alpha^*_{j,k} = \begin{cases} \alpha_{j,k} & (j,k) \in I \\ 0 & otherwise \end{cases}$$

For example, in linear thresholding, I is taken as the set of low frequencies; and in hard thresholding, I is defined as the set of all coefficients whose magnitudes are larger than a given tolerance, otherwise, it is smaller than the tolerance. Since orthonormal wavelets form an orthonormal basis of the L^2 space, it is obvious that the hard thresholding selection of I minimizes the L^2 error between the compressed image $u(x)$ and the observed image $z(x)$.

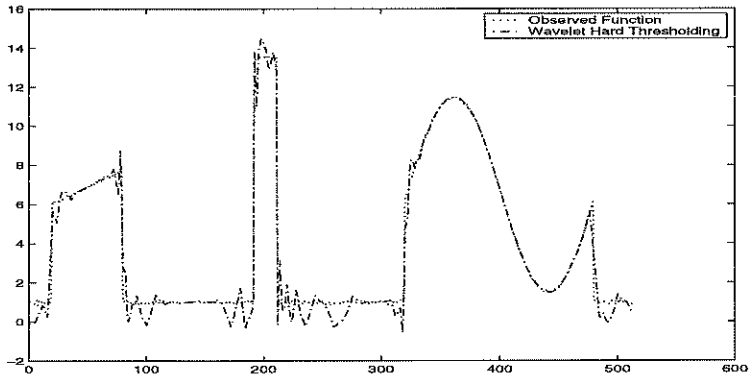

FIG. 1. *The observed function (dotted) has large jumps. The 4-level DB6 wavelet hard thresholding approximation (dash-dotted) is reconstructed by retaining the largest* 64 *coefficients. It generates oscillations at each jump.*

The hard thresholding approximations introduce oscillations at the edges, although they are optimal in the L^2 space. This is due to the fact that the L^2 norm minimization does not penalize oscillations. In Fig 1, we show a 4-level Daubechies-6 (DB6) wavelet hard thresholding approximation (dash-dotted) to a discontinuous function (dotted). The approximation is reconstructed by retaining the largest 64 non-zero coefficients and truncating the other coefficients to zero. It is obvious that it generates severe oscillations at each jump. Fig 2 is a 2-D image containing four noisy squares with different sizes and intensities. We show its 4-level DB6 wavelet hard thresholding approximation in Fig 3. The approximation contains edge artifacts along the boundaries of the objects, while in the observed image, these objects have sharp edges.

Wavelet thresholding can cause oscillations near edges, consequently increasing the TV norm of the reconstructed image. To suppress these oscillations, we propose the following model to select the index set I, and modify the values of the retained wavelet coefficients $\beta_{j,k}$ such that the reconstructed image $u(\vec{\beta}, x)$ form a less severe oscillatory approximation:

$$\min_{\beta_{j,k},(j,k)\in I} F(u,z) = \lambda \int |\nabla_x u(\vec{\beta}, x)| dx + \frac{1}{2}\|u - z\|_2^2 \tag{3}$$

subject to

$$|I| = m, \tag{4}$$

where $u(\vec{\beta}, x)$ has wavelet transform:

$$u(\vec{\beta}, x) = \sum_{j,k} \beta_{j,k}\phi_{j,k}(x).$$

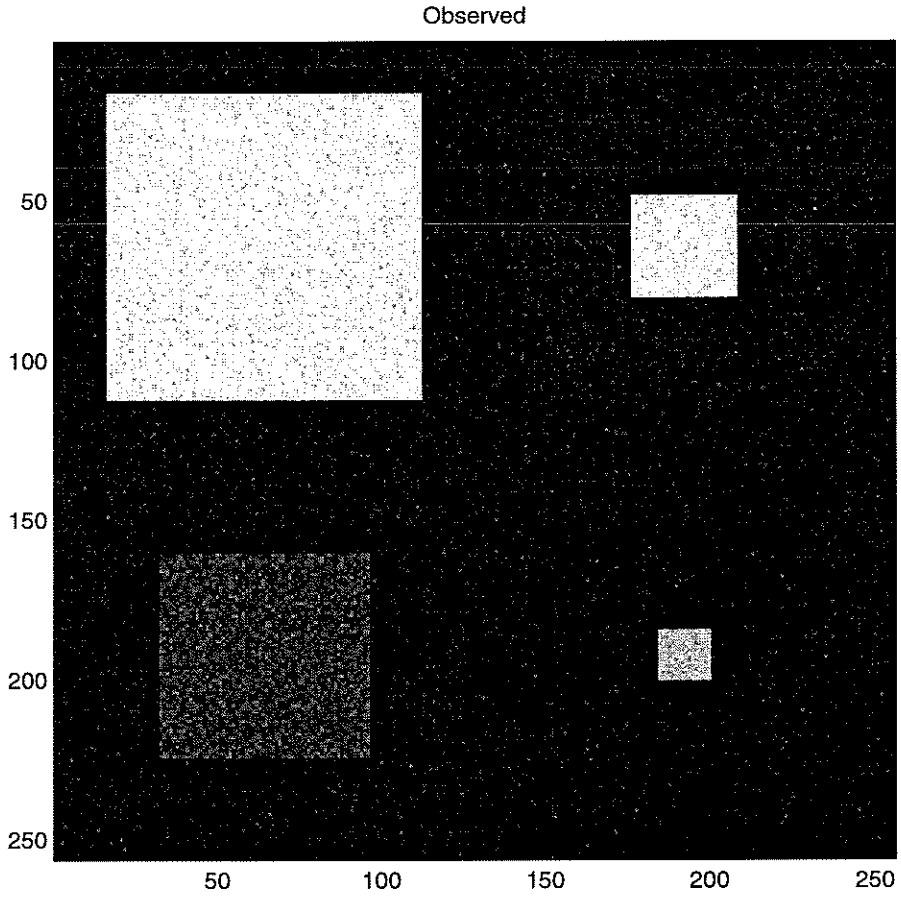

FIG. 2. *The observed image has features with sharp edges despite of the presence of noise.*

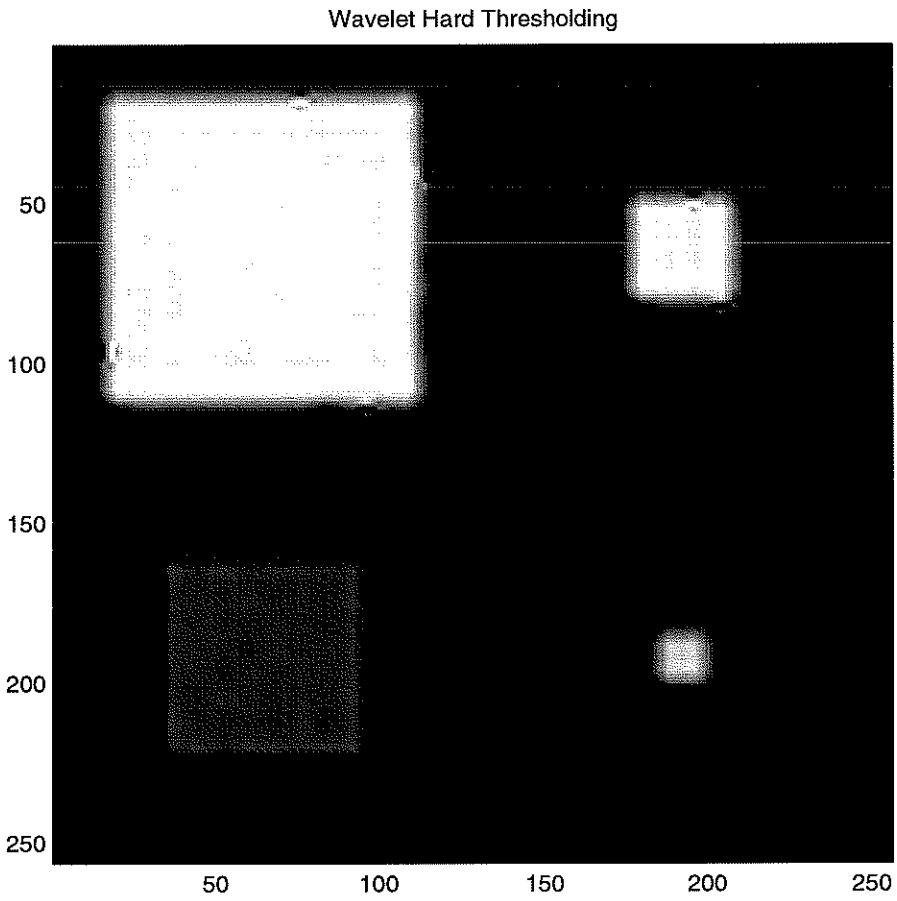

FIG. 3. *The 4-level DB6 wavelet hard thresholding reconstruction which retains the largest* $16x16$ *coefficients. Edge artifacts are clearly seen along the boundaries.*

Here we have $\beta_{j,k} = 0$ if $(j,k) \notin I$, $|I|$ represents the number of elements in I, m is a given integer, and λ the regularization parameter.

The first term in the objective functional reduces the oscillations of $u(x)$ by diminishing its TV norm. The second term is the standard L^2 fitting term which controls the difference between $u(x)$ and the observed image $z(x)$. The regularization parameter λ is used to balance the trade-off between the suppression of oscillations and the fitting term. When λ tends to zero, $u(x)$ goes to the standard hard thresholding approximation. On the other hand, when λ tends to infinity, the suppression term dominates the objective functional, and therefore $u(x)$ tends to a constant. As a TV regularization parameter, λ also controls the smallest scale of features which are preserved [27], i.e. for a given value of λ, there exists a size of feature such that the model treats all features smaller than this size as oscillations and eliminates them, while preserving features which are bigger than this critical scale. In practice, λ can be determined in many ways, for instance, using the L-curve technique [19] to select the best λ, or determining it by using a set of training images. In this paper, we do not discuss these approaches in detail, though we use the latter choice to select λ in our numerical experiments.

Compared to the approach proposed in [12], which uses the TV denoising method followed by standard wavelet thresholding to obtain high ratio compression for noisy data, the advantage of the proposed TV regularized wavelet compression model is that the TV regularized model can reduce the oscillations generated by wavelet thresholding as well as the noise, while TV denoising followed by standard thresholding may generate new oscillations after denoising. Also, the TV regularized wavelet compression model can directly work on wavelet coefficients, making it easier to be fit into practical compression schemes, especially for images given in a wavelet coefficient format (e.g. the upcoming wavelet based JPEG 2000 compression standard). In addition, the TV regularized model operates on a smaller number of coefficients (in the hard thresholding case). Potentially, it could be faster than TV denoising followed by standard thresholding.

Remark: The TV regularization term in the model can be replaced by the H_1 regularization term $\|\nabla u\|_2^2$, or other regularization terms. Compared to the TV term, these other norms usually smooth out sharp edges in the reconstructed images. We will show a comparison in our numerical experiments in section 5.

The TV regularized model for wavelet thresholding is a nonlinear integer optimization problem which in general cannot be solved efficiently. There are two crucial tasks in finding the global minimizer of (3): selecting the index set I; and modifying the retained coefficients $\beta_{j,k}$, $(j,k) \in I$. The major difficulty is in the selection of the index set I, because there are too many combinations for possible I. In fact, for each selected I, which forms a subspace of the L^2 space, there is a local minimizer in this subspace. And the global minimizer is among these local minimizers. On the other hand, both the magnitude and the location of

wavelet coefficients reflect the significance of corresponding features, although they may not necessarily determine the index set I which contains the global solution. So, we can use them to approximate the optimal set I. We will consider several relaxations of such approximations in the next section. After the set I has been determined, finding the minimizer in such a subspace becomes a convex unconstrainted optimization problem. We will address some numerical methods to solve these problems in section 4.

3. Relaxations of General TV Model. In this section, we consider several relaxations of the TV regularized wavelet compression model and give their Euler-Lagrangian equations. The purpose is to reduce the difficulty in determining the index set in the integer constraint (4).

3.1. The Standard Hard Thresholding. A simple way to select the index set I is to choose it according to the magnitude of the coefficients. In other words, we can simply use the standard hard thresholding nonzero coefficient index set I_H to approximate I. Then the TV model is simplified as:

$$\min_{\beta_{j,k},(j,k)\in I_H} F(u,z) = \lambda \int |\nabla_x u(\vec{\beta},x)|dx + \|u-z\|_2^2. \tag{5}$$

As we mentioned before, once we restrict the index set to I_H, the minimization problem becomes convex and unconstrained, and has an unique solution $u(x)$ in this subspace. The solution $u(x)$ satisfies the Euler-Lagrangian equation in wavelet space:

$$-\lambda \int \nabla_x \left(\frac{\nabla_x u}{|\nabla_x u|} \right) \phi_{j,k}(x)dx + 2(\beta_{j,k} - \alpha_{j,k}) = 0, \quad (j,k) \in I_H. \tag{6}$$

3.2. Smooth Approximations of the Constraint. Another way to relax the integer constraint $|I| = m$ is to approximate $|I|$ by smooth functions so that we can apply standard techniques for continuous optimizations. Notice that we have the fact:

$$|I| = ||\vec{\beta}||_0,$$

where $||\cdot||_0$ is defined as the number of non-zero elements in the vector. Olshausen and Field [22] proposed using $\sum_{j,k} \log(1+\beta_{j,k}^2)$ to approximate the 0-norm in controlling the number of non-zero patches in sparse images. Donoho [17] showed that the p-norm function $\|\cdot\|_p^p$ forms a more accurate approximation to the integer constraint.

Using these smooth approximations, we can relax the integer constraint to continuous constraints:

$$(\sum_{j,k} \log(1+\beta_{j,k}^2) - m)^2 \leq \gamma^2, \tag{7}$$

or

$$(\sum_{j,k} |\beta_{j,k}|^p - m)^2 \leq \gamma^2, \tag{8}$$

where γ is a given small positive number used for controlling the number of non-zero components in the coefficients.

Since both approximations are smooth, we can easily convert the constrained problems to unconstrained problems by introducing the Lagrangian multiplier τ:

$$\min_{\beta_{j,k},(j,k)\in I_H} F(u,z) = \lambda \int |\nabla_x u(\vec{\beta},x)|dx + \|u-z\|_2^2 + \tau(\sum_{j,k} \log(1+\beta_{j,k}^2) - m)^2, \tag{9}$$

or

$$\min_{\beta_{j,k},(j,k)\in I_H} F(u,z) = \lambda \int |\nabla_x u(\vec{\beta},x)|dx + \|u-z\|_2^2 + \tau(\sum_{j,k} |\beta_{j,k}|^p - m)^2. \tag{10}$$

Then it is easy to obtain the corresponding Euler-Lagrangian equations in wavelet space as:

$$-\lambda \int \nabla_x \left(\frac{\nabla_x u}{|\nabla_x u|}\right) \phi_{j,k}(x)dx + 2(\beta_{j,k} - \alpha_{j,k}) + 2\tau(\sum_{j,k} \log(1+\beta_{j,k}^2) - m)\frac{\beta_{j,k}}{1+\beta_{j,k}^2} = 0, \tag{11}$$

or

$$-\lambda \int \nabla_x \left(\frac{\nabla_x u}{|\nabla_x u|}\right) \phi_{j,k}(x)dx + 2(\beta_{j,k} - \alpha_{j,k}) + 2\tau(\sum_{j,k} |\beta_{j,k}|^p - m)\frac{\beta_{j,k}}{|\beta_{j,k}|^{2-p}} = 0. \tag{12}$$

Remark: In these two approximations, since the log-function $\sum_{j,k} \log(1+\beta_{j,k}^2)$ and p-norm function $\sum_{j,k} |\beta_{j,k}|^p$ $(p \leq 1)$ are not convex functions, there may exist many local minimizers which are also solutions to the Euler-Lagrangian equations.

4. Numerics. To find the solutions for the relaxations of the TV regularized wavelet compression model, we want to solve the associated Euler-Lagrangian equations (6), (11), and (12). In fact, many numerical methods for similar equations in physical space have been proposed in literature, such as Rudin, Osher and Fatemi's time marching scheme [25], Vogel and Omen's fixed-point iterative method [28], and Chan, Golub and Mulet's primal-dual method [8]. All these methods can be adapted to the wavelet space. Here, we use the simple fixed-point iterative method as an example to show some numerical aspects involved in the computation.

The fixed-point iterative method linearizes the nonlinear terms in the Euler-Lagrangian equations with approximations from the previous iteration. We denote $D_{x,+}$ $(D_{x,-})$ as the

forward (backward) finite differences in physical space. Then the fixed-point schemes to the three relaxations are:

$$(13) \qquad -\lambda \int D_{x,-}\left(\frac{D_{x,+}u^{n+1}}{\sqrt{|D_{x,+}u^n|^2+\epsilon_1}}\right)\phi_{j,k}(x)dx + 2(\beta_{j,k}^{n+1}-\alpha_{j,k}) = 0, \quad (j,k)\in I_H,$$

and

$$(14) \qquad -\lambda \int D_{x,-}\left(\frac{D_{x,+}u^{n+1}}{\sqrt{|D_{x,+}u^n|^2+\epsilon_1}}\right)\phi_{j,k}(x)dx$$
$$+2(\beta_{j,k}^{n+1}-\alpha_{j,k}) + 2\tau(\sum_{j,k}\log(1+(\beta_{j,k}^n)^2)-m)\frac{\beta_{j,k}^{n+1}}{1+(\beta_{j,k}^n)^2} = 0,$$

and

$$(15) \qquad -\lambda \int D_{x,-}\left(\frac{D_{x,+}u^{n+1}}{\sqrt{|D_{x,+}u^n|^2+\epsilon_1}}\right)\phi_{j,k}(x)dx$$
$$+2(\beta_{j,k}^{n+1}-\alpha_{j,k}) + \tau(\sum_{j,k}|\beta_{j,k}^n|^p-m)\frac{\beta_{j,k}^{n+1}}{(|\beta_{j,k}^n|+\epsilon_2)^{2-p}} = 0,$$

respectively, where $u^n = u(x,\vec{\beta}^n)$, ϵ_1 and ϵ_2 are small positive numbers which are used to prevent blow-up in regions where $\nabla u = 0$, or $\beta_{j,k} = 0$. Equations (13), (14) and (15) are linear equations in the unknowns $\beta_{j,k}$, which we solve by Conjugate Gradient (CG) without preconditioning.

We note that the unknowns $\beta_{j,k}^{n+1}$ are in wavelet space but the finite difference operators $D_{x,+}$ and $D_{x,-}$ are defined in physical space. To compute them, we need to transform the data from wavelet space back to physical space. Then after calculating the finite differences, we transform the data back to wavelet space.

Remark: On the other hand, since wavelet transforms are local, we can in fact directly compute the finite difference terms locally in wavelet space so that it is not necessary to transform the data back and forth between the two spaces. In our numerical examples, we do not use this method, so we will not discuss it in detail.

5. Examples. In this section, we will show some 1-D and 2-D examples to demonstrate the improvement in images of the TV regularized wavelet compression models for wavelet thresholding. In all computations, we use the fixed-point iterative schemes (13), (14) and (15) introduced in section 4. We choose the parameters ϵ_1 and ϵ_2 as 10^{-8}, and $\tau = 10^{-5}$ in all situations.

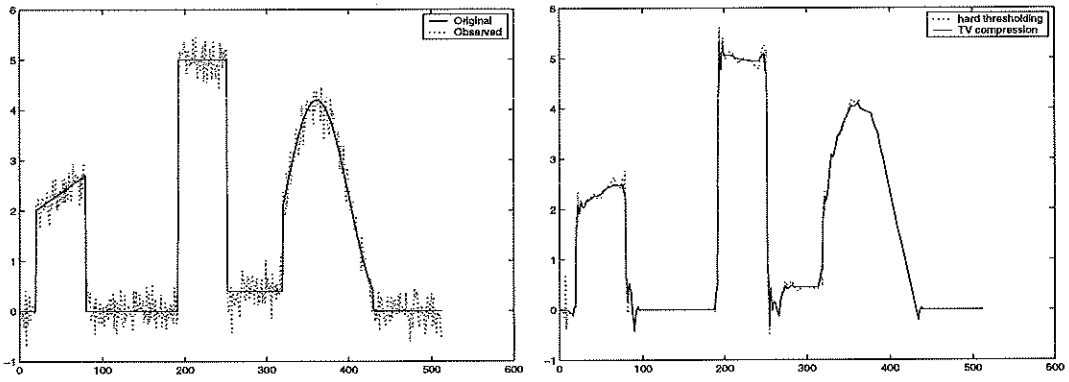

FIG. 4. *Left: The observed function (dotted) and the original noise free function (solid). Right:* $\lambda = 0.001$, *As* λ *becomes smaller, the TV norm approximation (solid) tends to the hard threshold approximation.*

In the first 1-D example, we show a sequence of images (Fig 4 (right) to 6 (left)) which are computed by hard thresholding relaxation (13) of the observed function shown in Fig 4 (left). We use the 4-level Daubechies-4 (DB4) wavelet transform, and take $\lambda = 0.001, 0.01, 0.1$, and 1 respectively. In each picture, we show the standard wavelet hard thresholding approximations (dotted line) restored by retaining the largest $m = 50$ coefficients (with respect to the original 519 coefficients), and the TV regularized wavelet compression approximations (solid) which are also reconstructed by these 50 non-zero coefficients with perturbed values. We notice that when λ tends to 0, the TV regularized model results are getting closer to the standard wavelet hard thresholding approximation which has more oscillations at the discontinuities. As λ increases, the TV norm of the reconstructed images decrease, specifically the oscillations at discontinuities are suppressed. When $\lambda = 1$, the reconstructed images are almost flat. All features of the observed image have been lost. This reflects the domination of the TV norm regularization over the fitting term.

More importantly, from these figures, we notice that when $\lambda = 0.01$, the TV regularized approximation has fewer oscillations than that of standard hard thresholding but still retains the silent features of the image and keep the sharp edges. For all images generated using larger λ, some features have been altered, while on the other end, for all images generated using smaller λ, the edge artifacts are still significant, although all features are preserved.

For the next 1-D example, we display in Fig 6 (right) the reconstructed image using the H_1 regularization $||\nabla u||_2^2$ term instead of the TV term in (2) with $\lambda = 0.0002$. The approximation (solid) is also reconstructed by using $m = 50$ non-zero coefficients. Comparing to the previous examples, it is obvious that H_1 regularization smears all sharp edges because it doesn't allow the existence of discontinuities.

The next 1-D example shows another sequence of images (Fig 7 to 8 (left)) which correspond to the TV regularized wavelet compression models with the hard thresholding (5),

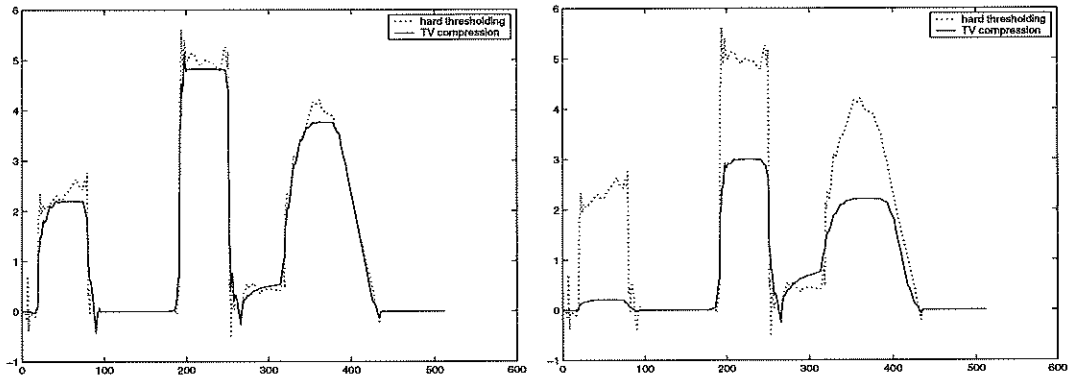

FIG. 5. *Left:* $\lambda = 0.01$, *The TV norm reconstructed approximation (solid) has fewer oscillations at the discontinuities than that of the wavelet hard thresholding approximation (dotted). Right:* $\lambda = 0.1$, *The TV norm hard thresholding approximation (solid) smooths the oscillations out but also alters the features.*

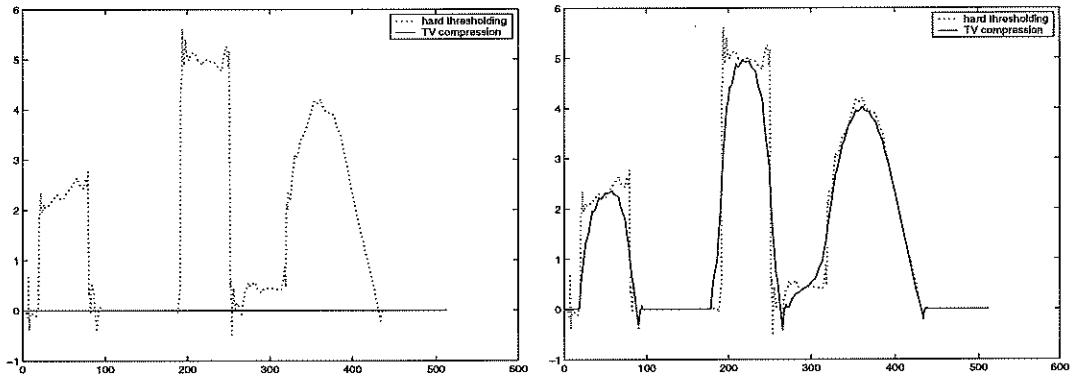

FIG. 6. *Left:* $\lambda = 1$, *The TV reconstructed approximation (solid) is almost a straight line which indicates that the regularization term dominates the objective functional. All features have been eliminated. Right:* $\lambda = 0.0002$, *the H-1 regularization approximation (solid) smooths all sharp edges in the reconstructed image.*

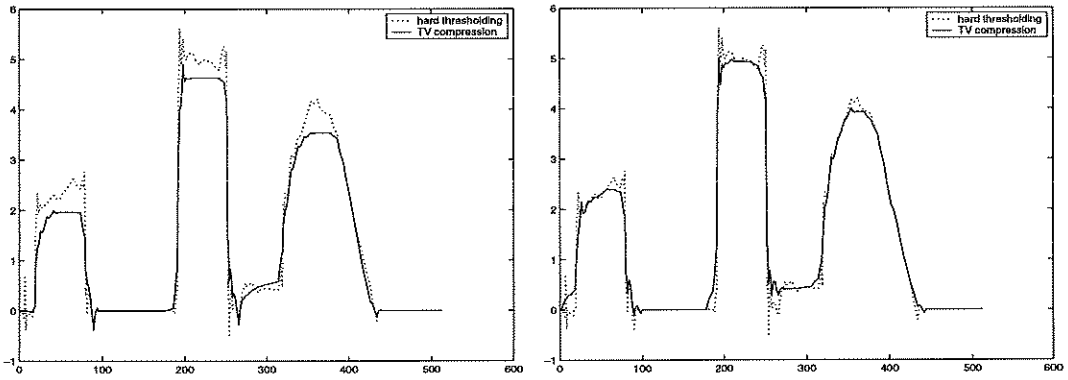

FIG. 7. *Left:* $\lambda = 0.02$, *TV norm hard thresholding approximations (solid) have fewer oscillations at discontinuities. It keeps all features. Right:* $\lambda = 0.02$, *The TV norm log function reconstructed approximation (solid) has fewer oscillations at the discontinuities than that of the wavelet hard thresholding approximation (dotted). It keeps all features as well.*

the log function (9), and the p-norm (10) approximations of the noisy image in the previous example with $\lambda = 0.02$. They are computed by the fixed-point scheme with 10 iterations. All of them are restored by keeping $m = 50$ non-zero wavelet coefficients. The three approximations retain the features of the observed function and have less severe edge oscillations at the discontinuities. Among them, the p-norm approximation fits the original function better than the two others in this example, although all are very similar.

We also compare the results of this TV regularized wavelet compression models with that of the procedure we described in [12] which is to denoise the image by the TV denoising model first, and followed by the standard hard thresholding compression (shown in Fig 8 (right)). The results are very similar to the approximations obtained in the previous example except that in the result of TV denoising followed by standard hard threshold, there exist more edge oscillations (at the right jump of the first bump) than the results of the TV regularized wavelet compression shown in the previous examples.

In the last 1-D example, we show another comparison of approximations obtained by different numbers of fixed-point iterations. The pictures in Fig 9 are calculated using $5, 10$, and 20 fixed-point iterations of (13) with $\lambda = 0.01$ respectively. The approximations are very close. This illustrates that the fixed-point scheme converges fast in the first few iterations.

The next example is for a 2-D comparison of the standard hard thresholding and the TV regularized wavelet compression model images. As shown in Fig 3), the standard 4-level DB6 wavelet compression by retaining the largest $m = 16 \times 16$ coefficients (the ratio of compression is 256:1) has obvious edge artifacts along the boundaries of the objects. In Fig 10, the image is obtained by solving the TV regularized hard thresholding approximation (6) with $\lambda = 0.05$. We perturb the values of the 16×16 non-zero coefficients retained in the

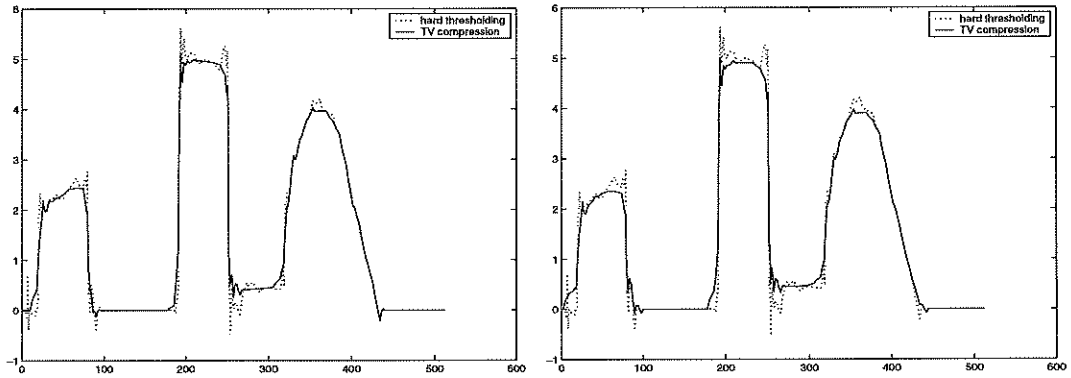

FIG. 8. *Left:* $\lambda = 0.02$, *the p-norm approximation (solid) keeps all features as well as eliminating most of the edge oscillations. Right: The approximation obtained by TV denoising followed by standard hard thresholding.*

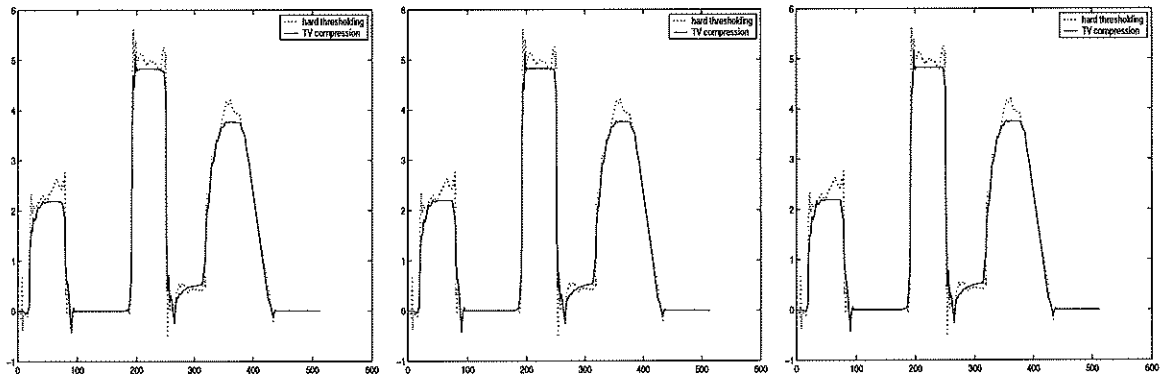

FIG. 9. $\lambda = 0.01$, *TV hard thresholding reconstruction after* 5 *(left),* 10 *(middle),* 20 *(right) fixed-point iterations. The difference between them are invisilbe.*

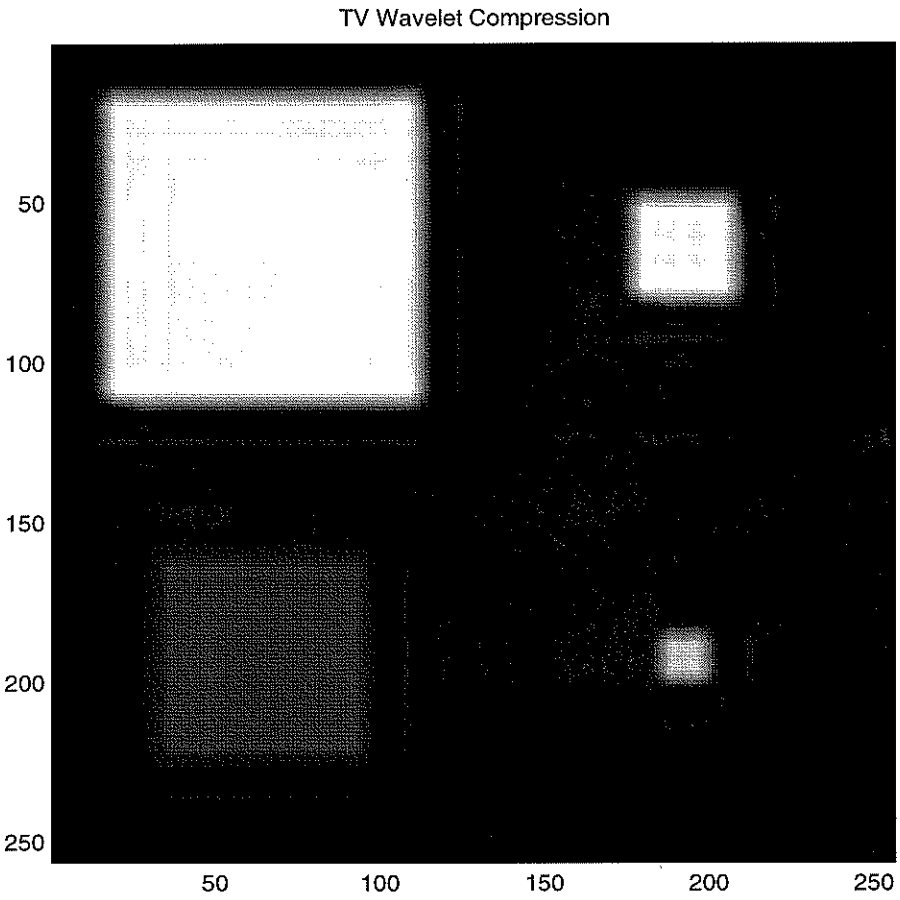

FIG. 10. *The TV regularized Compression with hard thresholding. It keeps the largest* 16×16 *coefficients. Compared to the images shown in Fig 2 and 3, less severe edge artifacts present in this image.*

standard hard thresholding. It is obvious that in this picture, the edge artifacts are less severe than in the standard case. Meanwhile, since the regularization parameter λ also controls the smallest size of features to preserve, in the TV regularized restored image, smaller features (such as the smallest square) are altered more than the large features, i.e. the intensities are lower than the standard approximation. In Fig 11, we show the cameraman image with Gaussian white noise. We display the 64×64 non-zero coefficient reconstruction calculated by standard hard thresholding in Fig 12, and the TV regularized wavelet compression model with hard thresholding in Fig 13. Compared to the standard hard thresholding image, the edge artifacts in the TV model approximations are much less severe.

6. Conclusion. , We have used the TV regularized model to select and modify the non-zero wavelet coefficients in the thresholding procedure. The resulting compressed images

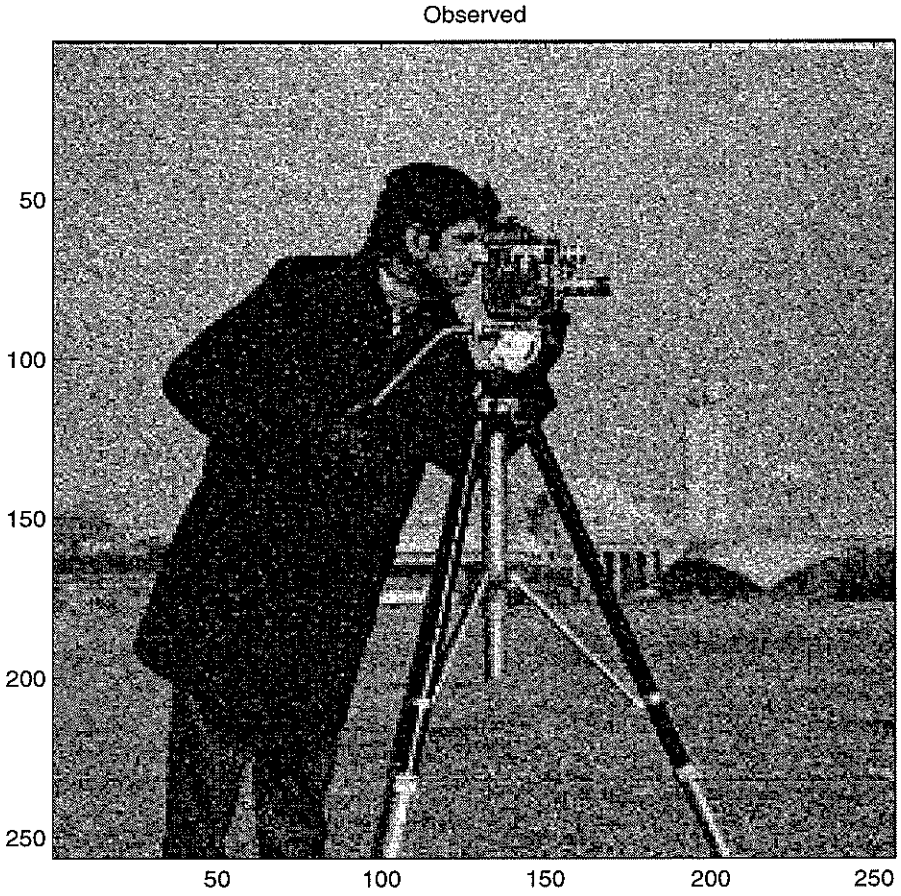

FIG. 11. *The noisy cameraman image.*

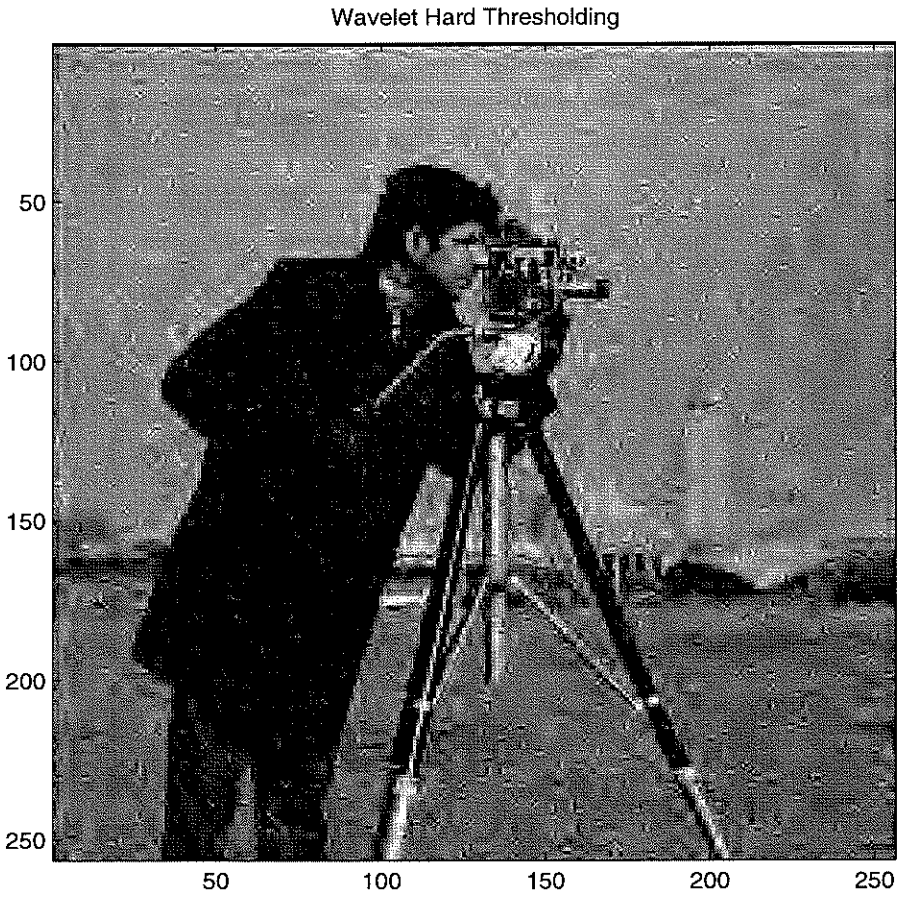

FIG. 12. *The standard hard thresholding approximation It keeps the largest* 64×64 *coefficients. Severe edge artifacts present in it.*

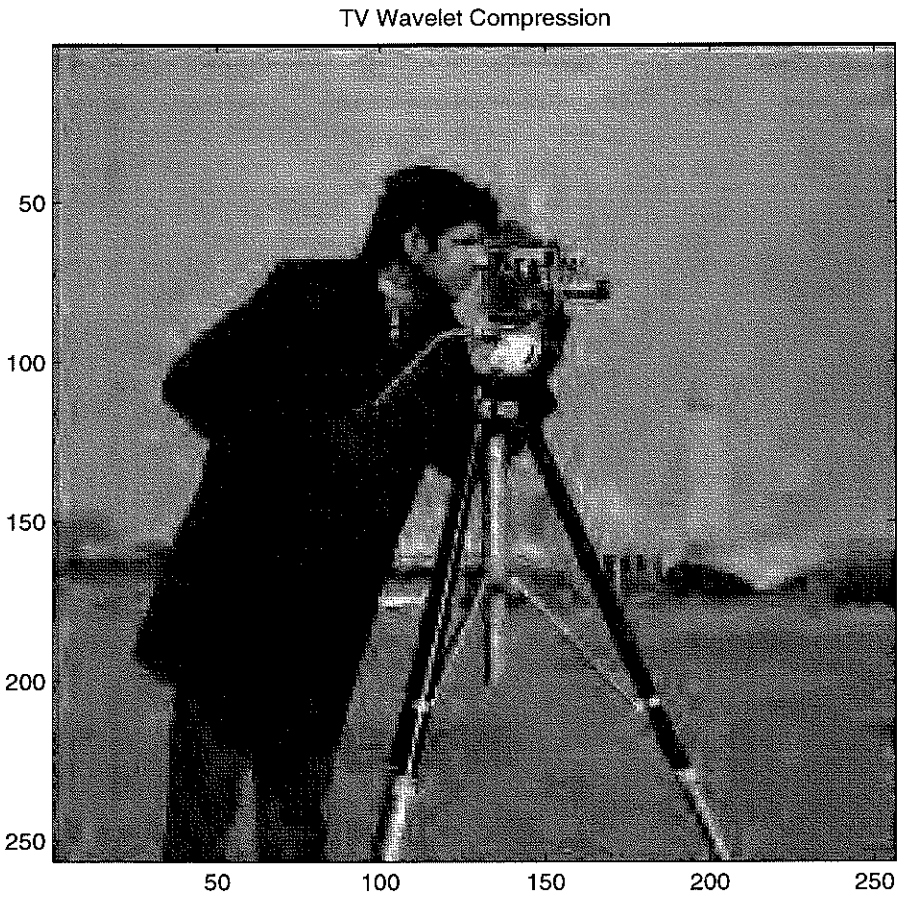

FIG. 13. *The TV norm hard thresholding. It keeps* 64×64 *nonzero coefficients. There are much fewer edge artifacts in the compressed images.*

contain less severe edge artifacts than those in the standard thresholding images, especially when large noise is present in the image. The model can directly operate on the wavelet coefficients, and therefore, can easily be embedded into practical compression schemes. More work needs to be done to improve the speed of convergence and to make the method more practical.

REFERENCES

[1] L. Alveraz, F. Guichard, P. L. Lions and J. M. Morel, *Axioms and Fundamental Equations of Image Processing*, Arch. Rational Mechanics and Anal., 16 (1993), pp200-257.

[2] L. Alveraz, J. M. Morel, *Formalization and Computational Aspects of Image Analysis*, Acta Numerica (1994), pp1-59.

[3] M. Bertalmio, G. Sapiro, V. Caselles and C. Ballester, *Image Inpainting*, Tech. Report, ECE-University of Minnesota, 1999.

[4] P. Blomgren and T. F. Chan, *Color TV: Total Variation Methods for Restoration of Vector Valued Images*, CAM Report, No. 96-5, Dept. of Math., UCLA, 1996.

[5] E. Candes and D. Donoho, *Curvelets: A Surprisingly Effective Nonadaptive Representation of Objects with Edges*, Tech. Report, Dept. of Stat., Stanford Univ., 1999.

[6] A. Chambolle, R. DeVore, N. Lee and B. Lucier, *Nonlinear Wavelet Image Processing: Variational Problems, Compression, and Noise Removal Through Wavelet Shrinkage*, IEEE Tran. Image Proc., Vol. 7, No. 3, Mar. 1998, pp319-333.

[7] A. Chambolle, P. L. Lions, *Image Recovery via Total Variational Minimization and Related Prolems*, Numer. Math., 76, 1997, pp167-188.

[8] T. F. Chan, G. H. Golub, and P. Mulet, *A Nonlinear Primal-Dual Method for Total Variation-Based Image Restoration*, in ICAOS'96, 12th International Conference on Analysis and Optimization of Systems: Images, Wavelets, and PDEs, Paris, June 26-28, 1996, number 219 in Lecture Notes in Control and Information Sciences, pp. 241-252.

[9] T. F. Chan and L. Vese, *Active Contour Without Edges* Submit to IEEE Tran. on Image Proc., 1998.

[10] T. F. Chan and C. K. Wong, *Total Variation Blind Deconvolution*, CAM Report, No. 96-44, Dept. of Math., UCLA, 1996.

[11] T. F. Chan and H. M. Zhou, *Adaptive ENO-Wavelet Transforms for Discontinuous functions*, CAM Report, No. 99-21, Dept. of Math., UCLA, Submit to SIAM Numer. Anal., 1999.

[12] T. F. Chan and H. M. Zhou, *Feature Preserving Lossy Image Compression Using Nonlinear PDE's*, SPIE Proceedings on Advanced Signal Processing Algorithms, Architectures, and Implementations VIII, Vol. 3461, F. T. Luk, ed., San Diego, California, July 1998, pp 316-327.

[13] P. Claypoole, G. Davis, W. Sweldens and R. Baraniuk, *Nonlinear Wavelet Transforms for Image Coding*, Correspond. Author: Baraniuk, Dept. of Elec. and Comp. Sci., also Submit to IEEE Tran. on Image Proc., Preprint, 1999.

[14] D. Donoho, *De-noising by Soft Thresholding*, IEEE Trans. Inf. Th. 41(1995), pp613-627.

[15] D. Donoho, *Wedgelets: Nearly-Minimax Estimation of Edges*, Tech. Report, Dept. of Stat., Stanford Univ., 1997.

[16] D. Donoho, *Orthonormal Ridgelets and Linear Singularities* Tech. Report, Dept. of Stat., Stanford Univ., 1998.

[17] D. Donoho, *Sparse Components of Images and Optimal Atomic Decompositions*, Tech. Report, Dept. of Stat., Stanford Univ., 1999.

[18] *Special Issue on Partial Differential Equations and Geometry-Driven Diffusion in Image Processing*

and Analysis, IEEE Tran. on Image Proc., Vol. 7, No. 3, Mar. 1998.

[19] P. C. Hensen, *The L-curve and Its Use in the Numerical Treatment of Inverse Problems*, Tech. Report, IMM-REP 99-15, Dept. of Math. Model., Tech. Univ. of Denmark, 1999.

[20] J. M. Morel and S. Solimini, *Variational Methods in Image Segmentation*, Birkhauser, 1994.

[21] D. Mumford and J. Shah, *Optimal Approximation by Piecewise Smooth Functions and Associated Variational Problems*, Comm, Pure Appl. Math. 42, 1989, pp577-685.

[22] B. A. Olshausen and D. J. Field, *Emergence of simple-cell receptive field properties by learning a sparse code for natural images*, Nature 281, pp607-609.

[23] A. Marquina and S. Osher, *Explicit Algorithms for a New Time Dependent Model Based on Level Set Motion for Nonlinear Deblurring and Noise Removal*, CAM Report, No. 99-5, Dept. of Math., UCLA, 1999.

[24] P. Perona and J. Malik, *Scale-space and edge detection using anisotropic diffusion*, IEEE T PATTERN ANAL. 12: (7), July, 1990, pp629-639.

[25] L. Rudin, S. Osher and E. Fatemi, *Nonlinear Total Variation Based Noise Removal Algorithms*, Physica D, Vol 60(1992), pp. 259-268.

[26] G. Sapiro and A. Tannenbaum, *Affine Invariant Scale-Space*, Internet. J. Comput. Vision, 11 (1993), pp. 25-44.

[27] David M. Strong, Peter Blomgren, and Tony F. Chan, *spatially Adaptive Local Feature-Driven Total Variation Minimizing Image Restoration*, Proceedings of SPIE, vol. 3167, 1997.

[28] C. Vogel and M. Oman, *Iterative Methods for Total Variation Denoising*, SIAM J. Sci. Comput., Vol. 17, 227-238, 1996..