UCLA
COMPUTATIONAL AND APPLIED MATHEMATICS

A Nonlinear Primal-Dual Method for Total Variation-Based Image Restoration

Tony F. Chan
Gene H. Golub
Pep Mulet

September 1995
CAM Report 95-43

Department of Mathematics
University of California, Los Angeles
Los Angeles, CA. 90024-1555

A NONLINEAR PRIMAL-DUAL METHOD FOR TOTAL VARIATION-BASED IMAGE RESTORATION[§]

TONY F. CHAN*, GENE H. GOLUB[†] AND PEP MULET[‡]

Abstract. We present a new method for solving total variation (TV) minimization problems in image restoration. The main idea is to remove some of the singularity caused by the non-differentiability of the quantity $|\nabla u|$ in the definition of the TV-norm *before* we apply a linearization technique such as Newton's method. This is accomplished by introducing an additional variable for the *flux* quantity appearing in the gradient of the objective function. Our method can be viewed as a primal-dual method as proposed by Conn and Overton [8] and Andersen [3] for the minimization of a sum of Euclidean norms. Experimental results show that the new method has much improved global convergence behaviour than the primal Newton's method.

1. Introduction. During some phases of the manipulation of an image some random noise and blurring is usually introduced. The presence of this noise and blurring makes difficult and inaccurate the latter phases of the image processing.

The algorithms for noise removal and deblurring have been mainly based on least squares. The output of these L^2-based algorithms will be a continuous function, which cannot obviously be a good approximation to our original image if it contains edges. To overcome this difficulty a technique based on the minimization of the Total Variation norm subject to some noise constraints is proposed in [14], where it is also proposed a time marching scheme to solve the associated Euler-Lagrange equations. Since this method can be slow due to stability constraints in the time step size, a number of alternative methods have been proposed, [17], [7], [11].

One of the difficulties of solving the Euler-Lagrange equations is the presence of a highly nonlinear and non-differentiable term, which causes convergence difficulties for Newton method even when combined with a globalization technique such as a line search. The idea of our new algorithm is to remove some of the singularity caused by the non-differentiability of the objective function *before* we apply a linearization technique such as Newton's method. This is accomplished by introducing an additional variable for the *flux* quantity appearing in the gradient of the objective function, which can be interpreted as the unit normal to the level sets of the image function. Our method can be viewed as a primal-dual method as proposed by Conn and Overton [8] and Andersen [3] for the minimization of a sum of Euclidean norms. Experimental results show that the new method has much improved global convergence behaviour than

* Department of Mathematics, University of California, Los Angeles, 405 Hilgard Avenue, Los Angeles, CA 90024-1555. E-mail address: chan@math.ucla.edu. Supported by grants NSF ASC-92-01266 and ONR-N00014-92-J-1890.

† Computer Science Department, Stanford University. E-mail: golub@sccm.stanford.edu.The work of this author was in part supported by the NSF :Grant # CCR-9505393.

‡ Department of Mathematics, University of California, Los Angeles, 405 Hilgard Avenue, Los Angeles, CA 90024-1555 and Department of Algebra, Universitat de València, Dr. Moliner, 50, 46100 Burjassot, Spain. E-mail address: mulet@math.ucla.edu. Supported by DGICYT grants EX94 28990695 and PB94-0987.

§ Paper presented at the AMS/SIAM Summer Workshop on *Linear and Nonlinear CG Methods*, July 9-13, 1995, Seattle.

the primal Newton's method. It is hoped that the new approach can be applied to other geometry-based PDE methods in image restoration, such as anisotropic diffusion [13], affine invariant flows [15] and mean curvature flows [2], since the same singularity caused by $|\nabla u|$ occurs in these methods as well.

The organization of this paper is as follows: in section 2 we introduce the problem, the nonlinear equations associated to it and discuss how to solve them. In section 3 we discuss a linearization technique based on introducing an auxiliary *flux variable*. In section 4 we introduce the constrained formulation for the problem and consider two variants of Newton method for solving the first order conditions. Finally, in section 5 we present some numerical results.

2. Total Variation Regularization. An image can be interpreted as either a real function defined on a bounded and open domain of $\mathbb{R}^2$, Ω, (for simplicity we will assume Ω to be a rectangle henceforth) or as a suitable discretization of this continuous image. The notation $||u||$ ($u \in L^2(\Omega)$) stands for the 2-norm of the function u, $||u|| = (\int_\Omega u^2\,dx\,dy)^{\frac{1}{2}}$, $|u|$ ($u = (u_1, \ldots, u_d)$ a vector function) denotes the function $(\sum_1^d u_i^2)^{\frac{1}{2}}$ and $||y||$ ($y \in \mathbb{R}^m$) denotes the 2-norm of the vector y.

Our interest is to restore an image which is contaminated with noise and/or blur. The restoration process should recover the edges of the image. Let us denote by u_0 the observed image and u the real image. The model of degradation we assume is $Ku + n = u_0$, where n is a Gaussian white noise, of which we assume to know its level measured in the 2-norm, and K is a (known) linear operator, the *blur operator* (usually a convolution operator).

In general, the problem $Ku = z$, with K a compact operator, is ill-posed, so it is not worth solving this equation (or a discretization of it), for the data is assumed to be inexact, and the solution would be highly oscillatory. But if we impose a certain *regularity* condition on the solution u, then the method becomes stable. We can consider two techniques of regularization: Tikhonov regularization and noise level constrained problem. We follow in this paper the constrained formulation. This turns our problem into

$$\begin{aligned}&\min_u \; R(u)\\ &\text{subject to } ||Ku - u_0|| = \sigma,\end{aligned}$$

where R is a certain functional which measures the irregularity of u in a certain sense. For example, $R(u) = ||u||, ||\Delta u||, ||\nabla u \cdot \nabla u||$ The drawback of using these functionals is that they do not allow discontinuities in the solution, and since we are interested in recovering features of the image, they are not suitable for our purposes.

In [14], it is proposed to use as regularization functional the so-called *Total Variation norm* or *TV-norm*:

$$TV(u) = \int_\Omega |\nabla u|\,dx\,dy = \int_\Omega \sqrt{u_x^2 + u_y^2}\,dx\,dy.$$

The solution to the previous problem for $R = TV$ can have discontinuities, thus allowing us to recover the edges of the original image. For this functional the problem can be written as

$$\begin{aligned}&\min\nolimits_u \int_\Omega \sqrt{u_x^2 + u_y^2}\,dx\,dy\\ &\text{subject to } ||Ku - u_0|| = \sigma.\end{aligned} \tag{1}$$

The Euler-Lagrange equation for this problem using Neumann boundary conditions is

$$0 = -\nabla \cdot \left(\frac{\nabla u}{|\nabla u|}\right) + \lambda K^*(Ku - u_0). \tag{2}$$

This equation is degenerate due to the presence of the term $1/|\nabla u|$. To overcome this difficulty we slightly perturb the Total Variation norm functional to become

$$\sqrt{|\nabla u|^2 + \beta} \tag{3}$$

where β is a small positive parameter, so now the problem is

$$\begin{aligned} &\min_u \int_\Omega \sqrt{u_x^2 + u_y^2 + \beta}\, dx\, dy \\ &\text{subject to } ||Ku - u_0|| = \sigma. \end{aligned} \tag{4}$$

and the equation is:

$$0 = -\nabla \cdot \left(\frac{\nabla u}{\sqrt{|\nabla u|^2 + \beta}}\right) + \lambda K^*(Ku - u_0). \tag{5}$$

The main difficulty that this equation poses is the linearization of the highly nonlinear term $-\nabla \cdot \left(\frac{\nabla u}{\sqrt{|\nabla u|^2+\beta}}\right)$.

A number of methods have been proposed to solve this problem. L. Rudin, S. Osher and E. Fatemi [14] have used a time marching scheme to reach the steady state of

$$u_t = \nabla \cdot \left(\frac{\nabla u}{\sqrt{|\nabla u|^2 + \beta}}\right) - \lambda K^*(Ku - u_0)$$

If the noise constraints are satisfied at each step then this method is equivalent to Rosen's projected gradient method, which can be slowly convergent, due to stability constraints. C. Vogel and M. Oman [17] have used the following fixed point iteration to solve the Euler-Lagrange equation

$$-\nabla \cdot \left(\frac{\nabla u^{k+1}}{\sqrt{|\nabla u^k|^2 + \beta}}\right) + \lambda K^*(Ku^{k+1} - u_0) = 0.$$

This method is robust but linearly convergent.

Due to the presence of the highly nonlinear term $\nabla \cdot \left(\frac{\nabla u}{\sqrt{|\nabla u|^2+\beta}}\right)$, Newton method does not work satisfactorily, in the sense that its domain of convergence is small. This is especially true if the regularizing parameter β is small. On the other hand, if β is relatively large then this term is well behaved. So it is natural to use a continuation procedure starting with a large value of β and gradually reducing it to the desired value. T. Chan, R. Chan and H. Zhou propose in [7] such an approach, a continuation algorithm based on Newton's iteration to solve equation (5). Although this method is locally quadratically convergent, the continuation step can be difficult to control.

	Newton iteration for $g(x) = 0$					Newton iteration for $f(x) = 0$				
$x_0 \backslash \beta$	10^{-1}	10^{-2}	10^{-3}	10^{-4}	10^{-5}	10^{-1}	10^{-2}	10^{-3}	10^{-4}	10^{-5}
1	7	6	4	3	5	*	9	6	6	*
2	6	5	2	4	6	*	8	4	*	*
3	6	4	3	5	6	10	7	5	*	*
4	5	4	3	5	7	9	6	*	*	*
5	5	3	4	6	7	8	5	*	*	*

TABLE 1

Comparison of the number of iterations required by Newton method to solve $f(x) = 0$ and $g(x) = 0$, for $a = 0.9999$, different β (horizontally) and different initial guesses x_0 (vertically). A $*$ means that the corresponding iteration has failed to converge.

3. Linearization based on the flux variable. We propose here a better technique to linearize the term $\nabla \cdot \left(\frac{\nabla u}{|\nabla u|}\right)$. This technique bears some similarity to techniques from primal-dual optimization methods and gives a better global convergence behavior than that of the usual Newton continuation method applied to this problem.

The method is based on the following simple observation. While the singularity and non-differentiability of the term $w = \nabla u / |\nabla u|$ is the source of the numerical problems, $|w|$ itself is *smooth* because it is in fact the unit normal vector to the level sets of u. The numerical difficulties arise only because we linearize it in the wrong way.

The idea of the new method is to introduce

$$w = \frac{\nabla u}{\sqrt{|\nabla u|^2 + \beta}}$$

as a new variable and replace (5) by the following system of nonlinear partial differential equations:

$$\begin{aligned} -\nabla \cdot w + \lambda K^*(Ku - u_0) &= 0 \\ w\sqrt{|\nabla u|^2 + \beta} - \nabla u &= 0. \end{aligned} \tag{6}$$

We can then linearized this (u, w) system by Newton's method, for example. This approach is similar to the technique of introducing a *flux variable* in the *mixed finite element method* of [4].

The hope is that the (u, w) is somehow better behaved than the u system. Although at this point we do not have a theory to support this. One is of course the numerical results that we will present later in this paper. What we will do now is to give a scalar example that can explain the better convergence behavior of the flux variable approach compared to the standard one. We compare Newton method applied to the equivalent equations $f(x) = a - \frac{x}{\sqrt{x^2+\beta}} = 0$, which resembles $w - \frac{\nabla u}{\sqrt{|\nabla u|^2+\beta}} = 0$, and to $g(x) = a\sqrt{x^2 + \beta} - x = 0$, which resembles $w\sqrt{|\nabla u|^2 + \beta} - \nabla u = 0$ where $a \approx 1$, $\beta \approx 0$. In Fig 1 we can see that the graph of f is close to be horizontal in almost all the x-axis, whereas the graph of g resembles a relatively non horizontal line in almost all the x-axis. The results of Newton method applied to both systems with different initial guesses shown in Table 1 suggest that the domain of convergence of g is much bigger than that of f.

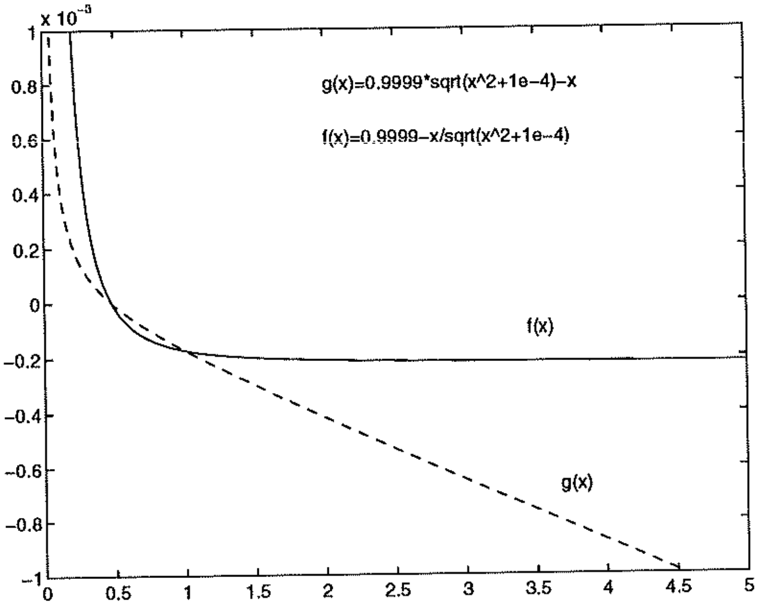

FIG. 1. *Plot of $f(x)$ and $g(x)$. The y-axis has been scaled by a factor of* 10^{-3}.

4. The Discrete Constrained Formulation. The techniques described in this section also apply to the Tikhonov regularization approach for this problem, [17], [7]. Here we treat the more general constrained case.

The images u (resp. u_0) are assumed to be discretized in $m = m_h \times m_v$ (m_h, horizontal size, m_v vertical size) square pixels of dimension $h \times h$ which we order row wise in a vector y (resp. y_0). We use forward differences with Neumann boundary condition for the discretization of ∇u, and a standard quadrature rule for the discretization of the blurring operator. The discretization of (P_β) we obtain in this way is

$$\begin{aligned} &\min_y \sum_{i=1}^{m} h^2 \sqrt{\|\frac{1}{h} A_i^T y\|^2 + \beta} \\ &\text{subject to } \frac{1}{2}(h^2 \|By - y_0\|^2 - \sigma^2) = 0, \end{aligned}$$

i.e.

$$(P_\beta) \qquad \begin{aligned} &\min_y \sum_{i=1}^{m} \sqrt{\|A_i^T y\|^2 + \beta_h} \\ &\text{subject to } \frac{1}{2}(\|By - y_0\|^2 - \sigma_h^2) = 0. \end{aligned}$$

(which is called again (P_β) as no confusion can arise), where $A_i^T y = (y_{i+1} - y_i, y_{i+m_h} - y_i)^T$ for interior pixels and modified accordingly to Neumann boundary conditions for boundary pixels, $\beta_h = h^2\beta$, B is a discretization of the blurring operator K and $\sigma_h = \sigma/h$. If no confusion can arise, we will drop the subscript in σ_h and β_h.

The problem (P_β) is a convex problem with differentiable objective and constraint functions. So, we can obtain the solution y_β of (P_β) by solving the first order conditions for (P_β). Under mild conditions on B, it can be seen that the solution y of (P_β) is unique and $y_\beta \to y$, the solution of (P_0), as $\beta \to 0$, see [1].

4.1. Primal Newton method. The Lagrangian for (P_β) is

$$\mathcal{L}(y,\lambda) = \sum_{i=1}^{n}(||A_i^T y||^2 + \beta)^{\frac{1}{2}} + \frac{1}{2}\lambda(||By - y_0||^2 - \overline{\sigma}^2)$$

and the Karush-Kuhn-Tucker conditions

$$\frac{\partial \mathcal{L}}{\partial y} = \sum \frac{A_i A_i^T y}{\eta_i} + \lambda B^T(By - y_0) = 0 \tag{7}$$

$$\frac{\partial \mathcal{L}}{\partial \lambda} = \frac{1}{2}(||By - y_0||^2 - \overline{\sigma}^2) = 0 \tag{8}$$

where $\eta_i = \sqrt{||A_i^T y||^2 + \beta}$ (note that (7) corresponds to a discretization of (5)). If we use Newton method to solve this system then we have to solve at each step the following linear system:

$$\begin{bmatrix} C & By - y_0 \\ (By - y_0)^T & 0 \end{bmatrix} \begin{bmatrix} \Delta y \\ \Delta \lambda \end{bmatrix} = - \begin{bmatrix} \sum \frac{A_i A_i^T y}{\eta_i} + \lambda B^T(By - y_0) \\ \frac{1}{2}(||By - y_0||^2 - \overline{\sigma}^2) \end{bmatrix}$$

where

$$\begin{aligned} C &= AE^{-1}FA^T + \lambda B^T B \\ A &= \begin{bmatrix} A_1 & \dots & A_m \end{bmatrix} \\ E &= \operatorname{diag}(\eta_i I_2)_{i=1,\dots,m} \\ F &= \operatorname{diag}(I_2 - \frac{A_i^T y y^T A_i}{\eta_i^2})_{i=1,\dots,m} \\ I_2 &= 2 \times 2 \text{ identity matrix.} \end{aligned}$$

We call this approach the *primal Newton method.*

4.2. Primal-dual Newton method. This approach can be regarded as introducing an auxiliary flux variable as in (6) before we linearize the problem using Newton method. It also has a dual formulation as noted in [8] and [3] for the unconstrained problem. There are several ways to introduce the dual variables in this setup. We follow the approach of [3] for simplicity's sake, although the approach of [8] can give much information about the dual problem.

We consider the following equivalent formulation of (P_β)

$$(P'_\beta) \qquad \min_y \sum_{i=1}^{m}(||z_i||^2 + \beta)^{\frac{1}{2}}$$

$$\text{subject to} \begin{cases} A_i^T y - z_i = 0 \\ \frac{1}{2}(||By - y_0||^2 - \overline{\sigma}^2) = 0 \end{cases} .$$

The Lagrangian for this problem is

$$\mathcal{L}(y,z,x,\lambda) = \sum_{i=1}^{n}(||z_i||^2 + \beta)^{\frac{1}{2}} + \sum_{i=1}^{n} x_i^T(A_i^T y - z_i) + \frac{1}{2}\lambda(||By - y_0||^2 - \overline{\sigma}^2)$$

where x and λ are Lagrange multipliers ($x = (x_i), x_i \in \mathbb{R}^2, z = (z_i), z_i \in \mathbb{R}^2$). Hence, the Karush-Kuhn-Tucker conditions for (P'_β) are

$$\frac{\partial \mathcal{L}}{\partial y} = \sum A_i x_i + \lambda B^T(By - y_0) = 0 \tag{9}$$

$$\frac{\partial \mathcal{L}}{\partial x_i} = \frac{z_i}{(||z_i||^2 + \beta)^{\frac{1}{2}}} - x_i = 0 \quad (i = 1, \ldots, n) \tag{10}$$

$$\frac{\partial \mathcal{L}}{\partial z_i} = A_i^T y - z_i = 0 \quad (i = 1, \ldots, n) \tag{11}$$

$$\frac{\partial \mathcal{L}}{\partial \lambda} = \frac{1}{2}(||By - y_0||^2 - \overline{\sigma}^2) = 0 \tag{12}$$

It is worth noting that if we eliminate z_i and x_i from the previous equations then we obtain (7). Our approach differs from the standard Newton one at this point: we allow

$$x_i \neq \frac{z_i}{\sqrt{||z_i||^2 + \beta}}$$

during the Newton iteration.

We eliminate z_i from (10) and multiply the second set of equations by $\sqrt{||A_i^T y||^2 + \beta}$ to obtain

$$\sum A_i x_i + \lambda B^T(By - y_0) = 0 \tag{13}$$

$$A_i^T y - \sqrt{||A_i^T y||^2 + \beta}\, x_i = 0 (i = 1, \ldots, n) \tag{14}$$

$$\frac{1}{2}(||By - y_0||^2 - \overline{\sigma}^2) = 0 \tag{15}$$

(The equations 13 and 14 correspond to a discretization of (6)) We now apply Newton method to this system.

We believe that the change performed to the second set of equations may be the key of the success of this method, see section 3 and [8]: we change an equation containing a highly nonlinear term by others which are much more linear.

The system we have to solve at each Newton step is the following:

$$\begin{bmatrix} A & \lambda B^T B & w \\ E & -\overline{F} A^T & 0 \\ 0 & w^T & 0 \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta y \\ \Delta \lambda \end{bmatrix} = - \begin{bmatrix} Ax + \lambda B^T w \\ Ex - A^T y \\ \frac{1}{2}(||w||^2 - \sigma^2) \end{bmatrix} \tag{16}$$

where $w = By - y_0$ and $\overline{F} = \mathrm{diag}(I_2 - \frac{x_i y^T A_i}{\eta_i})$. By first eliminating Δx, we can solve this system as follows:

$$\begin{bmatrix} \overline{C} & w \\ w^T & 0 \end{bmatrix} \begin{bmatrix} \Delta y \\ \Delta \lambda \end{bmatrix} = - \begin{bmatrix} AE^{-1}A^T y + \lambda B^T w \\ \frac{1}{2}(||w||^2 - \sigma^2) \end{bmatrix} \tag{17}$$

$$\Delta x = -x + E^{-1}A^T y + E^{-1} F A^T \Delta y$$

where $\overline{C} = AE^{-1}\overline{F}A^T + \lambda B^T B$. Note that $\overline{C} \to C$ as we approach the solution of the nonlinear system and also that the complexity of the solution of both methods is comparable.

We call this method the *primal-dual Newton method.*

5. Numerical results. We have compared our algorithms on a series of similarly shaped pictures (see Fig 5) to which we have added Gaussian noise and no blur ($B = I$). The parameter β_{min} is $2^{-13} \approx 1.2e-4$. The relative tolerances we have used are $10^{-6}/\sqrt{m}$ for the outer Newton iteration and $10^{-4}/\sqrt{m}$ for the inner conjugate gradient (CG) iteration. We have used a quasi-Newton method, replacing the coefficient matrix in (17) by its symmetrization (the superlinear convergence is preserved, since $\overline{C} \to C$, which is symmetric), then used block elimination for the $(n+1)\times(n+1)$ system (17) and then the preconditioned conjugate gradient method with ILU(1) as preconditioner for the $n \times n$ systems resulting from it.

The following is a pseudo code for the algorithm we have used in our comparison. This algorithm is not intended to be of practical use; we have used it to illustrate the improvement of the global convergence of the primal-dual Newton method compared to that of the primal Newton method.

Init: Choose a big enough β, $y = y_0 + 2\sigma$, $k = 0$

while $\beta \geq \beta_{min}$

if (Solve P_β for y by either variant of Newton method) successful

$\beta = \beta/2$

else

$\beta(k) = 2 * \beta$

$k = k + 1$

After the execution of this algorithm we obtain a sequence $\beta(i)$, $i = 0, \ldots, k$, which we may term the *best continuation sequence of parameters* β. In Fig 2 we plot the overall number of inner CG iterations required to reach β_{min} using the best continuation sequence in each case versus the size of the picture. In Fig 3 and 4 we show a more detailed picture of the best continuation sequence for some of the pictures. In the x axis we show the best continuation sequence and in the y axis the cumulative number of inner CG iterations required to arrive to the solution of subproblem P_β.

For illustrative purposes, we show in Figure 5, the pictures obtained at each step of the best continuation for the primal-dual Newton method for the original and noisy picture shown in Fig 5. The size is 50×50, $\sigma^2 \approx 204$ (equivalent to $SNR = \frac{|u-\overline{u}|_2}{\sigma} \approx 2.32$). We have displayed the noisy image scaling it to fit into the gray-scale range.

Finally, for the sake of comparison we have used the H^1-norm, $R(u) = \int (\nabla u)^2$ as regularization functional to denoise the previous noisy image; the result is also shown in Fig 5.

REFERENCES

[1] R. Acar and C. R. Vogel, *Analysis of total variation penalty methods for ill-posed problems*, Inverse Problems, 10 (1994), pp. 1217–1229.

[2] L. Alvarez, P. Lions, and J. Morel, *Image selective smoothing and edge detection by non-linear diffusion (ii)*, SIAM J. Numer. Anal., 29 (1992), pp. 845–866.

[3] K. D. Andersen, *Minimizing a sum of norms (large scale solution of symmetric positive definite linear systems)*, PhD thesis, Odense University, 1995.

[4] F. Brezzi, *A survey of mixed finite element methods*, in Finite elements, ICASE/NASA LaRC, Springer, 1988, pp. 34–49.

[5] P. H. Calamai and A. R. Conn, *A stable algorithm for solving the multifacility localtion problem involving euclidean distances*, SIAM Journal on Scientific and Statistical Computing, 1 (1980), pp. 512–526.

[6] ——, *A second-order method for solving the continuous multifacility location problem*, in Numerical analysis: Proceedings of the Ninth Biennial Conference, Dundee, Scotland, G. A.

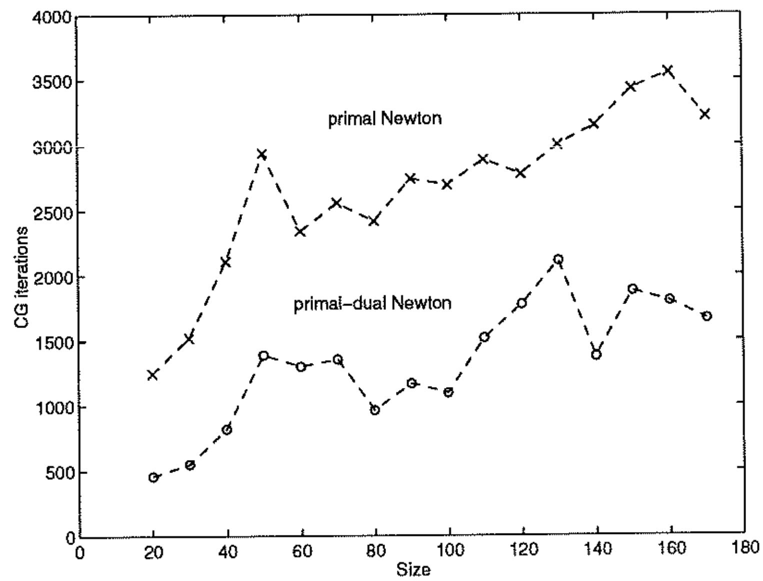

FIG. 2. *Comparison of work for primal and primal-dual methods.*

Watson, ed., vol. 912 of Lecture Notes in Mathematics, Springer-Verlag, 1982, pp. 1–25.

[7] R. H. CHAN, T. F. CHAN, AND H. M. ZHOU, *Continuation method for total variation denoising problems*, Tech. Rep. 95-18, University of California, Los Angeles, 1995.

[8] A. R. CONN AND M. L. OVERTON, *A primal-dual interior point method for minimizing a sum of euclidean norms.* preprint.

[9] D. DOBSON AND F. SANTOSA, *Recovery of blocky images from noisy and blured data*, Tech. Rep. 94-7, Center for the Mathematics of Waves, University of Delaware, 1994.

[10] G. GOLUB AND C. VAN LOAN, *Matrix computations*, 2^{nd} *ed.*, The Johns Hopkins University Press, 1989.

[11] Y. LI AND F. SANTOSA, *An affine scaling algorithm for minimizing total variation in image enhancement*, Tech. Rep. 12/94, Center for theory and simulation in Science and Engineering, Cornell University, 1994.

[12] M. E. OMAN, *Fast multigrid techniques in total variation-based image reconstruction.* to appear in the Preliminary Proceedings of the 1995 Copper Mountain Conference on Multigrid Methods.

[13] P. PERONA AND J. MALIK, *Scale space and edge detection using anisotropic diffusion*, IEEE Trans. Pattern Anal. Mach. Intelligence, 12 (1990), pp. 629–639.

[14] L. RUDIN, S. OSHER, AND E. FATEMI, *Nonlinear total variation based noise removal algorithms*, Physica D, 60 (1992), pp. 259–268.

[15] G. SAPIRO AND A. TANNENBAUM, *Area and length preserving geometric invariant scale-space*, in Proc. 3rd European conf. on Computer Vision, Stockholm, Sweden, May 1994., vol. LNCS Vol. 801, 1994, pp. 449–458.

[16] C. R. VOGEL, *A multigrid method for total variation-based image denoising.* in Computation and Control IV, conference proceedings to be published by Birkhauser.

[17] C. R. VOGEL AND M. E. OMAN, *Iterative methods for total variation denoising*, SIAM J. Sci. Statist. Comput., (to appear).

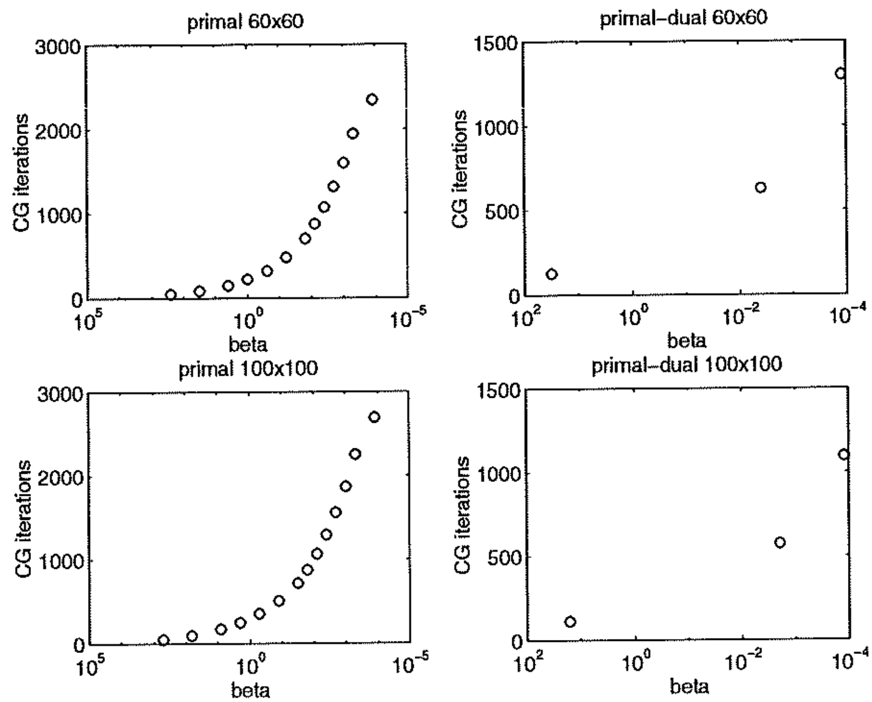

FIG. 3. *Comparison of best continuation sequences.*

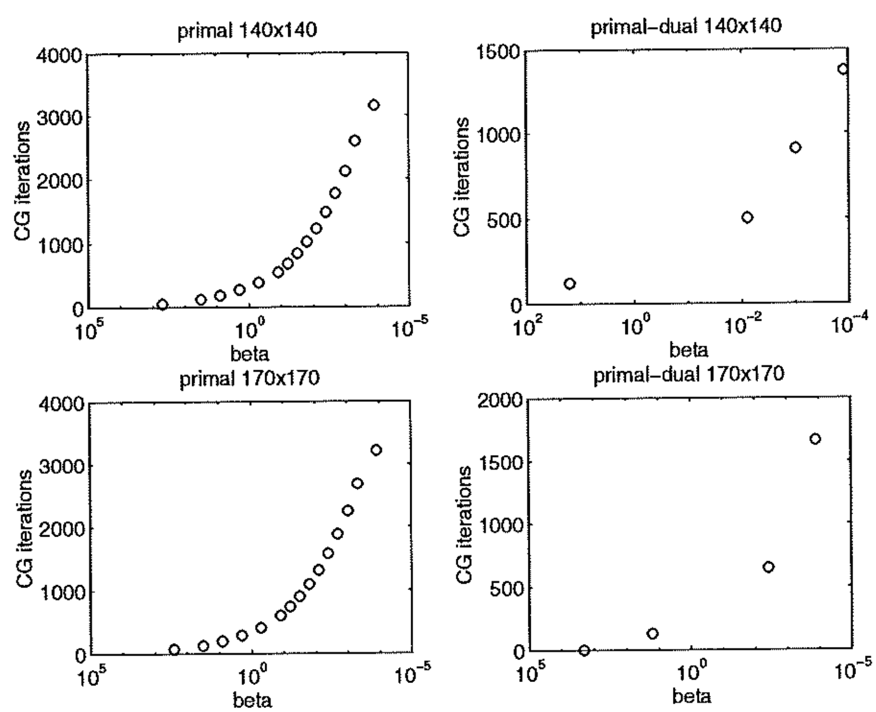

FIG. 4. *Comparison of best continuation sequences (continued).*

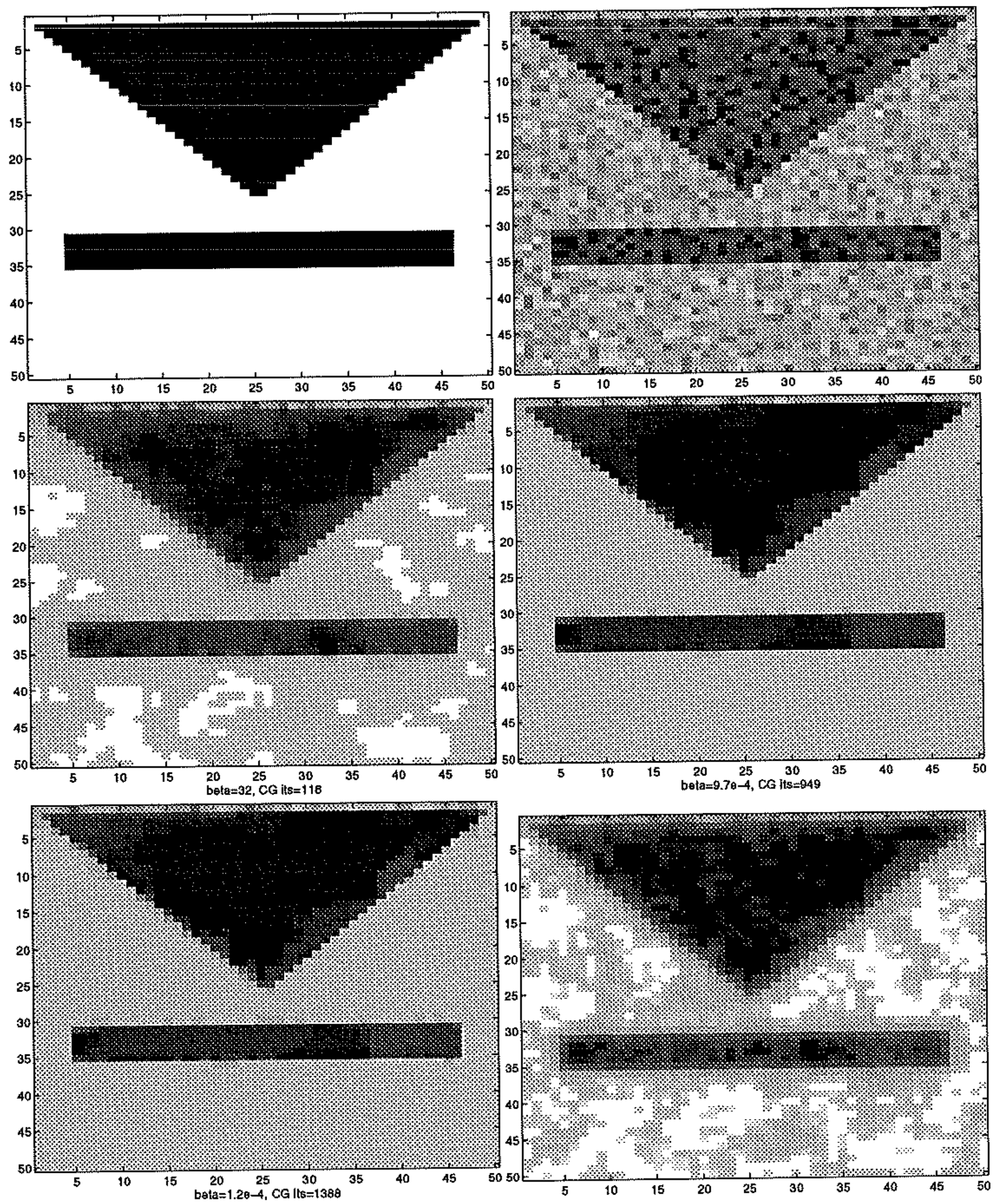

FIG. 5. *From left to right and from top to bottom: 1. Original image, size=* 50×50; *2. Noisy image: display using thresholding,* $\sigma \approx 14.3$, $SNR \approx 2.32$; *3. Image obtained in first step of best continuation; 4. Image obtained in second step of best continuation; 5. Image obtained for desired* $\beta \approx 1.2\mathrm{e}-4$; *6. Image obtained by using* H^1*-semi-norm as regularization functional*