UCLA
COMPUTATIONAL AND APPLIED MATHEMATICS

Modular Solvers for Constrained Image Restoration Problems

Peter Blomgren
Tony F. Chan

November 1997
CAM Report 97-52

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

Modular Solvers for Constrained Image Restoration Problems

Peter Blomgren and Tony F. Chan

Abstract—**Many problems in image restoration cam be formulated as either an unconstrained nonlinear optimization problem:**

$$\min_u R(u) + \frac{\lambda}{2}\|\mathbb{K}u - z\|^2$$

which is the *Tikhonov* [1] approach, where the regularization parameter λ is to be determined; or as a noise constrained problem:

$$\min_u R(u), \quad \text{subject to} \quad \frac{1}{2}\|\mathbb{K}u - z\|^2 = \frac{1}{2}|\Omega|\sigma^2,$$

where σ is the (estimated) variance of the the noise. In both formulations z is the measured (noisy, and blurred) image, and $\mathbb{K}$ a convolution operator.

In practice, it is much easier to develop algorithms for the unconstrained problem, and not always obvious how to adapt such methods to solve the corresponding constrained problem.

In this paper, we present a new method which can make use of *any* existing convergent method for the unconstrained problem to solve the constrained one. The new method is based on a Newton iteration applied to an extended system of nonlinear equations, which couples the constraint and the regularized problem. The existing solver is used in a block elimination algorithm.

The new modular solver enables us to easily solve the constrained image restoration problem; the solver automatically identifies the regularization parameter, λ, during the iterative solution process.

We present some numerical results. The results indicate that even in the worst case the constrained solver requires only about twice as much work as the unconstrained one, and in some instances the constrained solver can be even faster.

I. Introduction

The regularized image restoration problem consists of minimizing a regularity functional, $R(u)$, over a function space $X(\Omega)$, subject to constraints relating the resulting image u, to the measured image z. The constraints depend on the noise level, expressed as either the variance, σ^2, of the noise, or the signal-to-noise-ratio, *SNR*; and the blurring of the image, usually expressed as a convolution $\mathbb{K}u$:

$$\min_{u\in X(\Omega)} R(u) \quad \text{subject to} \quad \frac{1}{2}\|\mathbb{K}u - z\|_2^2 = \frac{1}{2}|\Omega|\sigma^2. \qquad (1)$$

It is reasonable to consider the constrained problem since estimates for the noise level are obtainable in most applications. There already exist some approaches for solving the constrained problem: Rudin-Osher-Fatemi [2] used the projected gradient method of Rosen [3] to control the tradeoff between regularization and fit to measured data, in the setting of an explicit time-marching algorithm. This approach converges quite slowly, due to a time-step restriction. Chan-Golub-Mulet [4] introduced a fully constrained primal-dual approach, which is quadratically convergent. While these approaches are very successful, they are extremely non-modular in the sense that the constraints and the regularizing functional are strongly coupled in the solver. Non-modularity in itself is not an argument against any method. However, it does make it hard to adapt the algorithm to modifications to the regularity functional, and/or the constraints. In our modular approach, we can combine any solver for the unconstrained problem, with any set of constraints.

The motivation for the modular approach is based on the observation that in many cases it is much more straightforward to design schemes for solving the corresponding unconstrained — Tikhonov regularized [1] — problem:

$$\min_{u\in X(\Omega)} R(u) + \frac{\lambda}{2}\|\mathbb{K}u - z\|_2^2, \qquad (2)$$

where λ is a regularization parameter, balancing the tradeoff between minimizing the functional and staying true to the measured image. If chosen correctly, λ is the *Lagrange multiplier* corresponding to the constraint in the constrained formulation (1).

For (1), we get:

$$-\nabla_i R(u) + \lambda\mathbb{K}^*(\mathbb{K}u_i - z_i) = 0, \quad i = 1, 2, \ldots, m. \qquad (3)$$

Finding the solution to the constrained problem comprises of simultaneously identifying the *Lagrange multiplier*, as well as solving the *Euler-Lagrange* equations.

The goal of this paper is to present a highly modular approach to solve the constrained problem. Given an efficient solver for the unconstrained problem (with λ given), for instance a fixed-point [5], primal-dual [4], or transform based algorithm, we show how we can design a solver for the corresponding constrained problem, without special knowledge of the functional or the already existing solver.

We access the unconstrained solver in the form of a black box iteration: $u^{n+1} = S(u^n, \lambda^n)$, which given the estimate u^n to the solution of (2) and a value λ^n approximating λ, returns an improved iterate u^{n+1}. In general it is not necessary for the solver to return an exact solution to (2); it is sufficient that the intermediate solves perform a few iterations of the unconstrained solver.

The modularity facilitates plug-and-play capabilities which enhances experimentation and performance eval-

This work is supported by the ONR under contract ONR N00017-96-1-0277, and by NSF grant DMS-9626755

University of California, Los Angeles. Department of Mathematics. 405 Hilgard Ave. Los Angeles, CA 90095-1555. {blomgren,chan}@math.ucla.edu

uation of different regularity functionals, discretization schemes, constraints, etc...

The modular solver is an adaptation of a general approximate Newton method for coupled nonlinear systems, introduced by Chan [6]. Since the solver is based on a Newton scheme, the overall solver converges quadratically, provided the supplied unconstrained solver has quadratic convergence properties.

In this paper, we present three applications of the modular solver: *(i)* Restoration of gray scale images, $u : \Omega \to \mathbb{R}$, where the regularity functional is the total variation norm of Rudin-Osher-Fatemi [2]. Further, we assume Gaussian "white" noise, with variance σ^2, and no blurring, *i.e.* $\mathbb{K} \equiv I$. We base the solver on the quadratically convergent primal-dual solver of Chan-Golub-Mulet [4]. *(ii)* Restoration of vector valued (RGB color) images, $u : \Omega \to \mathbb{R}^3$, where the regularity functional, $R(u) = \mathrm{TV}_{\mathrm{n,m}}(u)$, is the color-TV norm introduced in Blomgren-Chan [7]. Noise and blur are as in *(i)*. For this problem, we base the constrained solver on a fixed-point lagged diffusivity scheme for the unconstrained problem; an adaption of the scheme introduced by Vogel-Oman [5] for the intensity image case. *(iii)* Restoration of a gray scale image degraded by both noise and blur. We use the Total Variation regularization functional, a fixed-point solver using a cosine transform based preconditioner to solver the resulting linear systems [8].

The algorithm can easily be adapted to other regularization functionals, *e.g.* H^1-regularization, and modified, or multiple constraints.

II. The Modular Solver

As in Chan [6], we introduce the following notation:

$$G^i(u,\lambda) = -\nabla_i R(u) + \lambda \mathbb{K}^*(\mathbb{K}u_i - z_i), \quad i = 1,2,\ldots,m$$

$$G(u,\lambda) = \begin{bmatrix} G^1(u,\lambda) \\ G^2(u,\lambda) \\ \vdots \\ G^m(u,\lambda) \end{bmatrix}$$

$$N(u,\lambda) = \|\mathbb{K}u - z\|_2^2 - |\Omega|\sigma^2.$$

The KKT first order necessary condition for optimality of the constrained problem (1) — see for instance Nash-Sofer [9, chapter 14]) — can be written as the coupled nonlinear system:

$$\begin{bmatrix} G(u,\lambda) \\ N(u,\lambda) \end{bmatrix} = 0. \tag{4}$$

Our modular solver is based on a Newton iteration applied to this system. Assuming that the solution exists, and is regular enough that the Jacobian

$$J(u,\lambda) = \begin{bmatrix} G_u & G_\lambda \\ N_u & N_\lambda \end{bmatrix}$$

is nonsingular at the solution, at each step of the Newton iteration, we are faced with solving the following *linear* system:

$$\begin{bmatrix} G_u & G_\lambda \\ N_u & N_\lambda \end{bmatrix} \begin{bmatrix} \delta u \\ \delta\lambda \end{bmatrix} = - \begin{bmatrix} G \\ N \end{bmatrix}, \tag{5}$$

yielding the changes $(\delta u, \delta\lambda)$ in the Newton iterates.

We apply a block elimination algorithm to the system (*see* Keller [10]):

First we notice that $N_\lambda = 0$ and premultiplying the equation for δu by G_u^{-1} yields:

$$\begin{bmatrix} I & G_u^{-1}G_\lambda \\ N_u & 0 \end{bmatrix} \begin{bmatrix} \delta u \\ \delta\lambda \end{bmatrix} = - \begin{bmatrix} G_u^{-1}G \\ N \end{bmatrix}. \tag{6}$$

Now the key idea in the modular algorithm is to approximate the terms $w = G_u^{-1}G$, and $v = G_u^{-1}G_\lambda$ by calls to the unconstrained solver S. First, observe that w can be approximated by one call to the existing constrained solver, $w = S(u,\lambda) - u$. The reason this works is that $G_u^{-1}G$ corresponds to the Newton correction to u for the unconstrained problem (where λ is fixed).

Moreover, we can approximate v by another call to the solver: Differentiating $G(u,\lambda) = 0$ with respect to λ yields $G_u u_\lambda + G_\lambda = 0$. Hence $v = G_u^{-1}G_\lambda = -u_\lambda$. At convergence $-u_\lambda = -S(u,\lambda)_\lambda = -S_u u_\lambda - S_\lambda$. Provided S is sufficiently contractive, *i.e.* $\|S_u\| \ll 1$, it is reasonable to approximate $v \approx -S_\lambda$. In particular, if the solver S is a Newton solver, this approximation is exact, since $S_u = 0$. In practice, all reasonably fast solvers are contractive enough to make this approximation work. Finally, we use a finite difference approximation to S_λ, so that $v = \frac{S(u,\lambda) - S(u,\lambda+\epsilon)}{\epsilon}$, where $|\epsilon| \ll 1$.

Using the approximation v of $G_u^{-1}G_\lambda$ we rewrite (6):

$$\begin{bmatrix} I & v \\ N_u & 0 \end{bmatrix} \begin{bmatrix} \delta u \\ \delta\lambda \end{bmatrix} = - \begin{bmatrix} -w \\ N \end{bmatrix}. \tag{7}$$

Now, elimination of the $(2,1)$-block yields

$$\begin{bmatrix} I & v \\ 0 & -N_u v \end{bmatrix} \begin{bmatrix} \delta u \\ \delta\lambda \end{bmatrix} = - \begin{bmatrix} -w \\ N + N_u w \end{bmatrix}. \tag{8}$$

Thus we can summarize the algorithm as follows:

Algorithm: *Modular Solver*

Problem:
$\min R(u)$, subject to $\frac{1}{2}\|\mathbb{K}u - z\| = \frac{1}{2}|\Omega|\sigma^2$.

Assumption:
$u \leftarrow S(u,\lambda)$ is a convergent solver for
$\nabla R(u) + \lambda \mathbb{K}^*(\mathbb{K}u - z) = 0$.

1. Compute $w = S(u^n, \lambda^n) - u^n$.
2. Compute $v = [\epsilon]^{-1}\left[S(u^n,\lambda^n) - S(u^n, \lambda^n + \epsilon)\right]$.
3. Compute $\delta\lambda = [N_u v]^{-1}(N + N_u w)$, where
   - $N = \frac{1}{2}\left((\mathbb{K}u - z)^2 - |\Omega|\sigma^2\right)$,
   - $N_u = \mathbb{K}^*(\mathbb{K}u - z)$.
4. Compute $\delta u = w - v\,\delta\lambda$.
5. Update $\begin{cases} u^{n+1} &= u^n + \delta u \\ \lambda^{n+1} &= \lambda^n + \delta\lambda. \end{cases}$

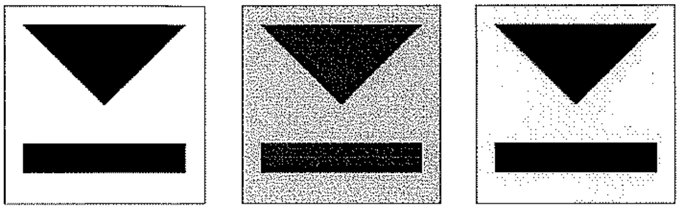

Fig. 1. The true, noisy (SNR = 16.4 dB), and the recovered (SNR = 30.2 dB images.

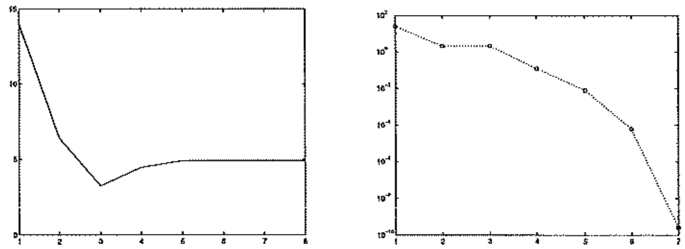

Fig. 2. **Left:** The regularization parameter, λ, as a function of the number of iterations; and **Right:** The residual of the constraint, *e.g.* $N(u,\lambda)$ as a function of the iterations.

Hence, each iteration requires two calls to the unconstrained solver, S. The solution obtained from the first call can be used as an initial guess for the second call, thus speeding up this call considerably*.

III. Numerical Results

We present three applications of the modular solver. First we reconstruct a gray scale image using a primal-dual solver for the unconstrained part of the problem. Since the primal-dual solver converges quadratically, we expect the overall convergence of the modular solver to be quadratic; the numerics show that this is indeed the case. The second application is a reconstruction of a color image. Here, the unconstrained module is a lagged diffusivity fixed point scheme, which is globally and linearly convergent. The final application is a restoration of an image degraded by non-trivial blur, and a small amount of noise. The unconstrained module for this application is also linearly convergent.

For stability reasons we damp the update $\delta\lambda$ so that λ never changes by more than one order of magnitude between iterations.

A. Gray Scale Reconstruction

For our first example, we use the total variation norm of Rudin-Osher-Fatemi [2] to restore a gray scale image. The regularity functional is defined by:

$$\mathrm{TV}_{\mathrm{n},1}(\Phi) \equiv \int_\Omega |\nabla\Phi|\,dx,$$

and the associated *Euler-Lagrange* equations are:

$$-\nabla\circ\left(\frac{\nabla u}{|\nabla u|}\right) + \lambda\mathbb{K}^*(\mathbb{K}u - z) = 0.$$

In the primal dual approach of Chan-Golub-Mulet [4] an auxiliary variable $w = \frac{\nabla u}{|\nabla u|}$, $\|w\| = 1$ is introduced. This leads to an enlarged system of equations which is less non-linear than the original one. Thus, the convergence properties of the expanded problem are more favorable.

The result of using the primal-dual algorithm on a simple test image can be seen in Figure 1. As can be seen in Figure 2, the correct *Lagrange multiplier* λ is identified in 5–6 iterations. After that, convergence is quadratic, as expected.

*For example, when S is the fixed point algorithm of Vogel-Oman [5], the first call may use 5–7 iterations, and the second 2 to reach the same precision.

B. Color Reconstruction

Our second example shows a reconstruction of a color image, where the regularizing functional is the Color-TV norm of Blomgren-Chan [7]:

$$\mathrm{TV}_{\mathrm{n,m}}(u) \equiv \left[\sum_{i=1}^{3}\left[\mathrm{TV}_{\mathrm{n},1}(u_i)\right]^2\right]^{1/2},$$

and the unconstrained module is a lagged diffusivity fixed point scheme, *cf.* Vogel-Oman [5]:

$$-\nabla\circ\left(\frac{\nabla u^{n+1}}{|\nabla u^n|}\right) + \lambda\mathbb{K}^*(\mathbb{K}u^{n+1} - z) = 0.$$

For a detailed discussion of the $\mathrm{TV}_{\mathrm{n,m}}$ norm, including comparisons to other possible extensions, see Blomgren-Chan [7]. Sochen-Kimmel-Malladi [11] relate the $\mathrm{TV}_{\mathrm{n,m}}$ norm to a general framework of flows.

The result of the reconstruction can be seen in figure 3. Figure 4 shows the evolution of the Lagrange multiplier, λ, the residual of the constraint, as well as the constraint of the coupled nonlinear system. We notice that the convergence is linear as expected, since the fixed point solver is only linearly convergent.

Figure 5 shows the evolution of λ for several initial guesses. Even though the modular method is based on a Newton iteration, which is not guaranteed to be globally convergent, we notice that the method is quite robust — typically it converges for initial guesses off by as much as 3 orders of magnitude.

C. Reconstruction with Nontrivial Blur

Our final example shows reconstruction of an image degraded by Gaussian blur, and a small amount of additive noise (SNR = 25.2 dB). Figure 6 shows the true, degraded, and recovered images, and figure 7 shows convergence statistics. The underlying unconstrained solver, based on a fixed-point algorithm [5] with cosine transform based preconditioning [8], converges linearly, a property which is inherited by the modular solver.

D. Work Comparison

Figure 8 shows a convergence comparison of the constrained and unconstrained methods (experiment 2, color restoration). The nonlinear residual is plotted against the

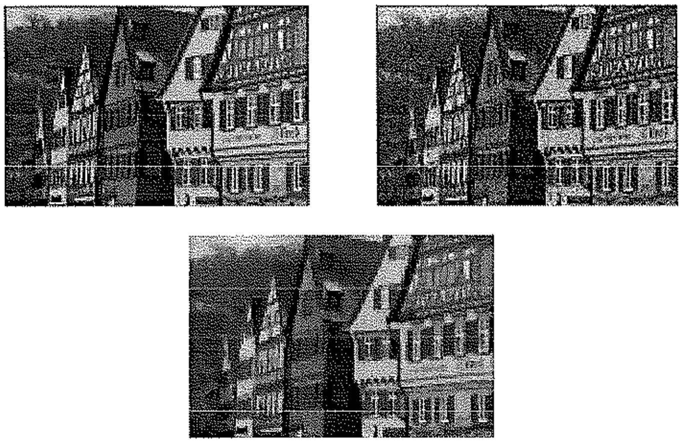

Fig. 3. The true, noisy, and the recovered images.

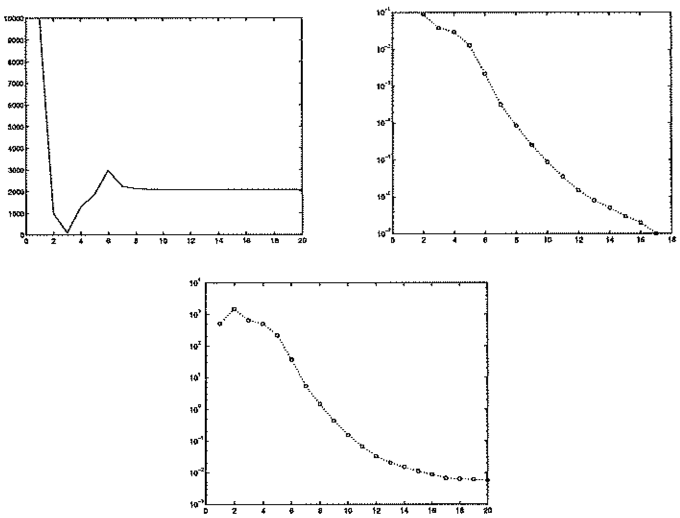

Fig. 4. **Upper-Left:** The regularization parameter, λ, as a function of the number of iterations. **Upper-Right:** The residual of the constraint, *e.g.* $N(u, \lambda)$ as a function of the iterations. **Center:** The residual of the coupled system.

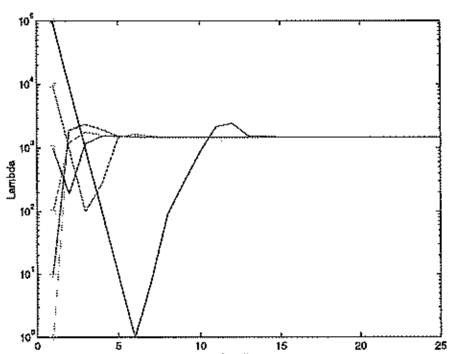

Fig. 5. Convergence of λ for different initial guesses.

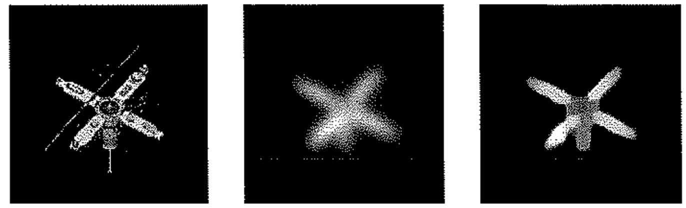

Fig. 6. The true, degraded, and the recovered images.

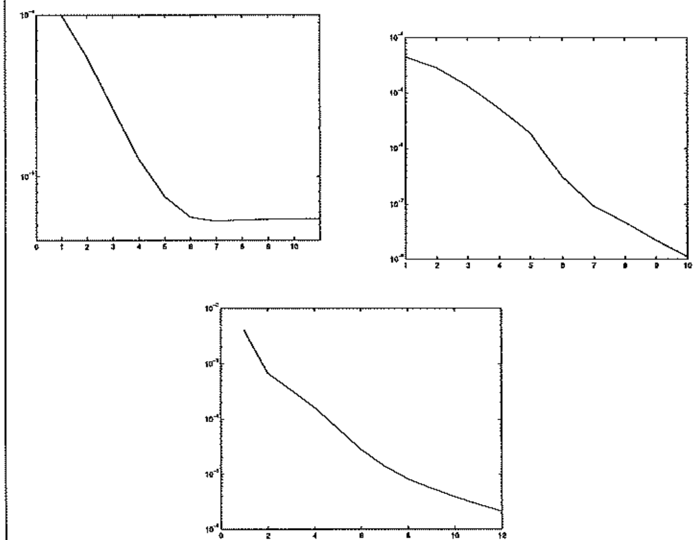

Fig. 7. **Upper-Left:** Convergence of the Lagrange multiplier. **Upper-Right:** The estimated error in the Lagrange multiplier; notice the linear convergence. **Center:** The residual of the non-linear system; again, the convergence is linear.

number of linear systems solved — the actual work performed. We notice that after the initial search for the correct value of the Lagrange multiplier λ, the constrained solver converges faster than the unconstrained one (which was given the correct multiplier λ).

Each iteration of the constrained solver is at most twice as expensive as an iteration for the unconstrained solver, since there are two calls to the unconstrained solver. However, since the second call involves a small perturbation of the Lagrange multiplier, the solution obtained from the first call can be used as an extremely accurate initial guess, thus speeding up the second call considerably.

In practice, the modular solver produces a solution using about the same, or less, amount of work compared with the unconstrained solver (which is given the correct multiplier).

IV. Conclusion

As indicated by the three applications presented, the modular solver is easy to implement. The experimental settings are quite different, nonetheless the same solver is successful in all three.

The easy *"plug-and-play"* facilitates easy experimentation and evaluation of different regularization models. Finding the correct Lagrange multiplier is the only way to make fair comparisons between different models.

The approach is very robust. In most cases the modular

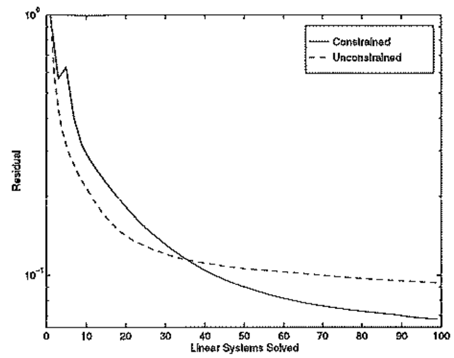

Fig. 8. Comparison of the residual for the constrained, and unconstrained solvers. Notice that the x-axis shows the total number of linear systems solved (the actual work), not the number of iterations.

solver converges for a wide range of initial guesses for the Lagrange multiplier.

The modular solver is efficient. In practice we can compute a solution to the constrained solution in the same amount of time it takes to compute the unconstrained solution.

References

[1] A. N. Tikhonov and V. Y. Arsenin, *Solutions of Ill-Posed Problems*, John Wiley, 1977.

[2] Leonid I. Rudin, Stanley Osher, and Emad Fatemi, "Nonlinear Total Variation Based Noise Removal Algorithms," *Physica D*, vol. 60, pp. 259–268, 1992.

[3] J. G. Rosen, "The Gradient Projection Methods for Nonlinear Programming, Part II, Nonlinear Constraints," *J. Soc. Indust. Appl. Math*, vol. 9, pp. 514, 1961.

[4] Tony F. Chan, Gene Golub, and Pep Mulet, "A Nonlinear Primal-Dual Method for TV-Based Image Restoration," in *ICAOS '96, 12th International Conference on Analysis and Optimization of Systems: Images, Wavelets, and PDEs, Paris, June 26–28, 1996*, M. Berger, R. Deriche, I. Herlin, J. Jaffre, and J. Morel, Eds., 1996, number 219 in Lecture Notes in Control and Information Sciences, pp. 241–252.

[5] C. R. Vogel and M. E. Oman, "Fast Total Variation-Based Image Reconstruction," in *Proceedings of the ASME Symposium on Inverse Problems*, 1995.

[6] Tony F. Chan, "An Approximate Newton Method for Coupled Nonlinear Systems," *SIAM Journal on Numerical Analysis*, vol. 22, no. 5, pp. 904–913, October 1985.

[7] Peter Blomgren and Tony F. Chan, "Color TV: Total Variation Methods for Restoration of Vector Valued Images," Tech. Rep. CAM 96-5, UCLA Department of Mathematics, February 1996.

[8] R. Chan, T. F. Chan, and C. K. Wong, "Cosine Transform Based Preconditioners for Total Variation Minimization Problems in Image Processing," in *Iterative Methods in Linear Algebra*, S. Margenov and P. Vassilevski, Eds., vol. 3 of *IMACS Series in Computational and Applied Math.*, pp. 311–329. IMACS, 1996.

[9] Stephen G. Nash and Ariela Sofer, *Linear and Nonlinear Programming*, McGraw-Hill, 1996.

[10] H. B. Keller, "Numerical Solution of Bifurcation and Nonlinear Eigenvalue Problems," in *Applications of Bifurcation Theory*, P. Rabinowitz, Ed., pp. 359–384. Academic Press, 1977.

[11] N Sochen, R Kimmel, and R Malladi, "From High Energy Physics to Low Level Vision," Tech. Rep. LBNL-39243 / UC-405, Ernest Orlando Lawrence Berkeley National Laboratory, Physics Division, Mathematics Department, August 1996.