Deflated Preconditioned Conjugate Gradients for Nonlinear Diffusion Image Enhancement

Shan, Xiujie; van Gijzen, Martin

DOI
10.1007/978-3-030-55874-1_45

Publication date
2021

Document Version
Final published version

Published in
Numerical Mathematics and Advanced Applications, ENUMATH 2019 - European Conference

Deflated Preconditioned Conjugate Gradients for Nonlinear Diffusion Image Enhancement

Xiujie Shan and Martin van Gijzen

Abstract Nonlinear diffusion equations have been successfully used for image enhancement by reducing the noise in the image while protecting the edges. In discretized form, the denoising requires the solution of a sequence of linear systems. The underlying system matrices stem from a discrete diffusion operator with large jumps in the diffusion coefficients. As a result these matrices can be very ill-conditioned, which leads to slow convergence for iterative methods such as the Conjugate Gradient method. To speed-up the convergence we use deflation and preconditioning. The deflation vectors are defined by a decomposition of the image. The resulting numerical method is easy to implement and matrix-free. We evaluate the performance of the method on a simulated image and on a measured low-field MR image for various types of deflation vectors.

1 Introduction

Many people have benefited from the development of the MRI scanner. However, MRI scanners are expensive and therefore unaffordable for many people in low-income countries. Thus developing a simple and affordable MRI system is urgently needed. The research described in this paper is part of the work to develop a low-field MRI machine for imaging the head of small infants to detect hydrocephalus, a disease that affects many newborns in Africa. A Halbach-array of permanent magnets was designed, optimized, and built [3, 6] to replace the expensive superconducting magnets that are used in conventional MRI systems. This simpler and inexpensive hardware yields more noisy images, which requires the use of denoising processing for medicine practice.

X. Shan
Delft University of Technology, Harbin Institute of Technology, Harbin, China

M. van Gijzen (✉)
Delft University of Technology, Delft, Netherlands
e-mail: M.B.vanGijzen@tudelft.nl

F. J. Vermolen, C. Vuik (eds.), *Numerical Mathematics and Advanced Applications ENUMATH 2019*, Lecture Notes in Computational Science and Engineering 139,
https://doi.org/10.1007/978-3-030-55874-1_45

The diffusion filtering method interprets pixel intensities as a physical quantity that spreads by a diffusion process in the image [2]. The most simple diffusion model for image denoising is standard heat diffusion. The solution of the model is equivalent to a Gaussian low-pass filter, which is also considered to be the filter in signal processing. The major drawback of this model is that it diffuses edges as well as noise. To overcome this, the constant diffusion coefficient is replaced by a coefficient that depends on the image gradient. This idea was first proposed in [7] by Perona and Malik. The resulting PM-model is given by:

$$\begin{aligned} \frac{\partial u}{\partial t} &= \nabla \cdot (c(\|\nabla u\|)\nabla u) \quad \text{in} \quad \Omega \times (0,T), \\ u(x,0) &= f \quad \text{in} \quad \Omega, \\ \frac{\partial u}{\partial \mathbf{n}} &= 0 \quad \text{on} \quad \partial\Omega \times (0,T), \end{aligned} \tag{1}$$

where Ω is the image domain, T is the stopping time, u is the pixel value (which is complex for MR images), f is the noisy image and c is a nonnegative monotonically decreasing function with $c(0) = 1$ and $c(+\infty) \to 0$. Because of the ill-posedness of the PM model, Catté et al. [1] have introduced a regularization method that makes the problem well-posed.

In this paper, we consider the following diffusion coefficient which was originally proposed in [7], modified with the technique in [1] to make the problem well-posed:

$$c(\|\nabla u\|) = e^{-(\|G_\sigma * \nabla u\|/K)^2}. \tag{2}$$

In this equation, G_σ is a Gaussian with standard deviation σ and K is a damping parameter.

We discretize Eq. (1) in space using the standard finite different method, see e.g. [1]. We use implicit Euler to discretize in time and take the diffusion coefficient corresponding the previous time step to linearize the equation. In every time step, we have to solve a large and sparse linear system

$$Au = b \tag{3}$$

where A is symmetric and positive definite. For such systems, the conjugate gradient (CG) method [4] is the method of choice. A classical result for the convergence of CG is that after k iterations the error is bounded by

$$\|u - u_k\|_A \leq 2\|u - u_0\|_A \left(\frac{\sqrt{\kappa} - 1}{\sqrt{\kappa} + 1}\right)^k \tag{4}$$

where $\kappa = \lambda_n/\lambda_1$ is the spectral condition number and the A-norm of u is given by $\|u\|_A = (u^T Au)^{1/2}$. The convergence is slow when the condition number κ is very

large. One way to improve this is to solve the preconditioned system $M^{-1}Au = M^{-1}b$, where M is a matrix that resembles the matrix A. To further speed up the convergence, one can use a deflation technique to map isolated extreme eigenvalues to zero, effectively removing them from the system. Nicolaides [5] chooses deflation vectors that correspond to subdomains: entries of the deflation vector are one for the nodes in its subdomain and others are zero. In [10], subdomain deflation is applied to Poisson problems with strong contrasts in the coefficient which results in a strong improvement of the convergence. This has motivated us to apply this technique to our problem. To define subdomains, we segment the image. Thresholding, region growing, and small patches are used for segmentation, leading to different ways to define the deflation vectors.

The structure of our paper is as follows. Section 2 describes the deflated and preconditioned CG method and we give three choices of the deflation vectors. The influence of preconditioner for the systems is also investigated by analyzing the eigenvalues in Sect. 2. Section 3 gives numerical experiments for the simulated Shepp-Logan image [9] and for a measured Shepp-Logan image. The comparison of the different deflation vectors is presented in Sect. 3 using a simulated and a measured Shepp-Logan image. We end with conclusions in Sect. 4.

2 PCG Methods with Subdomain Deflation

Deflation has been successfully applied to speed up the convergence of the Preconditioned Conjugate Gradient method (PCG) for a number of problem with strong variations in coefficients [8, 10]. The main idea [5] of this DPCG method is summarized below.

2.1 DPCG

The idea of deflation is to split the solution into two parts, one in the range of the deflation subspace $\mathcal{R}(Z)$ and one in its complement. In order to achieve this, we define the projector P by

$$P = I - AZ(Z^T AZ)^{-1}Z^T, \quad Z \in \mathbb{R}^{n\times m} \tag{5}$$

where $Z = [z_1\ z_2\ \cdots\ z_m]$ is the deflation matrix, which we assume to be of full rank. I is the identity matrix. Since $u = (I - P^T)u + P^T u$ we have

$$(I - P^T)u = Z(Z^T AZ)^{-1}Z^T Au = ZA_c^{-1}Z^T b \tag{6}$$

where $A_c = Z^T AZ$. Equation (6) is easy to calculate, we only need to calculate $P^T u$. Using $AP^T = PA$, we can solve the deflated system

$$PA\tilde{u} = Pb \tag{7}$$

for $\tilde{u}$ using the PCG method and then multiplying $\tilde{u}$ by P^T to obtain $P^T u$.

A common choice for the matrix Z, first proposed in [5], is based on a decomposition of the domain Ω. Decomposing domain Ω into m nonoverlapping subdomains Ω_i, $i = 1, 2, \cdots, m$, we choose vectors z_i for $i \in \{1, 2, \ldots, m\}$ such that $z_i = 1$ on $\overline{\Omega_i}$ and $z_i = 0$ on Ω_j, $j \neq i$, $j \in \{1, 2, \ldots, m\}$. With this special choice of Z, the technique for solving the system is referred to subdomain deflation.

We now give the DPCG algorithm for solving the system (3) as follows. Since the pixels values correspond to MR images they are complex valued. For this reason we have to take complex inner products. We therefore use conjugate transpose H instead of normal transpose T in the algorithm. The preconditioning matrix is denoted by M.

DPCG Algorithm

$A_c = Z^H AZ$
$P = I - AZ(A_c)^{-1} Z^H$
$r_0 = Pb - PAu_0$
$k = 0$
while $r_k \neq 0$ **do**
  Solve $z_k = M^{-1} r_k$
  $k = k + 1$
  if $k = 1$ **then**
    $p_1 = z_0$
  else
    $\beta_k = r_{k-1}^H z_{k-1} / (r_{k-2}^H z_{k-2})$
    $p_k = z_{k-1} + \beta_k p_{k-1}$
  end if
  $\alpha_k = r_{k-1}^H z_{k-1} / (p_k^H PAp_k)$
  $\tilde{u}_k = \tilde{u}_{k-1} + \alpha_k p_k$
  $r_k = r_{k-1} - \alpha_k PAp_k$
end while
$u = Z(A_c)^{-1} Z^H b + P^H \tilde{u}_k$.

2.2 *Three Different Choices for the Deflation Vectors*

We use DPCG to solve Eq. (3). We construct the matrix Z by segmenting the image into small images in three different ways: using thresholding, region growing, and

same size patches. For the thresholding and region growing method, we expect that by choosing the interface at edges in the image, i.e., at the location of the jumps in the coefficients, the convergence of the iteration method can be improved. The third technique of same size patches corresponds to the method described in [5]. It does not make use of the image structure, but has the advantage that it is easy to implement. Below we describe the segmentation methods in more detail.

Thresholding

The thresholding method is frequently used for image segmentation. It is a simple and effective segmentation method for images with different intensities [2]. Assuming that the intensity values of image $|f|$ are between 0 and 1, we divide [0, 1] into subintervals I_k. The image is segmented by dividing it into (not necessarily connected) regions with pixel intensities in the same subinterval.

Region Growing

Region growing (RG) segments the image into connected regions with pixel intensities in the same subinterval. To this end, neighbouring pixels are examined, starting from an initial seed point, to determine whether the pixel neighbors should be added to the same region based on a growing condition. The region growing condition we use is as follows: let $|f(i_0, j_0)| \in I_k$ and pixel (i, j) be a neighbour of (i_0, j_0). Then if $|f(i, j)| \in I_k$, the two pixels belong to the same region.

Region Growing

Divide the interval $I = [0, 1]$ into parts I_k, $k = 1, \ldots, s$
for $k = 1 : s$ **do**
  while *stack is empty* **do**
    1 Search image sequentially, find the first pixel (i_0, j_0) that belongs to I_k that does not belong to a segment and set (i_0, j_0) to be seed point.
    2 For all neighbour pixels (i, j) of (i_0, j_0)
    if (i, j) *is not visited and satisfies the region growing condition* **then**
      Add pixel (i, j) to the stack.
    end if
    3 Take a new pixel from the stack and return it to step 2 as (i_0, j_0)
  end while
end for

Same Size Patches

The square domain Ω (image) with resolution $m \times n$ is segmented into $s \times r$ subdomains of the same size (patches), where $m/s, n/r$ are integers.

2.3 Preconditioner

Subdomain deflation works well if the system matrix contains a few small eigenvalues. In order to achieve this, deflation has to be combined with a suitable preconditioning technique. A simple preconditioner that can achieve this is diagonal scaling. This is illustrated in Fig. 1. The left panel shows the spectrum of the unpreconditioned matrix for the simulated Shepp-Logan image considered in numerical experiments. The right panel shows the spectrum of the preconditioned matrix. Clearly, diagonal scaling maps most eigenvalues to values close to one, with the exception of a few eigenvalues that are mapped to small values.

3 Experimental Results

In this section, we evaluate our method on two images: a simulated Shepp-Logan phantom (128×128) and a measured low-field MR image (128×128). Comparisons between CG, PCG, DPCG with the three different deflation methods are presented. Our results correspond to one time step of implicit Euler. For the time step, we take $\tau = 0.06$ and the damping parameter in the diffusion coefficient is $K = 3$. For the CG, PCG and DPCG iterations, initial gues is $u^0 = 0$ and as

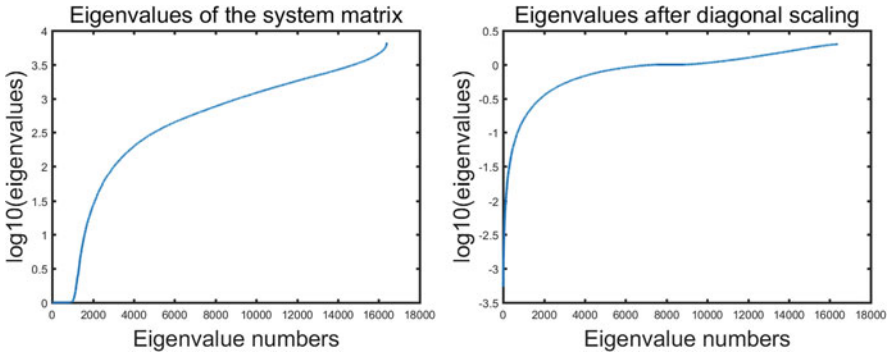

Fig. 1 From left to right: eigenvalues of the system matrix and eigenvalues after diagonal scaling. The eigenvalues are displayed in the log scale

convergence criterion we use $\|r_k\| \leq tol \cdot \|r_0\|$ with $tol = 10^{-5}$. All numerical experiments are carried out using Matlab R2016b on a standard laptop computer.

3.1 Simulated Shepp-Logan Image

Simulated Shepp-Logan image degraded with Gaussian noise with zero mean and variance 0.005 has been tested. The denoising results of the diffusion model are given in Fig. 2. We only show the denoising result of CG because all denoising results based on different numerical algorithms are the same (as they should be). Table 1 shows that CG and PCG need more iterations to converge than the deflated methods. Region growing based DPCG takes more time because of the clustering algorithm. Compared to thresholding segmentation, region growing seems to be more sensitive to noise.

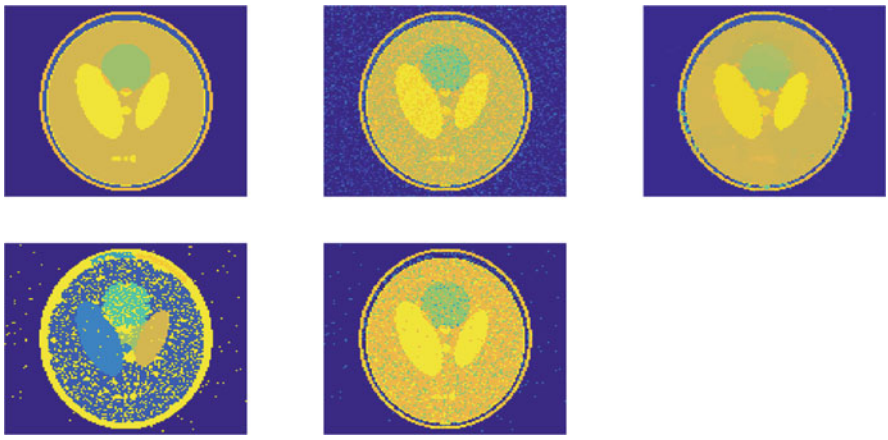

Fig. 2 First row from left to right: original image, noisy image and CG result. Second row from left to right: segmentation (Region growing) and segmentation (Thresholding)

Table 1 Comparisons for simulated Shepp-Logan

Methods	CG	PCG	RG-DPCG	Patches-DPCG(8^2)	Thres-DPCG
Iterations	455	349	230	212	212
Time [a](s)	0.27	0.20	1.21	0.34	0.16

[a]Timings are obtained using Matlab's `cputime` routine. These include the time to segment the image and to construct the deflation matrix

3.2 *Measured Shepp-Logan Image*

In this section, we test our algorithms on an image of 128×128 pixels acquired with the low-field MRI system described in [6]. Results of this Shepp-Logan image are given in Figs. 3 and 4.

From Fig. 3, we know that the diffusion model achieves a good result for denoising. However, due to the strong noise, segmentation of region growing and thresholding result in many small regions. We observe in Fig. 4 that patches-DPCG achieves the fastest convergence. In the above experiments, we use 4^2 patches to construct the deflation vectors. In Table 2, we investigate how the number of DPCG iterations and solution times depend on the number of patches. The number of iterations is reduced considerably for the three Patches-DPCG methods compared to standard CG and PCG. For this example, 4^2 patches yields the fastest solution time.

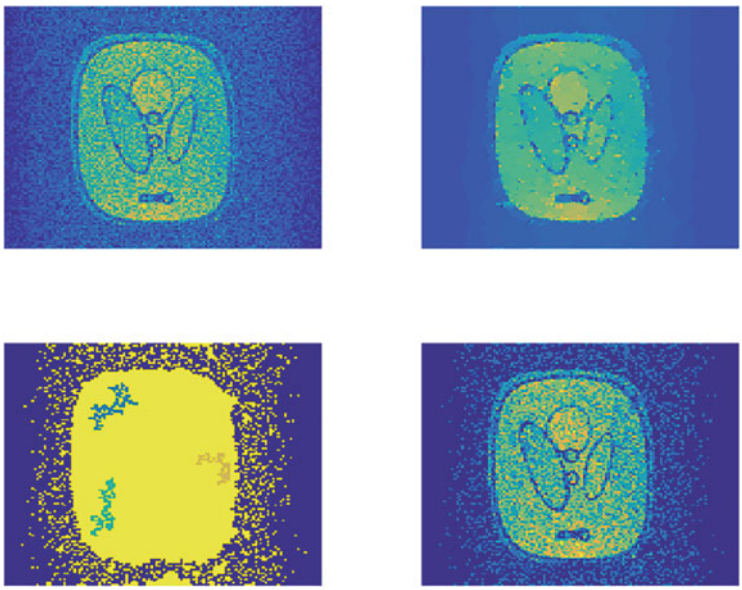

Fig. 3 First row from left to right: original image, DPCG (RG). Second row from left to right: segmentation (RG) and segmentation (Thresholding)

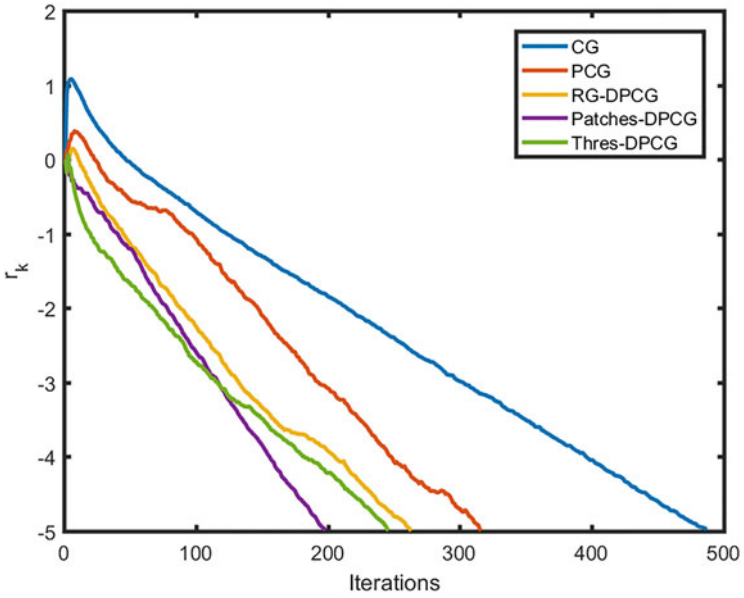

Fig. 4 Residual r_k for the measured MRI Shepp-Logan phantom image

Table 2 Different patches-DPCG methods, results for measured Shepp-Logan

Methods	CG	PCG	Patches-DPCG(4^2)	Patches-DPCG(8^2)	Patches-DPCG(16^2)
Iterations	487	316	198	188	162
Time (s)	0.35	0.25	0.23	0.37	0.88

4 Conclusions

We studied the DPCG method to solve the diffusion equation for image denoising. We used three different ways to construct the deflation vectors. The algorithm is tested on a simulated and a measured image. The deflation method works well for image denoising and the DPCG method converges faster than CG and PCG. Comparing the patch-based DPCG with region growing and thresholding-DPCG, we conclude that patches-DPCG achieves the best convergence and is not sensitive to noise.

Acknowledgments The authors thank the LUMC team for providing the low-field MR image. This work is partly funded by NWO WOTRO under grant W07.303.101 and by the Chinese Scholarship Council.

References

1. F. Catté, P-L. Lions, J-M. Morel, and T. Coll. Image Selective Smoothing and Edge Detection by Nonlinear Diffusion. *SIAM Journal on Numerical Analysis*, 29(1):182–193, 1992.
2. T. Chan and J. Shen. *Image Processing and Analysis: Variational, PDE, Wavelet, and Stochastic Methods*. Society for Industrial and Applied Mathematics Philadelphia, PA, USA, 2005.
3. M.L. de Leeuw den Bouter, M.B. van Gijzen, and R.F. Remis. Conjugate Gradient Variants for L_p-regularized Image Reconstruction in Low-field MRI. *SN Applied Sciences*, 2019.
4. M.R. Hestenes and E. Stiefel. Methods of Conjugate Gradients for Solving Linear Systems. *J. Research Nat. Bur. Standards*, 49:409–436, 1952.
5. R.A. Nicolaides. Deflation of Conjugate Gradients with Applications to Boundary Value Problems. *SIAM Journal on Numerical Analysis*, 24(2):355–365, 1987.
6. T. O'Reilly, W.M. Teeuwisse, and A.G. Webb. Three-dimensional MRI in a Homogenous 27 cm Diameter Bore Halbach Array Magnet. *Journal of Magnetic Resonance*, 307:106578, 08 2019.
7. P. Perona and J. Malik. Scale-space and Edge Detection Using Anisotropic Diffusion. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12(7):629–639, 1990.
8. G. Rohit, D. Lukarski, M.B. van Gijzen, and C. Vuik. Evaluation of the Deflated Preconditioned CG method to solve Bubbly and Porous Media Flow Problems on GPU and CPU. *International Journal for Numerical Methods in Fluids*, 80:666–683, 09 2015.
9. L.A. Shepp and B.F. Logan. The Fourier Reconstruction of a Head Section. *IEEE Transactions on Nuclear Science*, 21(3):21–43, 1974.
10. C. Vuik, A. Segal, and J.A. Meijerink. An Efficient Preconditioned CG Method for the Solution of a Class of Layered Problems with Extreme Contrasts in the Coefficients. *Journal of Computational Physics*, 152(1):385–403, 1999.