

Accelerated graph-based spectral polynomial filters

Knyazev, A.; Malyshev, A.

TR2015-106 September 2015

Abstract

Graph-based spectral denoising is a low-pass filtering using the eigendecomposition of the graph Laplacian matrix of anisotropic signal. Polynomial filtering avoids costly computation of the eigendecomposition by projections onto suitable Krylov subspaces. Polynomial filters can be based, e.g., on the bilateral and guided filters. We propose constructing accelerated polynomial filters by running flexible Krylov subspace based linear and eigenvalue solvers such as the Block Locally Optimal Preconditioned Conjugate Gradient (LOBPCG) method.

2015 IEEE International Workshop on Machine Learning for Signal Processing (MLSP)

This work may not be copied or reproduced in whole or in part for any commercial purpose. Permission to copy in whole or in part without payment of fee is granted for nonprofit educational and research purposes provided that all such whole or partial copies include the following: a notice that such copying is by permission of Mitsubishi Electric Research Laboratories, Inc.; an acknowledgment of the authors and individual contributions to the work; and all applicable portions of the copyright notice. Copying, reproduction, or republishing for any other purpose shall require a license with payment of fee to Mitsubishi Electric Research Laboratories, Inc. All rights reserved.

ACCELERATED GRAPH-BASED SPECTRAL POLYNOMIAL FILTERS

Andrew Knyazev and Alexander Malyshev

Mitsubishi Electric Research Labs (MERL), 201 Broadway, 8th floor, Cambridge, MA 02139, USA.
email: knyazev@merl.com, malyshev@merl.com

ABSTRACT

Graph-based spectral denoising is a low-pass filtering using the eigendecomposition of the graph Laplacian matrix of a noisy signal. Polynomial filtering avoids costly computation of the eigendecomposition by projections onto suitable Krylov subspaces. Polynomial filters can be based, e.g., on the bilateral and guided filters. We propose constructing accelerated polynomial filters by running flexible Krylov subspace based linear and eigenvalue solvers such as the Block Locally Optimal Preconditioned Conjugate Gradient (LOBPCG) method.

Index Terms— Image denoising, spectral polynomial filter, graph Laplacian, Krylov subspace method

1. INTRODUCTION

In this note, we deal with noise removal from a given noisy signal, which is a basic problem in signal processing, with applications, e.g., in image denoising [1]. Apart from the trivial application of removing noise prior to presenting the image to a human observer, pre-smoothing an image and noise removal may help to improve performance of many image-processing algorithms, such as compression, enhancement, segmentation etc. A noise removal operation is often referred to as a filter.

Modern denoising algorithms endeavor to preserve the image details while removing the noise. One of the most popular denoising filters is the bilateral filter (BF), which smooths images while preserving edges, by taking the weighted average of the nearby pixels. The weights depend on both the spatial distance and photometric similarity, thus providing local adaptivity to the input image. Bilateral filtering was introduced in [2, 3, 4] as an intuitive tool without theoretical justification. Since then, connections between the BF and other well-known filtering techniques such as anisotropic diffusion, weighted least squares, Bayesian methods, kernel regression and non-local means have been explored, see, e.g., survey [5].

A convenient way to represent images is by graphs [6], especially when considering images over irregular grids and treating non-local interactions between pixels. A single application of the bilateral filter to an image may be interpreted as

a vertex domain transform on a graph with pixels as vertices, intensity values of the pixels as the graph signal and filter coefficients as link weights that capture the similarity between the vertices. The BF transform is a nonlinear anisotropic diffusion, cf. [7, 8, 9], determined by entries of the graph Laplacian matrix, which are related to the BF weights. In the linear case, the solutions of time-dependent anisotropic diffusion problems are represented by operator exponentials or semigroups, which can be approximated by operator polynomials.

Eigenvalues and eigenvectors of the graph Laplacian matrix allow us to apply the Fourier analysis to the graph signals or images as in [6] and perform frequency selective filtering operations on graphs, similar to those in traditional signal processing. Expensive computation of eigenvectors is avoided in polynomial spectral filters, which are fully implemented in the vertex domain. For example, the spectral filters in [10, 11] are based on optimal polynomials constructed by means of the Chebyshev approximation and conjugate gradient algorithm.

A recent newcomer in the field of filtering is the guided filter (GF) proposed in [12, 13] and included into the image processing toolbox of MATLAB. According to our limited experience, the GF filter is significantly faster than the BF filter. The authors of [12] additionally advocate that GF is gradient preserving and avoids the gradient reversal problems in contrast to BF, which is not gradient preserving.

Iterative application of smoothing filters like BF and GF can be interpreted as matrix power transforms, in general case, nonlinear, or, equivalently as explicit integration in time of the corresponding anisotropic diffusion equation. In the present paper, we continue the approach of [10, 11] and propose acceleration of the iterative smoothing filters based on the polynomial approximations implicitly constructed in the conjugate gradient (CG) algorithm and in the LOBPCG, which is a leading eigensolver for large symmetric matrices [14]. LOBPCG has been successfully applied to image segmentation in [15].

The remainder of the paper is organized as follows. Section 2 gives main formulas of the original bilateral filter in the graph-based notation. Section 3 provides a similar description of the guided filter from [12]. Section 4 introduces a basic Fourier calculus on graphs and applies it to a simplified frequency analysis of iterations with the graph Lapla-

cian matrix. Section 5 describes the CG acceleration of the smoothing filters. Section 6 gives a brief description of the LOBPCG method and suggests it as an acceleration for BF and GF. Our numerical experiments in Section 7 demonstrate improved performance of our accelerated filters in the one-dimensional case mostly for the BF filter. Similar improvements are expected when accelerating the GF filter.

2. BILATERAL FILTER AS A VECTOR TRANSFORM

The bilateral filter transforms an input image x into the output image y by the weighted average of the pixels of x :

$$y_i = \sum_j \frac{w_{ij}}{\sum_j w_{ij}} x_j. \quad (1)$$

Let p_i denote the geometrical position of a pixel i . Then

$$w_{ij} = \exp\left(-\frac{\|p_i - p_j\|^2}{2\sigma_s^2}\right) \exp\left(-\frac{|x_i - x_j|^2}{2\sigma_r^2}\right), \quad (2)$$

where σ_s and σ_r are the filter parameters [4], $\|p_i - p_j\|$ is a spatial distance between pixels i and j . For color images, the photometric distance $|x_i - x_j|$ can be computed in the CIE-Lab color space as suggested in [4].

The BF weights w_{ij} determine an undirected graph $G = (\mathcal{V}, \mathcal{E})$, where the vertices $\mathcal{V} = \{1, 2, \dots, N\}$ are the pixels of the input image and the edges $\mathcal{E} = \{(i, j)\}$ connect the “neighboring” pixels i and j . The adjacency matrix W of the graph G is symmetric and has the entries $w_{ij} \geq 0$. Let D be the diagonal matrix with the nonnegative diagonal entries $d_i = \sum_j w_{ij}$. In the matrix notation, the bilateral filter operation (1) is the vector transform with W depending on a guidance image (we use x as a guidance image in (2))

$$y = D^{-1}Wx = x - D^{-1}Lx, \quad (3)$$

where

$$L = D - W. \quad (4)$$

is the Laplacian matrix for the graph G . Gershgorin’s theorem from matrix analysis guarantees a smoothing effect of the bilateral filter, when $\max_i d_i \leq 2$.

The bilateral filtering can be used iteratively. There are two ways to iterate BF: (1) by changing the weights w_{ij} at each iteration using the result of the previous iteration as a guidance image, or (2) by using the fixed weights at each iteration as calculated from the initial image as a guidance image for all iterations. The first alternative results in a nonlinear filter, where the BF graph changes at every iteration and we can only provide separate spectral interpretation for each stage. The second alternative generates a linear filter, the Laplacian matrix is fixed for all iterations, and it is possible to provide a spectral interpretation of the whole cascaded operation. The

second way is also faster to compute since the BF weights are computed only once in the beginning.

The fastest implementations of BF have the arithmetical complexity $O(N)$, where N is the number of pixels or voxels in the input image [16, 17, 18].

3. GUIDED FILTER AS A VECTOR TRANSFORM

Algorithm 1 Guided Filter

Input: x, g, ρ, ϵ

Output: y

$$\begin{aligned} \text{mean}_g &= f_{\text{mean}}(g, \rho) \\ \text{mean}_x &= f_{\text{mean}}(x, \rho) \\ \text{corr}_g &= f_{\text{mean}}(g \cdot * g, \rho) \\ \text{corr}_{gx} &= f_{\text{mean}}(g \cdot * x, \rho) \\ \text{var}_g &= \text{corr}_g - \text{mean}_g \cdot * \text{mean}_g \\ \text{cov}_{gx} &= \text{corr}_{gx} - \text{mean}_g \cdot * \text{mean}_x \\ a &= \text{cov}_{gx} / (\text{var}_g + \epsilon) \\ b &= \text{mean}_x - a \cdot * \text{mean}_g \\ \text{mean}_a &= f_{\text{mean}}(a, \rho) \\ \text{mean}_b &= f_{\text{mean}}(b, \rho) \\ y &= \text{mean}_a \cdot * g + \text{mean}_b \end{aligned}$$

Algorithm 1 is a pseudo-code of the guided filter from [12], where $f_{\text{mean}}(\cdot, \rho)$ denotes a mean filter with the window width ρ . The constant ϵ determines the smoothness degree of the filter: the larger ϵ the larger a smoothing effect. The arithmetical operations $\cdot *$ and $/$ are the componentwise multiplication and division.

The input image is x , the output image is y . The guidance image g has the same size as x . If g coincides with x , then the guided filter is called self-guided. The arithmetical complexity of GF equals $O(N)$, where N is the number of pixels in x .

The guided filter operation of Algorithm 1 is the transform

$$y = W(g)x, \quad (5)$$

where the implicitly constructed symmetric matrix $W(g)$ has the following entries, see [12]:

$$W_{ij}(g) = \frac{1}{|\omega|^2} \sum_{k: (i,j) \in \omega_k} \left(1 + \frac{(g_i - \mu_k)(g_j - \mu_k)}{\sigma_k^2 + \epsilon}\right). \quad (6)$$

Here ω_k is the neighborhood around pixel k of width ρ , where the mean filter $f_{\text{mean}}(\cdot, \rho)$ is applied, and $|\omega|$ denotes the number of pixels in ω_k , the same for all k . The values μ_k and σ_k^2 are the mean and variance of the image g in ω_k .

The graph Laplacian matrix is obtained from the matrix W in the standard way. According to [12], the values $d_i = \sum_j w_{ij}$ equal 1. Therefore, the diagonal matrix D equals the identity matrix I , and the graph Laplacian matrix is the symmetric nonnegative definite matrix $L = I - W$.

The guided filtering can also be used iteratively. Similar to the BF filter, the guided filter is applied iteratively: (1)

either using the result of the previous iteration as a guidance image, or (2) by using the initial image as a guidance image for all iterations. The first alternative results in a nonlinear GF filter. The second alternative generates a linear GF filter, and the Laplacian matrix remains fixed at each iteration.

4. SPECTRAL INTERPRETATION OF THE LOW-PASS FILTERS

The graph eigenstructure is the eigenvalues and eigenvectors of the graph Laplacian matrix L , which is symmetric and non-negative definite. In certain situations, the normalized Laplacian matrix $(\text{diag } L)^{-1/2} L (\text{diag } L)^{-1/2}$ may be more suitable than L . The spectral factorization of L is the matrix decomposition

$$L = U \Lambda U^T, \quad (7)$$

where the diagonal elements λ_i of the diagonal matrix Λ and columns u_i of the orthogonal matrix $U = [u_1, \dots, u_n]$ are, respectively, the eigenvalues of L and corresponding eigenvectors of the unit 2-norm.

Similar to the classical Fourier transform, the eigenvectors and eigenvalues of the graph Laplacian matrix L provide the oscillatory structure of graph signals. The eigenvalues $\lambda_1 \leq \dots \leq \lambda_i \leq \dots \leq \lambda_n$ can be treated as graph frequencies. The corresponding eigenvectors u_i of the Laplacian matrix L are generalized eigenmodes and demonstrate increasing oscillatory behavior as the magnitude of the graph frequency increases. The Graph Fourier Transform (GFT) of an image x is defined by the matrix transform $\hat{x} = U^T x$, the inverse GFT is the transform $x = U \hat{x}$.

Let us consider the BF vector transform (3). In numerical analysis, the linear transformations $(D^{-1}W)^k$ are called the power iterations with the amplification matrix $D^{-1}W$ or simple iterations for the equation $Lx = 0$ with the preconditioner D . Application of the transform (3) preserves the low frequency components of x and attenuates the high frequency components, cf. [10, 11]. It is also well-known that the Krylov subspaces well approximate the eigenvectors corresponding to the extreme eigenvalues. Thus the projections onto the suitable Krylov subspaces would be an appropriate choice for high- and low-pass filters. The Krylov subspace methods are efficient owing to their low cost, reasonably good convergence and simple implementation without painful parameter tuning. The convergence can be accelerated by the aid of good preconditioners.

5. CG ACCELERATION

Since the graph Laplacian matrix L is symmetric and non-negative definite, the first candidate to replace the simple iteration $x_{k+1} = x_k - D^{-1}Lx_k$ is the preconditioned conjugate gradient (CG) method for the system of homogeneous linear equations $Lx = 0$ with the preconditioner matrix D . To avoid

over-smoothing, only few iterations of the preconditioned CG method must be performed. A large number of iterations for $Lx = 0$ produces piecewise constant images and is better applicable to the image segmentation problems. Good references for the Krylov subspace methods are the books [19, 20].

Algorithm 2 Preconditioned CG(k_{\max}) for $Lx = 0$

Input: L, x_0, k_{\max} , preconditioner D
Output: x
 $x = x_0$
 $r = -Lx$
for $k = 1, \dots, k_{\max}$ **do**
 $s = D^{-1}r$
if $k = 1$ **then**
 $p = s$
else
 $\beta = (s^T(r - r_{old})) / (s_{old}^T s_{old})$
 $p = s + \beta p$
endif
 $q = Lp$
 $\alpha = (s^T r) / (p^T q)$
 $x = x + \alpha p$
 $r_{old} = r$
 $s_{old} = s$
 $r = r - \alpha q$
endfor

Algorithm 2 is a slightly modified standard preconditioned CG algorithm formally applied to the system of linear equations $Lx = 0$. It contains the formula for β that is different from that in the MATLAB implementation of the preconditioned CG. Such a formula converts the algorithm into a flexible variant, which possesses better convergence properties, when the input matrices L and D may change; see, e.g., [21, 22].

The CG iterations are the polynomial filters, i.e., represented in the form $x_k = p_k(L)x_0$, where the coefficients of the polynomial $p_k(\lambda)$ of degree k depend on the matrix L and initial vector x_0 . In the preconditioned case, the filter has the form $x_k = p_k(D^{-1}L)x_0$, and the coefficients of the polynomial $p_k(\lambda)$ depend on $D^{-1}L$ and x_0 .

6. LOBPCG ACCELERATION

A more powerful alternative to the Krylov subspace solvers for linear systems as the polynomial spectral filters are the eigensolvers with preconditioning. Since the graph Laplacian matrix L is symmetric, we propose the symmetric eigensolver LOBPCG with preconditioning for construction of low-pass filters. LOBPCG is a very efficient eigensolver and gives a fast solution to the spectral image segmentation problem, see [15]. To avoid over-smoothing during noise removal, LOBPCG must perform only few iterations.

Algorithm 3 below is a standard non-blocked version of the preconditioned LOBPCG algorithm, which smooths the

input signal x_0 with respect to the eigenmodes of the Laplacian matrix L . If necessary, Algorithm 3 can be modified to obey the constraint that the vectors x_k must be orthogonal to the vector e with all components equal to 1. The vector e is the eigenmode corresponding to the zero eigenvalue of L .

Algorithm 3 LOBPCG method as a low-pass filter

Input: L, D, x_0 , and a preconditioner T

Output: x_{kmax}

$p_0 = 0$

for $k = 0, \dots, k_{max} - 1$ **do**

$$\lambda_k = (x_k^T L x_k) / (x_k^T D x_k)$$

$$r = L x_k - \lambda_k D x_k$$

$$w_k = T r$$

use the Rayleigh-Ritz method for the pencil $L - \lambda D$ on the trial subspace $\text{span}\{w_k, x_k, p_k\}$

$$x_{k+1} = w_k + \tau_k x_k + \gamma_k p_k$$

(the Ritz vector for the minimum Ritz value)

$$p_{k+1} = w_k + \gamma_k p_k$$

endfor

7. NUMERICAL EXPERIMENTS WITH THE KRYLOV SUBSPACE-BASED POLYNOMIAL ACCELERATION OF THE FILTERS

Our experiments have been done in MATLAB. We have compared performance of the bilateral filter versus the CG accelerated BF (BF-CG) filter versus the self-guided GF filter for several one-dimensional signals. The clean 1D signal is piecewise linear with steep ascents and descents. The noisy signal is obtained from the clean signal by adding the Gaussian noise $0.1 \cdot \text{randn}(N, 1)$, where we test $N = 500$ in Figure 1 and $N = 1000$ in Figure 2, where N is number of samples in the signal discretized on a uniform grid.

Figure 1 and Figure 2 display our test results comparing 500 iterations of the self-guided BF with the window width equal to 3, 20 iterations of the self-guided GF with the MATLAB default window width 5, and, finally, 20 iterations of the CG accelerated BF, using the noiseless signal as a guidance and the diagonal matrix D as the preconditioner.

In Figures 1 and 2, we observe high quality denoising by all tested filters, both in terms of the average fit of the noiseless signal and of preserving the signal details. We notice that GF has a tendency to round sharp corners compared to BF and BF-CG, which is not surprising, since GF performs extensive smoothing on relatively wider neighborhoods. Interestingly, GF error is strongly discontinuous on the signal edges, while BF and BF-CG provide more conservative approximations. Increasing the number of samples apparently makes the BF and BF-CG errors smoother as can be seen comparing Figure 1 and Figure 2.

Miraculously all three filters give nearly the same error with the reported parameters, which requires further investigation and explanation.

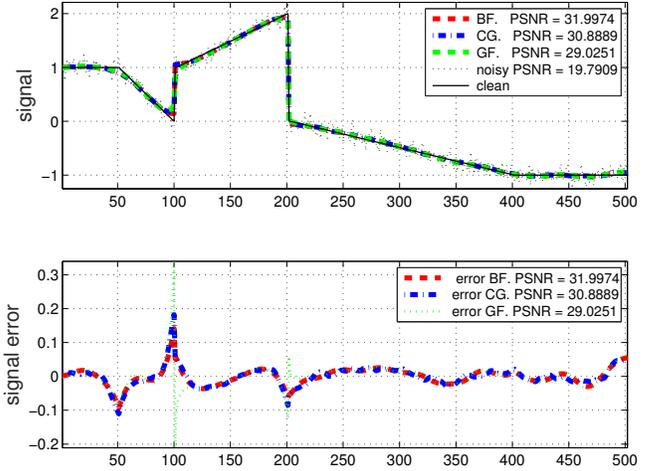


Fig. 1. Filters in 1D case: 500 points

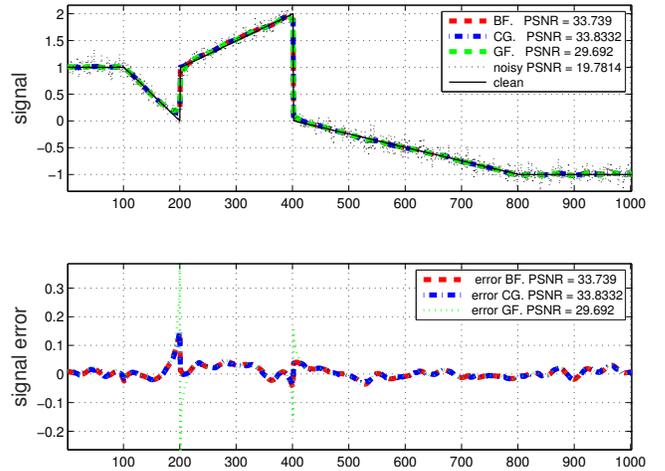


Fig. 2. Filters in 1D case: 1000 points

Our second series of tests deals with image denoising. We use the 5-point stencil to generate the lattice for the graph Laplacian matrix and the edge weights are determined by the BF filter with $\sigma_r = 0.1$.

Twenty iterations are performed using the CG and LOBPCG accelerated filters with and without the constraint. The quality of denoising is similar improving PSNR from approximately 20 for the noisy image to approximately 21 for the filtered images. The edges are well preserved but one can notice forming salt and pepper noise, which gets more extensive if the number of iterations is increased. Figure 4 compares the BF filter with the window half-width 5 and bilateral filter with the standard deviation $\sigma = 0.1$.



Fig. 3. Noisy versus clear image: gaussian noise, mean = 0, variance = 0.01, PSNR = 20.1



Fig. 4. BF filter versus CG

Conclusions

We propose accelerating iterative smoothing filters such as the bilateral and guided filters, by using the polynomials, generated by the conjugate gradient-type iterative solvers for linear systems and eigenvalue problems. This results in efficient parameter-free polynomial filters in the Krylov subspaces that approximate the spectral filters corresponding to the graph Laplacians, which are constructed using the guidance signals. Our numerical tests for one-dimensional signals demonstrate the typical behavior of the proposed accelerated filters and explain the motivations behind the construction. The tests



Fig. 5. LOBPCG with versus without the constraint



Fig. 6. LOBPCG with the constraint versus CG

showing image denoising reveal that the proposed filters are competitive. Our future work will concern accelerating the guided filters, testing the influence of non-linearities on the filter behavior, and incorporating efficient preconditioners.

8. REFERENCES

- [1] P. Milanfar, “A tour of modern image filtering: new insights and methods, both practical and theoretical,” *IEEE Signal Processing Magazine*, vol. 30, no. 1, pp. 106–128, 2013.
- [2] V. Aurich and J. Weule, “Non-linear gaussian filters performing edge preserving diffusion,” in *Mustererkennung 1995, 17. DAGM-Symposium*, Bielefeld, Germany, 1995, pp. 538–545.
- [3] S. M. Smith and J. M. Brady, “SUSAN — A new approach to low level image processing,” *Int. J. Comput. Vision*, vol. 23, no. 1, pp. 45–78, 1997.
- [4] C. Tomasi and R. Manduchi, “Bilateral filtering for gray and color images,” in *Proc. IEEE International Conference on Computer Vision*, Bombay, 1998, pp. 839–846.
- [5] S. Paris, P. Kornprobst, J. Tumblin, and F. Durand, “Bilateral filtering: Theory and applications,” *Foundations and Trends in Computer Graphics*, vol. 4, no. 1, pp. 1–73, 2009.
- [6] D.I. Shuman, S.K. Narang, P. Frossard, A. Ortega, and P. Vandergheynst, “The emerging field of signal processing on graphs: Extending high-dimensional data analysis to networks and other irregular domains,” *IEEE Signal Processing Magazine*, vol. 30, no. 3, pp. 83–98, 2013.
- [7] P. Perona and J. Malik, “Scale-space and edge detection using anisotropic diffusion,” in *Proceedings of IEEE Computer Society Workshop on Computer Vision*, Miami, FL, 1987, pp. 16–27.

- [8] P. Perona and J. Malik, "Scale-space and edge detection using anisotropic diffusion," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 12, no. 7, pp. 629–639, 1990.
- [9] D. Barash, "Fundamental relationship between bilateral filtering, adaptive smoothing, and the nonlinear diffusion equation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 6, pp. 844–847, 2002.
- [10] A. Gadde, S.K. Narang, and A. Ortega, "Bilateral filter: Graph spectral interpretation and extensions," in *Proceedings of the 20th IEEE International Conference on Image Processing (ICIP)*, Melbourne, Australia, 2013, pp. 1222–1226.
- [11] D. Tian, A. Knyazev, H. Mansour, and A. Vetro, "Chebyshev and conjugate gradient filters for graph image denoising," in *Proceedings of IEEE International Conference on Multimedia and Expo Workshops (ICMEW)*, Chengdu, 2014, pp. 1–6.
- [12] K. He, J. Sun, and X. Tang, "Guided image filtering," *IEEE transactions on pattern analysis and machine intelligence*, vol. 35, no. 6, pp. 1397–1409, 2013.
- [13] K. He and J. Sun, "Fast guided filter," Tech. Rep., 2015, arXiv:1505.00996v1.
- [14] A. Knyazev, "Toward the optimal preconditioned eigensolver: locally optimal block preconditioned conjugate gradient method," *SIAM J. Scientific Computing*, vol. 23, no. 2, pp. 517–541, 2001.
- [15] A. Knyazev, "Modern preconditioned eigensolvers for spectral image segmentation and graph bisection," in *Proc. 3rd International Conf. on Data Mining (ICDM)*, Melbourne, Australia, 2003, pp. 1–4.
- [16] Q. Yang, K.-H. Tan, and N. Ahuja, "Real-time $O(1)$ bilateral filtering," in *IEEE Conf. Computer Vision Pattern Recognition (CVPR)*, Miami, FL, 2009, pp. 557–564.
- [17] K. N. Chaudhury, "Acceleration of the shiftable $O(1)$ algorithm for bilateral filtering and nonlocal means," *IEEE Transactions on Image Processing*, vol. 22, no. 4, pp. 1291–1300, 2013.
- [18] K. Sugimoto and S.-I. Kamata, "Compressive bilateral filtering," *IEEE Transactions on Image Processing*, vol. 24, no. 11, pp. 3357–3369, 2015.
- [19] A. Greenbaum, *Iterative Methods for Solving Linear Systems*, SIAM, Philadelphia, PA, 1997.
- [20] H. A. van der Vorst, *Iterative Krylov Methods for Large Linear Systems*, Cambridge University Press, Cambridge, UK, 2003.
- [21] A. V. Knyazev and I. Lashuk, "Steepest descent and conjugate gradient methods with variable preconditioning," *SIAM J. Matrix Analysis and Applications*, vol. 29, no. 4, pp. 1267–1280, 2008.
- [22] H. Bouwmeester, A. Dougherty, and A. V. Knyazev, "Nonsymmetric multigrid preconditioning for conjugate gradient methods," Tech. Rep., 2012, arXiv:1212.6680.