MITSUBISHI ELECTRIC RESEARCH LABORATORIES
Cambridge, Massachusetts

Accelerated graph-based spectral polynomial filters

Andrew Knyazev (knyazev@merl.com)
Alexander Malyshev (malyshev@merl.com)

MLSP2015
IEEE International Workshop on
MACHINE LEARNING FOR SIGNAL PROCESSING
September 17-20, 2015 **Boston, USA**

Contents:

- Bilateral filter (BF)
- Guided filter (GF)
- Spectral properties of the iterated BF and GF
- Preconditioned conjugate gradient (PCG) acceleration
- Locally optimal block preconditioned conjugate gradient (LOBPCG) acceleration
- Numerical results

Bilateral filter (BF)

A discrete function $x[j]$, $j \in \{1, 2, \dots, N\}$, is an input signal for the bilateral filter. The output signal $y[i]$ is a weighted average of the signal values $x[j]$:

$$y[i] = \sum_j \frac{w_{ij}}{\sum_j w_{ij}} x[j].$$

Every index i has a spatial position p_i, and a spatial distance $\|p_i - p_j\|$ is determined for all pairs i and j.
The weights w_{ij} are defined by means of a guidance signal $g[i]$:

$$w_{ij} = \exp\left(-\frac{\|p_i - p_j\|^2}{2\sigma_d^2}\right) \exp\left(-\frac{(g[i] - g[j])^2}{2\sigma_r^2}\right), \qquad (1)$$

where σ_d and σ_r are the filter parameters. When $g = x$, the bilateral filter is nonlinear and called self-guided.

Iterated bilateral filter

- The weights w_{ij} are the entries of a symmetric nonnegative matrix W. Let us denote by D the diagonal matrix with the positive diagonal entries $d_i = \sum_j w_{ij}$. Then BF is the vector transform $y = D^{-1}Wx$. The symmetric nonnegative defined matrix $L = D - W$ is referred as to a Laplacian matrix. The spectrum of $D^{-1}W$ is real, and the eigenvalues corresponding to the highest oscillations lie near 0.
- The BF transform $y = D^{-1}Wx$ can be applied iteratively,
  1. by changing the weights w_{ij} at each iteration using the result of the previous iteration as a guidance signal g, or
  2. by using the fixed weights, calculated from the initial signal as a guidance signal, for all iterations.

  The former results in a nonlinear filter, the latter generates a linear filter, which may be faster, since the BF weights are computed only once at the very beginning.

Guided filter (GF)

Algorithm 1 Guided Filter (GF)

Input: x, g, w, ϵ

Output: y

$$mean_g = f_{mean}(g, w)$$
$$mean_x = f_{mean}(x, w)$$
$$corr_g = f_{mean}(g. * g, w)$$
$$corr_{gx} = f_{mean}(g. * x, w)$$
$$var_g = corr_g - mean_g. * mean_g$$
$$cov_{gx} = corr_{gx} - mean_g. * mean_x$$
$$a = cov_{gx}./(var_g + \epsilon)$$
$$b = mean_x - a. * mean_g$$
$$mean_a = f_{mean}(a, w)$$
$$mean_b = f_{mean}(b, w)$$
$$y = mean_a. * g + mean_b$$

Guided filter (GF)

$f_{mean}(\cdot, w)$ is a mean filter with the window width w. The constant ϵ determines the smoothness degree: the larger ϵ the larger smoothing effect. The dot preceded operations $.*$ and $./$ denote the componentwise multiplication and division. Arithmetical complexity of the GF algorithm can be $O(N)$.

GF is $y = Wx$, where the entries of the symmetric matrix $W(g)$ are

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

The windows ω_k of width w around all k have the number of pixels $|\omega|$.
The values μ_k and σ_k^2 are the mean and variance of g over ω_k.
Since $d_i = \sum_j w_{ij} = 1$, the graph Laplacian matrix equals $L = I - W$.
The eigenvalues of $L(g)$ are real nonnegative with the low frequencies accumulated near 0 and high frequencies near 1. Similar to BF, the guided filter can be applied iteratively.

Spectral properties of the graph-based filters

The spectral factorization of the symmetric nonnegative graph Laplacian

$$L = U \Lambda U^T$$

is determined by the diagonal matrix Λ with the diagonals λ_i and the orthogonal matrix $U = [u_1, \ldots, u_n]$ with the columns u_i.
The eigenvalues $\lambda_1 \leq \cdots \leq \lambda_i \leq \cdots \leq \lambda_n$ 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 $\widehat{x} = U^T x$, the inverse GFT is the transform $x = U\widehat{x}$.

Spectral properties of the iterated low-pass filters

Let us consider the iterated BF or GF vector transforms. 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 $D^{-1}W$ preserves the low frequency components of x and attenuates the high frequency components. 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.

Preconditioned Conjugate Gradient acceleration of a smoothing filter

Algorithm PCG($k_{\max}$) with $l_{\max}$ restarts

Input: x_0, $k_{\max}$, l_{max}
Output: x
$x = x_0$
for $l = 1, \ldots, l_{\max}$ **do**
  $r = W(x)x - D(x)x$
  for $k = 1, \ldots, k_{\max} - 1$ **do**
    $s = D^{-1}(x)r$; $\gamma = s^T r$
    if $k = 1$ **then** $p = s$ **else** $\beta = \gamma/\gamma_{old}$; $p = s + \beta p$ **endif**
    $q = D(x)p - W(x)p$; $\alpha = \gamma/(p^T q)$
    $x = x + \alpha p$; $r = r - \alpha q$; $\gamma_{old} = \gamma$
  endfor
endfor

Locally Optimal Block Preconditioned Conjugate Gradient acceleration of a smoothing filter

Algorithm 3 LOBPCG method

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

Output: x_{kmax}

$p_0 = 0$

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

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

$r = Lx_k - \lambda_k D x_k$

$w_k = Tr$

use the Rayleigh-Ritz method for the pencil $L - \lambda D$

on the trial subspace $span\{w_k, x_k, p_k\}$

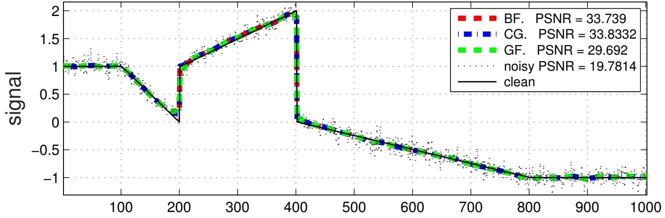
$x_{k+1} = w_k + \tau_k x_k + \gamma_k p_k$

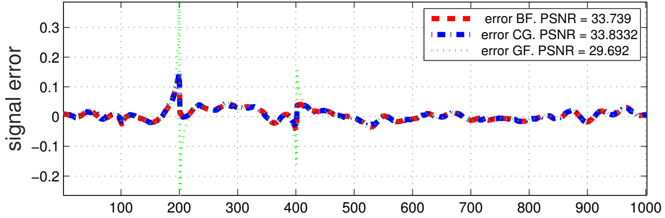
(the Ritz vector for the minimum Ritz value)

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

endfor

```
noisy = clean + randn(size(clean))*0.1
```





BF versus GF

LOBPCG versus PCG

Conclusions

- The bilateral and guided filters are written in the guided form.
- PCG acceleration can be applied to iterated smoothing filters in the guided form, formally solving $Lx = 0$ by PCG.
- LOBPCG acceleration can be also applied to iterated filters.
- PCG and LOBPCG algorithms considerably accelerate low-pass filtering without quality degradation.

References I

C. Tomasi and R. Manduchi, Bilateral filtering for gray and color images, in *Proc. IEEE International Conf. Comp. Vision*, Bombay, 1998, pp. 839–846.

J. Sun, K. He and X. Tang, Guided image filtering, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 35(6):1397–1409, 2013.

D. Tian, A. Knyazev, H. Mansour, and A. Vetro, Chebyshev and conjugate gradient filters for graph image denoising, in *Proc. IEEE International Conf. Multimedia and Expo Workshops (ICMEW)*, Chengdu, 2014, pp. 1–6.

References II

A. Knyazev, Toward the optimal preconditioned eigensolver: locally optimal block preconditioned conjugate gradient method, *SIAM J. Scientific Comp.*, 23(2):517–541, 2001.