ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

Global Image Denoising

B.Sreenuvasulu
PG Scholar,
Department of ECE
S.S.C.E.T, Engineering College, AP, India.

M.Janakidevi
Associate Professor,
Department of ECE
S.S.C.E.T, Engineering College, AP, India.

Abstract:

Mostly the shortcoming of the patch based methods by developing a paradigm for accurately global filtering where each pixel is estimated from all pixels in the image. The two objectives are- firstly, we furnish a statistical analysis of the implemented global filter based on the spectral decomposition of the corresponding operator and have to find out the effect of truncation of it, secondly, develop an approximation to the spectral components using the Nystrom extension. Using these two objectives, illustrate that this global filter can be implemented well by sampling a comparatively small percentage of the pixels in the image. Implemented system show that our approach can successfully globalize any existing denoising filters to approximate each pixel by means of all pixels in the image, consequently improving the best patch-based methods.

1.1 Introduction

Digital image processing is electronic data processing on a 2-D array of numbers. The array is a numeric representation of an image. A real image is formed on a sensor when an energy emission strikes the sensor with sufficient intensity to create a sensor output. An image may be defined as a two-dimensional function, f(x, y), where x and y are spatial (plane) coordinates, and the amplitude of f at any pair of coordinates (x, y) is called the intensity or gray level of the image at that point. When x, y, and the amplitude values of f are all finite, discrete quantities, we call the image a digital image. The field of digital image processing refers to processing digital images by means of a digital computer. A digital image is composed of a finite number of elements, each of which has a particular location and value. These elements are referred to as picture elements, image elements, pels, and pixels. Pixel is most widely used to denote the elements of a digital image.Images play the single most important role in human perception. Humans are limited to the visual band of the electromagnetic (EM) spectrum, imaging machines cover almost the entire EM spectrum, ranging from gamma to radio waves. They can operate on images generated by sources that humans are not accustomed to associating with images.

These include ultrasound, electron microscopy, and computer-generated images. Thus, digital image processing encompasses a wide and varied field of applications.Digital image processing is the use of computer algorithms to perform image processing on digital images. Digital image processing has the same advantages over analog image processing as digital signal processing has over analog signal processing it allows a much wider range of algorithms to be applied to the input data, and can avoid problems such as the build-up of noise and signal distortion during processing . Image processing is a subclass of signal processing concerned specifically with pictures. Improve image quality for human perception and/or computer interpretation.

1.2 Fundamental Steps in Digital Image Processing

1.2.1 Image Acquisition

An image is captured by a sensor (such as a monochrome or color TV camera) and digitized. If the output of the camera or sensor is not already in digital form, an analog-to-digital converter digitizes it. Generally, the image acquisition stage involves preprocessing, such as scaling.

1.2.2 Image Enhancement
Image enhancement is among the simplest and most appealing areas of digital image processing. Basically, the idea behind enhancement techniques is to bring out detail that is obscured, or simply to highlight certain features of interest in an image. A familiar example of enhancement is when we increase the contrast of an image because “it looks better.” To bring out detail is obscured, or simply to highlight certain features of interest in an image.

1.2.3 Image Restoration
Image restoration is an area that also deals with improving the appearance of an image. Image restoration is objective, in the sense that restoration techniques tend to be based on mathematical or probabilistic models of image degradation. Enhancement, on the other hand, is based on human subjective preferences regarding what constitutes a “good” enhancement result. Improving the appearance of an image tend to be based on mathematical or probabilistic models of image degradation.

1.2.4 Image Compression
Image Compression deals with techniques for reducing the storage required saving an image, or the bandwidth required transmitting it. Although storage technology has improved significantly over the past decade, the same cannot be said for transmission capacity. This is true particularly in uses of the Internet, which are characterized by significant pictorial content. Image compression is familiar to most users of computers in the form of image file extensions, such as the jpg file extension used in the JPEG (Joint Photographic Experts Group) image compression standard.

1.2.5 Image Segmentation
Image segmentation is the process of partitioning a digital image into multiple segments (sets of pixels, also known as super pixels). The goal of segmentation is to simplify and/or change the representation of an image into something that is more meaningful and easier to analyze. Image segmentation is typically used to locate objects and boundaries (lines, curves, etc.) in images. More precisely, image segmentation is the process of assigning a label to every pixel in an image such that pixels with the same label share certain visual characteristics.

The result of image segmentation is a set of segments that collectively cover the entire image, or a set of contours extracted from the image (see edge detection). Each of the pixels in a region are similar with respect to some characteristic or computed property, such as color, intensity, or texture. Adjacent regions are significantly different with respect to the same characteristic(s). When applied to a stack of images, typical in medical imaging, the resulting contours after image segmentation can be used to create 3D reconstructions with the help of interpolation algorithms like marching cubes.Edge detection is a well-developed field on its own within image processing. Region boundaries and edges are closely related, since there is often a sharp adjustment in intensity at the region boundaries. Edge detection techniques have therefore been used as the base of another segmentation technique. The edges identified by edge detection are often disconnected. To segment an object from an image however, one needs closed region boundaries. The desired edges are the boundaries between such objects. Segmentation methods can also be applied to edges obtained from edge detectors. Lindeberg and Li developed an integrated method that segments edges into straight and curved edge segments for parts-based object recognition, based on a minimum description length (MDL) criterion that was optimized by a split-and-merge-like method with candidate breakpoint obtained from complementary junction cues to obtain more likely points at which to consider partitions into different segment.

4.1 Global image denoising
Most on hand up to date image denoising algorithms are based on exploiting similarity between a relatively modest number of patches. These patch-based methods

ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

are decisively reliant on patch matching, and their performance is controlled by the capability to consistently determine adequately similar patches. As the number of patches increases, a point of retreating proceeds is reached where the performance improvement due to more patches is offset by the lower likelihood of ending sufficiently close matches. The complete effect is that while patch-based methods, such as BM3D, are excellent overall, they are ultimately limited in how well they can do on (larger) images with increasing complexity. In this proposed method, first, we give a statistical analysis of our proposed global filter, based on a spectral decomposition of its corresponding operator, and we study the effect of truncation of this spectral decomposition. Second, we derive an approximation to the spectral (principal) components using the Nyström extension. Using these, we demonstrate that this global filter can be implemented efficiently by sampling a fairly small percentage of the pixels in the image.The Nyström method was originally introduced as a method for finding mathematical solution to Eigen decomposition problems. The Nyström extension has been helpful for different applications such as image editing [17], manifold learning [15], and image segmentation [16]. Conveniently, in our global filtering structure, the filter matrix is not a full-rank local filter and thus can be strictly approximated with a low-rank matrix with the help of the Nyström method.Our effort to this line of research is to bring in a new global denoising filter, which takes into relation all informative parts of an image. Definitely, with this global filter, the idea of patch-based processing is no longer limiting, and we are capable to show that performances of the on hand patch-based filters are improved.The block diagram of the proposed global image denoising (GLIDE) construction is illustrated in Fig 1. As it can be seen, after applying a pre-filter on the noisy image, a small fraction of the pixels are sampled to be fed to the Nyström method. Then, the global filter is approximated through its Eigen values and eigenvectors. The final estimate of the image is constructed through reduction of the filter Eigen values.

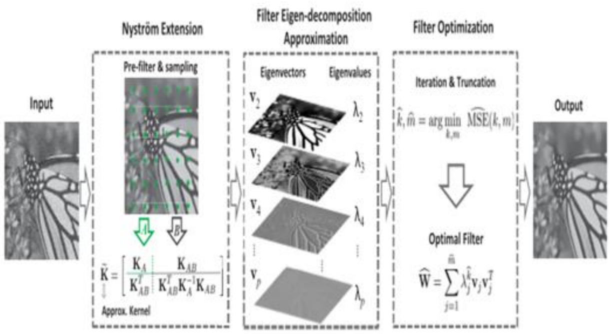

Fig 4.1: GLIDE's pipeline

4.2 Statistical analysis of global filter

With z_i representing the i^{th} underlying pixel, our measurement model for the denoising problem is:

$y_i = z_i + e_i$, for $i = 1,\ldots,n$, --- (4.1)

where y_i is the noisy pixel value and e_i denotes the additive noise. The vectorized measurement model for recovering the underlying pixels $z = [z_1,\ldots,z_n]^T$ is given by (15). Most spatial domain filters can be represented through the following non-parametric restoration framework [2], [6]:

$\hat{z}_i = \arg\max_{z_i} \sum_{j=1}^{n} [z_i - y_j]^2 K_{ij}$ -- **(4.2)**

where the kernel function K_{ij} measures the similarity between the samples y_i and y_j, and $\hat{z}_i$ denotes the i^{th} estimated pixel.Minimizing equation (16) gives a normalized weighted averaging process in which some data-adaptive weights are assigned to each pixel:

$\hat{z}_i = w_i^T y$,-- (4.3)

where the weight vector w_i is

$w_i = \frac{1}{\sum_{j=1}^{n} K_{ij}} [K_{i1}, K_{i2},\ldots,K_{in}]^T$------------------------- (4.4)

in which $[K_{i1}, K_{i2},\ldots,K_{in}]$ denotes the i^{th} row of the symmetric kernel matrix K. The filtering process for all the pixels can be represented by stacking the weight vectors together:

ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

^z=Wy = VSV^{T}y-- (4.5)

Where W=[w_1^T w_2^T ……w_n^T]^T ,called positive, row-stochastic filter matrix W is used to estimate denoisined signal ^z. The eigenvectors V =[v_1,v_2,.,v_n] specify a complete orthonormal basis for R^n and S = diag[λ1,.,λn] contains the eigenvalues in decreasing order $0 \le \lambda_n \le \cdots < \lambda_1 = 1$. This implies that the image y is first projected onto the eigenvectors of W, then each mode of the projected signal is reduced by its equivalent eigenvalue, and finally after mapping back to the signal domain, the recovered signal ^z is produced.The computational burden of constructing and decomposing such a large matrix as W is prohibitively high. However, the Nyström approximation, combined with our statistical analysis allows a resourceful solution. Before proceeding to the filter approximation, the behavior of the filter (in terms of MSE) is analyzed.

MSE of the image

From (16) we can show that each row of W can be expressed as:

$$\mathbf{w}_i^T = \sum_{j=1}^{n} \lambda_j \mathbf{v}_j(i) \mathbf{v}_j^T, \quad \text{-------------------------------------} \textbf{(4.6)}$$

where $v_j(i)$ denotes the i-th entry of the j-th eigenvector. Then each estimated pixel ^z_i has the following form:

$$\widehat{z}_i = \sum_{j=1}^{n} \lambda_j \mathbf{v}_j(i) \mathbf{v}_j^T \mathbf{y}, \quad \text{-------------------------------------} \textbf{(4.7)}$$

Bias of this estimate can be expressed as:

$$\text{bias}(\widehat{z}_i) = z_i - \text{E}(\widehat{z}_i) = \sum_{j=1}^{n} \mathbf{v}_j(i)\mathbf{v}_j^T \mathbf{z} - \sum_{j=1}^{n} \lambda_j \mathbf{v}_j(i) \mathbf{v}_j^T \mathbf{z}$$

$$= \sum_{j=1}^{n} (1-\lambda_j)\mathbf{v}_j(i) b_j \quad \text{------------------------} \textbf{(4.8)}$$

where b =V^T z=[b1,...,bn]^T contains the projected signal in all modes. The variance term also has the following form:

$$\text{var}(\widehat{z}_i) = \sigma^2(\mathbf{w}_i^T \mathbf{w}_i) = \sigma^2 \Big(\sum_{j=1}^{n} \lambda_j \mathbf{v}_j(i) \mathbf{v}_j^T\Big)\Big(\sum_{j'=1}^{n} \lambda_{j'} \mathbf{v}_{j'}(i) \mathbf{v}_{j'}\Big)$$

$$= \sigma^2 \sum_{j=1}^{n} \lambda_j^2 \mathbf{v}_j(i)^2 \quad \text{--------------------} \textbf{(4.9)}$$

where in the last equation we have v^T_j $v_{j'}$ = $\delta_{jj'}$ for the orthonormal basis functions. Overall, the MSE of the i-th estimated pixel is:

$$\text{MSE}_i = \text{bias}(\widehat{z}_i)^2 + \text{var}(\widehat{z}_i)$$

$$= \Big(\sum_{j=1}^{n} (1-\lambda_j) \mathbf{v}_j(i) b_j\Big)^2 + \sigma^2 \sum_{j=1}^{n} \lambda_j^2 \mathbf{v}_j(i)^2. \quad \text{------} \textbf{(4.10)}$$

This expression can be used to analyze the framework given in (4.7). The estimated MSE of the whole image is given by:

$$\text{MSE} = \sum_{i=1}^{n} \text{MSE}_i = \sum_{i=1}^{n} \text{bias}(\widehat{z}_i)^2 + \sum_{i=1}^{n} \text{var}(\widehat{z}_i) \quad \text{----------} \textbf{(4.11)}$$

Reminding the orthonormality of the eigenvectors V^T V = VV^T = I, the variance term can be written as:

$$\sum_{i=1}^{n} \text{var}(\widehat{z}_i) = \sum_{i=1}^{n} \sigma^2 \sum_{j=1}^{n} \lambda_j^2 \mathbf{v}_j(i)^2$$

$$= \sigma^2 \sum_{j=1}^{n} \lambda_j^2 \sum_{i=1}^{n} \mathbf{v}_j(i)^2 = \sigma^2 \sum_{j=1}^{n} \lambda_j^2 \quad \text{------------} \textbf{(4.12)}$$

where in the last equation $\sum_{i=1}^{n} \mathbf{v}_l(i)\mathbf{v}_j(i) = 0$ with l ≠ j.

ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

From what we have for the squared bias and variance we caconclude:

$$\mathrm{MSE} = \sum_{j=1}^{n}(1-\lambda_j)^2 b_j^2 + \sigma^2\lambda_j^2. \quad \text{--------------------}$$

(4.13)

Apparently MSE is a function of the latent signal, noise, filter eigen values and eigenvectors. The filter eigen values are the shrinkage factors which directly tune the filtering performance.

4.3 Filter Approximation

Our objective is different because in the end we need to approximate the filtering matrix W, hence we first review what is done in [16] and then adapt it to the approximation we need to effect here. In the following, the Nyström approach for approximating the similarity (affinity) matrix K is used first and then the Sinkhorn method is applied to estimate the eigen-decomposition of the symmetric, doubly-stochastic filter W. Since the approximated eigenvectors are not exactly orthog- onal, finally an orthogonalization procedure is employed to obtain an orthonormal approximation for eigen-decomposition of W. These steps are shown in Algorithm 1 and we will discuss them in more details below:

4.3.1 Nystrom Approximation

This method is a numerical approximation for estimating the eigenvectors of the symmetric kernel matrix

$$\mathbf{K} = \mathbf{\Phi}\mathbf{\Pi}\mathbf{\Phi}^T \quad \text{--}$$

----- **(4.14)**

where $\mathbf{\Phi} = [\boldsymbol{\phi}_1, \cdots, \boldsymbol{\phi}_n]$ represents the orthonormal eigenvectors and $\mathbf{\Pi} = [\pi_1, \pi_2, \ldots, \pi_n]$ contains the eigen values of K. Nyström [13] suggests that instead of computing all the entries of K, we can sample our data points and estimate the leading eigenvectors of the matrix K and, as a result, an approximation ˜K can then be built from those estimated eigenvectors.Having p pixels in a sampled subimage A, we can compute the p × p kernel matrix K_A which represents the similarity weights of pixels in A. We also define the subimage B containing the rest of (n−p) pixels, followed by the p×(n−p) matrix K_{AB}, which contains the kernel weights between pixels in A and B. The similarity matrix K in block form is therefore:

$$\mathbf{K} = \begin{bmatrix} \mathbf{K}_A & \mathbf{K}_{AB} \\ \mathbf{K}_{AB}^T & \mathbf{K}_B \end{bmatrix} \quad \text{--------------------------------------}$$

----- **(4.15)**

where K_B denotes the (n − p)×(n − p) similarity weights between pixels in the subimage B. As can be seen, (4.3) can be thought of as a permutation of the old K. Nyström suggests the following approximation for the first p eigenvectors of K:

$$\tilde{\mathbf{\Phi}} = \begin{bmatrix} \mathbf{\Phi}_A \\ \mathbf{K}_{AB}^T\mathbf{\Phi}_A\mathbf{\Pi}_A^{-1} \end{bmatrix} \quad \text{------------------------------------}$$

--- **(4.16)**

where $\mathbf{K}_A = \mathbf{\Phi}_A\mathbf{\Pi}_A\mathbf{\Phi}_A^T$. Intuitively, we can say that the first p entries of $\mathbf{\Phi}$ are computed exactly, and the (n − p) remaining ones are approximated by a weighted projection of K_{AB} over the eigenvectors of K_A. Then the approximated similarity matrix will be:

$$\begin{aligned}\tilde{\mathbf{K}} &= \tilde{\mathbf{\Phi}}\mathbf{\Pi}_A\tilde{\mathbf{\Phi}}^T \\ &= \begin{bmatrix} \mathbf{\Phi}_A \\ \mathbf{K}_{AB}^T\mathbf{\Phi}_A\mathbf{\Pi}_A^{-1} \end{bmatrix}\mathbf{\Pi}_A\left[\,\mathbf{\Phi}_A^T \;\; \mathbf{\Pi}_A^{-1}\mathbf{\Phi}_A^T\mathbf{K}_{AB}\,\right] \\ &= \begin{bmatrix} \mathbf{K}_A & \mathbf{K}_{AB} \\ \mathbf{K}_{AB}^T & \mathbf{K}_{AB}^T\mathbf{K}_A^{-1}\mathbf{K}_{AB} \end{bmatrix} \quad \text{----------}\end{aligned}$$

(4.17)

Comparing (4.15) and (4.17) it can be seen that the huge matrix K_B is approximated by $\mathbf{K}_{AB}^T\mathbf{K}_A^{-1}\mathbf{K}_{AB}$.

A key aspect of the Nyström approximation is the sampling procedure in which the columns (or rows) of the original K are selected. The Nyström method was first introduced by a uniform distribution sampling over data [12]. Efficiency of the uniform sampling has been explored in many practical applications [15], [16]. More recently, theoretical aspects of nonuniform

ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

sampling techniques on real-world data sets have been studied [24], [25]. In general, these nonuniform sampling procedures are biased toward selection of the most informative points of the data-set. However, due to the imposed complexity of the nonuniform distribution updating procedure, practical application of these adaptive methods is limited. In the current framework, our data are images which contain a high degree of spatial correlation between pixels. This leads us to use spatially uniform sampling instead of the random sampling procedure. Spatially uniform sampling is a simple but effective approach in which the spatial distance of the samples are always equally fixed.To study the performance of the Nyström approximation, we evaluate the relative accuracy defined in [25]:

Relative Accuracy= $\frac{\|\mathbf{K}-\mathbf{K}_{(r)}\|_F}{\|\mathbf{K}-\tilde{\mathbf{K}}_{(r)}\|_F} \times 100$ --------------------- **(4.18)**

where K and $K_{(r)}$ are the actual kernel and its exact rank-r approximation. The approximated kernel $\tilde{K}_{(r)}$ is reconstructed by using r leading eigenvectors from the Nyström method. The relative accuracy is lower bounded by zero and will ideally approach 100%. We fixed r = 50 to capture about 90% of the spectral energy of the global kernel for each image. The samples are uniformly selected over the image lattice, and the relative accuracy is averaged for 20 sampling realizations. It can be seen that while higher sampling percentage leads to smaller error in the approximated kernel matrix, a saturation point is reached beyond 20% sampling density. Furthermore, for a fixed sampling rate the error depends on the contents of the underlying image. Surprisingly, textured images with high frequency components such as Mandrill produce less error compared to smooth images like House. This observation is consistent with results of [26] where it is shown that the error of the Nyström approximation is proportional to coherency of the kernel eigenvectorsOne could assume that at this point we can easily compute our approximated W and we are done! But as discussed earlier, statistical analysis of this filter needs access to its eigen- decomposition. Constructing a huge W matrix and then computing its eigenvectors is too expensive. Instead, in the following we explore an efficient way to find the eigenvectors of W.

4.3.2 Sinkhorn

The filter W is the row-normalized kernel matrix K:

$\mathbf{W} = \mathbf{D}^{-1}\mathbf{K}$ --- **(4.19)**

where $\mathbf{D} = \mathrm{diag}[\sum_{j=1}^{n} K_{1j}, \sum_{j=1}^{n} K_{2j}, .., \sum_{j=1}^{n} K_{nj}]$.

We approximate the matrix W with a doubly-stochastic (symmetric) positive definite matrix, using Sinkhorn's algorithm [20]. Based on this method, given a positive valued matrix K, there exist diagonal matrices R = diag(r) and C = diag(c) such that Wsym = RKC. Since we have estimated the leading eigenvectors of K, there is no need to compute RKC. Instead, as can be seen in Algorithm. 1, Wsym is approximated by its two sub-blocks W_A and W_{AB} where:

$\mathbf{W}_{sym} = \begin{bmatrix} \mathbf{W}_A & \mathbf{W}_{AB} \\ \mathbf{W}_{AB}^T & \mathbf{W}_B \end{bmatrix}$ -------------------------------- **(4.20)**

Again, the Nyström method could give the approximated eigenvectors, but the only minor problem is that these eigen- vectors are not quite orthogonal. In the following we discuss an approximation of the orthogonal eigenvectors.

BIBLIOGRAPHY

[1] R. Manduchi and C. Tomasi , "Bilateral filtering for gray and color images," in Proc. 6th ICCV, Jan. 1998, pp. 836–846.

[2] P. Milanfar, H. Takeda, and S. Farsiu, "Kernel regression for image processing and reconstruction," IEEE Trans. Image Process., vol. 16, no. 2, pp. 349–366, Feb. 2007.

[3] B. Coll ,A. Buades and, J. M. Morel , "A review of image denoising algorithms, with a new one," Multiscale Model. Simul., vol. 4, no. 2, pp. 490–530, 2005.

[4] K. Egiazarian, V. Katkovnik, A. Foi, and K. Dabov, "Image denoising by sparse 3-D transform-domain collaborative filtering," IEEE Trans. Image Process., vol. 16, no. 8, pp. 2080–2095, Aug. 2007.

[5] P. Milanfar and P. Chatterjee , "Patch-based near-optimal image denois- ing," IEEE Trans. Image Process., vol. 21, no. 4, pp. 1635–1649, Apr. 2012.

[6] P. Milanfar, "A tour of modern image filtering: New insights and methods, both practical and theoretical," IEEE Signal Process. Mag., vol. 30, no. 1, pp. 106–128, Jun. 2013.

[7] D. Zhang, W. Dong, L. Zhang, and G. Shi, "Two-stage image denoising by principal component analysis with local pixel grouping," Pattern Recognit., vol. 43, pp. 1531–1549, Apr. 2010.

[8] M. Elad and M. Aharon, "Image denoising via sparse and redundant representations over learned dictionaries," IEEE Trans. Image Process., vol. 15, no. 12, pp. 3736–3745, Dec. 2006.

[9] P. Chatterjee and P. Milanfar, "Is denoising dead?" IEEE Trans. Image Process., vol. 19, no. 4, pp. 895–911, Apr. 2010.

[10] F. Durand, B. Nadler, W. T. Freeman, and A. Levin, "Patch complex- ity, finite pixel correlations and optimal denoising," in Proc. ECCV, Oct. 2012, pp. 73–86.

[11] P. Milanfar, H. Talebi, and X. Zhu, "How to SAIF-ly boost denois- ing performance," IEEE Trans. Image Process., vol. 22, no. 4, pp. 1470–1485, Apr. 2013.

[12] M. Seeger and C. Williams, "Using the Nyström method to speed up kernel machines," in Proc. Adv. NIPS, 2001, pp. 682–688.

[13] E. Nyström, "Über die Praktische Auflösung von linearn ingtegra- glechungen mit anwendungen auf Randwertaufgaben der potentialthe- orie," Commentationes Phys. Math., vol. 4, no. 15, pp. 1–52, Apr. 1928.

[14] C. T. Baker, The Numerical Treatment of Integral Equations. Oxford, U.K.: Clarendon Press, 1977.

[15] C. Fowlkes, A. Talwalkar, and H. Rowley, "Large-scale manifold learn- ing," in Proc. IEEE Conf. CVPR, Jun. 2008, pp. 1–8.

[16] J. Malik,, F. Chung, C. Fowlkes, and S. Belongie "Spectral grouping using the Nyström method," IEEE Trans. Pattern Anal. Mach. Intell., vol. 26, no. 2, pp. 214–225, Feb. 2004.

[17] R. Fattal, D. Lischinski, and Z. Farbman "Diffusion maps for edge- aware image editing," ACM Trans. Graph., vol. 29, no. 6, pp. 1–145, 2010.

[18] E. Seneta, Non-negative Matrices and Markov Chains. New York, NY, USA: Springer-Verlag, 1981.

[19] C. R. Johnson and R. A. Horn, Matrix Analysis. Cambridge, U.K.: Cambridge Univ. Press, 1991.

[20] P. Milanfar, "Symmetrizing smoothing filters," SIAM J. Imag. Sci., vol. 6, no. 1, pp. 263–284, 2012.

[21] C. Stein, "Estimation of the mean of a multivariate normal distribution," Anna. Statist., vol. 9, no. 6, pp. 1135–1151, 1981.

[22] G. Wahba, "The fast Monte-Carlo cross-validation and CL procedures: Comments, new results and applications to image recovery problems-comments," Comput. Stat., vol. 10, no. 3, pp. 249–250, 1995.

[23] T. Blu, M. Unser, and S. Ramani, , "Monte-Carlo sure: A black-box opti- mization of regularization parameters for general denoising algorithms," IEEE Trans. Image Process., vol. 17, no. 9, pp. 1540–1554, Sep. 2008.