On learning optimized reaction diffusion processes for effective image restoration

Yunjin Chen[1,2], Wei Yu[1], Thomas Pock[1,3]
[1]ICG, Graz University of Technology [2]PDL, National University of Defense Technology
[3]Digital Safety & Security Department, AIT Austrian Institute of Technology GmbH

Image restoration is a long-standing problem in low-level computer vision with many interesting applications. The goal of this work is to propose a simple but effective approach with both high computational efficiency and high restoration quality. To that end, we extend conventional nonlinear reaction diffusion models by several parametrized linear filters as well as several parametrized influence functions.

In the fully discrete case $u \in \mathbb{R}^N$, the conventional Perona-Malik type nonlinear diffusion process [5] is formulated as the following discrete PDE with an explicit finite difference scheme

$$\frac{u_{t+1} - u_t}{\Delta t} = -\sum_{i=\{x,y\}} \nabla_i^\top \Lambda(u_t) \nabla_i u_t \doteq -\sum_{i=\{x,y\}} \nabla_i^\top \phi(\nabla_i u_t), \tag{1}$$

where matrices ∇_x and $\nabla_y \in \mathbb{R}^{N \times N}$ are finite difference approximation of the gradient operators in x-direction and y-direction, respectively and Δt denotes the time step. $\Lambda(u_t) \in \mathbb{R}^{N \times N}$ is defined as a diagonal matrix

$$\Lambda(u_t) = \mathrm{diag}\left(g\left(\sqrt{(\nabla_x u_t)_p^2 + (\nabla_y u_t)_p^2}\right)\right)_{p=1,\cdots,N},$$

where function g is known as edge-stopping function. If ignoring the coupled relation between $\nabla_x u$ and $\nabla_y u$, the P-M model can be also written as the second formula on the right side in (1), where $\phi(\nabla_i u) = (\phi(\nabla_i u)_1, \cdots, \phi(\nabla_i u)_N)^\top \in \mathbb{R}^N$ with function $\phi(z) = zg(z)$, known as influence function.

By introducing more linear filters and adjustable influence functions, our proposed nonlinear reaction diffusion model is formulated as

$$\frac{u_t - u_{t-1}}{\Delta t} = -\underbrace{\sum_{i=1}^{N_k} K_i^{t\top} \phi_i^t(K_i^t u_{t-1})}_{\text{diffusion term}} - \underbrace{\psi(u_{t-1}, f_n)}_{\text{reaction term}}, \tag{2}$$

where $K_i \in \mathbb{R}^{N \times N}$ is a highly sparse matrix, implemented as 2D convolution of the image u with the filter kernel k_i, i.e., $K_i u \Leftrightarrow k_i * u$, K_i is a set of linear filters and N_k is the number of filters. The specific formulation for the reaction term $\psi(u)$ depends on applications. In our work, instead of making use of the well-chosen filters and influence functions, we train the nonlinear diffusion process for specific image restoration problem, including both the linear filters and the influence functions.

In this paper, we train our models for two representative image restoration problems: (1) image denoising with Gaussian noise and (2) JPEG blocking artifacts reduction. We use a loss minimization scheme to train the model parameters $\Theta_t = \{\lambda^t, \phi_i^t, k_i^t\}$ for each stage t of the diffusion process, given S training samples $\{f_n^{(s)}, u_{gt}^{(s)}\}_{s=1}^S$, where $f_n^{(s)}$ is a noisy input and $u_{gt}^{(s)}$ is the corresponding ground truth clean image. The training task is to minimize the cost function

$$\mathcal{L}(\Theta_{1,\cdots,T}) = \sum_{s=1}^{S} \ell(u_T^{(s)}, u_{gt}^{(s)}) = \sum_{s=1}^{S} \frac{1}{2}\|u_T^{(s)} - u_{gt}^{(s)}\|_2^2, \tag{3}$$

where the loss function only depends on u_T (the output of the final stage T).

For image denoising, the diffusion equation is given as

$$u_t = u_{t-1} - \left(\sum_{i=1}^{N_k} \bar{k}_i^t * \phi_i^t(k_i^t * u_{t-1}) + \lambda^t(u_{t-1} - f_n)\right). \tag{4}$$

We trained the above diffusion model for the Gaussian denoising task. In our work, we mainly considered two trained reaction diffusion (TRD) models, namely $\mathrm{TRD}_{5\times5}^T$ and $\mathrm{TRD}_{7\times7}^T$, where $\mathrm{TRD}_{m\times m}^T$ denotes a nonlinear diffusion process of stage T with filters of size $m \times m$.

The performance of the trained models is summarized in Table 1, together with a selection of recent state-of-the-art denoising algorithms. The run time performance is presented in Table 2.

In summary, our proposed nonlinear diffusion process has several remarkable benefits as follows:

This is an extended abstract. The full paper is available at the Computer Vision Foundation webpage.

Method	σ = 15	σ = 25	St.	$\sigma = 15$: $\mathrm{TRD}_{5\times5}$	$\sigma = 15$: $\mathrm{TRD}_{7\times7}$
BM3D [2]	31.08	28.56	2	31.14	31.30
LSSC [4]	31.27	28.70	5	31.30	**31.42**
EPLL [7]	31.19	28.68	8	31.34	**31.43**
opt-MRF [1]	31.18	28.66		$\sigma = 25$	
–	–	–		$\mathrm{TRD}_{5\times5}$	$\mathrm{TRD}_{7\times7}$
WNNM [3]	31.37	28.83	2	28.58	28.77
$\mathrm{CSF}_{5\times5}^5$ [6]	31.14	28.60	5	28.78	**28.92**
$\mathrm{CSF}_{7\times7}^5$ [6]	31.24	28.72	8	28.83	**28.95**

Table 1: Average PSNR (dB) on 68 images from for image denoising with $\sigma = 15, 25$.

Method	256^2	512^2	1024^2	2048^2	3072^2
BM3D [2]	1.1	4.0	17	76.4	176.0
$\mathrm{CSF}_{7\times7}^5$ [6]	3.27	11.6	40.82	151.2	494.8
WNNM [3]	122.9	532.9	2094.6	–	–
$\mathrm{TRD}_{5\times5}^5$	0.51	1.53	5.48	24.97	53.3
	0.43	0.78	2.25	8.01	21.6
	0.005	0.015	0.054	0.18	0.39
$\mathrm{TRD}_{7\times7}^5$	1.21	3.72	14.0	62.2	135.9
	0.56	1.17	3.64	13.01	30.1
	0.01	0.032	0.116	0.40	0.87

Table 2: Run time comparison for image denoising (in seconds) with different implementations. (1) The run time results with gray background are evaluated with the single-threaded implementation on Intel(R) Xeon(R) CPU E5-2680 v2 @ 2.80GHz; (2) the blue colored run times are obtained with multi-threaded computation using Matlab *parfor* on the above CPUs; (3) the run time results colored in red are executed on a NVIDIA GeForce GTX 780Ti GPU. We do not count the memory transfer time between CPU/GPU for the GPU implementation (if counted, the run time will nearly double)

1) It is conceptually simple as it is just a time-dynamic nonlinear reaction diffusion model with trained filters and influence functions;
2) It has broad applicability to a variety of image restoration problems. In principle, all existing diffusion based models can be revisited with appropriate training;
3) It yields excellent results for several tasks in image restoration, including Gaussian image denoising, and JPEG deblocking;
4) It is computationally very efficient and well suited for parallel computation on GPUs.

[1] Y. Chen, T. Pock, R. Ranftl, and H. Bischof. Revisiting loss-specific training of filter-based MRFs for image restoration. In *GCPR*, 2013.

[2] K. Dabov, A. Foi, V. Katkovnik, and K. O. Egiazarian. Image denoising by sparse 3-d transform-domain collaborative filtering. *IEEE TIP*, 16 (8):2080–2095, 2007.

[3] S. Gu, L. Zhang, W. Zuo, and X. Feng. Weighted nuclear norm minimization with application to image denoising. In *CVPR*, 2014.

[4] J. Mairal, F. Bach, J. Ponce, G. Sapiro, and A. Zisserman. Non-local sparse models for image restoration. In *ICCV*, 2009.

[5] P. Perona and J. Malik. Scale-space and edge detection using anisotropic diffusion. *IEEE TPAMI*, 12(7):629–639, 1990.

[6] U. Schmidt and S. Roth. Shrinkage fields for effective image restoration. In *CVPR*, 2014.

[7] D. Zoran and Y. Weiss. From learning models of natural image patches to whole image restoration. In *ICCV*, 2011.