https://doi.org/10.2352/EI.2024.36.10.IPAS-245

Image Restoration via Collaborative Filtering and Deep Learning

Wenzhu Xing, Igor Shevkunov, Vladimir Katkovnik, Karen Egiazarian; Center for Signal Processing, Tampere University, Korkeakoulunkatu 1; Tampere, Finland

Abstract

In this paper, we investigate the challenge of image restoration from severely incomplete data, encompassing compressive sensing image restoration and image inpainting. We propose a versatile implementation framework of plug-and-play ADMM image reconstruction, leveraging readily several available denoisers including model-based nonlocal denoisers and deep learning-based denoisers. We conduct a comprehensive comparative analysis against state-of-the-art methods, showcasing superior performance in both qualitative and quantitative aspects, including image quality and implementation complexity.

Introduction

Numerous stochastic approximation iterative techniques have been advanced to tackle the challenge of restoring images from severely incomplete data [2],[9], e.g. compressive sensing and image interpolation. One such method, which incorporates regularization through BM3D filtering, was initially proposed in [6] for compressive sensing and later extended for image deblurring [5] and image denoising [13]. These approaches fall under the category of plug-and-play (PnP) methods [10], where image denoisers serve as regularizers. The Alternating Direction Method of Multipliers (ADMM), utilizing off-the-shelf denoisers as implicit image priors, stands out as the most widely used PnP method.

Over the past decade, numerous variations of nonlocal image restoration methods have emerged [5, 6, 7, 12]. These methods employ a range of nonlocal image denoisers, such as BM3D [6], BM3D-SVD, and BM3D-Wiener, to address tasks like image deblurring and super-resolution. They notably enhance the performance of state-of-the-art methods like IDD-BM3D [5] and GSR [7]. The recently introduced HSSE image restoration method [12] adopts a similar approach, employing two denoisers within the Plug-and-Play (PnP) Alternating Direction Method of Multipliers (ADMM) framework. Specifically, it utilizes BM3D-SVD on both internal and external patches. A comparative analysis featured in a recent overview [11] highlights the superiority of HSSE over other state-of-the-art methods. This advantage is particularly evident when compared to techniques utilizing image sparsity, self-similarity, and deep learning priors. One drawback commonly associated with most PnP ADMM methods, including HSSE, is their slow convergence speed, often necessitating hundreds or even thousands of iterations.

In this paper, we introduce an efficient PnP ADMM algorithm that combines neural network and nonlocal sparsity priors for image restoration. Our approach leverages a deep learning-based denoiser as an external image prior, complemented by BM3D-SVD incorporating soft-thresholding filtering as an internal image prior.

We conduct a comparative evaluation of our proposed method against state-of-the-art techniques for compressive sensing and image inpainting. Results demonstrate that our approach yields reconstructed images of superior quality while requiring significantly fewer computations, typically by a factor of tens or hundreds.

Image Restoration Based on Nonlocal Collaborative Filtering and Neural Networks

In this section, we introduce a PnP ADMM image restoration framework that integrates two distinct image priors: nonlocal group sparsity and a neural network prior.

Let $\mathbf{X}$ be a grayscale image of size $N\times M$ and the observations are given by the equation $\mathbf{Y}=H(\mathbf{x})+\varepsilon$, where $H(.)$ represents a degradation operation (e.g. blur), ε is the additive zero-mean i.i.d. Gaussian noise with the standard deviation σ_n.

For image restoration we will use image patches. Let $\mathbf{x}_{(i,j)}$ be a (reference) image patch of size $n\times m$, taken from the intersection of i-th row and j-th column ($i=1,...,N; j=1,...,M$) of image $\mathbf{X}$, and let $X_{(i,j)}=\{\mathbf{x}_{(i,j,0)},\mathbf{x}_{(i,j,1)},...,\mathbf{x}_{(i,j,K-1)}\}$, be a group of patches consisting from the patch $\mathbf{x}_{(i,j,0)}=\mathbf{x}_{(i,j)}$ and $K-1$ most similar patches to it, all taken from the neighborhood of $\mathbf{x}_{(i,j)}$ of size $L\times L$. The group of patches $X_{(i,j)}$ can be written as a 3D array of size $n\times m\times K$.

A problem of reconstruction of image X from observation Y can be formulated as the following unconstrained optimization problem:

$$\hat{X}=argmin_X\frac{1}{2\sigma_n^2}||Y-H(X)||_2+\tau f(X),$$

where the first summand is a fidelity term, and a second summand is a regularization term involving image priors $f(X)$ with a regularization parameter τ.

Here, we integrate a deep neural network prior with the nonlocal group sparsity prior within a unified regularization framework. We formalize the image reconstruction process by leveraging the vectorized representation of the variable:

$$\begin{aligned}(\hat{X},(\widehat{A}_{i,j})_{i=1,j=1}^{N,M}) = argmin_{(A_{i,j})_{i=1,j=1}^{N,M},X}\Big(\frac{1}{2\sigma_n^2}||(Y-HX||_2^2\\ +\sum_{i,j}\frac{1}{2\sigma_{i,j}^2}||\mathbf{R}_{(i,j)}X-\mathbf{D}_{(i,j)}\mathbf{A}_{(i,j)}||_2^2\\ +\tau f(X)+\sum_{i,j}\frac{1}{\rho_{i,j}}||\mathbf{A}_{(i,j)}||_1\Big).\end{aligned} \quad (1)$$

Here $\mathbf{A}_{i,j}$ are vectors of the sparse spectral representation of the groups, and $\mathbf{D}_{i,j}$ the sparsifying synthesis operators (orthonormal matrices), such that for the true data $\mathbf{X}_{i,j} = \mathbf{D}_{i,j}\mathbf{A}_{i,j}$. The matrices $\mathbf{R}_{i,j}$ are the grouping matrices such that $\mathbf{X}_{i,j} = \mathbf{R}_{i,j}X$.

To facilitate the optimization on X, we apply alternating direction method of multipliers (ADMM) algorithm [1], and introduce auxiliary variable Z with the constraint $X = Z$. Then the problems can be can be rewritten as the following four step optimization problem.

1: Group-wise spectral analysis of Z

$$(\widehat{A}_{i,j})_{i=1,j=1}^{N,M} = argmin_{(A_{i,j})_{i=1,j=1}^{N,M}} \Big(\sum_{i,j} \frac{1}{2\sigma_{i,j}^2} ||\mathbf{R}_{(i,j)}Z - \mathbf{D}_{(i,j)}\mathbf{A}_{(i,j)}||_2^2 + \sum_{i,j} \frac{1}{\rho_{i,j}} ||\mathbf{A}_{(i,j)}||_1 \Big). \quad (2)$$

2: Estimation of Z

$$\widehat{Z} = argmin_{(Z)} \Big(\sum_{i,j} \frac{1}{2\sigma_{i,j}^2} ||\mathbf{R}_{(i,j)}Z - \mathbf{D}_{(i,j)}A_{(i,j)}||_2^2 + \frac{\mu}{2} ||X - Z - \Lambda||_2^2 \Big). \quad (3)$$

3: Estimation of X

$$\widehat{X} = argmin_{(X)} \Big(\frac{1}{2\sigma_n^2} ||(Y - HX||_2^2 + \frac{\mu}{2} ||X - Z - \Lambda||_2^2 + \tau f(X) \Big). \quad (4)$$

4: Update the Lagrange variable Λ

$$\Lambda = \Lambda - (X - Z). \quad (5)$$

These optimization problems allow the following solutions

Step 1: For each (i,j), the solution for (2) is soft-thresholding of the variable $\mathbf{D}_{(i,j)}^T \mathbf{R}_{(i,j)} Z$

$$(\widehat{A}_{i,j}) = \mathbf{SOFT}(\mathbf{D}_{(i,j)}^T \mathbf{R}_{(i,j)} Z). \quad (6)$$

Step 2: Solution for the quadratic criterion in (3) gives:

$$\widehat{Z} = \Big(\sum_{i,j} \frac{1}{\sigma_{i,j}^2} \mathbf{R}_{(i,j)}^T \mathbf{R}_{(i,j)} + \mu \mathbf{I} \Big)^{-1} \Big(\sum_{i,j} \frac{1}{\sigma_{i,j}^2} \mathbf{R}_{(i,j)}^T \mathbf{D}_{(i,j)} \mathbf{A}_{(i,j)} + \mu (X - \Lambda) \Big). \quad (7)$$

Step 3: If $\tau = 0$ the solution for (4) is as follows:

$$\widehat{X} = \Big(\frac{1}{\sigma_n^2} H^T H + \mu \mathbf{I} \Big)^{-1} (H^T Y + \mu (z + \Lambda)). \quad (8)$$

Implementation of the algorithm. If we consider problems 1 and 2 as the analysis and aggregation stages of the BM3D algorithm [4], we observe that this algorithm not only provides solutions but also facilitates data-adaptive data grouping, specifically through the synthesis of grouping matrices $\mathbf{R}i, j$. Additionally, it computes group-wise weights, denoted as noise standard deviations $\sigma i, j^2$, reflecting the Gaussian noise present in observations. These weights play a crucial role in aggregating group-wise patch estimates. Notably, an improved calculation method for these weights is introduced in the updated version of BM3D [8]. To regularize the estimate of X, we employ a deep neural network denoiser. For implementing the derived algorithm, we utilize both the nonlocal collaborative filter BM3D and a deep neural network denoiser. The neural network operator is denoted as Φ. It's worth noting that in this implementation, we adopt BM3D-SVD [11] with soft thresholding as is, akin to the denoiser employed in HSSE [12], without any intervention in the codes that adjust the weights according to the formulas (7)-(8).

Step-by-step iterations of the developed algorithm, which we call ADMM-NN3D, following cascaded Neural Network and three- dimensional (3-D) collaborative filtering (NN3D) framework [3], are given below:

1) Filtering of the $(i-1)-$iteration estimate with deep denoiser: $\tilde{X}^{(i-1)} = \Phi(\hat{X}^{(i-1)})$.
2) Update of the estimate : $\hat{X}^{(i)} = \frac{\alpha \hat{X}^{(i-1)} + \beta \tilde{X}^{(i-1)}}{\alpha + \beta}$
3) Filtering with the nonlocal collaborative denoiser: $\tilde{X}^{(i)} = BM3D(\hat{X}^{(i)})$,
4) Update of the filtered estimate:

$$\hat{V}^{(i)} = \frac{\alpha \tilde{X}^{(i)} + \mu(\hat{X}^{(i-1)} - \Lambda^{(i-1)})}{\alpha + \mu}$$

5) Filtering $V^{(i)}$ with the deep denoiser Φ:

$$\tilde{V}^{(i)} = \Phi(\hat{V}^{(i)}).$$

6) Update the reconstructed image estimate

$$\hat{X}^{(i)} = (H^T H + \mu I)^{-1} (H^T Y + \mu(\tilde{V}(i) + \Lambda^{(i-1)})),$$

7) Update the Lagrangian multiplier term

$$\Lambda^{(i)} = \Lambda^{(i-1)} - (\hat{X}^{(i)} - \hat{V}^{(i)}),$$

where $\hat{V}^{(i)}$ is the auxiliary estimate at the $i^t h$ stage of the iterative algorithm, and $\hat{X}^{(i)}$, is the output at i^{th} stage of the algorithm.

Experiments

In this section, we present experimental results conducted on Set 12, which comprises twelve commonly used grayscale images of dimensions 256x256 and 512x512 pixels [14]. These results aim to demonstrate the efficacy of the proposed algorithm for both image inpainting and compressive sensing (CS) image reconstruction. Following established practices in prior research, block-based compressive sensing is employed across all tested methods, utilizing a block size of 32×32.

For each given CS subrate, the corresponding measurement matrix is constructed by generating a random Gaussian matrix and subsequently orthogonalizing its rows [15]. In the case of image inpainting, we consider image restoration with randomly missing pixels, with varying percentages of missing pixels.

The primary parameters of the BM3D-SVD denoiser are configured similarly to those utilized in other methods, such as

HSSE: each patch size is set to 7 × 7, the number of similar patches is fixed at K = 60, noise standard deviation is set to $\sigma_n = \sqrt{2}$, and the ADMM balance factor is adjusted to $\eta = 0.001$ for inpainting and $\eta = 0.003$ for compressive sensing. Additionally, we employ the Drunet denoiser [14] as the neural network denoiser $F(\cdot)$.

For the proposed method ADMM-NN3D, the number of iterations ranges from 40 to 70, depending on the subrate, with lower subrates requiring a larger number of iterations. Notably, this iteration count is significantly lower than that used in state-of-the-art SSR and HSSE methods, where iteration counts varied from 500 to 2000.

The results of image compressed sensing methods and image inpainting are tabulated in Tables 1 and 2, respectively. It is evident from these tables that the proposed method consistently outperforms state-of-the-art techniques.

Table 1: PSNR values of compressed sensing methods on Set12 for different subrates.

CS Methods	0.1	0.2	0.3	0.4	Average
SDA	23.68	25.84	27.28	28.49	26.32
ReconNet	24.56	26.31	29.33	30.71	27.73
IRCNN	26.31	30.64	33.15	34.99	31.27
SSR	27.57	30.84	32.90	34.91	31.56
GSR	26.85	30.91	33.43	35.46	31.66
HSSE	27.71	31.24	33.67	35.71	32.08
CREAM	27.31	31.66	34.14	36.11	32.31
ADMM-NN3D	**28.72**	**32.59**	**34.76**	**36.35**	**33.11**

Table 2: PSNR values of image inpainting methods on Set12 for different percentage of missing pixels.

Inpainting Methods	90 %	80 %	70 %	60 %	Average
SSR	25.27	28.42	30.59	32.35	29.16
HSSE	25.32	28.77	31.06	32.91	29.52
ADMM-NN3D	**25.92**	**29.67**	**32.07**	**34.47**	**30.53**

The results of the compared methods are drawn from the studies referenced in [12] and [15]. Figure 1-2 illustrates examples of image inpainting and CS reconstruction using HSSE and our proposed method, showcasing a notable enhancement in reconstruction quality with a PSNR improvement of over 3 dB.

Conclusions

In this paper, we introduced a novel ADMM image restoration method that incorporates nonlocal self-similarity (BM3D-SVD) and deep neural network priors. Our experiments showcase that the proposed method surpasses state-of-the-art techniques in both speed (requiring 10-20 times fewer computations) and reconstructed image quality for both image inpainting and compressed sensing image reconstruction tasks.

Figure 1: 'Monarch' image, compressed sensing image restoration, rate 0.1, HSSE (500 iterations), PSNR=26.5, SSIM = 0.883, the proposed method (70 iterations), PSNR=29, SSIM = 0.918.

References

[1] Manya V Afonso, José M Bioucas-Dias, and Mário AT Figueiredo. Fast image recovery using variable splitting and constrained optimization. *IEEE transactions on image processing*, 19(9):2345–2356, 2010.

[2] Stanley H Chan, Xiran Wang, and Omar A Elgendy. Plug-and-play admm for image restoration: Fixed-point convergence and applications. *IEEE Transactions on Computational Imaging*, 3(1):84–98, 2016.

[3] Cristóvão Cruz, Alessandro Foi, Vladimir Katkovnik, and Karen Egiazarian. Nonlocality-reinforced convolutional neural networks for image denoising. *IEEE Signal Processing Letters*, 25(8):1216–1220, 2018.

[4] Kostadin Dabov, Alessandro Foi, Vladimir Katkovnik, and Karen Egiazarian. Image denoising by sparse 3-d transform-domain collaborative filtering. *IEEE Transactions on image processing*, 16(8):2080–2095, 2007.

[5] Aram Danielyan, Vladimir Katkovnik, and Karen Egiazarian. Bm3d frames and variational image deblurring. *IEEE Transactions on Image Processing*, 21(4):1715–1728, 2012.

[6] Karen Egiazarian, Alessandro Foi, and Vladimir Katkovnik. Compressed sensing image reconstruction via recursive spatially adaptive filtering. In *2007 IEEE International Conference on Image Processing*, volume 1, pages I – 549–I – 552, 2007.

[7] Xiaolei Lu and Xuebin Lü. Admm for image restoration based on nonlocal simultaneous sparse bayesian coding. *Signal Processing: Image Communication*, 70:157–173, 2019.

[8] Matteo Maggioni, Vladimir Katkovnik, Karen Egiazarian, and Alessandro Foi. Nonlocal transform-domain filter for volumetric data denoising and reconstruction. *IEEE transactions on image processing*, 22(1):119–133, 2012.

[9] Yu Sun, Zihui Wu, Xiaojian Xu, Brendt Wohlberg, and Ulugbek S Kamilov. Scalable plug-and-play admm with convergence guarantees. *IEEE Transactions on Computational Imaging*, 7:849–863, 2021.

[10] Singanallur V. Venkatakrishnan, Charles A. Bouman, and Brendt Wohlberg. Plug-and-play priors for model based reconstruction. In *2013 IEEE Global Conference on Signal and Information Processing*, pages 945–948, 2013.

[11] Zhiyuan Zha, Bihan Wen, Xin Yuan, Saiprasad Ravishankar, Jiantao Zhou, and Ce Zhu. Learning nonlocal sparse and low-rank models for image compressive sensing: Nonlocal sparse and low-rank modeling. *IEEE Signal Processing Magazine*, 40(1):32–44, 2023.

[12] Zhiyuan Zha, Bihan Wen, Xin Yuan, Jiantao Zhou, Ce Zhu, and Alex Chichung Kot. A hybrid structural sparsification error model for image restoration. *IEEE Transactions on Neural Networks and Learning Systems*, 33(9):4451–4465, 2022.

[13] K. Zhang, W. Zuo, and L. Zhang. Deep plug-and-play super-resolution for arbitrary blur kernels. In *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1671–1681, 2019.

[14] Kai Zhang, Yawei Li, Wangmeng Zuo, Lei Zhang, Luc Van Gool,

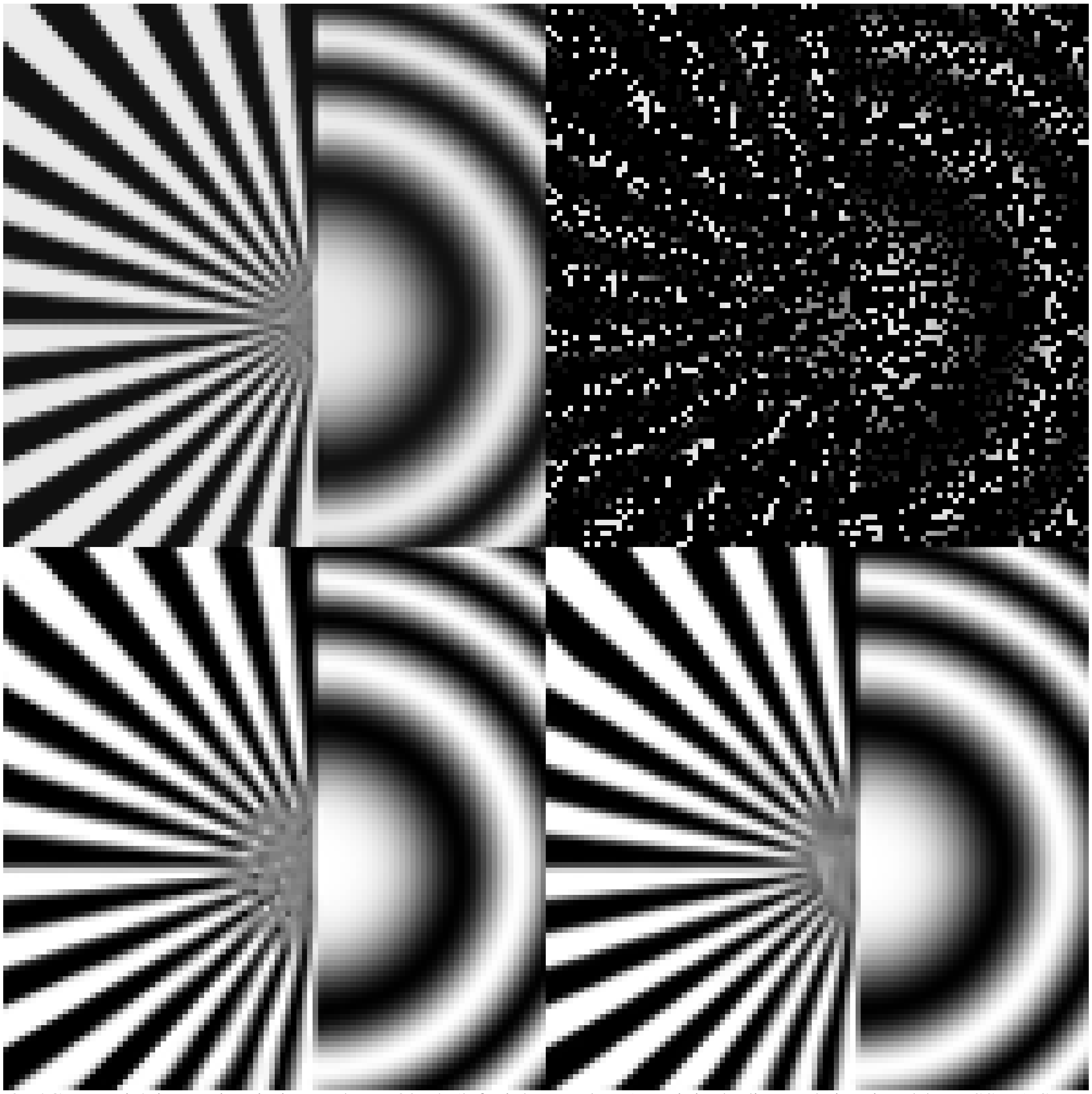

Figure 2: 'Geometric' image inpainting, subrate 80 % (left-right, up-down): original, distorted, inpainted by HSSE (PSNR=27.88; FSIM=0.98; SSIM=0.976), and inpainted by the proposed method (PSNR = 30.48, FSIM=0.99, SSIM=0.986).

and Radu Timofte. Plug-and-play image restoration with deep denoiser prior. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 44(10):6360–6376, 2021.

[15] Chen Zhao, Jian Zhang, Ronggang Wang, and Wen Gao. Cream: Cnn-regularized admm framework for compressive-sensed image reconstruction. *IEEE Access*, 6:76838–76853, 2018.

Author Biography

Wenzhu Xing received her Bachelor's degree in software engineering from the Sichuan University (2016), and her Master's degree in Master of Science (Technology) from Tampere University (2020). Since then she has been a researcher in the Center for Signal Processing in Tampere University, Finland. Her work has focused on computer vision and image processing tasks.

Igor Shevkunov is a senior researcher from Tampere University since 2017. He received his Ph.D. in Optics from St.Petersburg State University, Russia, in 2013. He is the author of more than 50 refereed papers. His main research interests are digital holography, phase retrieval, and interferometry with the development of state-of-the-art imaging and processing techniques.

Vladimir Katkovnik , Ph.D. (1964), D.Sc. (1974), with Tampere University (Tampere, Finland) since 2003; published 6 books and over 300 papers. The research interests include stochastic image/signal processing, linear and nonlinear filtering, nonparametric estimation, computational imaging, nonstationary systems, and time-frequency analysis.

Karen Egiazarian (Eguiazarian) is a Professor from Tampere University (Tampere, Finland) since 1999, leading the Computational Imaging group. He has published over 700 papers. His main research interests include image/video restoration, computational imaging, digital holography, and remote sensing. Prof. Egiazarian was EiC of the Journal of Electronic Imaging (2016-2020).