2022 21st International Symposium on Distributed Computing and Applications for Business Engineering and Science (DCABES)

Group-based Sparse Coding with Adaptive Dictionary Learning for Image Denoising

Jiaying Wang, Meiqing Wang*, Hang Cheng, Rong Liu
College of Mathematics and Statistics
Fuzhou University
Fuzhou, China
Email: 851399016@qq.com;
mqwang@fzu.edu.cn;
hcheng@fzu.edu.cn;
liu_r@fzu.edu.cn

Fei Chen
College of Computer Science and Data Science
Fuzhou University
Fuzhou, China
Email: chenfei314@fzu.edu.cn

***Abstract*—Group sparse coding (GSC) is a powerful mechanism that has achieved great success in many low-level vision tasks, showing great potential in image denoising. Traditional group sparse coding generally uses overcomplete dictionaries and l_1-norm to regularize sparse coefficients. But this is only an estimate of the solution, which cannot obtain a sparse solution and has a high computational cost. In this paper, we use a GSC framework with adaptive dictionary learning for image denoising. In order to improve the accuracy of obtaining sparse coefficients, the dictionary used in this paper is learned from the input image, which can be obtained by applying SVD once for each patch group. Then use ADMM algorithm to solve the objective function. Experimental results show that the PSNR value of our approach not only is competitive with many advanced image denoising methods but also achieves better visual effects.**

***Keywords*-group-based sparse coding; adaptive dictionary learning; weighted l_p-norm minimization; ADMM; image denoising.**

I. INTRODUCTION

Image, as an information carrier, have powerful information-containing capabilities, which leads to higher requirements for image quality. In real life, images obtained by machines are often not perfect. Due to the limitations of imaging methods and conditions and external interference, digital image signals will inevitably be polluted by noise signals. Important information of the research target in the image is often disturbed or covered by noise signals. It increases the difficulty of the subsequent research and processing of the image, so denoising the image is an important task of image processing. Image denoising is a process of recovering a clean image X from an observed noisy image Y as precisely as possible, while preserving important details such as edge and feature information. The degradation model can usually be expressed as $Y = X + e$, where $Y \in \mathbb{R}^N$ is the observed noisy image, and $X \in \mathbb{R}^N$ denotes the unknown original image, $e \in \mathbb{R}^N$ is an additive white Gaussian noise with standard deviation σ_e and zero mean. Mathematically, the image denoising problem is ill-posed, and its solution is often not unique. In order to successfully reconstruct a clean image X from an observed noisy image Y, an image prior model is required.

Sparsity and non-local self-similarity are important properties of natural images, which can be uesd as priors, the most representative works are sparse representation-based schemes. Traditional patch-based sparse coding (PSC) has been used widely in many image restoration tasks and achieved good results. This scheme encodes image patches as sparse linear combinations of atoms in an overcomplete redundant dictionary, which is usually fixed or learned from natural images.

However, patch-based sparse coding models usually assume the independence of image patches without considering the correlation between similar patches. The latest progress in Group-based Sparse Coding (GSC) is to use groups of similar image patches as the basic unit for subsequent processing, and show great potential in various image processing tasks.

Inspired by the work of Zha et al. [1] and Wang et al. [2] in sparse coding, in this paper we use a group-based sparse coding with adaptive dictionary learning for image denoising. To make this denoising approach tractable, the alternating direction multiplier method (ADMM) [3] is adopted to solve the large-scale l_p-norm minimization problem. Furthermore, we use an adaptive dictionary learning method that adaptively learns the corresponding dictionary according to the different input images. The experiment show that the PSNR results of our method is not only better than the classical denoising methods, such as BM3D and EPLL but also better than many state-of-the-art denoising methods.

II. GROUP-BASED SPARSE CODING

Group sparse coding (GSC) is a powerful mechanism for fusing the inherent local sparsity and non-local self-similarity (NSS) of natural images [1], [2]. Based on PSC, GSC considers the correlation between image patches. Specifically, first to divide the observed noisy image $Y \in \mathbb{R}^N$ into n overlapping patches of size $\sqrt{s} \times \sqrt{s}$, and vectorize each image patch, denoted as $y_i \in R^s,\ i = 1, 2, \ldots, n$. Then, a window with a fixed size of $C \times C$ is used to traverse the entire image to search and select similar matching patches to form a set of similar image patches of y_i. The search for similar patches usually

2473-3636/22/$31.00 ©2022 IEEE
DOI 10.1109/DCABES57229.2022.00036

uses KNN [4]. Next, stack all the patches in the set to form a matrix $Y_i \in \mathbb{R}^{s\times m}$.

For each image patch group Y_i, given a dictionary D_i, usually a overcomplete dictionary fixed in advance or learned through PCA, the group sparse coefficient A_i can be obtained by,

$$\hat{A}_i = \arg\min_{A_i} \left(\frac{1}{2{\sigma_e}^2} \|Y_i - D_i A_i\|_2^2 + \lambda \|A_i\|_0 \right) \quad (1)$$

where $\|\|_2$ and $\|\|_0$ are the Euclidean norm and the l_0-norm respectively, and λ denotes the regularization parameter.

Once the coefficient $\{A_i\}_{i=1}^n$ are obtained, the original image X can be reconstructed with $\hat{X} = DA$. However, since the l_0-norm minimization is a discontinuous optimization problem, solving the above equation is NP-hard. In order to make it tractable, it is often replaced by other norm minimization.

III. Group-based Sparse Coding with Adaptive Dictionary Learning

A. Modeling of Weighted l_p-norm Minimization

In the work of Zha et al. [1], it is proved by theoretical analysis and experiments that using weighted l_p-norm minimization to replace l_0-norm minimization has the best effect on image restoration, so here in this work, we use l_p-norm to replace l_0-norm. Therefore, Eq. (1) is replaced by the following minimization problem,

$$\hat{A}_i = \arg\min_{A_i} \left(\frac{1}{2{\sigma_e}^2} \|Y_i - D_i A_i\|_2^2 + \lambda \|W_i A_i\|_p \right) \quad (2)$$

where W_i denotes the weight of each patch group Y_i, which is used to strengthen the representation ability of A_i for each image patch group Y_i.

B. Adaptive Dictionary Learning

Dictionary learning is a large scale problem, requires a large amount of computation. Traditional dictionary learning is to learn an over-complete dictionary to achieve the purpose of sparse representation. In order to improve the accuracy of obtaining sparse coefficients, we use an adaptive dictionary learning approach. The dictionary is learned from the patch group Y_i of the input image, so that the structure is not too complicated and easy to learn.

Different from the work of Wang et al. [2], we construct the dictionary by simple singular value decomposition (SVD), which saves the computational cost and time. Specifically, apply singular value decomposition to Y_i,

$$Y_i = U_i \Delta_i V_i^T = \sum_{j=1}^{t} \delta_{i,j} u_{i,j} v_{i,j}^T \quad (3)$$

where $\Delta_i = diag\left(\delta_{i,1}, \delta_{i,2}, \ldots, \delta_{i,t}\right)$ is a diagonal matrix, t represents the minimum value of s and m, and $u_{i,j}, v_{i,j}$ denotes the j-th column of U_i and V_i, respectively.

Then, the dictionary atom $d_{i,j}$ is defined as,

$$d_{i,j} = v_{i,j}^T u_{i,j}, \ j = 1, 2, \ldots, t \quad (4)$$

In this way, we learn an adaptive dictionary $D_i = [d_{i,1}, d_{i,2}, \ldots, d_{i,t}]$ for each patch group Y_i. Note that in this way, the entire dictionary learning process requires only a simple SVD for each image patch group.

C. Solving the Weighted l_p-norm Minimization by ADMM

The alternating direction multiplier method (ADMM) [3] was proved only needs a small memory to reach the convergence condition, so it is popular for most large-scale problems. Note that Eq. (2) is a large-scale optimization problem and is non-convex, which is difficult to solve.

In this paper, we apply ADMM to Eq. (2) to make it tractable. First, an auxiliary variable Z is introduced to transform Eq. (2) into another equivalence constraint,

$$\begin{aligned} &\hat{A}_i = \arg\min_{Z_i, A_i} \left(\frac{1}{2{\sigma_e}^2} \|Y_i - Z_i\|_2^2 + \lambda \|W_i A_i\|_p \right), \\ &s.t. \ Z_i = D_i A_i \end{aligned} \quad (5)$$

The subscript i is omitted in the derivation below for conciseness. Next, convert the above equation into the following three iterations:

$$Z^{k+1} = \arg\min_{Z} \left(\frac{1}{2{\sigma_e}^2} \|Y - Z\|_2^2 + \frac{\rho}{2} \left\|Z - DA^k + B^k\right\|_2^2 \right) \quad (6)$$

$$A^{k+1} = \arg\min_{A} \left(\lambda \|WA\|_p + \frac{\rho}{2} \left\|Z^k - DA + B^k\right\|_2^2 \right) \quad (7)$$

$$B^{k+1} = B^k + \left(Z^{k+1} - DA^{k+1}\right) \quad (8)$$

where ρ is the penalty parameter of ADMM. It can be observed that Eq. (8) is just a simple update, so solving Eq. (5) is divided into two minimization sub-problems. Fortunately, each subproblem has a valid solution. The superscript k is omitted for convenience.

Note that Eq. (6) is a least squares (LS) problem and has a closed-form solution as follows,

$$\hat{Z} = \left(I_N + \rho\sigma_e^2 I_N\right)^{-1} \times \left(Y + \rho\sigma_e^2 \left(DA - B\right)\right) \quad (9)$$

where I is an N-dimensional identity matrix consistent with the size of the observed image.

To get the solution to Eq. (7), we apply the generalized soft-thresholding (GST) algorithm [5]. The closed-form solution can be obtain by,

$$A_i = GST\left(y_i, \tau w_i, p\right) \quad (10)$$

where $\tau = \frac{\lambda snm}{\rho\sigma_e N}$, y_i and w_i are the vectorization of the similarity patch group matrix Y_i and its corresponding weight matrix W_i respectively.

Since large numbers in the sparse coefficient A_i usually contain important information in the image Y, such as edges and textures. To preserve these information in each iteration, we usually scale down the values of elements in A_i moderately proportional to each iteration. Inspired by [1], w_i and λ are set to $w_i = \frac{1}{|y_i| + \varepsilon}$, $\lambda = \frac{2\sqrt{2}\sigma_e^2}{\varphi_i + \theta}$ in each iteration, where φ_i is the estimated variance of the vector y_i, ε and θ are small positive constants.

The group-based sparse coding framwork with adaptive dictionary learning for image denoising we used is summarized in **Algorithm 1**.

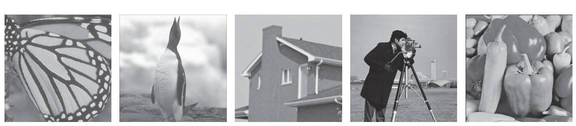
Algorithm 1: The GSC framework with Adaptive Dictionary Learning

---

Input: The observation Y.
Initialization: X^0, σ_e, N, n, s, m, C, ε, θ, $MaxIter$, p, ρ
for $k = 0$ *to* $MaxIter$ **do**
  while *stopping criterion not met* **do**
    Update λ^{k+1} by computing $\lambda = \frac{2\sqrt{2}\sigma_e^2}{\varphi_i + \theta}$;
    Iterative regularization
    $Y^{k+1} = X^k + \lambda^{k+1}\left(Y^k - X^k\right)$;
    for *Each patch group* Y_i **do**
      Construct dictionary D_i by Eq. (4);
      Update τ by computing $\tau = \frac{\lambda snm}{\rho \sigma_e N}$;
      Update w_i by computing $w = \frac{1}{|y| + \varepsilon}$;
      Update A_i by computing Eq. (10);
      Update B_i by computing Eq. (8);
      Reconstruct $\hat{X}_i$ by $\hat{X}_i = D_i A_i$;
    end
    Collect $\hat{X}_i$ to form the denoised image $\hat{X}$.
  end
end
Output: The reconstructed clean image $\hat{X}$.

IV. Experiments

To validate the effectiveness of the algorithm mentioned above, in this section we conduct some experiments to compare it with some traditional, benchmark, and state-of-the-art methods for image denoising, including BM3D [6], EPLL [7], LRJS [8], WNNM [9] and WSNM [10]. The images used in the experiment are shown in Fig. 1.



(a) (b) (c) (d) (e)

Figure 1. The five image for denoising experiments

The parameters of the our approach are set as follows: for each image patch, the patch size $\sqrt{s} \times \sqrt{s}$ is set to 8×8, and the number of similar patches to be searched m is set to 40, the window to search similar patches $C \times C$ is set to 25×25, the small positive constant (ε, θ) are set to (0.1, 0.3). For parameters (p, ρ), there are different settings for different noise levels. Specifically, (p, ρ) are set to (1, 0.03), (0.95, 0.003), (0.9, 0.001) and (0.85, 0.0005) for $\sigma_e \leq 20$, $20 < \sigma_e \leq 50$, $50 < \sigma_e \leq 70$ and $70 < \sigma_e \leq 100$ respectively. The experiments were conducted on Matlab R2016a with a machine have Intel Core i7-6700 CPU @ 3.40GHz and 16GB memory.

Due to limited space, the PSNR and visualization results of only five images under four different noise levels are presented in this section, $i.e.$, σ_e=20, 30, 50, and 90. Table I presents all the PSNR results of our experiment, the best results are highlighted in bold, and the visualization results are given in Fig. 2.

As can be seen from Table I, our denoising method is competitive, and the PSNR value is better than other competing methods in most cases, compared to BM3D, EPLL, LRJS, WNNM and WSNM, the average PSNR is improved by 0.504dB, 0.633dB, 0.332dB, 0.103dB, 0.068dB, respectively. For some images, the approach used in this paper cannot obtain the highest value of PSNR. However, as can be seen from Fig. 2, although WSNM obtains a higher PSNR value, the visual effect is not as good as ours. It can be seen that other denoising methods used for comparison still produce some undesirable artifacts when denoising the image. In contrast, the denoising scheme we use not only preserves intact edge details but also eliminates unwanted visual artifacts.

V. Conclusion

Different from traditional group sparse coding, this paper use an image denoising framework based on group sparse coding using adaptive dictionary learning, which uses a simple SVD to learn the dictionary from the input image, reduces the amount of computation while improving the accuracy of obtaining sparse coefficients. The weighted l_p-norm is used to replace the l_0-norm. In order to make the scheme tractable, the ADMM algorithm is used to solve the weighted l_p-norm minimization problem. Experimental results show that the image denoising framework we used in this paper outperforms many other state-of-the-art denoising methods.

References

[1] Zha Z, Yuan X, Wen B, et al. A benchmark for sparse coding: When group sparsity meets rank minimization[J]. IEEE Transactions on Image Processing, 2020, 29: 5094-5109.

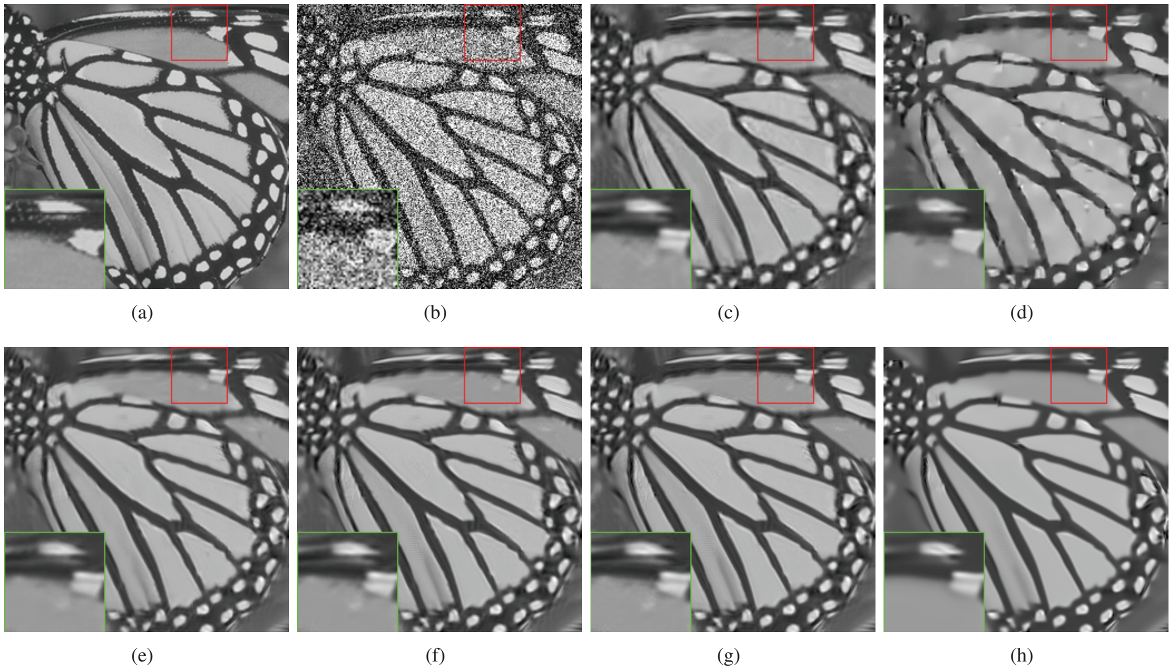
[2] Wang Q, Zhang X, Wu Y, et al. Nonconvex Weighted ℓ_p Minimization Based Group Sparse Representation Framework for Image Denoising[J]. IEEE Signal Processing Letters, 2017, 24(11): 1686-1690.

[3] He B, Liao L Z, Han D, et al. A new inexact alternating directions method for monotone variational inequalities[J]. Mathematical Programming, 2002, 92(1): 103-118.

[4] Sun S, Huang R. An adaptive k-nearest neighbor algorithm[C]//2010 seventh international conference on fuzzy systems and knowledge discovery. IEEE, 2010, 1: 91-94.

[5] Voronin S, Chartrand R. A new generalized thresholding algorithm for inverse problems with sparsity constraints[C]//2013 IEEE International Conference on Acoustics, Speech and Signal Processing. IEEE, 2013: 1636-1640.

[6] Dabov K, Foi A, Katkovnik V, et al. Image denoising by sparse 3-D transform-domain collaborative filtering[J]. IEEE Transactions on image processing, 2007, 16(8): 2080-2095.



(a) (b) (c) (d)

(e) (f) (g) (h)

Figure 2. Denoising perfprmance comparison of $Butterfly$ by different methods with $\sigma_e = 70$: (a) Original image; (b) Noisy image; (c) BM3D [6] (PSNR=23.75dB); (d) EPLL [7] (PSNR=23.74dB); (e) LRJS [8] (PSNR=24.11dB); (f) WNNM [9] (PSNR=24.36dB); (g) WSNM [10] (PSNR=24.51dB); (h) Ours (PSNR=24.47dB).

Table I
PSNR(DB) Results of Different Denoising Methods

	σ_e= 20						σ_e= 50					
	BM3D	EPLL	LRJS	WNNM	WSNM	OURS	BM3D	EPLL	LRJS	WNNM	WSNM	OURS
Butterfly	30.08	30.21	31.14	31.24	**31.24**	31.17	25.37	25.59	26.21	26.32	26.37	**26.39**
Penguin	33.57	33.51	33.31	33.60	33.60	**33.75**	28.78	28.79	28.49	28.99	29.00	**29.56**
House	33.77	33.08	33.92	34.01	34.04	**34.12**	29.69	29.02	30.23	30.32	30.37	**30.52**
Cameraman	30.48	30.37	30.43	**30.75**	30.73	30.67	26.12	26.09	26.32	26.42	26.44	**26.51**
Peppers	31.28	31.25	30.40	31.54	31.55	**31.57**	26.68	26.75	26.75	26.91	26.93	**26.98**
Average	31.84	31.68	32.04	32.23	32.23	**32.25**	27.33	27.25	27.60	27.79	27.82	**27.99**
	σ_e= 70						σ_e= 90					
	BM3D	EPLL	LRJS	WNNM	WSNM	OURS	BM3D	EPLL	LRJS	WNNM	WSNM	OURS
Butterfly	23.75	23.74	24.11	24.36	**24.51**	24.47	22.46	22.26	22.90	23.13	23.28	**23.31**
Penguin	27.29	27.30	27.00	27.39	27.39	**27.41**	26.28	26.34	26.02	26.34	26.26	**26.16**
House	27.91	27.32	28.20	28.60	28.69	**28.76**	26.47	26.03	26.91	27.27	**27.32**	27.28
Cameraman	24.61	24.59	24.71	24.85	24.89	**24.97**	23.52	23.44	23.64	23.78	23.80	**23.86**
Peppers	25.07	25.03	25.01	25.26	25.32	**25.37**	23.87	23.76	23.78	23.98	24.03	**24.29**
Average	25.73	25.60	25.81	26.09	26.16	**26.20**	24.52	24.37	24.65	24.90	24.94	**24.98**

[7] Zoran D, Weiss Y. From learning models of natural image patches to whole image restoration[C]//2011 international conference on computer vision. IEEE, 2011: 479-486.

[8] Zha Z, Wen B, Yuan X, et al. Low-Rank Regularized Joint Sparsity for Image Denoising[C]//2021 IEEE International Conference on Image Processing (ICIP). IEEE, 2021: 1644-1648.

[9] Liu X, Jing X Y, Tang G, et al. Image denoising using weighted nuclear norm minimization with multiple strategies[J]. Signal Processing, 2017, 135: 239-252.

[10] Xie Y, Gu S, Liu Y, et al. Weighted Schatten p-norm minimization for image denoising and background subtraction[J]. IEEE transactions on image processing, 2016, 25(10): 4842-4857.