

## Research Article

# Signal Reconstruction Based on Probabilistic Dictionary Learning Combined with Group Sparse Representation Clustering

Bin Liang<sup>1</sup> and Shuxing Liu<sup>2</sup>

<sup>1</sup>Jiuzhou Polytechnic, Xuzhou 221116, China

<sup>2</sup>Southwest China Institute of Electronic Technology, Chengdu 610036, China

Correspondence should be addressed to Bin Liang; [liangb@jzp.edu.cn](mailto:liangb@jzp.edu.cn)

Received 10 November 2020; Revised 29 November 2020; Accepted 3 December 2020; Published 12 December 2020

Academic Editor: Yi-Zhang Jiang

Copyright © 2020 Bin Liang and Shuxing Liu. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In order to make full use of nonlocal and local similarity and improve the efficiency and adaptability of the NPB-DL algorithm, this paper proposes a signal reconstruction algorithm based on dictionary learning algorithm combined with structure similarity clustering. Nonparametric Bayesian for Dirichlet process is firstly introduced into the prior probability modeling of clustering labels, and then, Dirichlet prior distribution is applied to the prior probability of cluster labels so as to ensure the analyticity and conjugation of the probability model. Experimental results show that the proposed algorithm is not only superior to other comparison algorithms in numerical evaluation indicators but also closer to the original image in terms of visual effects.

## 1. Introduction

Dictionary learning is an important model of image and communication signal processing [1]. It is a powerful tool for signal and image data classification, compression, denoising, repair, and even superresolution [2]. With the improvement of object detection technology, dictionary learning has become more and more important in the field of perception, image remote sensing, and meteorological and military reconnaissance [3].

In recent years, the nonparametric Bayesian methods have attracted much attention in dictionary learning. Compared with the traditional comprehensive dictionary learning algorithm represented by the method of optimal directions (MOD) [4] and K-singular value decomposition (K-SVD) [1, 4, 5], the existing dictionary learning algorithms have shown significant advantages. This superiority is mainly reflected in three aspects: (1) the number of dictionary atoms, signal sparsity, and regularization parameter values can be automatically inferred during the dictionary learning process so that there is no need to preset the initial values of these parameters and avoid improper initialization. (2) There is a strict theory to prove the convergence of the

dictionary algorithm and the optimality of the solution, but K-SVD and MOD still lack the support of theoretical proof. (3) Dictionary learning is based on the probabilistic graph model, which has a clear and open structure and is suitable for introducing various regularized probabilistic prior constraints [6].

In the field of signal reconstruction, many excellent algorithms have been proposed by scholars at home and abroad. Zhou et al. introduced nonparametric Bayesian image denoising and compression reconstruction and proposed the beta process factor analysis algorithm (BPFA) [7], which is a successful attempt in the field of dictionary learning, and have achieved good application effect. However, the BPFA algorithm does not take into account the image similarity and variability of similarity and variability of global structural, so there is still space of promotion to improve the reconstruction characteristics in the dictionary learning. Subsequent improved algorithms mainly focus on improving the structural characteristics of nonparametric Bayesian dictionary learning, for example, Dirichlet process-beta process factor analysis and subsequent improved algorithms mainly focus on improving the structural characteristics of nonparametric Bayesian dictionary learning,

and algorithms such as Dirichlet process-beta process factor analysis and dependent hierarchical beta process (DHBP) [8]. These algorithms introduce spatial correlation to nonparametric Bayesian modeling. The sparse representation of images under dictionary learning can reflect a certain similarity in adjacent space, which significantly improves the application performance of the dictionary in denoising, compression perception, etc., but the computation complexity is quite high and the running time is too long [9].

Nevertheless, the existing NPB-DL algorithm generally models the data samples in the training set as independent distribution, and the data structure information contained in the correlation between the samples is ignored [10], which restricts the further improvement of data fitting ability in NPB-DL dictionary. With its excellent clustering/classification capabilities, NPB is naturally suitable for mining the structural similarity in the sample data. The DP-BPFA algorithm and PSBP-BPFA algorithm are typical representatives of this work. The DP-BPFA algorithm performs spatial clustering of data through the Dirichlet process, thus realizing the mining of nonlocal similarity of data structures. However, it ignores the local similarity of data structures and cannot guarantee the spatial smoothness of the clustering results. The PSBP-BPFA algorithm introduces the spatial location information of the data into the Gaussian kernel function and then builds a multiprobit regression model and applies it to the stick-breaking construction of the prior probability of the clustering label in the form of a cumulative probability density function. Therefore, the local similarity constraint of the structure is introduced into the spatial clustering, and the spatial smoothness of the clustering is taken into account to a certain extent. The probability model of the PSBP-BPFA algorithm does not have conjugation and analytical closed form [11], so it is impossible to use the variational Bayesian (VB) method for efficient deterministic inference, while the traditional sampling strategy for inferential learning of regression coefficients costs a lot of computation cost. In addition, the algorithm also lacks a selection mechanism for the width of kernel. If an additional layer of probability model is added to the width, it will continue to increase the computational cost.

Since structural similarity and variability have important influence on sparse signal representation, literature [12] proposed a multidictionary learning algorithm on the basis of Dirichlet process for multisource data, which solved the sparseness problem well in the use of structural similarity. However, multidictionary learning is suitable for situations where the data source is not a single source. For the clustering effect of the internal structure of single source data, its applicability and robustness are inadequate. Under the traditional structure dictionary, the sparse representation of the image has blocking artifact, which is not conducive to improve the ability of dictionary to express the structural features of signals. Since the support sets of the sparse representation of similar image patches are usually similar, this priori is conducive to the joint sparse reconstruction under the multiobservation vector model, thus reducing the number of observations and improving the reconstruction performance, which helps to improve the ability of the

dictionary to express the signal structure characteristics, and the support set of the sparse representation of similar image blocks is usually similar, which is conducive to the joint under the multiobservation vector model. It can be seen from the abovementioned analysis that dictionary learning is based on the similarity clustering of image patches, and the reconstruction ability will be improved by introducing the block structure characteristics into the sparse representation as the regularization constraint of dictionary learning. Therefore, nonparametric Bayesian dictionary learning keeps its advantages over traditional structure dictionary because of its adaptability to image structure [13].

In order to make full use of nonlocal and local similarity and improve the efficiency and adaptability of the NPB-DL algorithm, this paper proposes a signal reconstruction algorithm based on dictionary learning algorithm combined with structure similarity clustering. Nonparametric Bayesian for Dirichlet process is firstly introduced into the prior probability modeling of clustering labels, and then, Dirichlet prior distribution is applied to the prior probability of cluster labels, and the MRF is used to parameterize it so as to ensure the analyticity and conjugation of the probability model. Experimental results show that the proposed algorithm is not only superior to other comparison algorithms in numerical evaluation indicators but also closer to the original image in terms of visual effects.

## 2. Nonparametric Bayesian for Dirichlet Process

Bayesian method plays an important role for probability statistics in the data analysis, and it is also a reliable method to solve uncertain problems. Bayesian method is usually used to model a certain distribution of the observed data and obtain the prior information of the sample parameters from the actual samples dataset. However, these prior assumptions cannot fully describe the characteristics of the unknown sample. Therefore, the nonparametric Bayesian modeling method has the parameter model of infinite dimension parameter space to represent the prior knowledge [7, 13–15].

In order to infer from the observed data, it is necessary to model the data generation mechanism. Usually, due to the lack of clear knowledge of data generation mechanism, only very general assumptions can be made, leaving a large part of unspecified mechanisms. In other words, the distribution of data is not determined by a limited number of parameters. This kind of nonparametric model can prevent from serious error specification in the process of the data generation mechanism, so the knowledge space with tiny details is distributed in infinite space and is not available for nonparametric problems. The commonly used nonparametric Bayesian modeling methods include Dirichlet process, beta process, and their hybrid model. This paper is mainly based on the nonparametric Bayesian model of Dirichlet process to improve the representation ability of dictionary learning and enhance the accuracy of signal reconstruction [16, 17].

Dirichlet process can be regarded as the standardization process of completely random measures, and a unified

framework is established. It can be considered that the framework is useful for understanding the commonly used prior behavior and developing new models. In fact, even if completely random measures are very complex probabilistic tools, their use in nonparametric Bayesian reasoning can lead to intuitive posterior structures. DP can be defined as follows: given the measurement space  $(\Theta, B)$ , it is assumed that  $G_0$  is a probability measure of  $(\Theta, B)$ . The distribution  $g$  of Dirichlet process is a random probability measure based on  $(\Theta, B)$ .  $(\Psi_1, \dots, \Psi_r)$  is defined as any finite partition for  $\Theta$ . Therefore, the random vector  $(G(\Psi_1), \dots, G(\Psi_r))$  is denoted as follows:

$$(G(\psi_1), \dots, G(\psi_r)) \sim \text{Dir}(\alpha_0 G_0(\psi_1), \dots, \alpha_0 G_0(\psi_r)), \quad (1)$$

where  $\alpha_0$  is a weight coefficient with positive number.

The concentration parameter  $\alpha_0$  controls the degree of dispersion of the distribution. The sampling position  $\theta_i$  becomes more concentrated with the increase of  $\alpha_0$  and more dispersed with the decrease of  $\alpha_0$ . When the value of  $\alpha$  tends to infinity,  $G_0$  becomes infinitely concentrated. The positions of each sampling point are almost different, and the weight of each position is almost close to 0. When  $\alpha$  approaches 0,  $G_0$  will be very discrete. There is only one sampling point on the  $G_0$  distribution, and the weight is close to 1. As shown in Figure 1, the atoms of  $G$  distribution become more and more concentrated with the increase of  $\alpha$  value. The discreteness of  $G_0$  is also called the centralization of  $G_0$ .

### 3. Probabilistic Modeling Based on Clustering Structure Similarity

All the  $M \times M$  pixel patches in the image are processed by column stack to form a training set  $(x_i)_{i=1}^N$  composed of  $N$  vectors with  $P$ -dimensional, where  $P = M \times M$ . If the unknown number of atom  $d_k$  in the dictionary is  $K$ -order truncation and its value is a large enough positive integer, the  $K$  is usually set as 256 or 512, so the dictionary is written as  $D = [d_1, d_2, \dots, d_K] \in R^{P \times K}$ . In other words, as for any data sample  $x_i$ , it can be expressed as the following equation [17]:

$$x_i = D(s_i \Theta z_i) + \varepsilon_i, \quad (2)$$

where  $s_i \in R^{K \times 1}$  is the weight vector; the binary vector  $z_i \in R^{K \times 1}$  is the sparse mode indicator vector;  $\Theta$  is the product with element-by-element; and  $\varepsilon_i \in R^{P \times 1}$  is denoted as the residual. The sparse pattern indicator vector  $z_i$  obeys the Bernoulli distribution parameterized by the probability of using atoms  $\pi_i \in R^{K \times 1}$ , as shown in equation (3), where  $z_{ik}$  and  $\pi_{ik}$  are the first  $k$ -th elements of  $z_i$  and  $\pi_i$ , respectively.

$$z_i \underset{i.i.d.}{\sim} \prod_{k=1}^K \text{Bernoulli}(z_{ik} | \pi_{ik}). \quad (3)$$

Spatial clustering of all data samples in the training set through the NPB probability model can realize the use of nonlocal similarity structures. The initial value of the

number of clusters is truncated into a larger positive integer, and samples belonging to the same cluster have the same probability of atom usage. This process can be expressed as the Dirichlet process; it can be described in

$$G_i \sim DP(a_0, G_0), \quad (4)$$

where  $G_i$  is the prior measure of cluster  $x_i$ ;  $G_0$  is the initial measurement; and  $\alpha_0$  is the scale parameter. According to the discrete property of Dirichlet process, equation (3) can be further expressed as the discrete weighting form of point measure in

$$G_i = \sum_{l=1}^L \beta_{il} \cdot \delta(\pi_i - \pi_l^*). \quad (5)$$

The atom usage probability  $\pi_l^*$  corresponds to a point measure and is obtained from the sampling process of base measure  $G_0$ . The mixed weight  $\beta_{il}$  is a function of scale parameter  $\alpha_0$ , which represents that the sample  $\pi_i$  belongs to cluster  $\pi_l$  and the atom usage probability is  $\pi_i = \pi_l^*$ . Therefore, the beta distribution of equation (6) is used as the concrete form of point-measurement, where  $\alpha_0 = b_0 = 1$ ,  $\pi_{lk}^*$  is the  $k$ th element of  $\pi_l^*$ ;  $\alpha_0$  is set to  $1/L$ . Obviously, the sum of mixed weights  $\beta_{i1}, \beta_{i2}, \dots, \beta_{iL}$ , should be 1, so the Dirichlet prior distribution for  $\beta_{i1}, \beta_{i2}, \dots, \beta_{iL}$  is shown in

$$\pi_i^* \underset{i.i.d.}{\prod_{k=1}^K} \text{Beta}\left(\pi_{lk}^* \middle| \frac{a_0}{K, b_0(K-1)/K}\right), \quad (6)$$

$$\{\beta_{il}\}_{l=1}^L \sim \text{Dirichlet}(\beta_{i1}, \beta_{i2}, \dots, \beta_{iL} | \lambda_{i1}, \lambda_{i2}, \dots, \lambda_{iL}), \quad (7)$$

where parameters  $\lambda_{i1}, \lambda_{i2}, \dots, \lambda_{iL}$  are all functions of  $\alpha_0$ . It can be seen from the abovementioned analysis that on the basis of the Dirichlet mixed model of equations (4)–(7), a cluster label is assigned to each sample  $x_i$ , and it is assumed that the multidistribution of equation (7) is obeyed, and the corresponding distribution parameters  $\beta_{i1}, \beta_{i2}, \dots, \beta_{iL}$  are the mixed weight of the model.  $\pi_i$  corresponding to  $x_i$  can be expressed as equation (9). It can be seen that no matter how large the number of samples in the training set is, the values of  $\pi_i$  is always limited. In fact, with the adaptive inference of the final cluster number, the number of inferential learning  $\pi_i$  will be significantly less than  $L$ .

$$t_i \underset{i.i.d.}{\text{Multi}}(t_i = 1, t_i = 2, \dots, t_i = L | \beta_{i1}, \beta_{i2}, \dots, \beta_{iL}), \quad (8)$$

$$\pi_i = \pi_{t_i}^*, t_i = l, l \in \{1, 2, \dots, L\}. \quad (9)$$

The abovementioned analysis is based on clustering to realize the utilization of nonlocal similarity of data structure.

In order to take into account the local similarity, this paper establishes MRF in the prior probability distribution of cluster labels to improve the spatial smoothness of clustering. The specific method is to take the pixel patch corresponding to any data sample  $x_i$  as the center, the corresponding samples of 33 neighborhood pixel patches are stacked as the dataset  $\Omega_i$ , and the cluster labels of the samples in the set are smoothed in spatial domain to obtain the MRF

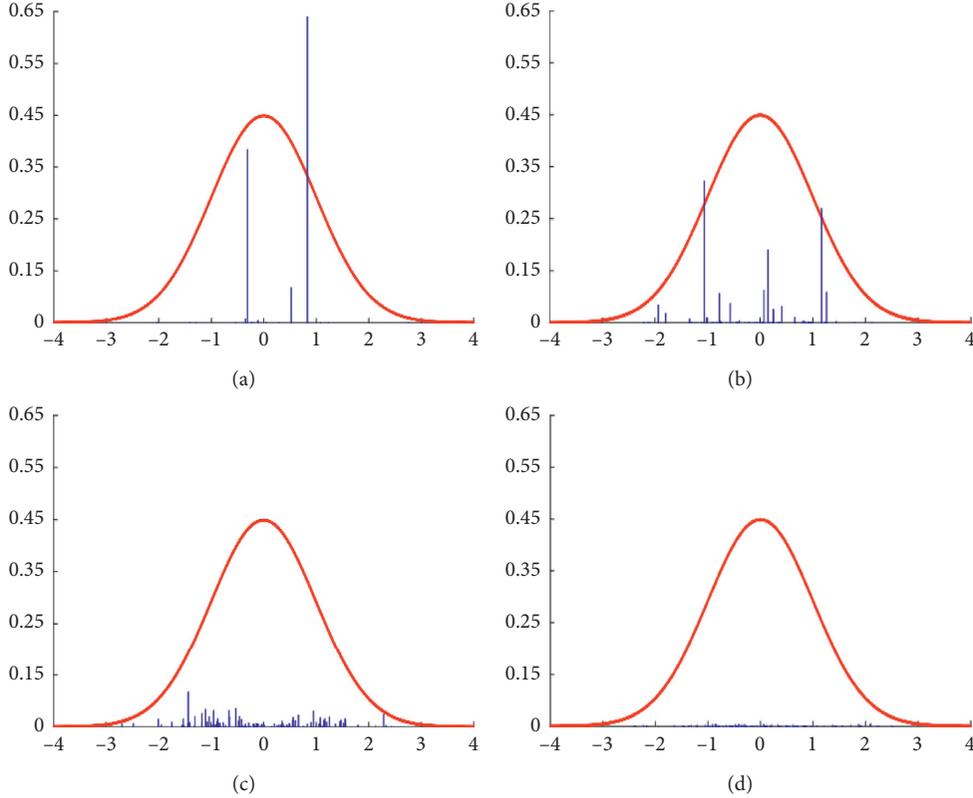


FIGURE 1: Comparison of distribution under the different parameter  $\alpha$ . (a)  $\alpha = 0.5$ , (b)  $\alpha = 6$ , (c)  $\alpha = 50$ , and (d)  $\alpha = 300$ .

factor  $G_{il}$ , as shown in equation (10), and then, the prior parameter  $\lambda_{i1}, \lambda_{i2}, \dots, \lambda_{iL}$ , of mixed weights  $\beta_{i1}, \beta_{i2}, \dots, \beta_{iL}$ , can be denoted as the product of  $\alpha_0$  and  $G_{i1}, G_{i2}, \dots, G_{iL}$ , namely  $\lambda_{il} = \alpha_0 G_{il}$ .

$$G_{il} = \frac{1}{|\Omega_i|} \sum_{x_m \in \Omega_i} I(t_m = l). \quad (10)$$

The larger the  $G_{il}$  value, the more concentrated the clustering is in the neighborhood data system, the greater the corresponding hybrid weight  $\beta_{il}$ , and the higher the probability that the sample belongs to the main cluster  $x_i$ , which makes the data sample and most of the neighborhood samples belong to the same cluster with a higher probability, thus reflecting the local similarity of the structure and improving the spatial smoothness of the clustering results. Compared with the PSBP algorithm, the first advantage of the probability model is the analyticity and conjugation, which can be used for more efficient deterministic inference. In addition, MRF is constructed by smoothing the neighborhood clustering labels so that the local similarity constraint on the clustering prior probability is only generated by a small number of samples in the neighborhood and does not need to calculate the kernel function value for all samples one by one like PSBP, which is also conducive to saving calculation expenses. The atoms  $d_k$ , weights  $s_i$ , and residuals  $\varepsilon_i$  in equation (2) follow the prior distribution provided by BPFA framework, where  $\gamma_s$  and  $\gamma_\varepsilon$  are precision parameters

of weight and residual, respectively;  $c_0, d_0, e_0$ , and  $f_0$  are super parameters for gamma distribution.

#### 4. Group Sparse Representation for Dictionary Learning

In Section 3, we mainly discussed probabilistic modeling based on clustering structure similarity. As we all know, there are a lot of similar redundant components in the image, and its sparse representation in the traditional structure dictionary also has similarity, which is reflected in the similarity of support sets [18]. Thus, the global similarity measure is introduced into the dictionary learning framework to improve the ability of the dictionary to express the image structure features [19]. This paper adopts clustering structure similarity method which is a simple and effective unsupervised clustering method. Under the condition that the data are not classified or labeled, the corresponding clustering centers are randomly selected according to the preset number of clusters. By calculating the similarity measure between each image patch and each cluster center, the cluster attribution of image patches is determined. Then, the average value of all image patches in each cluster is calculated. The abovementioned process is executed iteratively until each cluster center has no obvious deviation from the cluster center of the previous iteration or reaches the upper limit of the predetermined number of iterations.

However, the purpose of signal reconstruction is to recover the clear image  $x_i$  from the degraded image  $y_i$ . Therefore, the optimization problem is transformed into the following objective function:

$$\hat{\alpha}_G = \arg \min \frac{1}{2} \|\Phi D_G \circ \alpha_G\|_2^2 + \lambda \|\alpha_G\|_0, \quad (11)$$

where  $\alpha_G$  represents the concatenation of sparse matrices  $x_{G_k}$  of all similar patches in an image  $x$ ;  $D_G$  is the concatenation of the sparse dictionary  $D_{G_k}$  obtained by adaptive learning of a group of similar patches;  $D_G \circ \alpha_G$  represents the reconstructed image;  $\circ$  represents an abstract operation that is the multiplication of the sparse basis and the sparse coefficient. The objective function of dictionary learning in equation (11) has two optimization variables, the dictionary, and the coefficient matrix. According to the optimization theory, other variables are mostly fixed and one of the variables is updated in turn during the iterative operation. Therefore, the learning strategy is also used to first fix the dictionary and update the coefficient matrix; then, the coefficient matrix is fixed so as to update the dictionary [20].

## 5. Experimental Results and Analysis

This paper focuses on the performance of function optimization and image denoising to evaluate the effectiveness of the proposed reconstruction algorithm. Gray-scale images such as Barbara, Lena, Boath, House, and Peppers in the field of image processing are used as experimental objects, and MATLAB for simulation analysis is used [21]. The experimental environment of this experiment: Pentium-IV processor with 2.4 GHz, 2 GB memory, Windows 7, and MATLAB 7.10 programming environment.

**5.1. Parameter Setting.** In the dictionary learning stage, 105 pairs of  $9 \times 9$  pixel image patch pairs are randomly selected for dictionary learning, and the maximum overlap between patches is 4. Singular value decomposition is used to initialize dictionary elements and parameter values. The initial values of hyperparameter are  $c = d = e = 10^{-6}$ ,  $c_0 = 2$ , and  $\eta_0 = 0.55$ . The dictionary size is 2048. The first 1500 samples are optimized and the last 1500 samples approximate the dictionary and the posterior distribution of the parameters. In the experiment, the size of image patch  $S \times S$  is set to  $9 \times 9$ ; the size of nonlocal search window is set to  $31 \times 31$ ; the number of similar blocks  $K$  is set to 16; the number of clustering classes  $C$  is set to 40; the number of iterations is set to 9; and the residual compensation coefficient is set to 0.03. In the reconstruction phase, the number of iterations is 40. In this paper, only gray image is reweighted on the basis of nonparametric Bayesian model; since color image is composed of three channels of gray image, its processing flow is the same. The resolution of all images is the same, which is  $512 \times 512$ ; the processing uses sliding-window method to divide the image into 200000 patches. The image patches with overlapping resolution of 8 pixels  $\times$  8 pixels and the gray values of all image patches are processed by column stack to form the dictionary dataset. According to the flow of

the algorithm proposed in this paper, the number of atoms in the dictionary is set to 512.

**5.2. Image Quality Assessment.** Objective quantitative evaluation is the quantitative analysis of reconstruction results using the algorithm model, which has the characteristics of objectification, standardization, accuracy, quantification, and simplification [22]. The selected quantitative index belongs to the full reference image quality evaluation method. The pixel statistics, information entropy, structural information, and other parameters between the processed images of reconstruction algorithms and the high-quality images are evaluated objectively [23]. In our experiment, the peak signal-to-noise (PSNR) and structural similarity index measurement (SSIM) are used to quantitatively describe the performance of different comparison algorithms. In addition, the average running time is used to measure the efficiency of the comparison algorithms. PSNR is used to count the difference between the evaluated image and the reference image and describes the reconstruction performance of the algorithm. The mathematical expression is written as follows:

$$\text{PSNR} = 10 \log \frac{255^2 / \sum_{i=1}^M (x(i, j) - y(i, j))^2}{M \times N}. \quad (12)$$

SSIM is designed to measure the structural clarity of the image, whose value range is 0~1. The larger the value is, the better the effect of image reconstruction will be. The equation is denoted as follows:

$$\text{SSIM}(x, y) = \frac{2u_x u_y + C_1}{u_x^2 + u_y^2 + C_1} \times \frac{2\sigma_x \sigma_y + C_2}{\sigma_x^2 + \sigma_y^2 + C_2} \times \frac{\sigma_{xy} + C_3}{\sigma_x \sigma_y + C_3}. \quad (13)$$

**5.3. Quantitative Analysis for Image Reconstruction.** In order to facilitate the comparative analysis, the experiment is carried out from the function optimization and image denoising based on dictionary learning algorithms [24]. The involved comparison algorithms in the experiment include BPFA [5], DP-BPFA [8], PSBP-BPFA [12], SSIM-BPFA [13], and K-SVD [6]. Before the comparison experiment, the clustering effect of our proposed algorithm in this paper is given in the dictionary learning process, where three images of Lena, Peppers, and Mandrill were selected for comparative testing, as shown in Figure 2. The clustering is based on the difference of the probability of using atoms, and the image patches of the same cluster are obtained by taking the same pixel value, where  $L$  is the number of clusters inferred by our proposed algorithm. The number of atoms selected from left to right are 4, 8, and 12. It can be seen that the more the atoms, the better the clustering effect. When the number of atoms is 12 in dictionary, a lot of artifacts appear in the clustering results. When the number of atoms is 8, the effect is the best.

This experiment adds Gaussian white noise with standard deviation  $\sigma^2 = 5, 15, 25, 20,$  and  $35$  to the natural image and uses the abovementioned comparison reconstruction

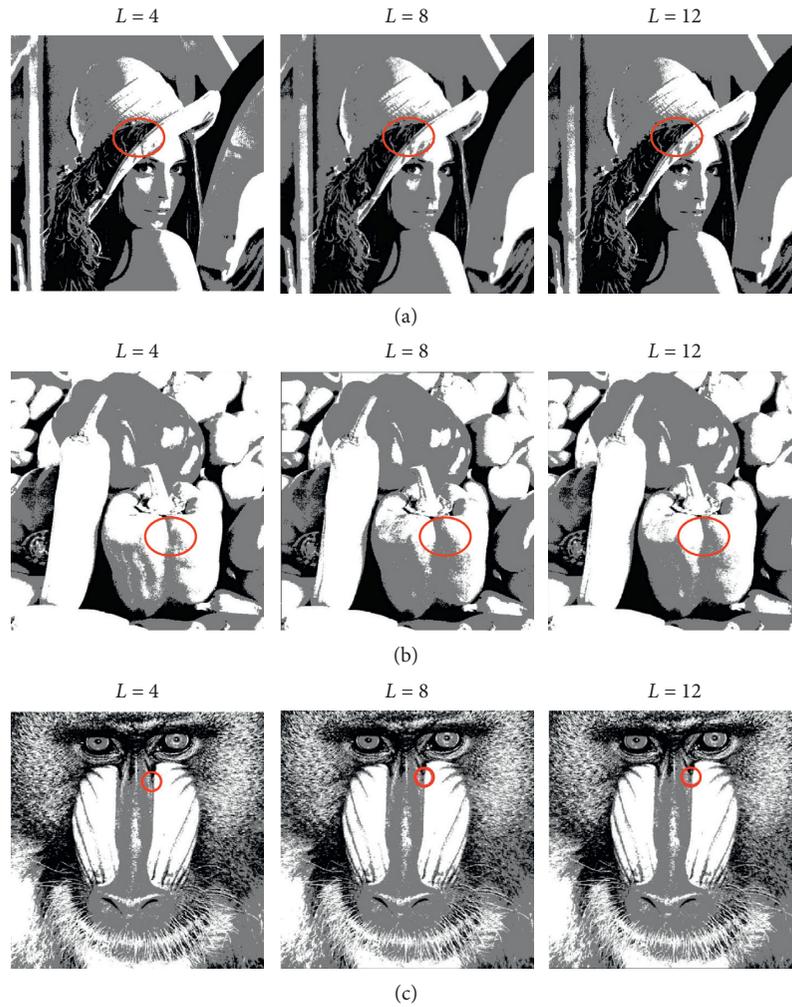


FIGURE 2: Clustering effect under different number of atoms. (a) Lena; (b) Pepper; (c) Mandril.

algorithms and the proposed algorithm to perform the image reconstruction [25–32]. For the experimental results, this paper also compares these reconstruction results from the subjective and objective aspects. The subjective evaluation mainly selects the house with more smooth areas and the Barbara image, with rich texture information [26, 28], and analyzes the performance of the reconstruction method from the visual effect of the image; the objective evaluation uses the PSNR and SSIM to measure the performance of the reconstruction effect from the objective indicators. Figure 3 shows the reconstruction effect of different comparison algorithms on different noisy images. The overall denoising PSNR of the six algorithms from low to high is BPFA algorithm, K-SVD algorithm, DP-BPFA algorithm, PSBP-BPFA algorithm, SSIM-BPFA algorithm, and our proposed algorithm in this paper. The difference or gap between K-SVD and BPFA is very small. They have not used the structure feature of the sample data; the gap between DP-BPFA and PSBP-BPFA is relatively small; they use the structural information in the data, but this texture structural information is not fully utilized. The SSIM-BPFA algorithm introduces the block structure feature of sparse representation on the basis of data

clustering preprocessing, and its data fitting accuracy is significantly improved compared with the previous four algorithms [29].

From the overall reconstruction effect in Figure 3, it can be seen that the reconstruction image shown in Figure 3(c) has severe speckle noise; the reconstruction image shown in Figure 3(d) has severe scratches; Figure 3(e) is the denoised image of PSBP-BPFA algorithm; Figure 3(f) has slight scratches on the edge of the image; the denoised image shown in Figure 3(g) has slight speckle noise; the denoising effect of the denoising image in Figure 3(h) shown is better, especially the smooth area of the image. From the above-mentioned analysis, it can be seen that for the house image with a noise level of 35, the proposed reconstruction algorithm has a better reconstruction effect in the smooth area of the image, and the detail retention ability has obvious advantages over other methods.

In order to further analyze the effect of these methods on the preservation of the details of the house image, Figure 4 is an enlarged image of the partial details. Figure 4 shows the reconstruction results of the enlarged area of all the comparison algorithms. Due to space constraints, we only chose

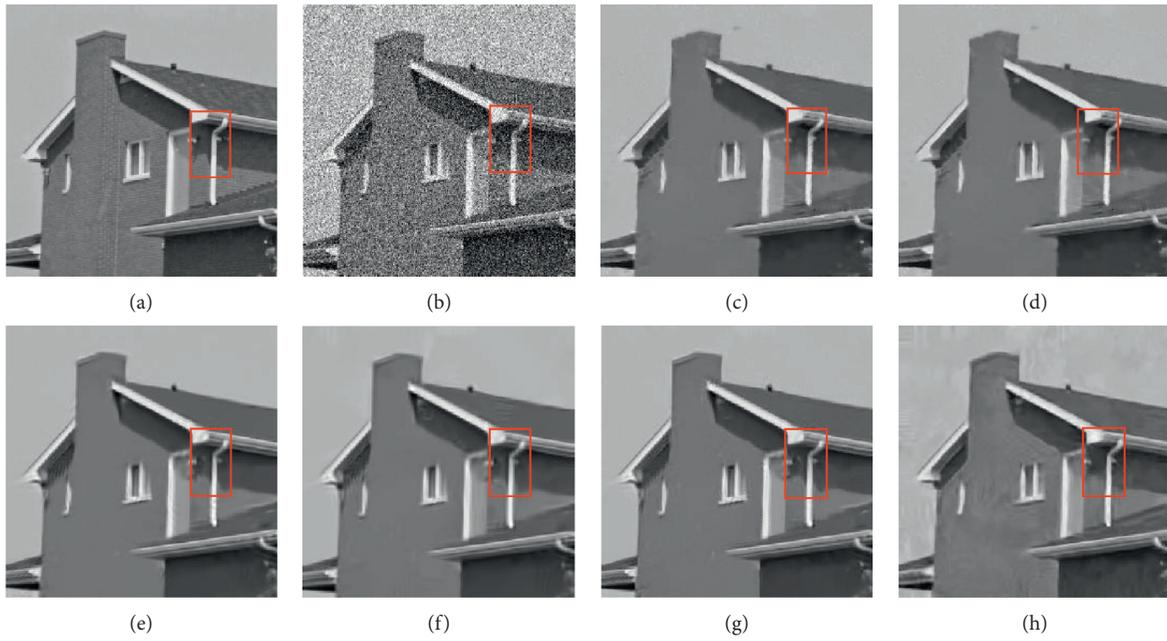


FIGURE 3: Reconstruction effect of different algorithms for house with  $\sigma^2 = 35$ . (a) Raw image. (b) Noisy image. (c) BPFA. (d) DP-BPFA. (e) PSBP-BPFA. (f) SSIM-BPFA. (g) Proposed algorithm. (h) K-SVD.

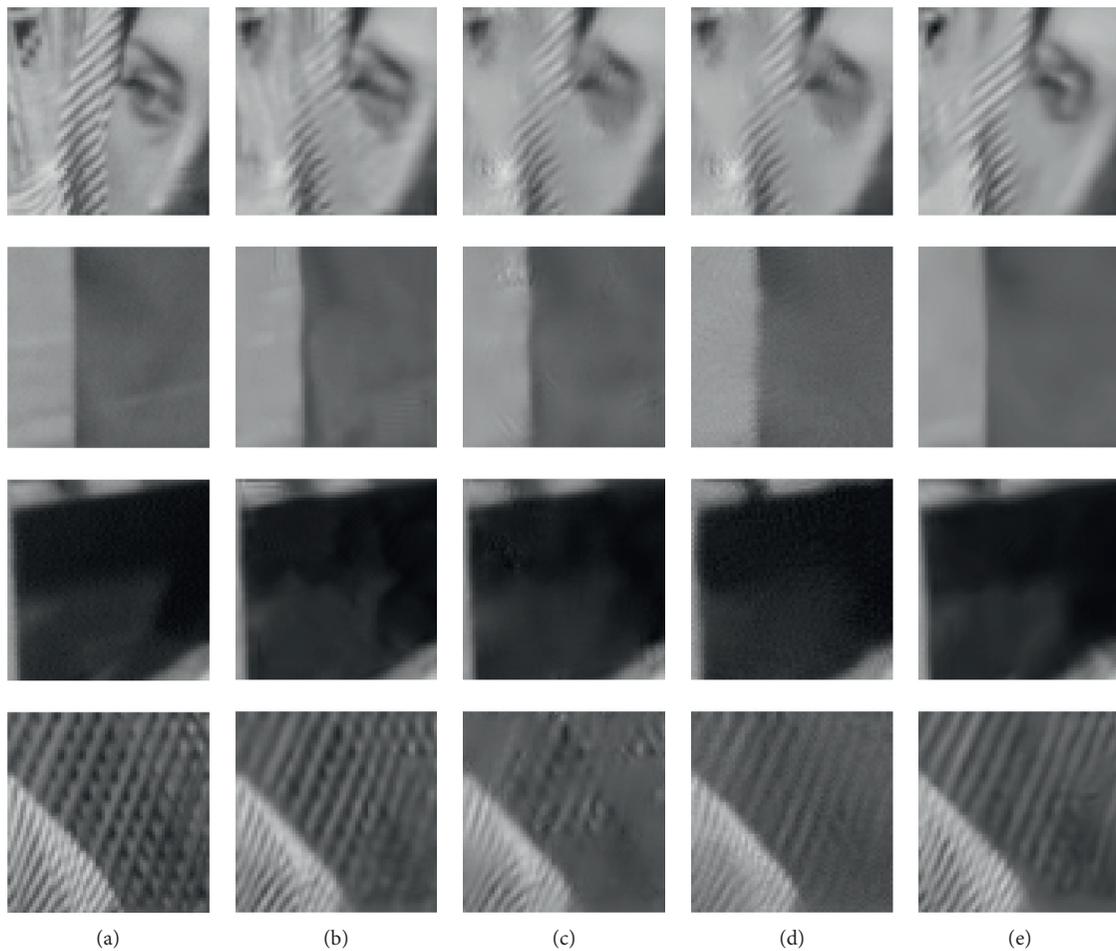


FIGURE 4: The local detail and structure comparison for different algorithms. (a) Raw image. (b) DP-BPFA. (c) PSBP-BPFA. (d) SSIM-BPFA. (e) Proposed algorithm.

TABLE 1: The SSIM and PSNR results for different standard algorithms.

Image	$\sigma$	BPFA		K-SVD		SSIM-BPFA		PSBP-BPFA		DP-BPFA		Proposed algorithm	
		SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR
Barbara	5	0.94	37.06	0.96	38.11	0.96	38.33	0.95	38.48	0.95	36.61	0.96	38.51
	10	0.92	33.16	0.93	34.42	0.94	34.97	0.90	34.97	0.92	33.03	0.94	34.93
	20	0.86	30.26	0.88	30.81	0.92	31.74	0.84	31.57	0.85	29.05	0.91	31.59
	25	0.82	29.08	0.85	29.58	0.88	30.72	0.81	30.47	0.80	27.27	0.89	30.51
	50	0.67	25.66	0.71	25.58	0.79	27.23	0.69	27.06	0.57	23.20	0.80	27.18
Pepper	5	0.96	35.15	0.98	36.66	0.98	36.59	0.89	36.70	0.98	35.85	0.98	36.71
	10	0.94	30.99	0.96	32.42	0.96	32.43	0.87	32.57	0.96	31.55	0.96	32.62
	20	0.88	27.25	0.92	28.45	0.92	28.87	0.78	28.78	0.92	28.14	0.93	28.87
	25	0.85	26.14	0.89	27.25	0.91	27.70	0.74	27.62	0.90	26.99	0.91	27.78
	50	0.71	22.96	0.77	23.25	0.83	24.53	0.56	24.25	0.77	22.34	0.83	24.55
Lena	5	0.91	37.94	0.94	38.62	0.94	38.71	0.95	38.69	0.92	37.61	0.95	38.61
	10	0.90	34.36	0.91	35.51	0.91	35.94	0.89	35.83	0.90	35.12	0.91	35.58
	20	0.83	31.61	0.86	32.38	0.87	33.07	0.81	32.90	0.86	32.37	0.87	32.77
	25	0.79	30.36	0.84	31.36	0.86	32.08	0.77	31.87	0.84	31.34	0.86	31.63
	50	0.67	27.32	0.76	27.75	0.79	29.05	0.63	28.87	0.55	25.15	0.80	28.81
House	5	0.89	37.01	0.95	37.50	0.95	37.80	0.95	37.89	0.93	36.24	0.95	37.86
	10	0.87	33.11	0.90	33.55	0.90	33.94	0.90	34.06	0.88	33.09	0.90	33.97
	20	0.78	29.72	0.81	30.05	0.83	30.54	0.81	30.64	0.81	30.12	0.83	30.60
	25	0.72	28.38	0.77	29.08	0.80	29.62	0.82	29.63	0.78	29.22	0.80	29.59
	50	0.59	25.33	0.66	26.08	0.70	26.81	0.70	26.69	0.56	24.38	0.70	26.71

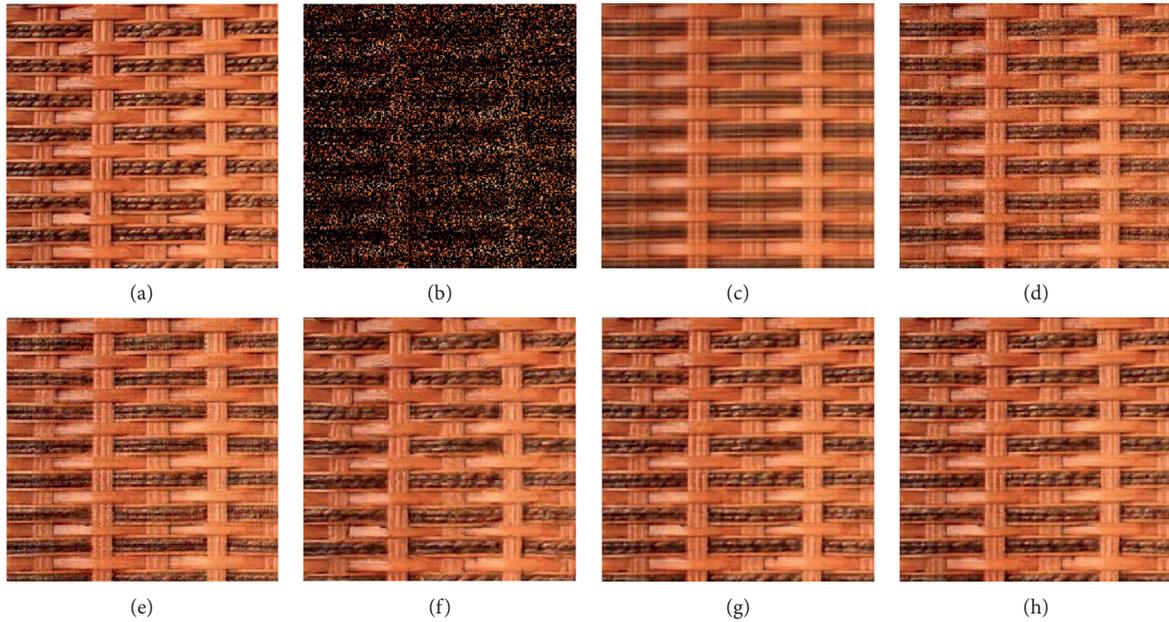


FIGURE 5: The reconstruction results of different algorithms for missing data in the image.

the best representative algorithm for comparison. It can be seen from the Barbara detail map that DP-BPFA and PSBP-BPFA have speckle noise [30, 31]; the detail image of figure DP-BPFA is severely scratched; the image texture information in PSBP-BPFA is relatively smooth and lost more information; for the left eye in Barbara, the image of SSIM-BPFA has slight scratches, but the structure is better preserved; our proposed algorithm has better reconstruction effects. The

SSIM and PSNR results for different standard algorithms are shown in Table 1. Experimental results show that the proposed algorithm obtains more reliable reconstruction results for natural images. Although the proposed algorithm achieves higher reconstruction probability and accuracy than traditional algorithms, its essence is to use correlation as a similarity measure to guide the choice of atoms, which will undoubtedly be restricted by RIP. It is worth exploring to find

other more effective similarity measures, such as Euclidean distance and Chebyshev distance [32].

Compared with the SSIM-BPFA algorithm, the advantage of our proposed algorithm in this paper is that it is more adaptive, and the clustering effect is more matched with dictionary learning, so it reflects the advantages of reconstruction. Since the signal-to-noise ratio in denoising applications has no significant impact on the SSIM value and algorithm running time, especially the SSIM values of several algorithms are close to 1, the experimental results of these two aspects will be reflected in the image reconstruction experiment to observe the impact of missing rates on them [33]. Figure 5 shows the experimental results of interpolation restoration of some images under 80% random pixel loss. Figure 5(b) is the damaged image after pixel missing. The remaining results are DP-BPFA, PSBP-BPFA, SSIM-BPFA, and our proposed algorithm in this paper, the PSNR value of our proposed algorithm in this paper is higher than other algorithms in general, and the visual effect is relatively the best. Even if the image data loss rate is from 10% to 90%, the average PSNR and SSIM results of the comparison algorithm for repair are lower than the proposed algorithm in this paper, where the experimental results are calculated from the average value after 50 independent experiments.

Figure 5 shows that the PSNR of all algorithms decreases when the data missing rate increases. The K-SVD algorithm shows advantages over BPFA and DP-BPFA algorithms under the condition of low missing rates. The PSNR of the K-SVD algorithm under the condition of high missing rate is the lowest among all algorithms, which shows that the proposed algorithm is more robust than the K-SVD algorithm. Therefore, the proposed algorithm in this paper has better adaptability to reflect the accuracy advantage in the image reconstruction.

## 6. Conclusion

This paper proposes a signal reconstruction algorithm based on adaptive dictionary learning, which overcomes the shortcomings of traditional sparse-based methods that are too restrictive and not robust enough for natural images. In addition, it also greatly improves the problem that the sparse model is easy to lose the global structure under the reconstruction framework. In this paper, an alternate optimization algorithm is designed to solve the nonparametric model, and experiments have proved that our proposed algorithm has good convergence and stability. Experimental results show that the proposed algorithm is not only superior to other comparison algorithms in numerical evaluation indicators but also closer to the original image in terms of visual effects. Although our proposed algorithm has made some progress in reconstruction accuracy, it still has space for improvement. For example, designing evolutionary operators with richer prior information or exploring more concise coefficient solving methods will help to further improve computational efficiency and accuracy.

## Data Availability

The labeled dataset used to support the findings of this study are available from the corresponding author upon request.

## Conflicts of Interest

The authors declare no conflicts of interest.

## References

- [1] W. Dong, "Sparsity-based image denoising via dictionary learning and structural clustering," *Computer Vision & Pattern Recognition IEEE*, vol. 25, no. 19, pp. 157–168, 2019.
- [2] D. Daoguang, "Nonparametric Bayesian dictionary learning algorithm based on structural similarity," *Journal on Communications*, vol. 15, pp. 17–28, 2019.
- [3] M. T. Wang, "A calculation mechanism for similarity measure with clustering an unbalanced hierarchical terminology structure," *International Conference on Parallel Processing Workshops IEEE*, vol. 25, no. 9, pp. 217–228, 2015.
- [4] F. J. Dekker, M. A. Koch, and H. Waldmann, "Protein structure similarity clustering (PSSC) and natural product structure as inspiration sources for drug development and chemical genomics," *Current Opinion in Chemical Biology*, vol. 9, no. 3, pp. 232–239, 2005.
- [5] L. Li, Jorge Silva, M. Zhou, and C. Lawrence, "Online bayesian dictionary learning for large datasets," in *Proceedings of the International Conference on Acoustics, Speech and Signal Processing (ICASSP2012)*, Kyoto, Japan, March 2012.
- [6] Rodrigues and P. G. L. M. Joo, "Clustering biomolecular complexes by residue contacts similarity," *Proteins-structure Function & Bioinformatics*, vol. 80, pp. 1810–1817, 2012.
- [7] M. Zhou, H. Yang, G. Sapiro, D. Dunson, and C. Lawrence, "Landmark-Dependent Hierarchical Beta Process for Robust Sparse Factor Analysis," in *Proceedings of the ICML2011 Structured Sparsity Workshop*, Bellevue, WA, USA, June 2011.
- [8] M. Zhou, H. Yang, G. Sapiro, D. Dunson, and C. Lawrence, "Covariate-dependent dictionary learning and sparse coding," in *Proceedings of the International Conference on Acoustics, Speech and Signal Processing (ICASSP2011)*, Prague, Czech Republic, May 2011.
- [9] L. Li, M. Zhou, E. Wang, and C. Lawrence, "Joint dictionary learning and topic modeling for image clustering," in *Proceedings of the International Conference on Acoustics, Speech and Signal Processing (ICASSP2011)*, Prague, Czech Republic, May 2011.
- [10] M. Zhou, C. Wang, M. Chen et al., "Nonparametric Bayesian Matrix Completion," in *Proceedings of the IEEE Sensor Array and Multichannel Signal Processing Workshop (SAM2010)*, Jerusalem, Israel, October 2010.
- [11] J. Paisley, M. Zhou, G. Sapiro, and C. Lawrence, "Nonparametric image interpolation and dictionary learning using spatially-dependent dirichlet and beta process priors," in *Proceedings of the International Conference on Image Processing (ICIP)*, Hong Kong, China, September 2010.
- [12] M. Zhou, J. Paisley, and C. Lawrence, "Nonparametric learning of dictionaries for sparse representation of sensor signals," in *Proceedings of the 3rd IEEE Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP2009)*, pp. 237–240, Aruba, December 2009.
- [13] M. Zhou and X. Li, "A variable step-size for frequency-domain acoustic echo cancellation," in *Proceedings of the 2007 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA'07)*, pp. 303–306, New Paltz, NY, USA, October 2007.
- [14] J. Chen and L.-P. Chau, "Multiscale dictionary learning via cross-scale cooperative learning and atom clustering for visual

- signal processing,” *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 25, no. 9, pp. 1457–1468, 2015.
- [15] F. Xiang, Z. Wang, and X. Yuan, “Image reconstruction based on sparse and redundant representation model: local vs nonlocal,” *Optik-International Journal for Light and Electron Optics*, vol. 124, no. 18, pp. 3636–3641, 2013.
- [16] X. Song, “Collaborative representation based face classification exploiting block weighted LBP and analysis dictionary learning,” *Pattern Recognition*, vol. 88, pp. 127–138, 2019.
- [17] Y. Zhang, W. Chen, Y. Chen, and J. Duan, “Sparsity-based inverse half-toning via semi-coupled multi-dictionary learning and structural clustering Engineering Applications of Artificial Intelligence,” *Engineering Applications of Artificial Intelligence*, vol. 72, pp. 43–53, 2018.
- [18] G.-M. Zhang, T.-L. Mai, and C.-M. Chen, “Clustering and visualizing similarity networks of membrane proteins,” *Proteins: Structure, Function, and Bioinformatics*, vol. 83, no. 8, pp. 1450–1461, 2015.
- [19] P. Zhu, Q. Hu, C. Zhang, and W. Zuo, “Subspace clustering guided unsupervised feature selection,” *Pattern Recognition*, vol. 66, pp. 364–374, 2017.
- [20] D. Zhu, “On clustering heterogeneous data and clustering by compression,” *Annals of Daaam & Proceedings*, pp. 133–138, 2009.
- [21] D. Zhu and D. Carstoiu, “On clustering heterogeneous data and clustering by compression,” in *Proceedings of the Annals of DAAAM*, vol. 24, no. 18, pp. 133–138, Zadar, Croatia, 2009.
- [22] Z. Zha and X. Zhang, “Compressed sensing image reconstruction via adaptive sparse nonlocal regularization,” *The Visual Computer*, vol. 34, no. 1, pp. 117–137, 2018.
- [23] H. Liu, C. Zhang, and I.-M. Kim, “Coupled online robust learning of observation and dictionary for adaptive analog-to-information conversion,” *IEEE Signal Processing Letters*, vol. 26, no. 1, pp. 139–143, 2019.
- [24] T. Sun, B. Jia, G. Li, and E. Gao, “Compressive sensing method to leverage prior information for submerged target echoes,” *The Journal of the Acoustical Society of America*, vol. 144, no. 3, pp. 1406–1415, 2018.
- [25] C. Blondel, “Super-resolution CT image reconstruction based on dictionary learning and sparse representation,” *Entific Reports*, vol. 144, pp. 1406–1415, 2018.
- [26] N. Kumar and R. Sinha, “Improved structured dictionary learning via correlation and class based block formation,” *IEEE Transactions on Signal Processing*, vol. 66, no. 19, pp. 5082–5095, 2018.
- [27] D. Du, “Compressive sensing image recovery using dictionary learning and shape-adaptive DCT thresholding,” *Magnetic Resonance Imaging*, vol. 55, pp. 60–71, 2018.
- [28] F. Veshki and S. A. Vorobyov, “An efficient coupled dictionary learning method,” *IEEE Signal Processing Letters*, vol. 99, p. 1, 2019.
- [29] X. Meng, “Estimation of chirp signals with time-varying amplitudes,” *Signal Processing*, vol. 147, pp. 1–10, 2018.
- [30] S. Sardellitti, S. Barbarossa, and P. Di Lorenzo, “Graph topology inference based on sparsifying transform learning,” *IEEE Transactions on Signal Processing*, vol. 7, pp. 1–14, 2019.
- [31] J. Wang, J.-F. Cai, Q. Zhu, Y. Shi, and B. Yin, “Multi-direction dictionary learning based depth map super-resolution with autoregressive modeling,” *IEEE Transactions on Multimedia*, vol. 22, no. 6, pp. 1470–1484, 2020.
- [32] D. Xu, F. Xia, X. Yang, and C. Tang, “Joint dictionary and graph learning for unsupervised feature selection,” *Applied Intelligence*, vol. 50, no. 5, pp. 1379–1397, 2020.
- [33] T. Trigano, S. Vaknin, and D. Luengo, “Fast proximal optimization for sparse reconstruction with dictionaries based on translated waveforms,” *Signal Processing*, vol. 169, Article ID 107379, 2019.