Notice of Violation of IEEE Publication Principles

"Single-Image Super-Resolution Algorithm Based on Structural Self-Similarity and Deformation Block Features"
by Yuantao Chen, Jin Wang, Xi Chen, Mingwei Zhu, Kai Yang, Zhi Wang, and Runlong Xia
in IEEE Access, April 2019

After careful and considered review of the content and authorship of this paper by a duly constituted expert committee, this paper has been found to be in violation of IEEE's Publication Principles.

This paper is a translation and duplication of the content from the paper cited below. The original content was copied without attribution (including appropriate references to the original author(s) and/or paper title) and without permission.

Due to the nature of this violation, reasonable effort should be made to remove all past references to this paper, and future references should be made to the following article:

"Single image super resolution algorithm based on structural self-similarity and deformation block feature"
by Wen Xiang, Ling Zhang, Yunhua Chen, Qiumin Ji
in the Journal of Computer Applications (39) 1, June 2018

IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received March 25, 2019, accepted April 11, 2019, date of publication April 18, 2019, date of current version May 16, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2911892

Single-Image Super-Resolution Algorithm Based on Structural Self-Similarity and Deformation Block Features

YUANTAO CHEN[1,2], JIN WANG[1,2], XI CHEN[1,2], MINGWEI ZHU[1,2], KAI YANG[3], ZHI WANG[1,2], AND RUNLONG XIA[4]

[1]School of Computer and Communication Engineering, Changsha University of Science and Technology, Changsha 410114, China
[2]Hunan Provincial Key Laboratory of Intelligent Processing of Big Data on Transportation, Changsha University of Science and Technology, Changsha 410114, China
[3]Technical Quality Department, Hunan ZOOMLION Heavy Industry Intelligent Technology Corporation Ltd., Changsha 410005, China
[4]Hunan Institute of Scientific and Technical Information, Changsha 411105, China

Corresponding author: Jin Wang (jinwang@csust.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant U1836208, Grant 61811530332, and Grant 61811540410, in part by the Open Research Fund of Hunan Provincial Key Laboratory of Intelligent Processing of Big Data on Transportation under Grant 2015TP1005, in part by the Changsha Science and Technology Planning under Grant KQ1703018, Grant KQ1706064, and Grant KQ1703018-01, in part by the Research Foundation of Education Bureau of Hunan Province under Grant 17A007, in part by the Teaching and Reforming Project of Changsha University of Science and Technology under Grant JG1755, Grant JG1711, Grant JG1615, Grant JG201815, Grant CN1501, Grant XJT[2015]291 No.156, Grant XJT[2016]400 No.219, Grant XJT[2017]452 No.132, and Grant XJT[2018]436 No.193, and in part by the Changsha Industrial Science and Technology Commissioner under Grant 2017-7.

ABSTRACT To solve the problem of insufficient sample resources and poor noise immunity in single-image super-resolution (SR) restoration procedure, the paper has proposed the single-image SR algorithm based on structural self-similarity and deformation block features (SSDBF). First, the proposed method constructs a scale model, expands the search space as much as possible, and overcomes the shortcomings caused by the lack of a single-image SR training sample; Second, the limited internal dictionary size is increased by the geometric deformation of the sample block; Finally, in order to improve the anti-noise performance of the reconstructed picture, a group sparse learning dictionary is used to reconstruct the pending image. The experimental results show that, compared with state-of-the-art algorithms such as bicubic interpolation (BI), sparse coding (SC), deep recursive convolutional network (DRCN), multi-scale deep SR network (MDSR), super-resolution convolutional neural network (SRCNN) and second-order directional total generalized variation (DTGV). The SR images with more subjective visual effects and higher objective evaluation can be obtained through the proposed method. Compared with existing algorithms, the structural network converges more rapidly, the image edge and texture reconstruction effects are obviously improved, and the image quality evaluation, such as peak signal-noise ratio (PSNR), root mean square error (RMSE), and structural similarity (SSIM), are also superior and popular in image evaluation.

INDEX TERMS Super-resolution, deformation block features, block matching, dictionary learning.

I. INTRODUCTION

With the development of information technology, Low-resolution (LR) images have been difficult to meet the demand for high-definition images. High-resolution (HR) images have high pixel density, clear details and rich information, which can meet the practical application requirements of image analysis and image understanding. The single-image super-resolution (SR) algorithm can be divided into Bicubic Interpolation, Image Reconstruction and Dictionary Learning types [1]. While there are three kinds of methods from nature, the discussions on the subject have been limited to learning based on these methods, and the corresponding algorithms are divided into two categories based on the source of the training patch: learning based on external datasets and learning based on internal datasets.

Most popular single-image super-resolution algorithms had focused on learning the relationship between image

The associate editor coordinating the review of this manuscript and approving it for publication was Ke Gu.

patches from low-resolution (LR) and high-resolution (HR) images in the external datasets. The existing mapping algorithms for learning from LR to HR include nearest neighbor searching [2], manifold learning [3], sparse coding [4], local linear regression [5], convolutional networks [6], and so on. However, there are some weaknesses in learning LR/HR mapping from external datasets. Firstly, the number and type of training images needed to achieve satisfactory recovery results remain unclear. Secondly, the large-scale training datasets are usually required to learn enough expressive LR/HR dictionaries. In order to avoid the use of external datasets and related problems, a series of algorithms [7]–[14] had been emerged that are based on natural images and patches in the same existing images and cross-scales. The internal LR/HR patching datasets of these methods can be constructed using the scale space of the given image itself in the Pyramid Tower model. Compared with external dictionaries, internal dictionaries have been shown to contain more relevant training patches [11]. The scale space of Pyramid Tower model has been used to extract reproduced blocks at scale and cross-scale, meaning that another important attribute of the natural image has been extracted, namely the self-similarity of structure. The use of an internal sample dictionary will result in poor training results in the internal data training sample, but a blind increase in the number of samples will increase the time consumption of algorithm.

In recent years, more and more deep learning methods had achieved remarkable results in the research field of image super-resolution, benefiting from the powerful feature characterization [9] of deep learning, which is more effective than traditional methods. Dong *et al.* [6] had firstly proposed the application of the Super-Resolution using Convolutional Neural Networks (SRCNN) algorithm to super-resolution images. Compared with traditional methods, the simple network structure has obtained the ideal super-resolution consequence. However, there are limitations of the simple network structure. Firstly, it is dependent on the context information of small image blocks. Secondly, the training convergence is too slow and the time complexity is high. Thirdly, the simple network only can be used for a single-scale *SR*procedure. Dong *et al.* [10] had proposed the Fast Super-Resolution Convolutional Neural Network (FSRCNN) by reducing the speed training of network parameters. FSRCNN used eight layers of network structure, making it deeper than SRCNN. Moreover, instead of Bicubic Interpolation, the anti-coiling layer was used on the last layer of the network. Finally, FSRCNN has achieved success in the convergence and super-resolution reconstruction field. Based on slow convergence and shallow network of SRCNN and FSRCNN networks, Zhao *et al.* [12] had proposed an image super-resolution algorithm (EEDS) based on end-to-end and shallow convolutional neural networks that has achieved better performance than others. However, because the deep network cannot fully extract the features of an LR image in the feature extraction stage, the loss of useful information and long-term memory content during the reconstruction process becomes serious when the feature of upsampling process is nonlinear mapping, as this causes the deep network to reduce the effect of super-resolution restoration. At the same time, the shallow network master is the main problem. When restoring the main components of LR images, the use of too many parameters limits the fast convergence of the network. Kim *et al.* [13] had proposed a highly accurate single-image super-resolution method named Very Deep Networks for Super-Resolution (VDSR). By using a very deep convolutional network of VGG-net [14] in ImageNet classification, the model employs cascaded small filters in a deep-network structure, using twenty weight-layers to efficiently utilize the context information of large image region.

Moreover, Kim *et al.* [15] had proposed the Deep Recursive Convolutional Network (DRCN) for image super-resolution restoration. The DRCN network had used a very deep recursive layer (as many as sixteen recursions), as increasing the recursion depth can improve performance without the need to introduce additional parameters to additional convolutions. In order to prevent the explosion and disappearance of the gradient, as well as to reduce the difficulty of training, the recursive monitoring and skipping connection methods are far more effective than previous methods. Recently, Zhao *et al.* [16] had proposed the Gradual Upsampling Network (GUN) model, which is based on a deep convolutional neural network. This method uses a gradual process to simplify the direct *SR*problem into a multi-step sampling task that employs very small magnification at each step. The Enhanced Deep Residual Networks for Single-Image Super-Resolution (EDSR) and the Multi-Scale Deep Super-Resolution network (MDSR) were proposed by Lim *et al.* [17] among others. The model has been optimized by removing unnecessary modules from the residual network to significantly enhance the performance of the model. Moreover, by extending the size of the model to further improve the performance, MDSR can reconstruct HR images with different magnification factors using a network model. Tai *et al.* [18] had proposed a very deep Memory Network (MemNet) for image restoration, which introduces memory blocks consisting of a recursive unit and a gate control unit that mine persistent memory through an adaptive learning process. The representation and output from previous memory blocks are connected and sent to the gate control unit. The gate control unit is adaptive to control memory. How many previous states should be retained, and decide how many current states should be stored and achieve superior performance in super-resolution restoration tasks. The Second-order Directional Total Generalized Variation (DTGV) [30] not only maintains the edge protection and noise suppression advantages exhibited by Directional Total Variation (DTV) [28] during super-resolution reconstruction, but also mitigates step artifacts and reconstructs clearer image details through higher-order processing.

Therefore, the paper has proposed a single-image super-resolution algorithm based on Structural Self-similarity and Deformation Block Features (SSDBF). In the paper, we have

adopted the algorithm of internal dictionary learning to build the Pyramid Tower model in zoom space and extract scale model, thereby avoiding a series of problems [43]–[45] associated with external dictionary learning. In addition, the use of geometric transform image blocks greatly expands the searching space and enhances the degree of dictionary performance without resulting in excessive time consumption. Finally, we also employ the group sparse learning method to enhance anti-noise resistance for achieve improved effectiveness.

II. THE ESTIMATION MODEL WITH DEFORMATION

Let ω be the set of index pixels of LR input image I. For each object fragment, $C\,(p_i)$ centers on the position of $p_i = \left(p_i^x; p_i^y\right)$ in I. Our goal is to map $C\,(p_i)$ patch in the lower sampled image I_D to the transformation matrix M_i of the mapping object of its nearest neighbor. In order to estimate the transformation matrix without making the algorithm more complicated, we need to solve Nearest Neighborhood Finding (NNF) estimation problem based on the dense nearest neighbor searching operation. The transform domain is been parameterized using i image pixels in the input from the LR image. The objective function of the NNF estimation problem is shown in (1),

$$\min_{\{\varphi_i\}} \sum_{i\in\omega} E_a\,(p_i, \varphi_i) + E_p\,(p_i, \varphi_i) + E_s\,(p_i, \varphi_i)\,, \tag{1}$$

Here, φ_i is the unknown parameter set for constructing the transformation matrix M_i of the model that needs to be estimated. The objective function of the model includes three kinds of energy consumption: (1) **Presenting Energy**; (2) **Planar Energy**; (3) **Scale Energy**. We first describe the cost function on the energy field.

A. PRESENTING ENERGY E_a

The cost function measures the similarity between the sampling target and the source patch. In the RGB space, we use the weighted sum of the squared *Gaussian* distance as our metric.

$$E_a(p_i, \varphi_i) = \|W_i(C(p_i) - Q(p_i, \varphi_i))\|_2^2\,, \tag{2}$$

The matrix W_i is the Gaussian weight of $\sigma^2 = 3$, and $Q\,(p_i; \varphi_i)$ represents the block sampled from I_D by using the M_i matrix transformation with the φ_i parameter.

B. PLANAR ENERGY E_p

For artificial images, the models can usually use standard vanishing point detection technology to reliably locate the scene. The geometry of a detected 3D scene can be used to guide the patch searching space. The plane localization code is modified in the plane structure guidance algorithm, and the planar compatibility of the costing function is added to encourage the search for the probabilistic flat label of the source patch and the object patching [19].

$$E_p = -\lambda_p \log(\Pr[m_i|(t_i^x, t_i^y) \times \Pr[m_i|(p_i^x, p_i^y)]), \tag{3}$$

Here, $\Pr\left[m_{ij}\,(x; y)\right]$ is A-posteriori probability of assigning m_i to the location $(x; y)$ of an image pixel.

C. SCALE ENERGY E_s

As continuous geometric transformation is allowed, it can be observed that the nearest neighborhood usually converges to the ordinary solution: that is, to match the patching target with itself in the lower sample image I_D. However, this rarely eventuates. The insignificant solution leads to the traditional Bicubic Interpolation of the super-resolution image [20], [46]. By introducing the scale costing function E_s, we can avoid these trivial solutions in (4),

$$E_s = \lambda_s \max(0, SRF - S(M_i)), \tag{4}$$

Here, the SRF indicates the desired super-resolution reconstruction factor, such as $2\times$, $3\times$ or $4\times$, and the function $S\,(\bullet)$ indicates the scale estimation of the projection transformation matrix. The function model is approximated by the first-order Taylor series expansion to describe the source block scale [21], which is sampled by M_i in (5),

$$S(M_i) = \sqrt{\det(\begin{bmatrix} M_{1,1} - M_{1,3}M_{3,1} & M_{1,2} - M_{1,3}M_{3,2} \\ M_{2,1} - M_{2,3}M_{3,1} & M_{3,1} - M_{2,3}M_{3,2} \end{bmatrix})}, \tag{5}$$

Here, $M_{u;v}$ represents the value of column u and line v in the transform matrix M_i, while $M_{3,3}$ is normalized to one. Intuitively, if the scale of the patching source is too small, the model will suffer. Therefore, the maximizing method uses the search for a source patch similar to the object patch, and it has a larger scale in the input LR image space. Therefore, we can provide more high-frequency details for the single super-resolution image. When the size of the source patch is large enough, the penalty threshold is reduced to zero.

III. GROUP SPARSE REPRESENTATION METHOD

The group sparse representation method had used all patching pairs to learn the dictionary and their sparse group to capture the relationship between HR and LR patches. To train the ability of dictionary learning, we can firstly extract features from LR patches that are similar to HR patches [22]. The features extracted from the LR patch are two first-order image gradients and two-order image gradients, along with horizontal and vertical axis, namely: $[1, 0, -1]$, $[1, 0, -1]$, $[-1, 0, 2, 0, -1]$, $[-1, 0, 2, 0, -1]$. For each HR slice, each of the feature vectors is formed by raster scanning of the value pixel after subtracting the mean value of the film.

For each high-resolution/low-resolution patching pair, we form the cascade feature vector. Since the size of LR and HR patching features are different, the two feature vectors are normalized independently before they are connected to a single vector. All cascaded feature vectors are normalized by unit norm vectors for dictionary learning with group sparsity constraints.

Because of the design of features, it is possible for HR and LR feature vectors to be zero. In this case, these feature vectors were discarded. In order to make use of the group similarity between blocks, the K-means clustering method has

been used to cluster pairs with similar feature vectors. Feature selection is accomplished via an image gradient generated by low-resolution patching. Using the given dictionary D, we can solve the group sparsity coefficient of each clustering U_i.

$$\min_{A_i} \|A_i\|_{1,2} \, s.t. \, \|Y_i - DA_i\|_F \leq \sqrt{n_i}\delta, \tag{6}$$

In (6), $\|A\|_{1,2} = \sum_{k=1}^{n} \left\|R^K\right\|_2$ and R^K are the k^{th} line of sparse A. Y_i is the column direction feature vector in the cluster U_i. n_i is the column number of Y_i. $\|*\|F$ is the Frobenius norm, and δ indicates how the eigenvectors that control the reconstruction should be similar to the original eigenvector. The Spectral Projected Gradient for L1 minimization (SPGL1) toolbox is used to solve the optimization problem.

Because the sparse coefficient of the group is solved in separated clustering, the dictionary is needed before the above equations can be solved. A represents the union of all coefficients A_i and Y represents the union of all eigenvectors Y_i. The dictionary D is updated via the K-SVD algorithm [23].

$$D = \arg\min_{D} \|Y - DA\|_F \quad s.t. \left\|D_j\right\|_2 = 1 \forall j, \tag{7}$$

D_j is the j^{th} column of D. We iteratively solve the group sparse coefficients in (6) and (7) until A and D converge. The product of dictionary D and coefficient A contains not only all the clusters in each clustering, but also the feature vectors generated by block similarity in all clusters. We use these eigenvectors to generate the output high-resolution images.

IV. THE PROPOSED ALGORITHM USING SCALE MODEL AND DEFORMATION FEATURES

A. RESOURCE BLCOK ACQUISITION

1) FEATURE EXTRACTION OF DEFORMATION FEATURES

Given the low-resolution image I, fuzzy and two-sampling is first performed to obtain the down-sampling version of its n^{th} layer. n is set to six. For ease of expression, only two layers of the structure are used to express I_{D1} and I_{D2}.

Using I, I_{D1} and I_{D2}, the proposed algorithm for obtaining high-resolution I_H images has been composed of the following steps:

1) For each patch C in the low-resolution image I, we can calculate the transformation matrix M, which reduces C to the best matching patch Q_1 or Q_2 in the lower-sample image I_{D1} and I_{D2} (see Fig.1). In order to obtain the parameters of transformation, we use the modified block-matching algorithm [15] to estimate the nearest neighborhood between I and I_{D1} or I and I_{D2}.

2) We then extract Q_{H1} or Q_{H2} from image I, which is the HR version of source patching Q_1 or Q_2.

3) We use the inverse matrix of the computed transform matrix M to "converse" high-resolution patch Q_{H1} or Q_{H2} in order to obtain the self-evaluation C_{H1} or C_{H2} value, which is the HR version of the object patch C that we estimate. We paste PH in HR image I_H corresponding to LR patch C_{H1} or C_{H2}.

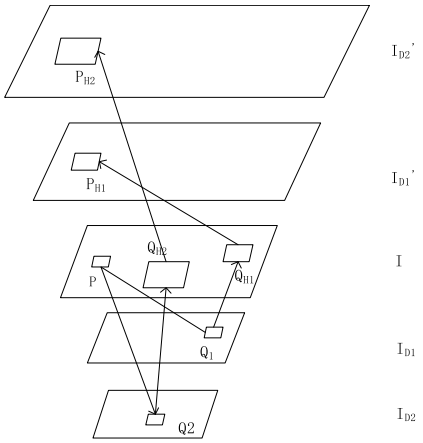


FIGURE 1. The example of two-level sampling of pyramid tower structure.

4) We repeat the above steps for all object patches to obtain the size of patch estimates for the high-resolution image I_H.

According to the similarity of $\exp\left(-\|C - Q_k\|^2/\sigma^2\right)$, the weights are calculated by averaging the overlap of the high-resolution slices in which σ controls the similarity.

2) DEFORMABLE FEATURE EXTRACTION

Previous research has proven that cross-scaling self-similarity performs better if appropriate geometric transformations are allowed. Accordingly, we can deform the resource blocks and extract the deformation characteristics of these blocks.

We now introduce how the transformation matrix M_i is designed and constructed based on the estimated parameter φ_i, to sample the source patch $Q(p_i; \varphi_i)$. Generally, the geometric transformation of patches can have as many as eight degrees of freedom (i.e., projection transformation). However, the affine transformation is inefficient at simulating the appearance changes of man-made images and structural scenes. In the algorithm, we can detect and localize the plane to calculate the planarization parameters. Fig.2 presents the visualization of vanishing point detection. Accordingly, we can parameterize M_i through $\varphi_i = (t_i; m_i)$, where $t_i = \left(t_i^x; t_i^y; t_i^s; t_i^q; t_i^a; t_i^b\right)$ is an affine motion parameter of the six dimensions and the index of the detected plane [24], [47].

The geometric transformation model $T_i(\varphi_i)$ has been adopted, and the explicit search strategy of Yang *et al.* [25] is combined with the perspective deformation estimation method of Barnes *et al.* [26] and the specific equation of geometric transformation model $T_i(\varphi_i)$. It can be expressed as (8),

$$T_i(\varphi_i) = H(p_i, t_i^x, t_i^y, m_i)S(t_i^s, t_i^\theta)A(t_i^\alpha, t_i^\beta), \tag{8}$$

The matrix H captures the perspective distortion of the object and the source patch location, as well as the plane parameters (as described in He *et al.* [24]). The Matrix S is

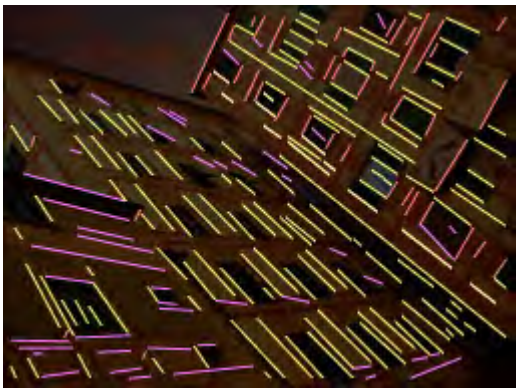


FIGURE 2. Vanishing point detection visualization.

as shown in (9),

$$S(t_i^s, t_i^\varphi) = \begin{bmatrix} t_i^s R(t_i^\varphi) & 0 \\ 0 & 1 \end{bmatrix}, \tag{9}$$

The similarity transformation A is captured by the scaling parameters S_i^s and the 2×2 rotation matrix $R\left(t_i^\varphi\right)$. The matrix $A(t_i^\alpha, t_i^\beta)$ is as shown in (10),

$$A(t_i^\alpha, t_i^\beta) = \begin{bmatrix} 1 & t_i^\alpha & 0 \\ t_i^\beta & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \tag{10}$$

The proposed combination transformation model is similar to the classical decomposition of the projection transformation matrix into three unique matrices: Similarity, Affinity and Perspective Collineation.

The method presented here is the general algorithm for the super-resolution image, which deals with both man-made and natural scenes in the framework. In addition, in the absence of any detected plane structure, the proposed algorithm automatically returns to the super-resolution image block-matching algorithm that only searches for the self-example of affine transformation.

The goal of the proposed method is to ''synthesize'' rather than ''decompose'' transform M_i for sampling source patches. The proposed equation can effectively decompose the position dependence of the object p_i and the source block position $\left(t_i^x; t_i^y\right)$ to estimate $H\left(p_i; t_i^x; t_j^y; m_i\right)$. The affine shape deformation parameters $\left(t_i^s; t_i^q; t_i^a; t_i^b\right)$ has the form of a deformation matrix. We can then use the piecewise smooth features in natural images to conduct effective nearest neighbor estimation.

The proposed method needs to estimate the nearest neighbor solution of seven dimensions on all overlapping objects. Unlike traditional methods based on self-sampling, it is only necessary here to estimate the 2D translation field. The solution space in the proposed equation is more difficult to search. Therefore, this method had modified the Block-Matching algorithm [27] according to the following detailed steps:

a: INITIALIZATION

The proposed method is not initialized randomly in block-matching, but uses zero displacement to initialize the nearest neighborhood; moreover, the scaling is equal to the required super-resolution reconstruction factor [27]. This was inspired by HaCohen *et al.* [19], who showed that a good self-example can usually be found in local areas. It was further found that the initialization strategy can provide a good start for faster convergence.

b: COMMUNICATION

This step effectively involves the propagation of good matches to neighbors. In contrast to the direct propagation transformation matrix M_i, the parameter $Q_i = (t_i; m_i)$ of propagation transformation matrix facilitates the affine shape transformation to remain unchanged for the source patch position.

c: RANDOMIZATION

After each iteration of propagation, random searching is performed to improve the current solution. At the same time, the random samples of plane exponents are drawn according to the posterior probability distribution. Random perturbation affine transformation and random sample position (coarse to fine) are used to search the best geometric transformation of the source patching and reduce the matching error.

B. CLASSIFICATION LEARNING OF RESOURCE BLOCK

Resource blocks are classified by using resource blocks that match the deformed blocks. One is used to build the high-resolution dictionary, while the other is for coefficient learning.

1) High-Resolution Dictionary

The dictionary resource block pair is used here. The dictionary resource block is composed of the C_H value and the low-resolution patch value C of the ''inverse transformation'' high-resolution patch by inverse matrix of the transformation matrix M. Firstly, we can use clustering sparse theory to cluster the dictionary resource blocks, as shown in Fig.3. Then, the K-SVD [23] algorithm had been used to learn the high-resolution dictionary.

2) SPARSE LEARNING

This takes advantage of deformed resource blocking pairing. The deformed resource block is composed of the low-resolution patching Q values of the high-resolution patching Q_H values and the transformation matrix M deformed by the transformation matrix M. Here, we can use group sparsity theory to group sparse pairs of deformed resource blocks. For each HR/LR patching, we can form a cascade feature vector. Before connecting them into a single vector, we independently normalize the two feature vectors. All cascaded feature vectors are normalized by unit norm vectors for dictionary learning with group sparsity constraints. The HR and

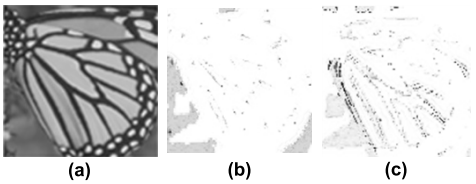


FIGURE 3. Structural clustering example. (a) Single original low-resolution image. (b) Clustering example. (c) Clustering example.

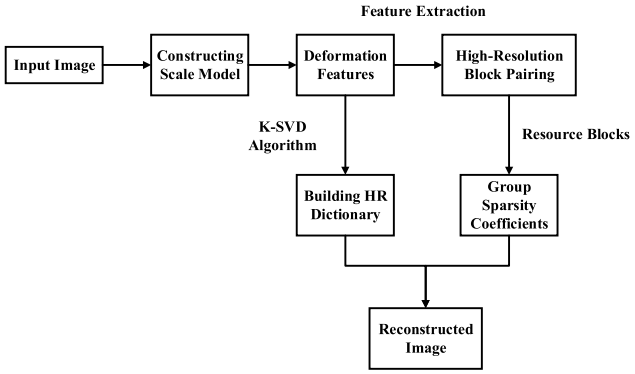


FIGURE 4. The flowchart of proposed algorithm.

LR block pairs with similar feature vectors are grouped by means of K-means clustering. Finally, we can use (6) and (7) to iteratively solve the group sparse coefficient A.

C. THE FLOW OF THE PROPOSED ALGORITHM

The proposed algorithm in the paper is as follows:

Step 1: Input the high-resolution image in the external datasets, reduce its scale, obtain the n-level low-resolution image, and build the scale model.

Step 2: Use the transformation matrix M to deform the scale model.

Step 3: Block-matching using the deformable block-matching algorithm; see part A of section IV.

Step 4: Extract C_H/C dictionary resource blocks for clustering, and use K-SVD algorithm [23] to build high-resolution dictionaries.

Step 5: Extract Q_H/Q resource blocks for clustering, and use (9) and (10) to learn group sparsity coefficients.

Step 6: Super-resolution reconstruction using high-resolution dictionary and group sparsity coefficient.

In order to illustrate the algorithm more clearly, a flowchart of the proposed algorithm can be found in Fig.4.

V. THE EXPERIMENTAL RESULTS AND ANALYSIS

A. THE RESULTS OF COMPARSION METHODS

In order to verify the effect of the proposed method, a large amount of simulation experiments had been carried out. We had selected four international open SR datasets: Set5, Set14, BSD100 and Urban100. The experiments were carried out under three common amplification factors ($2\times, 3\times$ or $4\times$). The *GPU*used was *NVIDIA GeForce*1080 T_i, the experimental environment was *Keras*, and *Python*3.5 and

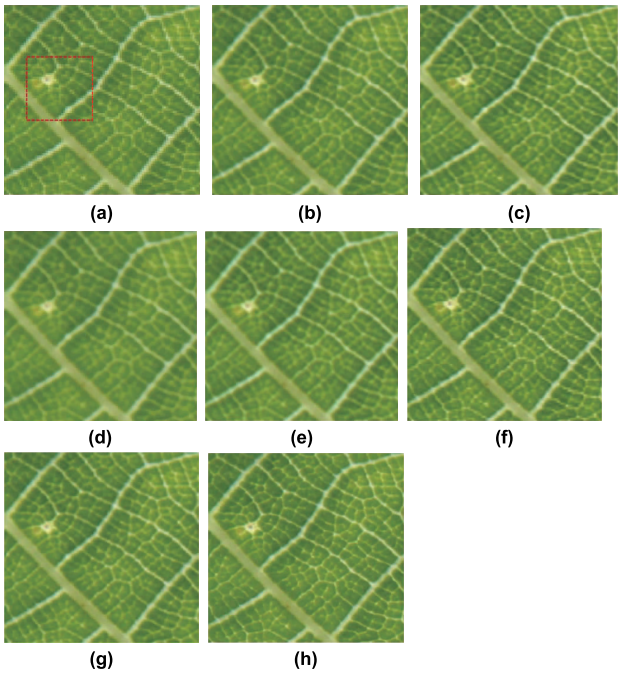


FIGURE 5. Results of algorithms on leaf image sample with $3\times$. (a) Original image. (b) Bicubic interpolation [20]. (c) Sparse coding [4]. (d) DRCN [15]. (e) MDSR [17]. (f) SRCNN [10]. (g) DTGV [30]. (h) SSDBF.

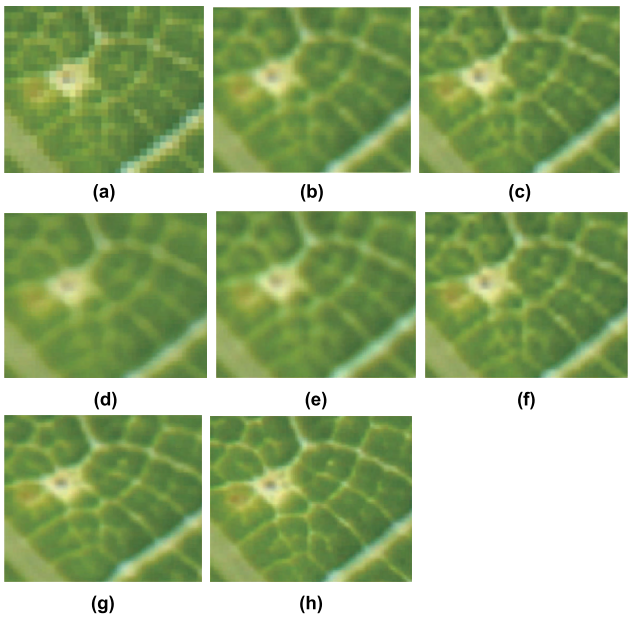


FIGURE 6. Results of algorithms on red box in leaf image sample of Fig.5. (a) Original image. (b) Bicubic interpolation [20]. (c) Sparse coding [4]. (d) DRCN [15]. (e) MDSR [17]. (f) SRCNN [10]. (g) DTGV [30]. (h) SSDBF.

*OpenCV*3.0 were used to carry out the simulation experiments.

The paper had selected some representative super-resolution reconstruction algorithms and some testing images to obtain comprehensive comparison results. The selected super-resolution algorithms have included: Bicubic Interpolation (BI) [20], Sparse Coding (SC) [4], Deep Recursive

TABLE 1. Average PSNR (dB), SSIM, and RMSE comparison results for different SR methods.

SR Methods	Leaf Image	*Set5*	*Set14*	*BSD100*	*Urban100*
	32.3989	31.0113	19.8387	26.1280	26.8098
Bicubic Interpolation	0.8799	0.8541	0.5623	0.8438	0.8125
	6.1178	7.1775	25.9780	12.5933	11.6426
	32.6823	31.0357	19.7078	25.9397	27.6240
Sparse Coding	0.8902	0.8659	0.5789	0.8425	0.8432
	5.9214	7.1574	26.3724	12.8693	10.6008
	32.5236	30.7315	21.9363	26.4834	27.6262
DRCN	0.8923	0.8832	0.5912	0.8587	0.8417
	6.2075	7.4842	20.4321	12.1302	10.7512
	32.4124	30.6725	21.9283	26.4744	27.5032
MDSR	0.8823	0.8640	0.5890	0.8637	0.8397
	6.1083	7.4630	20.4233	12.1010	10.7493
	32.9272	31.1052	21.5601	27.0230	28.1320
SRCNN	0.9092	0.8854	0.5957	0.8650	0.8656
	5.7568	7.1004	21.3076	11.3603	9.9986
	32.9864	31.5024	21.7342	27.2105	28.2467
DTGV	0.9153	0.9013	0.6104	0.8724	0.8845
	5.7177	6.7830	20.8848	11.1177	9.8675
	33.9464	31.6214	21.7732	27.2605	28.4467
SISR-SMDF	0.9256	0.9213	0.6314	0.8924	0.8935
	5.7297	6.8912	20.8918	11.1277	9.8825

Convolutional Network (DRCN) [15], Multi-Scale Deep Super-Resolution Network (MDSR) [17], Super-Resolution Convolutional Neural Network (SRCNN) [10] and Second-order Directional Total Generalized Variation (DTGV) [30]. Fig.5 and Fig.6 had shown a representative set of all experimental results. The experimental results of BI, DRCN and MDSR are both blurred, and no clear image edges can be obtained. The experimental results of SC, SRCNN and DTGV are slightly better and can produce sharper edges and image details. In contrast, the results of the proposed algorithm are the better and clearer than others, and the image edges can be obtained.

Furthermore, to compare the experimental results, we had used three commonly used indicators to quantitatively analyze the pros and cons of different experimental results, namely: Peak Signal-to-Noise (PSNR), Root Mean Square Error (RMSE) and Structural Similarity (SSIM) [31]–[35]. The PSNR was used to evaluate the performance of different image super-resolution methods. The higher the PSNR value of reconstructed image, the less the super-resolution image distortion and the higher the image quality. The SSIM values of the super-resolution image and the high-resolution image are closer to one, the more similar indicating the super-resolution image and the high-resolution image.

Tab. 1 had summarized the comparison of PSNR, RMSE and SSIM values for different image experiments with different super-resolution methods. It can be seen that the proposed method dominates the numerical comparison of PSNR, RMSE and SSIM, reflecting the effectiveness and superiority of the proposed method. It is worth noting that the proposed method has larger PSNR and SSIM values than DTGV [30]. It was reflected from the side that the proposed method can not only remove noise effects, but also reconstruct more clearer high-band textures than other methods.

B. THE EXPERIMENTAL RESULTS OF DATASETS

Tab. 2 - Tab. 4 have shown the PSNR, SSIM, and Generating Image Time for BI, DRCN, MSDR, SRCNN, DTGV, and the proposed method had reconstructed at different amplification factor. With the same testing dataset and the same amplification factor, the proposed method can reconstruct super-resolution images with higher PSNR and SSIM values in the shortest time. Fig.7 has shown partially reconstructed image of above several methods with baboon image. The currently displayed pictures are Butterfly (Set5) and Baboon (Set14), and the amplification factor are $2\times$, $3\times$ and $4\times$ times, respectively. It can be seen from the reconstructed image that the reconstructed image by the proposed method in the paper

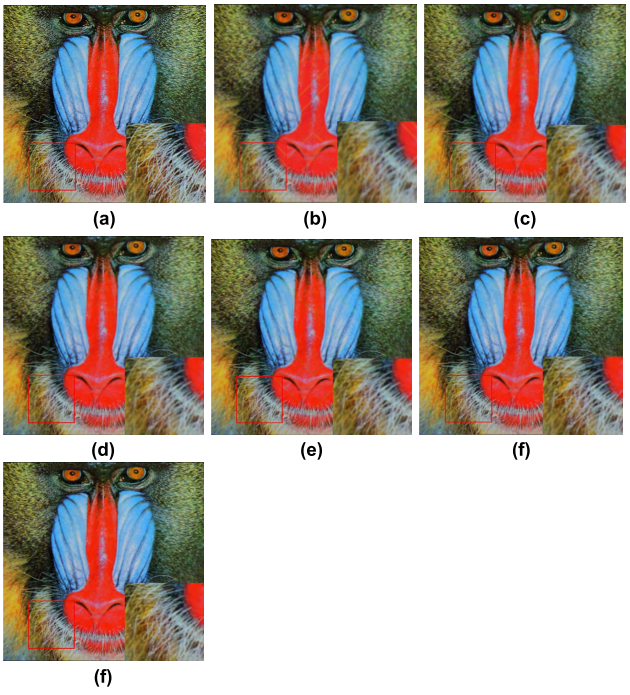


FIGURE 7. Reconstruction results of different methods on image baboon. (a) SR image. (b) Bicubic interpolation [20]. (c) DRCN [15]. (d) MDSR [17]. (e) SRCNN [10]. (f) DTGV [30]. (f) SSDBF.

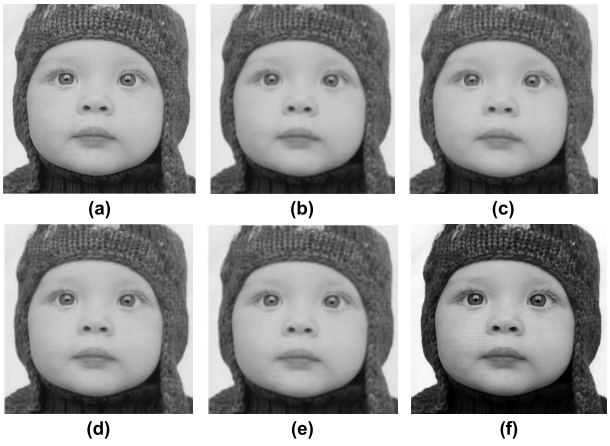


FIGURE 8. Results of algorithms on set5 dataset sample. (a) Original image. (b) Bicubic interpolation [20]. (c) SC [4]. (d) SRCNN [10]. (e) DTGV [30]. (f) SSDBF.

is the clearest and the texture is clearer than others. The following experimental results had shown that SSDBF can reduce the time complexity of the model and reconstruct the super-resolution image with good visual effects. The average PSNR of the Bicubic Interpolation method is the lowest at only 33.64db, while the best result generated by a comparison algorithm on the *Set5* database was about 36.28db. For its part, the proposed method yields a high PSNR value, reaching 36.52 db.

The validity of proposed method was verified by the experimental results. In order to evaluate the quantity of image reconstruction, we can compare the results

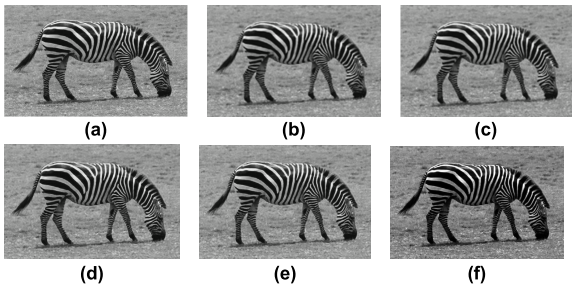


FIGURE 9. Results of algorithms on set14 dataset sample. (a) Original image. (b) Bicubic interpolation [20]. (c) SC [4]. (d) SRCNN [10]. (e) DTGV [30]. (f) The Proposed Method.

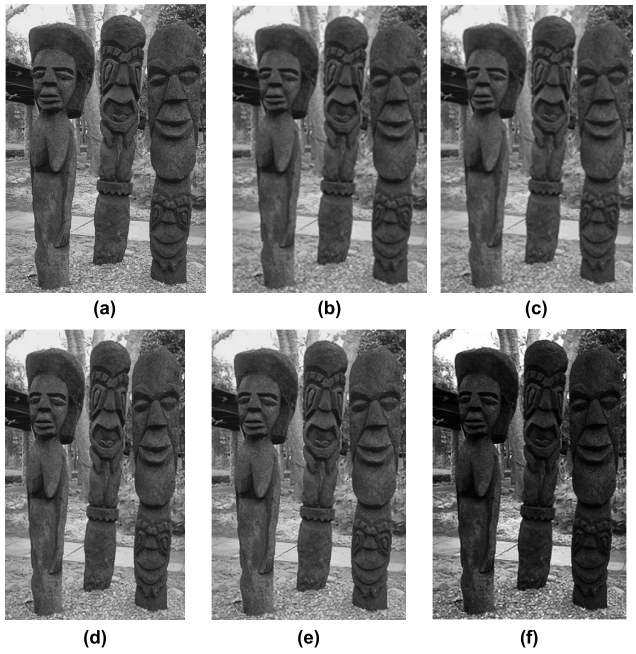


FIGURE 10. Results of algorithms on BSD100 dataset sample. (a) Original image. (b) Bicubic interpolation [20]. (c) SC [4]. (d) SRCNN [10]. (e) DTGV [30]. (f) SSDBF.

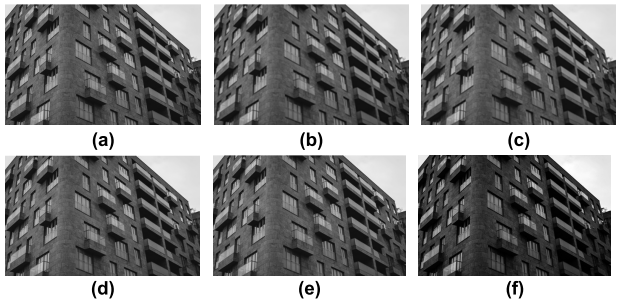


FIGURE 11. Results of algorithms on urban100 dataset sample. (a) Original image. (b) Bicubic interpolation [20]. (c) SC [4]. (d) SRCNN [10]. (e) DTGV [30]. (f) SSDBF.

from two perspectives: namely, Subjective Vision and Objective Evaluation. Taking the doll image in the *Set5* dataset as an example, Fig.8-Fig.11 present the subjective vision and the corresponding PSNR of different methods under two-times magnification. The four pictures of

TABLE 2. Average PSNR/dB comparison results for different SR methods.

Data set	Scale	*BI*	*DRCN*	*MDSR*	*SRCNN*	*DTGV*	*SISR-SMDF*
Set5	2×	35.78	36.24	36.28	37.63	37.74	37.78
	3×	30.47	32.62	33.67	33.83	33.82	34.08
	4×	29.07	30.07	30.08	31.54	31.54	31.79
Set14	2×	31.64	32.14	32.00	33.06	33.08	33.09
	3×	27.55	29.15	29.78	29.77	29.58	29.75
	4×	26.40	27.18	27.13	28.03	28.03	28.36
BSD 100	2×	30.77	31.11	31.11	31.95	31.95	32.03
	3×	27.22	28.34	28.83	28.74	28.74	28.94
	4×	26.61	26.71	26.70	27.10	27.10	27.49
Urban 100	2×	28.26	28.74	28.65	28.88	28.68	28.92
	3×	26.85	25.95	26.22	26.56	26.60	26.80
	4×	24.02	24.20	24.14	24.32	24.32	24.60

TABLE 3. Average SSIM comparison results for different SR methods.

Data set	Scale	*BI*	*DRCN*	*MDSR*	*SRCNN*	*DTGV*	*SISR-SMDF*
Set5	2×	0.931	0.955	0.959	0.959	0.950	0.956
	3×	0.869	0.909	0.921	0.912	0.910	0.923
	4×	0.811	0.863	0.883	0.884	0.879	0.890
Set14	2×	0.870	0.908	0.913	0.912	0.913	0.917
	3×	0.775	0.823	0.832	0.836	0.832	0.838
	4×	0.704	0.753	0.768	0.768	0.770	0.778
BSD 100	2×	0.844	0.889	0.896	0.898	0.892	0.896
	3×	0.741	0.786	0.799	0.797	0.801	0.812
	4×	0.670	0.712	0.726	0.725	0.723	0.739
Urban 100	2×	0.860	0.881	0.885	0.892	0.881	0.898
	3×	0.782	0.796	0.800	0.806	0.796	0.809
	4×	0.722	0.742	0.746	0.748	0.745	0.749

TABLE 4. Average generating image time/S comparison results for different SR methods.

Data set	Scale	*BI*	*DRCN*	*MDSR*	*SRCNN*	*DTGV*	*SISR-SMDF*
Set5	2×	0.00	2.18	0.12	0.64	0.20	0.08
	3×	0.00	2.24	0.14	0.45	0.21	0.12
	4×	0.00	2.11	0.15	0.33	0.24	0.13
Set14	2×	0.00	4.34	0.26	1.12	0.42	0.11
	3×	0.00	4.45	0.27	0.81	0.43	0.12
	4×	0.00	4.36	0.25	0.63	0.41	0.13
BSD 100	2×	0.00	2.57	0.17	0.84	0.19	0.07
	3×	0.00	2.57	0.20	0.65	0.22	0.07
	4×	0.00	2.59	0.28	0.46	0.21	0.06
Urban 100	2×	0.00	3.69	0.22	0.72	0.35	0.12
	3×	0.00	3.74	0.23	0.75	0.32	0.11
	4×	0.00	3.71	0.21	0.77	0.33	0.13

Fig.8-Fig.11 are representative of one of the datasets. Subjectively, we can see that the proposed method yields the best visual effect, which is better than that supplied by Sparse Dictionary (*SC*) [4], SRCNN [10] and DTGV [30]. In contrast to other methods, artifacts in reconstructed images are prone to artifacts, and some details cannot be recovered well. The local detail information restored by our proposed method is clear and delicate, and the overall effect is closer to that of the original image. Since the proposed method considers the self-similarity of the intrinsic image block structure, the visual restoration effect is the better and the edges are clearer than others.

VI. CONCLUSIONS

The super-resolution problem as it relates to the single image has considered from the angle of block deformation. Accordingly, a super-resolution reconstruction algorithm based on structural similarity and deformation block feature is proposed in this paper. The proposed algorithm obtains the characteristics of its own structural similarity through the scale decomposition of the image and extracts the features of the deformation block by transformation, thus avoiding a series of related problems such as those involving data type and the quantity of the external dataset. The resource block's data is been classified by the dictionary resource block and the deformation block resource, the high-resolution dictionary and the group sparsity coefficient are learned by different constraints, and the noise resistance of the image is enhanced. Compared with popular algorithms such as Bicubic Interpolation (BI), Sparse Coding (SC), Deep Recursive Convolutional Network (DRCN), Multi-Scale Deep Super-Resolution Network (MDSR), Super-Resolution Convolutional Neural Network (SRCNN) and Second-order Directional Total Generalized Variation (DTGV), the proposed algorithm can achieve a better super-resolution reconstruction effect.

The proposed method in the paper can not only avoid the over-fitting phenomenon of network deepening, but also obtain better reconstruction effect. The discrimination and sharpness of the details of the image texture are effectively improved, but there is still a gap between the edge reconstruction effect and the original high-resolution image. The algorithm effect is further improved to solve the blurring problem of the image edge region, and the super-resolution reconstruction effect of the single image has been improved.

ACKNOWLEDGEMENT

The authors would like to sincerely thank the anonymous reviewers. The authors also thank Dr. Ke Gu, Dr. Yan Gui, and Dr. Fei Yu for useful advises during this work.

REFERENCES

[1] M. Irani and S. Peleg, "Improving resolution by image registration," *CVGIP, Graph. Models Image Process.*, vol. 53, no. 3, pp. 231–239, May 1991.

[2] W. T. Freeman, T. R. Jones, and E. C. Pasztor, "Example-based super-resolution," *IEEE Comput. Graph. Appl.*, vol. 22, no. 2, pp. 56–65, Mar. 2002.

[3] H. Chang, D.-Y. Yeung, and Y. Xiong, "Super-resolution through neighbor embedding," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, Jun./Jul. 2004, p. 1.

[4] J. Yang, J. Wright, T. S. Huang, and Y. Ma, "Image super-resolution via sparse representation," *IEEE Trans. Image Process.*, vol. 19, no. 11, pp. 2861–2873, Nov. 2010.

[5] C.-Y. Yang and M.-H. Yang, "Fast direct super-resolution by simple functions," in *Proc. IEEE Int. Conf. Comput. Vis.*, Dec. 2013, pp. 561–568.
[6] C. Dong, C. C. Loy, K. He, and X. Tang, "Learning a deep convolutional network for image super-resolution," in *Proc. Eur. Conf. Comput. Vis.*, 2014, pp. 184–199.
[7] D. Zeng, Y. Dai, F. Li, R. Sherratt, and J. Wang, "Adversarial learning for distant supervised relation extraction," *Comput., Mater. Continua*, vol. 55, no. 1, pp. 121–136, 2018.
[8] L. Xiang, G. Zhao, Y. Li, W. Hao, and F. Li, "TUMK-ELM: A fast unsupervised heterogeneous data learning approach," *IEEE Access*, vol. 6, pp. 35305–35315, 2018.
[9] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," in *Proc. Int. Conf. Neural Inf. Process. Syst.*, 2012, pp. 1097–1105.
[10] C. Dong, C. C. Loy, K. He, and X. Tang, "Image super-resolution using deep convolutional networks," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 38, no. 2, pp. 295–307, Feb. 2016.
[11] R. Timofte, V. De, and L. Van Gool, "Anchored neighborhood regression for fast example-based super-resolution," in *Proc. IEEE Int. Conf. Comput. Vis.*, Dec. 2013, pp. 1920–1927.
[12] Y. Zhao, R. Wang, W. Jia, J. Yang, W. Wang, and W. Gao, "Local patch encoding-based method for single image super-resolution," *Inf. Sci.*, vols. 433–434, pp. 292–305, Apr. 2018.
[13] J. Kim, J. K. Lee, and K. M. Lee, "Accurate image super-resolution using very deep convolutional networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2016, pp. 1646–1654.
[14] V. F. Campana, K. F. Côco, E. O. T. Salles, and P. M. Ciarelli, "Modification in the SAR super-resolution model using the fractal descriptor LMME in the term regularizer," *IEEE Access*, vol. 6, pp. 39046–39062, 2018.
[15] J. Kim, J. K. Lee, and K. M. Lee, "Deeply-recursive convolutional network for image super-resolution," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2016, pp. 1637–1645.
[16] Y. Zhao, G. Li, W. Xie, W. Jia, H. Min, and X. Liu, "GUN: Gradual upsampling network for single image super-resolution," *IEEE Access*, vol. 6, pp. 39363–39374, 2018.
[17] B. Lim, S. Son, H. Kim, S. Nah, and K. M. Lee, "Enhanced deep residual networks for single image super-resolution," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. Workshops*, Jul. 2017, pp. 1132–1140.
[18] Y. Tai, J. Yang, X. Liu, and C. Xu, "MemNet: A persistent memory network for image restoration," in *Proc. IEEE Int. Conf. Comput. Vis.*, Oct. 2017, pp. 4549–4557.
[19] Y. HaCohen, R. Fattal, and D. Lischinski, "Image upsampling via texture hallucination," in *Proc. IEEE Int. Conf. Comput. Photography*, Mar. 2010, pp. 1–8.
[20] R. G. Keys, "Cubic convolution interpolation for digital image processing," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 29, no. 6, pp. 1153–1160, Dec. 1981.
[21] A. Singh and N. Ahuja, "Super-resolution using sub-band self-similarity," in *Proc. Asian Conf. Comput. Vis.*, 2014, pp. 552–568.
[22] M. Zontak and M. Irani, "Internal statistics of a single natural image," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2011, pp. 977–984.
[23] M. Aharon, M. Elad, and A. Bruckstein, "K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation," *IEEE Trans. Signal Process.*, vol. 54, no. 11, pp. 4311–4322, Nov. 2006.
[24] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2016, pp. 770–778.
[25] J. C. Yang *et al.*, "Image super-resolution as sparse representation of raw image patches," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2008. [Online]. Available: https://ieeexplore.ieee.org/document/4587647. doi: 10.1109/CVPR.2008.4587647.
[26] C. Barnes, E. Shechtman, A. Finkelstein, and D. B. Goldman, "Patchmatch: A randomized correspondence algorithm for structural image editing," *ACM Trans. Graph.*, vol. 28, no. 3, pp. 24–34, 2009.
[27] S. Darabi, E. Shechtman, C. Barnes, D. B. Goldman, and P. Sen, "Image melding: Combining inconsistent images using patch-based synthesis," *ACM Trans. Graph.*, vol. 31, no. 4, 2012, Art. no. 82.
[28] I. Bayram and M. E. Kamasak, "Directional total variation," *IEEE Signal Process. Lett.*, vol. 19, no. 12, pp. 781–784, Dec. 2012.
[29] W. Dong, L. Zhang, G. Shi, and X. Wu, "Image deblurring and super-resolution by adaptive sparse domain selection and adaptive regularization," *IEEE Trans. Image Process.*, vol. 20, no. 7, pp. 1838–1857, Jul. 2011.
[30] Z. H. Wu *et al.*, "Second-order directional total generalized variation regularization for image super-resolution," *Acta Electronica Sinica*, vol. 45, no. 11, pp. 2625–2632, 2017.
[31] X. Min, K. Ma, K. Gu, G. Zhai, Z. Wang, and W. Lin, "Unified blind quality assessment of compressed natural, graphic, and screen content images," *IEEE Trans. Image Process.*, vol. 26, no. 11, pp. 5462–5474, Nov. 2017.
[32] Z. Tang, Y. Zheng, K. Gu, K. Liao, W. Wang, and M. Yu, "Full-reference image quality assessment by combining features in spatial and frequency domains," *IEEE Trans. Broadcast.*, vol. 65, no. 1, pp. 138–151, Mar. 2019. doi: 10.1109/TBC.2018.2871376.
[33] K. Gu, L. Li, H. Lu, X. Min, and W. Lin, "A fast reliable image quality predictor by fusing micro- and macro-structures," *IEEE Trans. Ind. Electron.*, vol. 64, no. 5, pp. 3903–3912, May 2017.
[34] K. Gu, G. Zhai, X. Yang, and W. Zhang, "Using free energy principle for blind image quality assessment," *IEEE Trans. Multimedia*, vol. 17, no. 1, pp. 50–63, Jan. 2015.
[35] K. Gu, G. Zhai, W. Lin, X. Yang, and W. Zhang, "No-reference image sharpness assessment in autoregressive parameter space," *IEEE Trans. Image Process.*, vol. 24, no. 10, pp. 3218–3231, Oct. 2015.
[36] Y. Yang, Z. Lin, and S. Cohen, "Fast image super-resolution based on in-place example regression," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2013, pp. 1059–1066.
[37] J. M. Zhang *et al.*, "A fast object tracker based on integrated multiple features and dynamic learning rate," *Math. Problems Eng.*, vol. 2018, Dec. 2018, Art. no. 5986062. doi: 10.1155/2018/5986062.
[38] J. S. Park, J. W. Soh, and N. I. Cho, "High dynamic range and super-resolution imaging from a single image," *IEEE Access*, vol. 6, pp. 10966–10978, 2018.
[39] J. Sun, Y. Yang, N. N. Xiong, L. Dai, and X. Pen, "Complex network construction of multivariate time series using information geometry," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 49, no. 1, pp. 107–122, Jan. 2019.
[40] X. Jiang *et al.*, "Data fusion-based multi-object tracking for unconstrained visual sensor networks," *IEEE Access*, vol. 6, pp. 13716–13728, 2018.
[41] C. Wu, C. Luo, N. Xiong, W. Zhang, and T.-H. Kim, "A greedy deep learning method for medical disease analysis," *IEEE Access*, vol. 6, pp. 20021–20030, 2018.
[42] J. M. Zhang *et al.*, "Spatial and semantic convolutional features for robust visual object tracking," *Multimedia Tools Appl.*, vol. 77, pp. 1–21, Aug. 2018. doi: 10.1007/s11042-018-6562-81.
[43] Y. T. Chen *et al.*, "A novel online incremental and decremental learning algorithm based on variable support vector machine," *Cluster Comput.*, vol. 22, pp. 1–11, Jan. 2018. doi: 10.1007/s10586-018-1772-4.
[44] Y. Tu, Y. Lin, J. Wang, and J.-U. Kim, "Semi-supervised learning with generative adversarial networks on digital signal modulation classification," *Comput. Mater. Continua*, vol. 55, no. 2, pp. 243–254, 2018.
[45] Y. Gui and G. Zeng, "Joint learning of visual and spatial features for editpropagation from a single image," *Visual Comput.*, vol. 35, pp. 1–14, Jan. 2019. doi: 10.1007/s00371-019-01633-6.
[46] S. He, K. Xie, W. Chen, D. Zhang, and J. Wen, "Energy-aware routing for SWIPT in multi-hop energy-constrained wireless network," *IEEE Access*, vol. 6, pp. 17996–18008, 2018.
[47] Z. Liao, R. Zhang, S. He, D. Zeng, J. Wang, and H.-J. Kim, "Deep learning-based data storage for low latency in data center networks," *IEEE Access*, vol. 7, pp. 26411–26417, 2019.
[48] D. Cao *et al.*, "A relay-node selection on curve road in vehicular networks," *IEEE Access*, vol. 7, pp. 12714–12728, 2019.

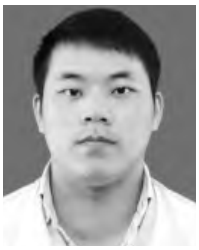
YUANTAO CHEN was born in 1980. He received the B.S. degree in computer science and technology from the Jianghan Petroleum Institute, the M.S. degree in geodetection and information technology from Yangtze University, and the Ph.D. degree in control science and engineering from the Nanjing University of Science and Technology, in 2014. He is currently an Associate Professor with the Changsha University of Science and Technology. His research interests include pattern recognition, image processing, and so on.

JIN WANG was born in 1979. He received the B.S. degree in electronic engineering and the M.S. degree in information engineering from the Nanjing University of Posts and Telecommunications, and the Ph.D. degree in computer engineering from Kyung Hee University, in 2010. He is currently a Professor with the Changsha University of Science and Technology. Since 2002, he has been engaged in wireless sensor networks, MANET, VANET research, and had accumulated and systematic research works in network energy-saving routing algorithms, data fusion, computational intelligence, intelligent medical treatment, and other fields.

XI CHEN was born in 1972. He received the master's degree in computer science from the Changsha University of Science and Technology, in 2007, where he is currently an Associate Professor. His research interests include artificial intelligence, image processing, big data processing, and so on.

MINGWEI ZHU was born in 1998. He is currently pursuing the degree with the Changsha University of Science and Technology. His main direction of study is Java and Android.

KAI YANG was born in 1980. He received the master's degree in mechanical engineering from Jilin University, in 2014. He is an Engineer with Technology Department, Hunan Zoomlion Intelligent Technology Company Ltd. His research interests include mechanical engineering, intelligent control technology, and so on.



ZHI WANG was born in 1993. He received the B.S. degree in tourism management from the Hunan University of Arts and Science. He is currently pursuing the master's degree with the Changsha University of Science and Technology. His research interests include image processing, pattern recognition, and so on.

RUNLONG XIA was born in 1982. He received the B.S. degree in electronic commerce from Hunan Normal University, in 2010. He is currently a Research Assistant with the Hunan Institute of Scientific and Technical Information. His research interests include news communication and public opinion analysis.