

Remote Sensing Image Super-resolution Based on Sparse Representation

Zhu Fuzhen^{1,a}, Liu Yue¹, Huang Xin¹ and Zhu Haitao²

¹ College of Electronic Engineering, Heilongjiang University, Harbin, 150080, China;

² Heilongjiang Duobaoshan Copper Industry Inc. , Heihe, 164300, China

Abstract. In order to obtain higher resolution remote sensing images with more details, an improved sparse representation remote sensing image super-resolution reconstruction(SRR) algorithm is proposed. First, remote sensing image is preprocessed to obtain the required training sample image; then, the KSVD algorithm is used for dictionary training to obtain the high-low resolution dictionary pairs; finally, the image feature extraction block is represented, which is improved by using adaptive filtering method. At the same time, the mean value filtering method is used to improve the super-resolution reconstruction iterative calculation. Experiment results show that, compared with the most advanced sparse representation super-resolution algorithm, the improved sparse representation super-resolution method can effectively avoid the loss of edge information of SRR image and obtain a better super-resolution reconstruction effect. The texture details are more abundant in subjective vision, the PSNR is increased about 1 dB, and the structure similarity (SSIM) is increased about 0.01.

1 Introduction

With the development of science, people need digital images with more informatio. Image super-resolution reconstruction (SRR) is a very important image processing technology. The higher resolution of image is, more detail information image has. People have paid more and more attention to the ecological environment. Remote sensing images have also received extensive attention as an important role for obtaining information on land features. SRR has been used widely in the fields of forest and grassland cover detection, wetland resource monitoring, and so on. It is very necessary to improve the resolution of remote sensing images. The image obtained in reality has a certain gap between the real image due to the influence of the imaging system, imaging environment, and other factors^[1]. That is to say, the image obtained by the physical imaging system is often getting low-resolution. The images required for research are all high-resolution, so improving the resolution of remote sensing images to obtain more information has become a hot topic of current research.

Image super-resolution reconstruction is defined as reconstructing a high-resolution image for one or more images in the same background. It is firstly proposed by Tsai^[2] in 1984. Since then, many scholars have conducted deeper research on image super-resolution reconstruction. There are mainly three methods for image super-resolution reconstruction. They are respectively based on interpolation, reconstruction, and learning^[3]. The interpolation method is based on

correlation function to insert missing pixel values on the basis of nearby pixels. The method mainly includes bilinear interpolation, cubic spline interpolation, and non-uniform interpolation^[4]. This method is simple. However, the reconstructed image is over smooth, and the reconstruction quality is not good. It has an obvious sawtooth effect. The more classical ones in the reconstruction-based methods have are POCS^[5], Iterative Back-projection (IBP)^[6], and the Probabilistic Method^[7], and so on. The algorithm is often constrained by a priori information. But there are also problems such as slow convergence and ensure solution. In recent years, a learning-based method has emerged. By training the image patch which similar to the reconstruction images, the functional relationship between low-resolution images and high-resolution images can be directly found through learning, so that the reconstruction of images can be realized. Among them, the algorithm proposed by Yang^[8] in reference[8] has made a great contribution to the research progress of image super-resolution, which refers to directly find the sparse coefficients that can collectively represent low- and high-resolution images in term of the learned dictionary, so as to complete the super-resolution reconstruction of the images.

In this paper, the reconstruction process is improved on the basis of Yang's algorithm in reference[8] to obtain higher spatial resolution image.

2 Sparse representation

^a Corresponding author: Email: zhufuzhen@hlju.edu.cn.

Zhu Fuzhen (1978-), doctor, master supervisor, research interests : image super-resolution, compressive sensing, neural network, deep learning, *et al.*

In the process of image SRR, low-resolution images are obtained by down-sampling and blur filtering from the high-resolution image, which is expressed as:

$$Y = DHX + n \quad (1)$$

where, D denotes a down-sample operator, H denotes a blur filter, n denotes the added noise information, X denotes a high resolution image, and Y denotes a low resolution image^[9]. The purpose of image reconstruction is to find out the relationship between Y and X , so as to get the desired result. There are many infinite solutions in such a definition, and such equation is called underdetermined equation. For the underdetermined equations, the additional regularization conditions are needed to constrain the problem.

Method in reference[8] is to construct an over-complete dictionary. The input signal $y \in R^m$, if there is $\alpha \in R^n$, causes $y = D\alpha$, signal y can be sparsely represented by the dictionary D . Where $D \in R^{m \times n}$ ($m \ll n$), so the equation is called an underdetermined equation. The equation has many infinite solutions.

For this equation, we need additional regularization^[10] conditions to constrain it. Choosing a rigorous convex function can guarantee that the equation has only one solution. The square European norm is the most typical, expressed as:

$$\min_{\alpha \in R^n} \|\alpha\|_0 \text{ s.t. } y = D\alpha \quad (2)$$

where, α denotes sparse coefficient, $\|\alpha\|_0$ denotes the number of non-zero elements which used to measure sparsity. Equation (2) is an ideal case. In practically, $D\alpha$ is an approximation to y . There is a certain error between them. This problem becomes a problem that minimizes the error, expressed as:

$$\min_{\alpha \in R^n} \|\alpha\|_0 \text{ s.t. } \|y - D\alpha\|_2^2 < \varepsilon \quad (3)$$

Solving equation (3) generally uses a greedy algorithm, which mainly includes Orthogonal Matching Pursuit, Progressive Orthogonal Matching Pursuit^[11] and Regular Orthogonal Matching Pursuit^[12]. It mainly iterated based on a certain similarity. The dictionary atom is trained to approximate sparse coefficient.

It was proved by Donoho^[13] that when α is sufficiently sparse, the norm l^0 in the formula can be replaced by a norm l^1 , expressed as:

$$\min_{\alpha \in R^n} \|\alpha\|_1 \text{ s.t. } \|y - D\alpha\|_2^2 < \varepsilon \quad (4)$$

Equation (4) can be expressed as follows using the Lagrange multiplier:

$$\min_{\alpha \in R^n} \|y - D\alpha\|_2^2 + \lambda \|\alpha\|_1 \quad (5)$$

where, λ denotes balanced parameters, Lasso algorithm can be used to solve the equation. The sparsity is judged by α . The more zero elements in α , the better sparseness in image.

3 Dictionary learning

There are various signals in reality, but the SRR image quality depends on a certain correlation between the trained dictionary and the reconstructed image. In order to get better high-resolution images, a series of images are trained and a learning dictionary is got. The dictionary can sparsely represent signals associated with it. The process is called dictionary training.

Given a combination of image pairs for training $p = \{x^h, x^l\}$, where $x^h = \{x^h_1, x^h_2, \dots, x^h_n\}$ represents a set of high-resolution image blocks, $x^l = \{x^l_1, x^l_2, \dots, x^l_n\}$ represents a set of low-resolution image blocks. The purpose of training dictionary is a dimension reduction operation. The object of training dictionary is a large database. The best feature can be found by training dictionary, and then the sparse representation can be obtained. Represent the desired results with as little data as possible. Therefore, it is important to obtain a correct sparse coefficient. Sparse representation is an ill-conditioned problem, so it is difficult to obtain a sparse coefficient practically. The respective sparse coding problem between low-resolution and high-resolution images can be expressed as:

$$D_h = \arg \min_{\{D_h, Z\}} \|x^h - D_h Z\|_2^2 + \lambda \|Z\|_1 \quad (6)$$

$$D_l = \arg \min_{\{D_l, Z\}} \|x^l - D_l Z\|_2^2 + \lambda \|Z\|_1 \quad (7)$$

Above formula can be expressed as:

$$\min_{D, A} \|x - DA\| + \lambda \left(\frac{1}{M} + \frac{1}{N} \right) \|A\|_1 \quad (8)$$

$$\text{where, } x = \begin{bmatrix} \frac{1}{\sqrt{M}} x^l \\ \frac{1}{\sqrt{N}} x^h \end{bmatrix}, D = \begin{bmatrix} \frac{1}{\sqrt{M}} D^l \\ \frac{1}{\sqrt{N}} D^h \end{bmatrix}.$$

Equation (8) is called joint dictionary training. Therefore, we can train and get two dictionaries by means of single dictionary training method.

4 Improved sparse representation image super-resolution reconstruction

Yang in reference[8] used the gradient operator and Laplace operator as filters for LRIs feature extraction. Four filters are used to extract the features.

$$\begin{aligned} O &= [1, 0, -1] & O_1 &= O^T \\ L &= [1, 0, -2, 0, 1] & L_1 &= L^T \end{aligned} \quad (9)$$

Laplace operator is more sensitive to the response of outliers and end points, but it will lose some edge information and increase noise in the image. In this paper, we use the adaptive filter. The different regions of an image have their own image characteristics. Given Z_{\min} denotes the minimum brightness value, Z_{\max} denotes the maximum brightness value, Z_{med} denotes the median in S_{xy} , Z_{xy} denotes the brightness value at

coordinates (x, y) . The adaptive median filter algorithm can be expressed as follows.

Level *A*: if $Z_{\min} < Z_{med} < Z_{\max}$, turn to level *B*; else increase window size. If the window size is less than or equal to S_{\max} , repeat level *A*; else output Z_{med} .

Level *B*: if $Z_{\min} < Z_{xy} < Z_{\max}$, output Z_{xy} ; else output Z_{med} . where, S_{\max} denotes the maximum size allowed in the adaptive filter window. Another option for the last step in level *A* is to output Z_{xy} .

Iterate computation is used in the process of SRR. Yang's method in reference[8] used a Gauss filter operator, but the Gauss filter operator is more sensitive to noise and discrete points. In this paper, the mean filter is chosen, which can get better smooth effect and faster converging speed compared with Gauss filtering. So mean filter is effective and simple. The mean filter method is to select a template for the current pixel being processed. The template is composed of several adjacent pixels and replacing the value of the original pixel with the mean value of the template. The process of mean filter can be expressed as:

$$g(x, y) = \frac{1}{m} \sum f(x, y) \quad (10)$$

where, $f(x, y)$ denotes pixel of image to be processed, $g(x, y)$ denotes the result, and m denotes the total number of pixels.

5 Experimental results and analysis

The process of image SRR can be represented as three stages. The input image is preprocessed to obtain the training samples. Then, the KSVD algorithm^[14] is used to learn and train the dictionary, and the dictionary is initialized in turn. The sparse coefficients are calculated and the training dictionary is iteratively updated to obtain the trained sparse dictionary. Finally, the improved sparse reconstruction method is used for SRR to obtain high-resolution.

The size of the trained dictionary was 1024, the image non-overlapping extraction size was 0.15, and the image magnification was 2. In this paper, 8 groups remote sensing images are selected and reconstructed with different algorithms in the test experiment. Two group experiments results are listed in this paper, as shown in Figure 1-2. When the number of iteration is 10 and the maximum coincidence is 5, the image reconstruction quality is the best.

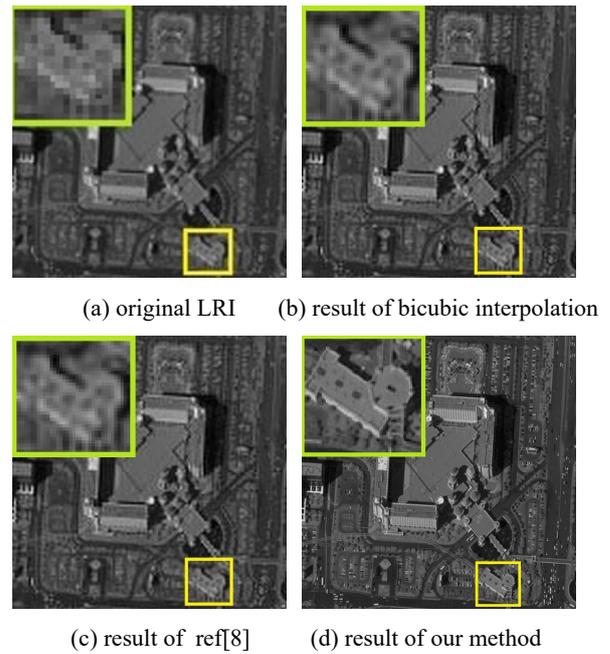


Figure 1. Comparison of the first group remote sensing image SRR with different algorithm

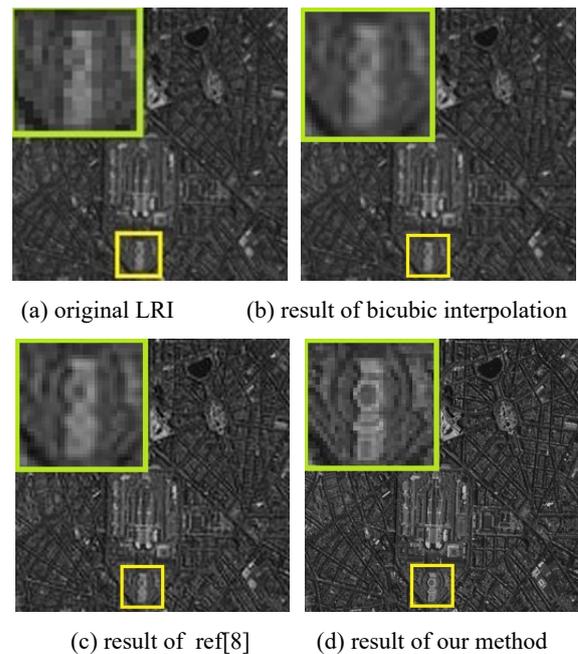


Figure 2. Comparison of the second group remote sensing image SRR with different algorithms

Comparing remote sensing image SRR results with different algorithms, experiments shows that result of the improved sparse representation image super-resolution reconstruction has clearer visual effects and richer high frequency texture details. In order to evaluate various image super-resolution reconstruction effects more intuitively, Peak Signal-to-noise Ratio (PSNR) and structural similarity (SSIM) are used, which are defined as equation (11) and (12) respectively. The statistic results of these two objective evaluation parameters are shown in Figure 3 and Figure 4.

$$PSNR = 10 \times \log_{10} \frac{255^2 MN}{\sum_{i=1}^M \sum_{j=1}^N (r_{ij} - f_{ij})^2} \quad (11)$$

where, M , N are the number of pixels contained in the image, r_{ij}, f_{ij} are the pixel values of the reconstructed high and low resolution images at the point (i, j) .

$$SSIM = \frac{4\mu_x\mu_y\sigma_{xy}}{(\mu_x^2 + \mu_y^2)(\sigma_x^2 + \sigma_y^2)} \quad (12)$$

where, μ_x and μ_y denote mean, σ_x and σ_y denote variance of image, σ_{xy} denotes covariance of image.

The higher PSNR is, the more details SRR image have, and the closer of SSIM is to 1, the better effect SRR image has.

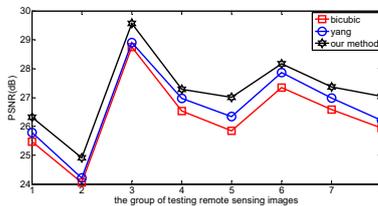


Figure 3. Comparison of PSNR with three algorithms

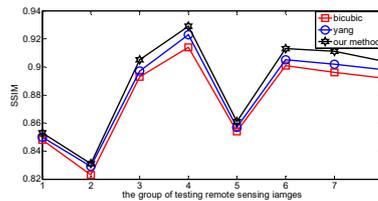


Figure 4. Comparison of SSIM with three algorithms

The experimental results show that the proposed method can get better remote sensing image SRR effect than other methods, not only in subjective visual effect but also in some objective valued parameters such as the peak signal-to-noise ratio and structural similarity .

6 Conclusion

In this paper, an improved sparse representation remote sensing image SRR algorithm is proposed. Specifically speaking, KSVD algorithm is used to train the sparse representation dictionary, and adaptive filtering method is used to improve the feature block extraction, and mean filtering is used to optimize the sparse reconstruction iteration method. The experiment results show that this improvement avoids the loss of edge information and obtains a higher quality SRR remote sensing image without increasing computation complexity. The subjective visual effect is better and the texture details are more abundant than other sparse representation SRR algorithm. The objective evaluation parameter is also increased, that is, PSNR is improved by about 1 dB, and the SSIM is increased by about 0.01 respectively.

Acknowledgment

This work is supported in part by the National Natural Science Foundation China (61601174), in part by the Postdoctoral Research Foundation of Heilongjiang Province (LBH-Q17150), in part by the Science and Technology Innovative Research Team in Higher Educational Institutions of Heilongjiang Province (No. 2012TD007), in part by the Fundamental Research Funds for the Heilongjiang Provincial Universities (KJCXZD201703) and in part by the Science Foundation of Heilongjiang Province of China (F2018026) .

References

1. J.W. Yang, H. Peng, L. Liu, et al. Remote sensing image restoration based on zero-norm regularized kernel estimation[J]. Optics and Precision Engineering, **22**, 8 (2014)
2. R.Y. Tsai, T.S. Huang. Multiframe image restoration and registration[J]. CVIP, **1**, 5 (1984)
3. J Jiang, X.S. Zhang. A review of super-resolution reconstruction algorithms[J]. Infrared Technology, **34**, 7 (2012)
4. S.C. Lin, C.T. Chen. Reconstructing vehicle license plate image from low resolution images using nonuniform interpolation method[J]. IJIP, **1**, 8 (2007)
5. H Huang, X Fan, C Qi, et al. Face Image Super-Resolution Reconstruction Based on Recognition and Projection onto Convex Sets[J]. Journal of Computer Research & Development, **42**, 8 (2005)
6. E.B. Castro, M. Nakano, G.S. Perez, et al. Improvement of image super-resolution algorithms using iterative back projection[J]. IEEE Latin America Transactions, **15**, 6 (2017)
7. K Donaldson, G.K. Myers. Bayesian super-resolution of text in video with a text-specific bimodal prior[J]. IJDAR, **7**, 9 (2005)
8. J. Yang, J. Wright, T.S. Huang, et al. Image super-resolution via sparse representation[J]. IEEE TIP, **19**, 13 (2010)
9. C.Z. Deng, W. Tian W, S.Q. Wang, et al. Super-resolution reconstruction of approximate sparsity regularized infrared images[J]. Optics and Precision Engineering, **6**, 7 (2014)
10. M.A. Davenport, M.B. Wakin. Analysis of orthogonal matching pursuit using the restricted isometry property[J]. IEEE Trans. on Information Theory, **56**, 7 (2010)
11. D. Needell, R. Vershynin. Signal recovery from incomplete and inaccurate measurements via ROMP[J]. J-STSP, **4**, 7 (2013)
12. T. Robert. Regression shrinkage and selection via the lasso: a retrospective[J]. Journal of the Royal Statistical Society, **73**, 10 (2011)
13. D.L. Donoho, Y. Tsaig, I. Drori, et al. Sparse solution of underdetermined systems of linear equations by stagewise orthogonal matching

- pursuit[J]. IEEE Trans. on Information Theory, **58**, 28 (2012)
14. M. Aharon, M. Elad, A. Bruckstein A. K-SVD: an algorithm for designing overcomplete dictionaries for sparse representation[J]. IEEE Trans. on Signal Processing, **54**, 12 (2006)