Journal of Image Processing Theory and Applications (2025)
Clausius Scientific Press, Canada

DOI: 10.23977/jipta.2025.080111
ISSN 2560-6239 Vol. 8 Num. 1

# *A fast restoration method for Hyperspectral image of NLRTAPC*

**Bo Li[a], Xue Yan[b], Xiaowei Yuan[c], Juan Deng[d], Wangyu Liao[d,*]**

*School of Information and Engineering, Sichuan Tourism University, Chengdu, 610100, China*
*[a]sctulibo2021@126.com, [b]13368545@qq.com, [c]44016042@qq.com, [d]aihui4888@gmail.com*
**Corresponding author: liaowangyu@vip.163.com*

***Keywords:*** Hyperspectral Image(HSI), Artificial Bee Colony Algorithm(ABC), Hadoop

***Abstract:*** Hyperspectral image(HSI) is an image cube with continuous narrow bands as the dimension, which has the advantages of high precision and multiple details. The most typical problem is that the HSI has a large capacity and is prone to mixed noise. When using the NLRTAPC model to denoise and restore HSI, a large amount of computer resources need to be consumed, so the efficiency is low when on a single machine system. In response to this, the paper proposes a fast processing method, this method takes Hadoop as the processing platform, deploys the NLRTAPC method on the nodes of Hadoop, and in addition, takes the Artificial Bee Colony algorithm(ABC) as the optimization tool, which is responsible for the optimal configuration of the number and parameters of the slices, and the NLRTAPC of each node runs in a distributed parallel manner to restore the slices. When all the nodes are finished, the final results of the nodes are spliced in order, so as to realize the rapid processing of the HSI. After simulation experiments and data analysis, the efficiency of this method is more than a dozen times higher than that of the single machine processing platform, showing the advantages of fast efficient and low error rate.

## 1. Introduction

HSI is a multi-dimensional image cube with several hundred to a thousand contiguous narrow bands, each of which can record the detailed details the earth's surface features. Cut into a number of two-dimensional matrix images by bands. So, a HSI can be simply considered as a collection of two-dimensional matrix images in the spectral dimension, which is an extension of a two- matrix image in the spectral dimension. HSI has the characteristics of high precision and rich details, and is often used in high-precision data collection, which is an important imaging means modern remote sensing technology.

However, in the process of high-spectral image imaging, it is inevitable to produce a variety of different types of noise due to its own hardware reasons and process, such as Gaussian noise, pulse noise, dead line noise and strip noise, etc., and various types of noise mix together, forming a complex mixed noise. Mixed noise leads to the degradation of HSI, directly affecting the accuracy of subsequent applications such as classification, unmixing, and ground object detection HSI, and therefore, many researchers have proposed restoration techniques to address mixed noise in HSI. Chen Y[1] et al. proposed a restoration model for HSI from the perspective of low-rank sparse and tensor decomposition, which has a denoising and reconstruction effect on certain noise; Fan H[2] et al. based on the low-rank theory, the tensor decomposition was used to achieve the restoration of HSI, and the reconstruction effect was good; He W[3] et al. proposed a low-rank restoration model for HSI based on total variation, which a certain suppression and reconstruction effect on mixed noise. These research results all start from the sparsity and low-rank of the HSI, proposed their own research model, and had a certain suppression effect on noise or mixed noise, providing reference for other research results. Moreover, Hz A[4] et al. proposed a restoration model based on the characteristics of non-convexity, low-rank, and self-similar of the HSI, and obtained good experimental results; some researchers proposed the restoration methods and models of the HSI from the perspective of low-rank weighting [5-8], which all achieved certain effects, suppressed some of the noise in the HSI , and had obvious effectiveness.

To sum up, a large number of researchers have proposed the methods and models of HSI reconstruction from the perspective of methods and models, among which the prominent models include: NLRTAPC, BM4D, NONLRMA, GSSTV, and WGLRTD, which have excellent HSI reconstruction performance among which the NLRTAPC model is particularly prominent and is often used as a typical case for comparative analysis. However, these models all operate on single machine system platforms, and generally have the disadvantages of low processing efficiency, long latency, and poor feasibility of experiments. They did not deeply study the restoration of high-spectrum images from the perspective of running efficiency, nor did they propose the restoration research of high-spectrum images based the big data platform, so, against the above shortcomings of the researchers, this paper proposes a fast restoration method for HSI of NLRTAPC.

## 2. Model and Method

The fast recovery method of high-spectrum image(HSI) of NLRTAPC proposed in this paper needs to solve three problems.

Firstly, the deployment of the Hadoop big data processing platform. This is the basic platform for the rapid recovery of high-spectral images, and all operations performed on Hadoop, and its distributed parallel function is the basic support for rapid processing. Secondly, the NLRTAPC is deployed on each node of the Hadoop platform for distributed processing, and each node uses the NLRTAPC model to process the HSI. When numerous nodes run NLRTAPC independently at the same time, the distributed parallel of NLRTAPC is realized. Thirdly, the artificial bee colony algorithm is deployed to the Hadoop platform to achieve the optimization of platform sharding and parameter configuration.

### 2.1 Deployment of Hadoop platform

The Hadoop big data platform, as the most core platform support for the rapid recovery operation of high-spectral images, provides the most important platform support for recovery. The number of nodes in Hadoop is large, and the nodes can run independently. When all nodes are running, the task of distributed parallel operation is realized, and its is much higher than that of the single machine system platform. Figure 1 is the 45th band of the public HSI, and it can be seen from the state of the image that it contains a of noise, that is, mixed noise, as shown in figure 1.

Figure 1. Noisy 45th band image

In order to quickly complete the restoration of HSI, it is necessary to deploy the Hadoop big data platform,so as to facilitate the distributed parallel execution of multiple nodes,and system architecture of the platform is shown in Figure 2.

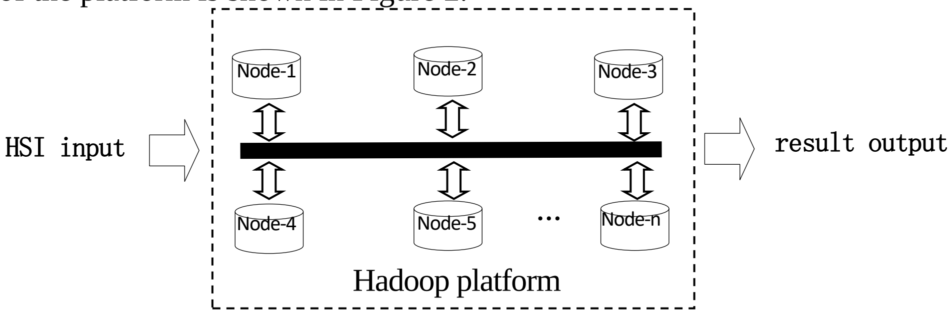


Figure 2. The structure of Hadoop platform

As seen in figure 2, each node is a computer with independent processing capability. When the HSI enters the Hadoop platform from the left port, the platform completes the tiling of the HSI, and the number of tiles is equal to the number of nodes in the platform, ensuring that each node can divided into tasks. Each node uses NLRTAPC to independently complete the data recovery processing, and the result(i) is output from the right port, i=1,2,3…n, then concatenate the all results in order to obtain the final processing result.

When each node deploys the NLRTAPC, the nodes can operate independently to run the restoration processing of HSI segmentation. The model of NLRTC is expressed by Eq. 1, as shown below.

$$\arg\min_{\boldsymbol{L},\boldsymbol{I},\boldsymbol{G}} \|\boldsymbol{L}\|_{\mu} + \beta\|pc(\boldsymbol{Y})\|_{1,p} + \lambda\|\boldsymbol{I}\|_{1} + \gamma\|\boldsymbol{G}\|_{\mathrm{F}}^{2} \quad s.t. \quad \boldsymbol{L} + \boldsymbol{I} + \boldsymbol{G} = \boldsymbol{Y} \tag{1}$$

The solution of equation.1 shows that the model can obtain the optimal solution, which indicates the correctness of the model, and the Alternating Multiplier method is used solve it, and its solution process is shown in table 1.

Table 1. Algorithm for solving the NLRTAPC

| Solving NLRTAPC model by Alternating Multiplier method |
|---|
| Input: Observation data $\boldsymbol{Y}$ ,Index set $\Omega$ ,parameter: $\beta > 0, \lambda > 0, \gamma > 0$ . |
| Initialize: $\boldsymbol{L} = \boldsymbol{Y}$ , $\boldsymbol{I}=0$, $\boldsymbol{G}=0$, $\boldsymbol{Z}=pc(\boldsymbol{Y})$, $\boldsymbol{W}=0$, $\boldsymbol{B}=0$ ,k=0,num=1000 and $\varepsilon=10^{-5}$ ,i=0: |
| (1) $\|\| L^{l+1} - L^{l} \|\|_F / \|\| L^{l} \|\|_F \geq \varepsilon$ or k<num; |
| (2)k ++; |
| (3) $\bar{Y}^k = fft(\boldsymbol{Y}^{(k)},[],3)$ ; |
| (4)Loop i=1, until the maximum number of bands of the observed data B; |
| (5)Calculate $L^{(i)}$ according to the known conditions; |
| (6)Calculate $Z^{(i)}$ according to the known conditions; |
| (7)Calculate $I^{(i)}$ according to the known conditions; |
| (8)Calculate $G^{(i)}$ according to the known conditions; |
| (9) Update the Lagrange multiplier and penalty parameter respectively according to formula (5-8); |
| (10) End loop; |
| (11) restore $\boldsymbol{L}$ , $\boldsymbol{I}$ , $\boldsymbol{G}$ , $\boldsymbol{Z}$ by fold(unfold(•)) , $\boldsymbol{Y}^{(k+1)} = \boldsymbol{L}^{(k+1)}$ ; |
| (12) $Y^{k+1} = ifft(Y^{k+1}, [], 3)$ ; |
| output: denoised $\boldsymbol{L}^{(k+1)}$ |

The solution of NLRTAPC can be realized by table 1, which shows that NLRTAPC is feasible and can be calculated to obtain the processing results high-resolution image segmentation.

## 2.2 Optimization of artificial bee colony(ABC)

When HSI is divided into segments, the workload of each node is dynamically changing, and the allocation of workload is completed by the ABC. The ABC allocates different workloads according to the different processing capabilities of each node, so that the workload and processing capability of the platform are optimized, and there is no situation where the is too small or overloaded. After using the artificial bee colony algorithm, figure 2 can be transformed into the form of figure 3.

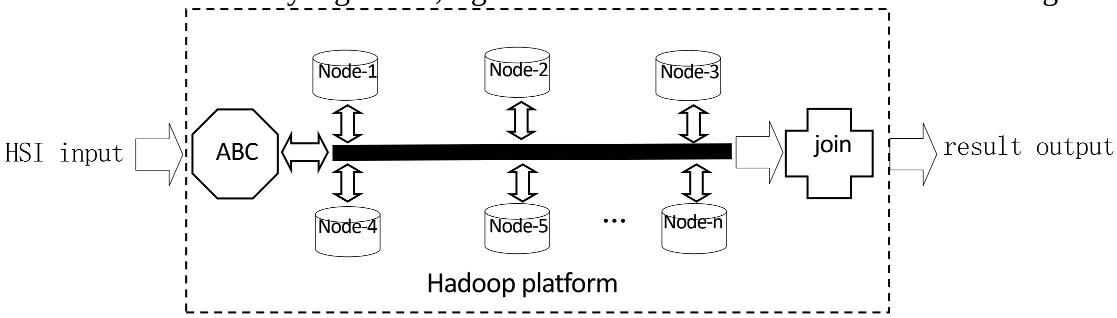


Figure 3. The structure of Hadoop with ABC deployed

In figure 3, the HSI enters into Hadoop platform from the left end of interface, artificial bee colony (ABC) divides the HSI into n parts according to the number of nodes-n, and each part of HSI is distributed to each node, which can be expressed by equation 2.

$$\text{Node(i)=NLRTAPC(HSI(i))} \tag{2}$$

$HSI(\mathrm{i}) = \frac{1}{\rho}(HSI)$ ,The value of $\rho$ is determined by ABC, which will compare and configure according to the parameters of each node, allocate the most suitable workload for each node, and its chart is shown in figure 4.

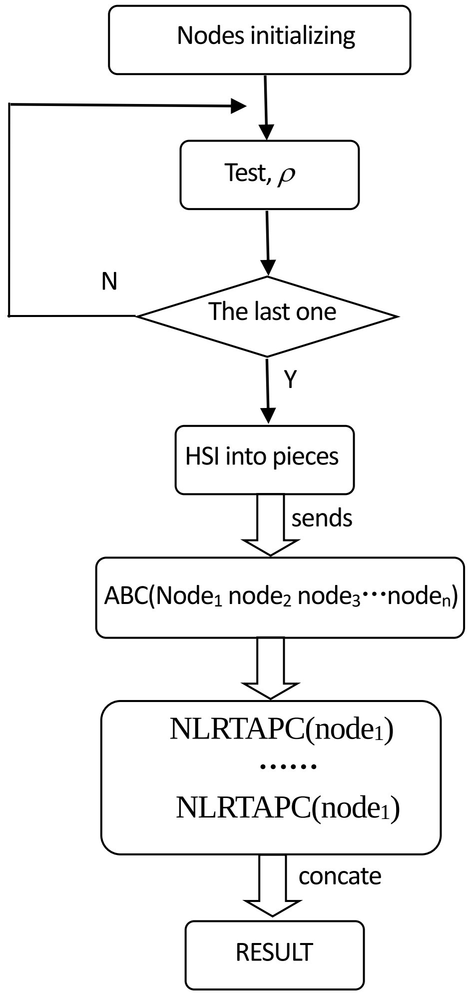


Figure 4. Flow diagram of system

From the process shown in figure. 4, the operation of the system is divided into 5 steps.

Step 1, Initialize all nodes in Hadoop, all nodes are in ready state.

Step 2, Complete the performance test of the node and obtain the performance parameter $\rho$ by loop.

Step3, Complete the HSI fragmentation and send under the control of ABC.

Step4, On each node, the NLRTAPC processing of HSI is completed in a distributed and parallel manner.

Step5, The results processed by NLRTAPC are spliced together, which implements rapid recovery.

The algorithm for implementing the system is shown in Table 2.

In table 2, ABC performs a performance test on the nodes, and the test results are returned in the form of parameters, and then the number of can be obtained according to the performance parameters and the number of image blocks, and the corresponding optimization can be completed. The optimal number of blocks is divided to each corresponding node. After node gets the block, it is denoised by using the NLRTAPC model. When all nodes execute NLRTAPC in a distributed and parallel way, the high- mixed denoising is realized. The running results of all nodes are stitched together, and the function of NLRTAPC is completed.

Table 2. System algorithm

| Node initialization, access node system parameters |
| --- |
| input:HSI,the number of nodes, Test data |
| Initialize:i=1,n=real(nodes)<br>For(i,i<=n,i++)<br>test(chr); |
| dly-time(i)=time(test);<br>$\rho$ =Min(dly-time);<br>$HSI(\mathrm{i})=\frac{1}{\rho}(HSI)$<br>$\sum_{i=1}^{n} HSI(\mathrm{i})=1$<br>End for<br>Split by band;<br>Distribute the slices to each node;<br>Do<br>pic(i)=NLRTAPC(HSI(i));<br>For (i=1;i<=n,i++)<br>Result=Concat pic(i)<br>Result+=pic(i)<br>End for<br>Get result; |

## 3. Experiment Analysis and Discussion

In the simulation experiment, the Hadoop platform contains 50 nodes, in order to balance the completion of the NLRTAPC task, the configuration of node is the same, the hardware is DELL Precision T3660, Intel i9 12900KF, 32G memory, the hard capacity is 2T, the graphics card is NVIDIA A4000, the network communication speed is 1000MPS, even so, the performance of node is slightly different, so the experiment is mainly divided into two parts.

### 3.1 ABC optimization partitioning

In the simulation experiment, ABC is responsible for two aspects of the work. Firstly, send test data to obtain the performance parameters $\rho$ of the node,which provides a direct basis for the fragmentation of HSI. Secondly, the obtained parameters are optimized to find the best parameter configuration, so that each node is allocated to the HSI fragment matching its performance. Taking the open HSI( Hydice) as an example, 400 bands are selected from it. When the number of nodes in the Hadoop platform is 10, the test data is now broadcast to all nodes. ABC allocates 400 bands to 10 nodes according to the calculated performance parameters and the allocation is shown in table 3.

From table 3, it can be seen that the number of bands of the hyperspectral image Hydice obtained at each node is different according to different performance parameters, and this allocation is calculated by ABC according to the optimal strategy, and the allocation in Table 3 is the optimal compared with other allocation methods, and the optimal allocation directly leads to the optimal performance. The data in Table 3 is represented by a scatter plot, as shown in figure 5.

Table 3. The number of bands assigned by ABC

| Node number | $\rho$ | Brands of Hydice |
|---|---|---|
| Hd-001 | 0.17 | 25 |
| Hd-002 | 0.25 | 36 |
| Hd-003 | 0.31 | 45 |
| Hd-004 | 0.20 | 29 |
| Hd-005 | 0.34 | 49 |
| Hd-006 | 0.19 | 28 |
| Hd-007 | 0.21 | 30 |
| Hd-008 | 0.32 | 46 |
| Hd-009 | 0.36 | 52 |
| Hd-010 | 0.41 | 60 |

As can be seen from figure 5, the position of the black dot represents the number of bands, and the distribution of the black dots is irregular. many bands are assigned is entirely determined by the performance parameter $\rho$ of ABC, which is calculated. The larger the $\rho$ of the node, the stronger the performance of the node, and more bands it is assigned, the heavier the processing task. On the contrary, when the ,$\rho$ of the node is smaller, the number of bands assigned to the node is less. In terms of the whole, the allocation quantity is guaranteed to be optimal.

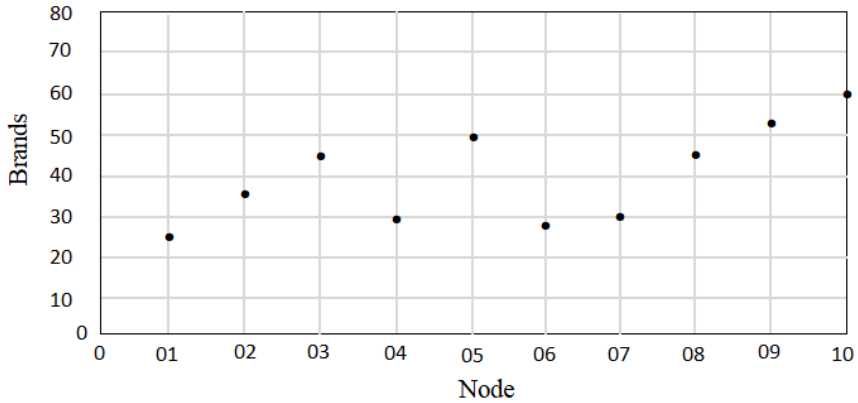


Figure 5. Band allocation chart

## 3.2 Distributed parallel processing of NLRTAPC

After deploying NLRTAPC on each node of Hadoop, the processing efficiency test of NLRTAPC can be carried out. Taking the open HSI(Hydice) as an example, 400 bands are selected from it. When the number of nodes in the Hadoop platform is 10-50, the running time consumed by the average allocation, ABC optimization allocation and single machine system for NLRTAPC processing is shown in Table 4.

Table 4. The superiority of ABC compared

| Hadoop | | | Single system |
|---|---|---|---|
| Number of nodes | Even(s) | ABC(s) | All in one |
| 10 | 321.09 | 215.62 | 3102.41 |
| 20 | 184.74 | 138.20 | 3112.08 |
| 30 | 127.42 | 97.03 | 3135.13 |
| 40 | 94.79 | 81.54 | 3109.47 |
| 50 | 72.18 | 49.03 | 3129.83 |

From the data in table 4, it can be seen that the efficiency of the Hadoop platform is much higher than that of the single machine system in the face of 400 bands of Hydice, and the more nodes, the more efficient it is. Under the same number of nodes, the efficiency of the ABC optimization allocation of bands is significantly higher than that of the average allocation of bands, and the curve diagram of the three methods is shown in figure 6.

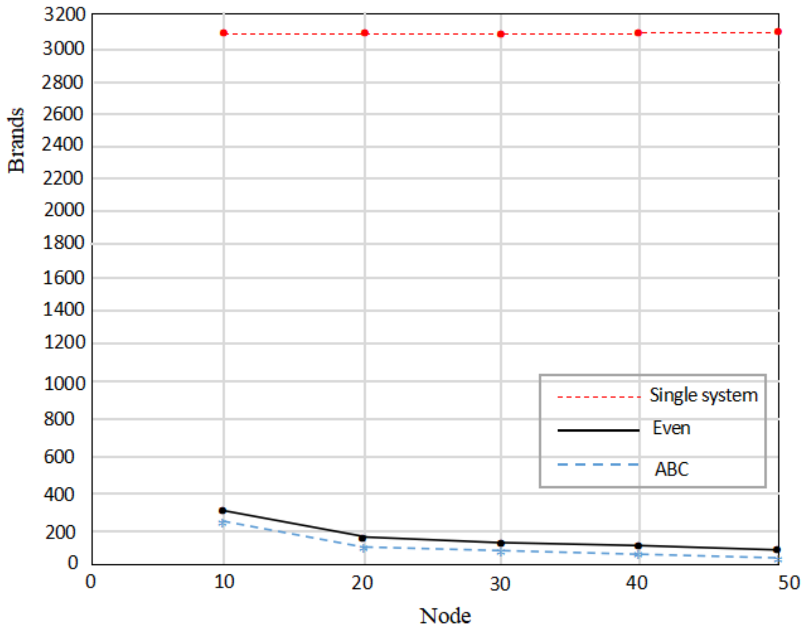


Figure 6. Curve comparison of 3 methods

It can be seen from figure. 6 that for the single machine system, the time it takes to process the HSI using NLRTAPC is the most, and compared with the Hadoop platform, the efficiency is more than 15 times lower, which fully affirms the high efficiency of the Hadoop platform; compared with the average allocation of bands, ABC is better when allocating bands, and its efficiency is also superior to EVEN, fully reflecting the superiority and high efficiency of ABC in processing HSI using NLRTAPC on the Hadoop platform.

## 4. Conclusion

NLRTAPC is a high-quality restoration method for HSI, which is widely used to remove mixed noise from HSI. It a better effect among the same kind of technologies. However, the efficiency of using NLRTAPC to restore HSI on a single machine system is extremely low, resulting unbearable latency, since the HSI is a kind of image cube with the characteristics of high precision, multiple details, and large capacity. Therefore, in response to this typical drawback, this paper proposes a solution that uses the artificial bee colony (ABC) as the optimization tool to optimize the slicing parameter configuration, so that the task and performance can be optimally matched. Moreover, replacing the single machine system with the Hadoop big data platform allows the nodes in Hadoop to process the assigned HSI in a distributed and parallel, which improves the efficiency by several dozen times.

From the data analysis results of simulation experiments, it can be seen that this method can significantly improve the recovery efficiency of NLRTAPC, and the total efficiency further increase with the increase of the number of nodes in Hadoop, and the efficiency is in direct proportion to the number of nodes, and the highest performance is achieved when the number of nodes is equal to the number of bands.

## Acknowledgement

This paper is supported by projects [2021]SCTUZK82, [2021]SCTUTP03, ZLGC2022B02 , 17ZA0286, 2025YFHZ0293and 24SDLYAQYB024.

## References

*[1] Chen Y, He W, Yokoya N, et al. Hyperspectral Image Restoration Using Weighted Group Sparsity-Regularized Low-Rank Tensor Decomposition[J]. IEEE Transactions on Systems, Man, and Cybernetics, 2019: 1-15.*
*[2] Fan H, Chen Y, Guo Y, et al. Hyperspectral image restoration using low-rank tensor recovery[J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2017, 10 (10): 4589-4604.*
*[3]He W, Zhang H, Shen H et al. Hyperspectral image denoising using local low-rank matrix recovery and global spatial–spectral total variation[J]IEEE J. Sel. Topics Appl. Earth Observ. RemoteSens, 2018, 11 (3): 713-729.*
*[4] Hz A, Xx A, Jn B. Hyperspectral image denoising via global spatial-spectral total variation regularized nonconvex local low-rank tensor approximation[J]. Signal Processing, 2020, 178.*
*[5] Ji T Y, Huang T Z, Zhao X L, et al. A non-convex tensor rank approximation for tensor completion[J]. Applied Mathematical Modelling, 2017, 48: 410-422.*
*[6] Jiang J, Yang J, Cui Y, et al. Mixed noise removal by weighted low rank model[J]. Neuro computing, 2015, 151 (2): 817-826.*
*[7] Jin Z, Wan Z, Jiao Y, et al. An Alternating Direction Method with Continuation for Nonconvex Low Rank Minimization[J]. Journal of Scientific Computing, 2016, 66 (2): 849-869.*
*[8] Shi Q, Cheung Y, Zhao Q, et al. Feature Extraction for Incomplete Data Via Low-Rank Tensor Decomposition With Feature Regularization[J]. IEEE Transactions on Neural Networks, 2019, 30 (6): 1803-1817.*