Scholar Publishing Group
Distributed Processing System
https://doi.org/10.38007/DPS.2022.030101
ISSN 2790-0916 Vol. 3, Issue 1: 1-9

Ship Image Classification Based on the Architecture of Imaging Spectral Data Distributed Processing System

Petkov Christopher*

Univ Batna2, Comp Sci Dept, LaSTIC Lab, Batna 05078, Algeria

**corresponding author*

Keywords: Hyperspectral Remote Sensing, Imaging Spectral Data, Distributed Processing System, Ship Image Classification

Abstract: With the rapid development of satellite remote sensing technology in China, it occupies an important position in the fields that require extremely high data accuracy, detail and in-depth information, and plays an increasingly important role in many aspects such as life, safety, and mineral resource detection. The application of imaging spectral images in various fields in China has just started. The purpose of this paper is to study the classification of ship images based on the structure of a distributed processing system for imaging spectral data. In this paper, referring to the feature estimation method and theory in traditional digital image processing system, a new technique of spectral image classification is proposed, and a new system is designed. Note that there are many similar or even duplicate group information in the image data, these feature information will waste a lot of time and computational resources in the process of repeated calculation, and cause the "Hughes" effect due to the increased participation in the amount of data in classification, accuracy will decrease. Therefore, it is necessary to select an appropriate area to participate in the calculation before distribution. To achieve the purpose of reducing time cost and improving accuracy and stability. In this paper, a selective-band distributed spectral data processing algorithm is proposed to solve the above performance and accuracy problems. Experiments show that the band selection algorithm in this paper consumes about one second less time than the traditional band non-selection algorithm, and the recognition accuracy rate reaches more than 97%.

1. Introduction

Imaging spectral images is mainly divided into two tasks: detection and classification. The traditional method of imaging spectral image ship detection mainly includes the following two stages: firstly, the image is segmented by sea and land; then, the ship is detected by the constant false alarm rate algorithm. The imaging spectral image ship classification method mainly includes the following three stages: first, the initial image is preprocessed to enhance the input image; then the feature parameters for classification are calculated; finally, the classification is based on the features. The traditional algorithm process of detection and classification is discontinuous, and the robustness is not high. In recent years, deep learning classification models, especially convolutional neural network methods, which are fully automatic and efficient, have achieved great success in the field of computer vision [1-2].

In the research of ship image classification based on the distributed processing system of imaging spectral data, many scholars have studied it and achieved good results. For example, Fassnacht FE used HYDICE data to compare different groups of priors and hyperspectral groups with group priors. The algorithm is chosen, the group prior is based on PCA parameters (maximum PCA variance and maximum PCA signal-to-noise ratio) and criteria (minimum error distribution, precision resolution, and orthogonality. View subspace) and group modification with group distribution The experimental results show that based on Orthogonal subspace sum distributions have the highest group distribution accuracy [3]. Akharov VM et al. utilize Choquet fuzzy integration to have the advantage of dealing with uncertain multi-source data. Combined with the IT maximum entropy basis, the interaction between bands and the separation index between objects, the hyperspectral data are grouped and selected, and the application of this method in the group separation is illustrated, which can improve the classification accuracy [4].

This paper firstly describes and analyzes the current situation of spectral data processing, briefly describes the current difficulties in spectral data processing, and designs a related distributed processing system for data preprocessing. On this basis, this paper chooses to use semi-supervised local the sparse embedding feature extraction algorithm processes the data. Finally, this paper designs relevant experiments to test the system in terms of time consumption and classification accuracy.

2. Research on Ship Image Classification Based on the Structure of Distributed Processing System for Imaging Spectral Data

2.1. Current Status of Spectral Data Processing

Although hyperspectral image data provides richer spectral and spatial information, in some targeted experiments, it seems that the increase in the amount of data does not bring much benefit to the data processing algorithm in the end, but makes the data processing algorithm time greatly increase. There are two main types of redundancy in hyperspectral images, one is spatial information redundancy, and the other is spectral information redundancy. Spatial information redundancy is mainly manifested in spectral images. Adjacent or similar objects usually show a strong coherence in space, while the values sampled by discrete pixels in spectral images are all different. This feature is not fully utilized, resulting in redundancy. Similarly, spectral information is redundant, and adjacent or similar bands usually do not show much difference in spectral reflectance, which means that a certain band can be completely fitted or predicted by its adjacent bands. So there is spectral information redundancy [5].

The huge amount of spectral data is compared with traditional digital images and multi-spectral images. Due to the development of technology, the imaging capability of imaging spectral cameras has been greatly improved, and the spectral images captured by it contain images of dozens to hundreds of bands. Compared with the image captured by the multispectral camera, the spectral band range collected is much larger, and the spectral resolution of the captured hyperspectral image is also greatly improved. Taking the internationally recognized hyperspectral AVIRIS test image, it captures a total of 224 bands of data, because only 200 bands are left after removing some unusable bands (such as after water absorption and some bands with abnormal reflectivity). If the image size of each band is 144 pixels × 144 pixels, and the pixel depth is estimated, the data space of such a small hyperspectral image has exceeded 80M bytes. Visible hyperspectral image data in the operation of such a huge amount of data, not only to consider the storage pressure and transmission efficiency, but also to the hyperspectral image processing algorithm will be a big challenge [6-7].

2.2. Ship Image Acquisition and Preprocessing

The ship image acquisition is completed in the cloud platform and the sharing system, and the remote sensing image acquisition method is used to collect the ship image in the MIRaaS cloud platform model. It should be noted that the production of ship remote sensing image collection is presented as a unique function. The image is processed by feature registration and noise reduction, and is produced, combined with a feature search method suitable for cloud distribution, using cloud computing. The platform extracts the execution mode of the image, uses the method of statistical feature analysis, analyzes the multi-information processing and resource planning of the ship image, makes the shared feature scanning table of the ship image, and distributes it to the binary similarity judgment T (xi+m, yi+n) ship image to obtain the pixels that can become corner points, create a space-space (Size-Space) to obtain the estimated value of the image similarity threshold o [8-9].

2.3. Algorithm Selection

This paper mainly uses semi-supervised local sparse embedding feature extraction algorithm for data processing. The algorithm is described as follows [10-11]:

Input: Hyperspectral data sample set X={x1, x2, x2,...,xN}∈RD.

Output: The low-dimensional subspace Z={Z1, Z2, Z3,..., ZN}∈Rd after feature extraction.

(1) The original hyperspectral dataset is divided into two parts: labeled samples Xlabel and unlabeled samples Xunlabel,

$$X = \{\mathrm{Xlabel}, \mathrm{Xunlabel}\} \tag{1}$$

(2) Divide the unlabeled data samples into over-complete dictionaries by groups;

(3) Solve the grouped sparse reconstruction weight matrix;

(4) Solve the global sparse reconstruction weight matrix;

(5) Solve the weight matrix of labeled data samples;

(6) Calculate the eigenvectors corresponding to the d largest eigenvalues of the S2-1S1 matrix, and then use the obtained eigenvectors,

Constitute a new set of projection vectors W;

(7) Finally, transform each coordinate of the original high-dimensional hyperspectral data set according to the space matrix, and calculate the projection
into a low-dimensional space,

$$z = W^T x \tag{2}$$

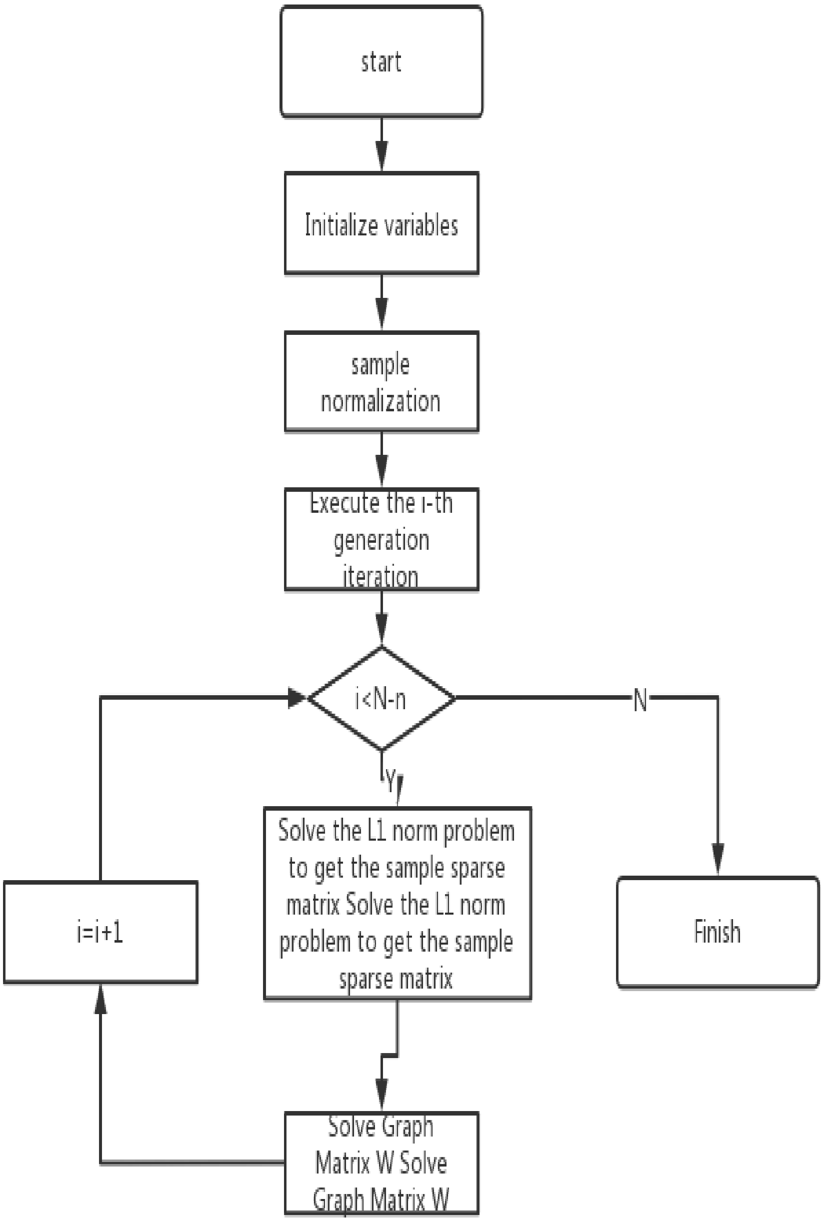

Figure 1. An algorithm flow diagram of the sparse representation coefficients

3. Research and Design Experiment of Ship Image Classification Based on Distributed Processing System Structure of Imaging Spectral Data

3.1. Theoretical Analysis of Distributed Cluster Computing Method

This step is to greatly reduce the computational time consumption in the first step. Because in the band selection process in the first step, although three new features representing pixels at different positions can be calculated while traversing each layer of hyperspectral image, but in the calculation process of traversing the entire hyperspectral image, this One step is very repetitive and time consuming. It is completely possible to split the same spectral image into layers and store them on different data nodes. Then perform parallel computing on different data nodes, as long as the results are finally written back to the same readable node, and then the next selection and classification operations are performed. It all depends on how many compute and data nodes there are. Although the MapReduce distributed computing framework is used to start the Map and Reduce processes, and to schedule nodes, it will consume a certain amount of time in the process of reading and writing back I/O. However, the number of nodes is still in a small range, and the calculation time can be greatly reduced by increasing the number of nodes. However, when it increases to a certain number, the scheduling cost and the time loss cost caused by frequent write-back will also increase, but the overall computing efficiency will decrease [12].

3.2. Experimental Design

This paper conducts experiments on the distributed processing system involved in this paper. The first is to test the time consumption that can be reduced by band selection, and the second is to test the accuracy of the relevant ship image classification and recognition.

4. Experimental Analysis of Ship Image Classification Based on the Structure of Imaging Spectral Data Distributed Processing System

4.1. Consumption Time Comparison

In this paper, for the input of the SVM-kernel classification algorithm, the time-consuming comparison table is calculated: the comparison experiment is a classification experiment without band selection, and the new algorithm proposed in this paper is a band selection algorithm that defines new features on the basis of its experiments. The experimental data are shown in Table 1.

Table 1. Time consumption is compared with the experimental algorithm without band selection

	1	2	3	4	5
Band selection	7.0	7.2	7.4	6.9	6.4
Band does not choose	5.9	5.5	5.6	5.8	5.4

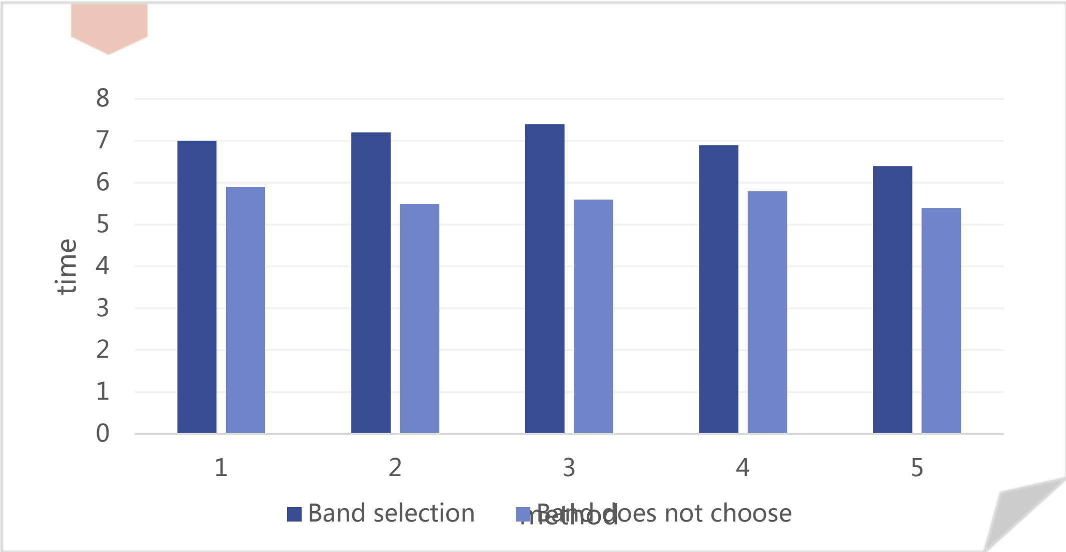

Figure 2. Comparing the time consumption of the band selection algorithm and the new algorithm without running in 5 experiments

It can be seen from Figure 2 that the consumption time of selecting a band is less than that of not selecting a band, about 1s less. This is because in the process of traditional hyperspectral classification experiments, the practice of dividing all spectral data into different feature vectors according to the bands will not only prolong the time of the entire classification process, but also cause images of the classification results due to the accuracy of the data itself.

4.2. Comparison of Recognition Accuracy

In order to verify the accuracy of the classification of ship images by the algorithm in this paper, this paper selects 500 ship images, uses the traditional algorithm and the algorithm in this paper to classify and recognize them, and records the recognition accuracy of each 100 images of the two. The data is shown in Table 2..

Table 2. Comparison of the identification accuracy of this algorithm and the traditional algorithm

	100	200	300	400	500
Traditional algorithm	81	80	76	77	76
The algorithm in this paper	97	98	98	98	97

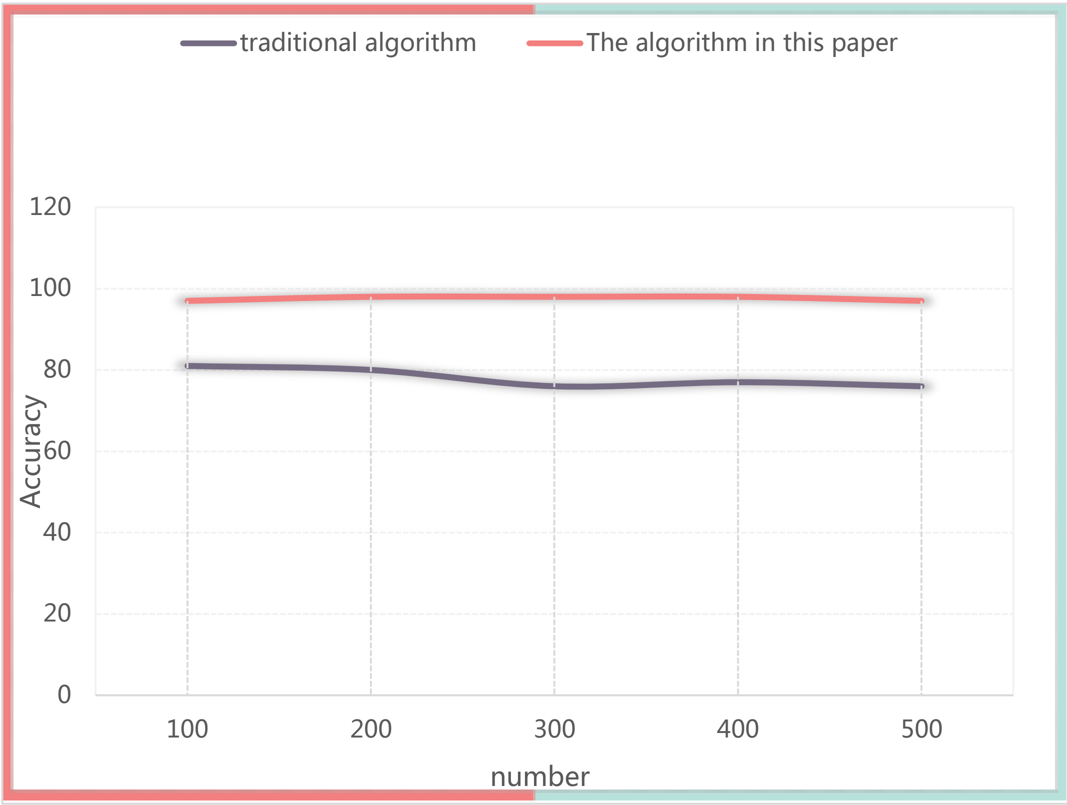

Figure 3. Comparison of ship recognition accuracy between the two different algorithms

It can be clearly seen from Figure 2 that the recognition accuracy of the algorithm in this paper is higher than that of the traditional algorithm, and the image classification and recognition accuracy has reached more than 97%. It is recommended to further improve the image algorithm in the future work. The recognition accuracy is 100%.

5. Conclusion

With the rapid development of human optical science, digital imaging technology and computer technology and performance, hyperspectral technology will gradually replace ordinary camera shooting technology and play an increasingly important role in daily life, security, mineral exploration and many other aspects. As one of the key directions, hyperspectral image processing technology will gradually become a popular research field. In the early stage of this paper, we learned about many hyperspectral image classification algorithms and their defects through investigation. In view of these defects, the theoretical feasibility analysis report of the new algorithm is put forward and sorted out. A hyperspectral image processing algorithm based on band selection is proposed. The algorithm first calculates the positional relationship features (including "viscosity direction" and "viscosity gradient value" and other features of the image data of all bands), and then calculates the band The correlation coefficient between them is selected, and the selected feature bands are sent to the classifier, which improves the accuracy of the classification algorithm and reduces the time required for classification. On the basis of improving the accuracy, the band selection process is segmented and moved into the distributed system, and the parallel advantage of the distributed system is used to greatly shorten the time-consuming calculation of multiple "sticky features" when the band selection algorithm is running. So that it can be put into use in real-world scenarios. The research shows that the method in this paper has better accuracy and lower misclassification rate for ship image classification.

Funding

This article is not supported by any foundation.

Data Availability

Data sharing is not applicable to this article as no new data were created or analysed in this study.

Conflict of Interest

The author states that this article has no conflict of interest.

References

[1] Scoba A N, Mikhaylov V K, Ayesh A, et al. The optimal placement of information resources on the nodes of a distributed information processing system based on a two-tier and three-tier client-server architecture. IOP Conference Series: Materials Science and Engineering, 2019, 483(1):012058 (7pp). https://doi.org/10.1088/1757-899X/483/1/012058

[2] Martinsons C, Hlayhel R. A method to measure the spectral responsivity of a photometer using optical excitations with arbitrary spectral distributions:. Lighting Research & Technology, 2022, 54(4):311-328. https://doi.org/10.1177/14771535211026322

[3] Fassnacht F E, Jana Müllerová, Schmidtlein S, et al. About the link between biodiversity and spectral variation. Applied Vegetation Science, 2022, 25(1):n/a-n/a. https://doi.org/10.1111/avsc.12643

[4] Zakharov V M, Shalagin S V, Eminov B F. Using the same type of IP-cores in the Virtex-6 family FPGA-architecture for distributed image processing. Journal of Physics: Conference Series, 2020, 1658(1):012078 (5pp).

[5] Golomolzin V V, Rublev A N, Kiseleva Y V, et al. Retrieval of Total Column Carbon Dioxide over Russia from Meteor-M No. 2 Satellite Data. Russian Meteorology and Hydrology, 2022, 47(4):304-314.

[6] Thimmaraja Y G, Nagaraja B G, Jayanna H S. A spatial procedure to spectral subtraction for speech enhancement. Multimedia Tools and Applications, 2022, 81(17):23633-23647.

[7] Goedhart J J, Papadakis V M. A machine learning classification methodology for Raman Hyperspectral imagery based on auto‐encoders. Journal of Raman Spectroscopy, 2022, 53(6):1126-1139. https://doi.org/10.1002/jrs.6339

[8] Kiyanagi R, Kofu M, Tatsumi K, et al. Optimization and inference of bin widths for histogramming inelastic neutron scattering spectra. Journal of Applied Crystallography, 2022, 55(3):533-543.

[9] Reddy H H C, I‐Hung Khoo, Moschytz G S. Structure‐induced low‐sensitivity design of sampled data and digital ladder filters using delta discrete‐time operator. International Journal of Circuit Theory and Applications, 2022, 50(6):2228-2251. https://doi.org/10.1002/cta.3248

[10] Men X, Wang A, Xu Y, et al. Vegetation detection based on spectral information and development of a low - cost vegetation sensor for selective spraying. Pest Management Science, 2022, 78(6):2467-2476. https://doi.org/10.1002/ps.6874
[11] Pujades E, Kalbacher T, Dietrich P, et al. From Dynamic Groundwater Level Measurements to Regional Aquifer Parameters— Assessing the Power of Spectral Analysis. Water Resources Research, 2022, 58(5):n/a-n/a.
[12] Legleiter C J, Sansom B J, Jacobson R B. Remote Sensing of Visible Dye Concentrations During a Tracer Experiment on a Large, Turbid River. Water Resources Research, 2022, 58(4):n/a-n/a. https://doi.org/10.1029/2021WR031396
[13] Hiller M, Tkach I, Wiechers H, et al. Distribution of Hyperfine Couplings in a Tyrosyl Radical Revealed by 263GHz ENDOR Spectroscopy. Applied Magnetic Resonance, 2021, 53(7-9):1015-1030.
[14] Wiesman A, Murman D, May P, et al. Spatio-spectral relationships between pathological neural dynamics and cognitive impairment along the Alzheimer's disease spectrum.. Alzheimer's & dementia (Amsterdam, Netherlands), 2021, 13(1):e12200. https://doi.org/10.1002/dad2.12200
[15] Nijs M, Smets T, Waelkens E, et al. A mathematical comparison of non-negative matrix factorization related methods with practical implications for the analysis of mass spectrometry imaging data.. Rapid communications in mass spectrometry : RCM, 2021, 35(21):e9181. https://doi.org/10.1002/rcm.9181
[16] Shailesh K R, Shailesh T. A technical note on digitizing color mapped spectral power distribution images. Color Research And Application, 2022, 47(3):541-554. https://doi.org/10.1002/col.22758
[17] Lucci V, Inskip J, Mcgrath M, et al. Longitudinal Assessment of Autonomic Function during the Acute Phase of Spinal Cord Injury: Use of Low-Frequency Blood Pressure Variability as a Quantitative Measure of Autonomic Function.. Journal of neurotrauma, 2021, 38(3):309-321. https://doi.org/10.1089/neu.2020.7286
[18] Yele V, Azam M, Wadhwani A. Synthesis, Molecular Docking and Biological Evaluation of 2-Aryloxy-N-Phenylacetamide and N'-(2-Aryloxyoxyacetyl) Benzohydrazide Derivatives as Potential Antibacterial Agents.. Chemistry & biodiversity, 2021, 18(4):e2000907. https://doi.org/10.1002/cbdv.202000907