Т. 87, № 6 ЖУРНАЛ ПРИКЛАДНОЙ СПЕКТРОСКОПИИ НОЯБРЬ — ДЕКАБРЬ 2020
V. 87, N 6 JOURNAL OF APPLIED SPECTROSCOPY NOVEMBER — DECEMBER 2020

METHOD FOR FINE PATTERN RECOGNITION OF SPACE TARGETS USING THE ENTROPY WEIGHT FUZZY-ROUGH NEAREST NEIGHBOR ALGORITHM

Qing-bo Li *, Yuan Wei, Wen-jie Li

*School of Instrumentation and Optoelectronic Engineering
at Precision Opto-Mechatronics Technology Key Laboratory of Education Ministry
at Beihang University, Beijing 100191, China; e-mail: qbleebuaa@buaa.edu.cn*

In space target recognition using spectral analysis technology, there is the problem that the composition or chemical properties of surface materials of the space target are similar. This problem leads to the high similarity of spectral curves and low accuracy of space target recognition. Similar object recognition is important in the study of actual space target observation. In this paper, an entropy weight fuzzy-rough nearest neighbor (EFRNN) algorithm is proposed to enhance the recognition accuracy of similar space targets, which is an improvement of the fuzzy-rough nearest neighbor algorithm. By introducing the feature weight determined using information entropy, the features of all the training samples are considered and quantified. Moreover, the proposed algorithm combined with fuzzy-rough set theory can overcome the fuzzy uncertainty caused by overlapping classes and the rough uncertainty caused by insufficient features, to a certain extent. The simulation results show that the proposed algorithm achieves very promising performance compared with existing algorithms. The EFRNN classifier yields an overall classification accuracy of 95.83%. The proposed algorithm is simple and efficient for similar space target recognition. Furthermore, the EFRNN algorithm does not require preset parameters and complex preprocessing.

Keywords: *fine pattern recognition, entropy weight, fuzzy-rough set, space targets.*

МЕТОД ТОЧНОГО РАСПОЗНАВАНИЯ ОБРАЗОВ КОСМИЧЕСКИХ ЦЕЛЕЙ С ПОМОЩЬЮ НЕЧЕТКО-ПРИБЛИЗИТЕЛЬНОГО АЛГОРИТМА БЛИЖАЙШЕГО СОСЕДА ПО ИНФОРМАЦИОННОЙ ЭНТРОПИИ

Q.-B. Li *, Y. Wei, W.-J. Li

УДК 543.42:537.591

*Школа приборостроения и оптоэлектронной инженерии Бейханского университета,
Пекин 100191, Китай; e-mail: qbleebuaa@buaa.edu.cn*

(Поступила 30 августа 2019)

Предлагается нечетко-приблизительный алгоритм ближайшего соседа по информационной энтропии (EFRNN) для повышения точности распознавания похожих космических целей, который является улучшением алгоритма нечеткого ближайшего соседа. Путем введения веса признака, определенного с использованием информационной энтропии, рассматриваются и количественно оцениваются характеристики всех обучающих выборок. Предложенный алгоритм в сочетании с теорией нечетких множеств может в определенной степени преодолеть нечеткую неопределенность, вызванную перекрытием классов, и грубую неопределенность, вызванную недостаточными функциями. Результаты моделирования показывают, что классификатор EFRNN дает общую точность классификации 95.83%. Предлагаемый алгоритм прост и эффективен для распознавания похожих космических целей, не требует предустановленных параметров и сложной предварительной обработки.

Ключевые слова: *точное распознавание образов, энтропийный вес, нечеткое множество, космические цели.*

Introduction. With an increasing number of spacecraft being sent into space, the recognition of space targets has become a topic of great interest in space research [1–6]. In the aerospace field, the term "space target" typically refers to various man-made aircraft and space debris that orbit the earth beyond the atmosphere [2]. At present, optical observation is mainly used to extract the characteristics of the target, including the size, shape, attitude, and orbit, through optical imaging technology and time sequential photometry for space object recognition [7]. Spectral characteristics are an important optical feature of space targets [8]. Space target recognition using spectral features can identify space targets without other characteristic information of space objects, such as geometry and orbit information. When the space target is far away from the observation device, the observed space target occupies very few image pixels, and even becomes a point target that lacks shape and size information [3]. In this case, recognition technology using spectral information has obvious advantages. Additionally, the spectral recognition method belongs to the area of single-frame detection, which does not require the multi-frame information of time sequences. Thus, it is robust to the speed of space target movement, and the amount of computation is relatively reduced, which can improve the recognition speed.

The basis of space target discrimination with spectral analysis technology is that different space object surface materials have different spectral curve features. Recently, some machine learning and deep learning methods have been used for space target recognition using spectral analysis methods. Cauquy et al. [4] used the artificial neural network method to process the central moments of space target spectra, and then the space target was identified according to the Spica database, which was provided by the Maui Surveillance Site in Hawaii, USA. Plemmons et al. [5] developed a nonnegative matrix factorization algorithm with novel smoothness constraints for unmixing reflectance spectra to identify space targets. Deng et al. [3] proposed a multi-scale convolutional neural network for feature learning and classification, which was used for space infrared point object discrimination. However, these methods aim at the recognition of general space targets, not similar space targets.

The reflectance spectra of space objects are highly similar in practical applications because of the similar composition and chemical properties of the surface materials of space objects [6]. Hence, the probability distribution of a certain class of space objects is difficult to determine, and it varies in multiple directions of feature space. In this case, it is difficult to discriminate space objects with high accuracy using traditional pattern recognition methods, such as the K-nearest neighbor (KNN) [9] algorithm and support vector machines [10]. The fuzzy-rough nearest neighbor (FRNN) algorithm is a generalization of the conventional KNN algorithm, and the classification efficiency is enhanced by exploiting fuzzy-rough uncertainty [11, 12]. However, the performance of the FRNN algorithm is unsatisfactory for similar space object discrimination. In this paper, an entropy weight FRNN (EFRNN) algorithm is proposed to discriminate space objects for which the reflectance spectra are highly similar using current observation devices. The proposed algorithm is a fine space object identification method, which has the advantages of simple operation, fast processing, and high accuracy.

Calculation. *Materials.* Four cuboid samples representing space objects were prepared in the simulation experiment. The six surfaces of each cuboid sample were composed of three materials in different proportions. Additionally, only one of the composition materials was different in any two samples among the four samples, which implies that the four cuboid samples were similar objects. The original spectra were obtained by the Headwall Photonics Hyperspec VNIR-N series spectrometer. Moreover, the spectra of each surface of the cuboid samples were collected in turn. Additionally, the distances between the spectrometer and the samples remained fixed. The spectral region of the spectrometer was between 400 and 1000 nm, and the spectral resolution was 2–3 nm. Hence, there was a total of 24 spectra with four classes assigned to the training set.

EFRNN method. The EFRNN algorithm is an improvement of the FRNN algorithm, and it introduces the concepts of information entropy weight and fuzzy-rough sets. Instead of k nearest neighbors, feature weights are determined based on the information entropy of all the training samples. Moreover, the fuzzy-rough sets can help to solve the fuzzy uncertainty caused by overlapping classes and the rough uncertainty caused by insufficient features, to some extent. Therefore, the classification accuracy for similar objects can be improved.

Suppose that the training set $\mathbf{X} = (\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_m)^T$ consists of m training samples with L classes. Each sample contains n features, that is, the ith sample can be represented as $\mathbf{x}_i = (x_{i1}, x_{i2}, \dots, x_{in})$, $i = 1, \dots, m$. Additionally, the class label of sample $\mathbf{x}_i$ is $y_i = c$, $c = 1, \dots, L$. The EFRNN algorithm processes the test sample $\mathbf{q} = (q_1, \dots, q_n)$ using the following steps:

Step 1. The feature weight of training set w is determined using information entropy, and is defined according the following formula:

$$z_{ij} = x_{ij} \Big/ \sum_{i=1}^{m} x_{ij}, \quad \beta = 1/\ln m,$$
$$H_j = -\beta \sum\nolimits_{i=1}^{m} z_{ij} \log(z_{ij}),$$
$$w_j = \frac{1 - H_j}{n - \sum\nolimits_{j=1}^{n} H_j}, \quad \forall j = 1, \cdots, n \tag{1}$$

where z_{ij} is the normalized data, β is an adjustment parameter, H_j is the total information entropy of the jth feature of the sample, and w_j is the information entropy weight of the jth feature of the sample.

Step 2. The information entropy weighted Euclidean distance d between training sample $\mathbf{x}_i$ and test sample $\mathbf{q}$ is calculated by the following formula:

$$d(\mathbf{x}_i, \mathbf{q}) = \sqrt{\sum_{j=1}^{n} w_j (x_{ij} - q_j)^2}, \tag{2}$$

where q_j is the jth feature of test sample $\mathbf{q}$.

Step 3. Considering all training samples, the class confidence value $o(c)$ with which test sample $\mathbf{q}$ can be classified to class c is determined by the following formula:

$$o(c) = \sum_{\mathbf{x_i} \in N} \frac{\mu_c(\mathbf{x_i}) \exp(-d^{1/(p-1)})}{|N|}, \tag{3}$$

where $\mu_c(\mathbf{x}_i)$ denotes the fuzzy membership of the ith sample in class c, p is a parameter that controls the overall weighting of the similarity, N is the training set, and $|N|$ is the cardinality of set N.

Step 4. After all the training samples have been considered, the class label of test sample q is determined as the class with maximum $o(c)$:

$$\text{label}(\mathbf{q}) = \arg\max_c o(c). \tag{4}$$

It should be noted that the conventional KNN algorithm assigns equal weight to the k nearest neighbors, and then the class label of the test sample is predicted based on only these k nearest neighbors. However, the EFRNN algorithm assigns weight based on the contribution of all training samples in the classification, and the contribution is measured using information entropy. Information entropy can objectively evaluate the importance of each feature and quantify it. Additionally, the EFRNN algorithm predicts the class label of the test sample with the class confidence value. In some papers [11–13], the class confidence value is also called the fuzzy-rough ownership function, which can quantify both fuzzy uncertainty and rough uncertainty. Therefore, the EFRNN algorithm overcomes the shortcomings of the conventional KNN algorithm, which are that the classification accuracy is affected by the value of k and the classification ability of overlapping classes is insufficient.

Results and discussion. In this study, a simulation was conducted to demonstrate the effectiveness of the EFRNN algorithm. The experimental process is described as follows.

Data preprocessing. The spectra of the four samples are shown in Fig. 1. Wavelet denoising was used to improve the signal-to-noise ratio of the spectra. The basic wavelet was the sym5 wavelet and the number of decomposition layers was 7. The spectral curves before and after wavelet denoising are shown in Fig. 2. Additionally, isometric feature mapping (ISOMAP) [14] was used to perform feature extraction and dimension reduction to increase the identification accuracy and computation speed.

Statistical analysis. As shown in Fig. 1, in the 400–800 nm region, the spectra of four samples are similar because they have similar spectral shapes and spectral intensities, which indicates that these four prepared cuboid samples can simulate space objects with similar reflectance spectra. A comparison of the classification accuracy of the EFRNN algorithm with that of the KNN and FRNN algorithms demonstrates the effect of the EFRNN algorithm.

After data preprocessing, all 24 spectra were used to verify the classification efficiency of the above algorithms. In the experiment, the control parameter for the EFRNN algorithm was assumed to be $p = 3$, and the fuzzy membership of the ith training sample in class c was assumed to be $\mu_c(\mathbf{x}_i) = 1$. Additionally, the parameter p and $\mu_c(\mathbf{x}_i)$ for the FRNN algorithm were assigned to the same values as those in the EFRNN algorithm. Because of the limited number of samples, the full cross-validation method was used. Table 1 shows the classification results for the KNN, FRNN, and EFRNN algorithms.

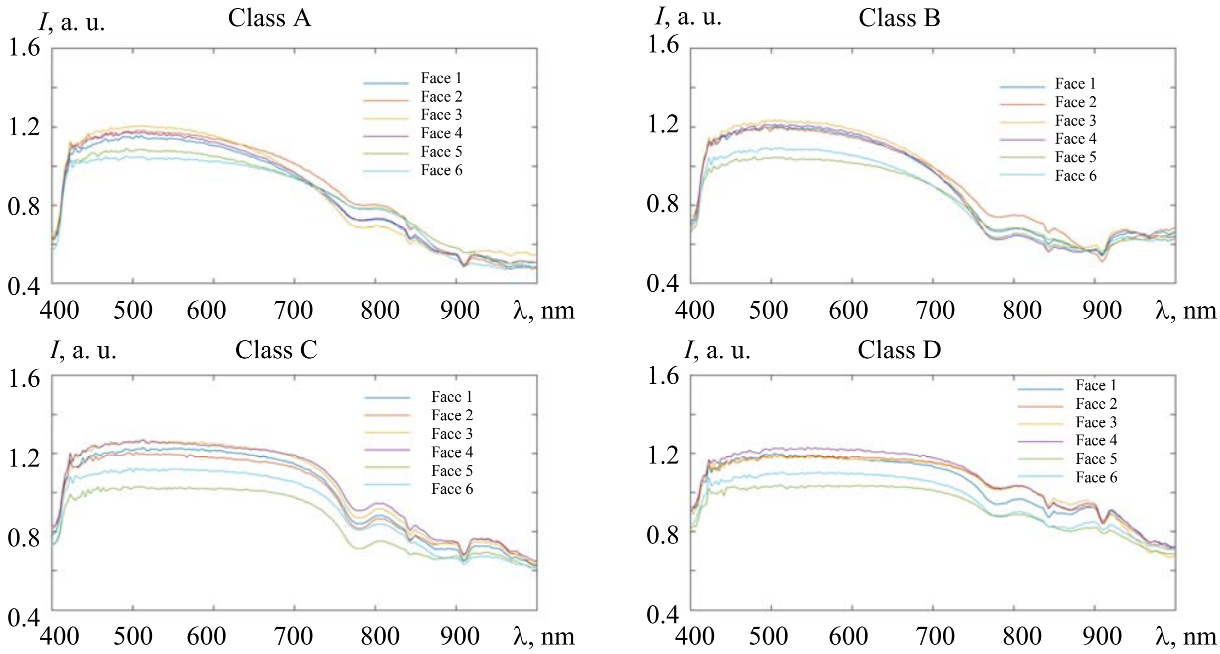

Fig. 1. Spectra of four class samples (Face1–Face6 represent the spectra of six surfaces of the cuboid sample).

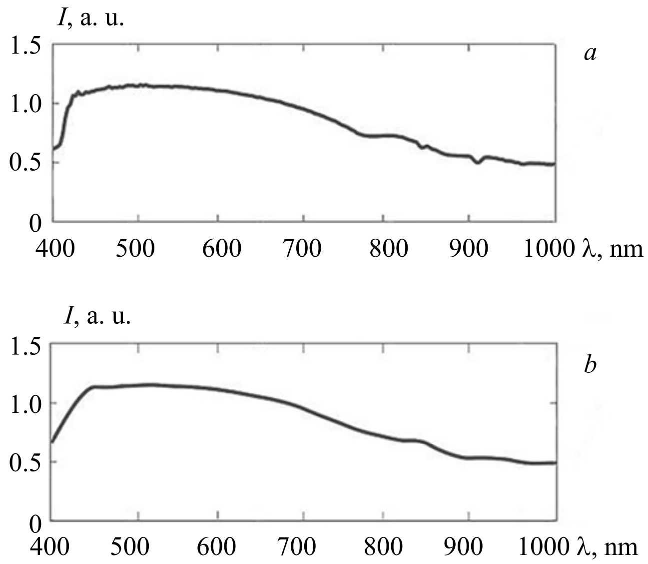

Fig. 2. Spectral curves before (a) and after wavelet (b) denoising.

TABLE 1. Comparison of the Classification Results for the KNN, FRNN, and EFRNN Algorithms

Class	KNN, %	FRNN, %	EFRNN, %
A	83.33	100	100
B	83.33	100	100
C	83.33	83.33	83.33
D	83.33	83.33	100
Overall	83.33	91.67	95.83

In Table 1, the classification accuracy of the EFRNN algorithm for classes A, B, and D reached 100%, and the overall classification accuracy for all samples reached 95.83%. Compared with 83.33% overall classification accuracy of the KNN algorithm, the EFRNN algorithm was clearly more capable in terms of similar object discrimination. Furthermore, the classification accuracy of the EFRNN algorithm was better, to a certain extent, than the 91.67% accuracy of the FRNN algorithm.

The EFRNN algorithm had the highest overall classification accuracy, followed by the FRNN and KNN algorithms. The accuracy of the KNN algorithm greatly depended on the choice of the optimal value of k, and the accuracy was high when the k value was appropriate, which was difficult to achieve. Furthermore, the EFRNN algorithm enhanced the overall classification accuracy of similar samples significantly. The feature weight measured using information entropy objectively evaluated the importance of every feature of spectral data while predicting the class of the test sample. Hence, instead of the weight distance calculated in the FRNN algorithm [11], the confidence value $o(c)$ was obtained by substituting the information entropy weighted Euclidean distance d into Eq. (3). The experimental results also confirmed that the proposed EFRNN algorithm can be used to recognize similar space targets.

Conclusions. Space targets occupy few pixels when the distance between them and observation devices is large. The advantage of spectral analysis technology is that it can discriminate space targets based on the reflectance spectral curves of their surface materials without information about their shape and size. To overcome the problem of the low classification accuracy of space targets with similar reflectance spectral curves in actual observation, an improved FRNN algorithm, that is, EFRNN, was proposed to enhance the accuracy and realize the fine pattern recognition of space targets.

The experimental results show that the proposed algorithm is simple and effective. The EFRNN algorithm does not require complicated preprocessing. Moreover, the spectral data collected by the spectrometer only requires wavelet denoising and feature extraction using ISOMAP. Then, these spectral data can constitute the training set for the classifier. Additionally, by exploiting fuzzy-rough set theory, the confidence value is used as the basis for evaluating the class of the test sample. It alleviates the fuzzy uncertainty caused by overlapping classes and the rough uncertainty caused by insufficient features, to a certain extent. Instead of determining k nearest neighbors like the KNN algorithm, the EFRNN algorithm introduces the feature weight measured using information entropy to quantify and synthesize all the spectral features of training samples. Hence, the information entropy weight reflects the contribution of all training sample information to classification. Combining fuzzy-rough set theory and the information entropy weight, the pattern recognition accuracy of the EFRNN algorithm for similar objects was better than those of the KNN and FRNN algorithms. Thus, the EFRNN algorithm is suitable for the practical application of space target fine pattern recognition.

Acknowledgments. The work is supported by the National Natural Science Foundation of China (Grant No. 61575015).

REFERENCES

1. Y. Han, H. Sun, J. Feng, L. Li, *Meas. Sci. Technol.*, **25**, No. 7, 075203 (2014).
2. Q. Li, K. Wu, Q. Gao, *Spectrosc. Spectral Anal.*, **36**, No. 12, 4067–4071 (2016).
3. Q. Deng, H. Lu, H. Tao, M. Hu, F. Zhao, *IEEE Access*, **7**, 28113–28123 (2019).
4. M. A. A. Cauquya, M. C. Roggemanna, T. J. Schulz, *Proc. SPIE*, **5428**, 48–57 (2004).
5. V. P. Pauca, J. Piper, R. J. Plemmons, *Linear Algebra Appl.*, **416**, No. 1, 29–47 (2006).
6. K. M. Jorgensen, *Using Reflectance Spectroscopy to Determine Material Type of Orbital Debris*, ProQuest Dissertations, University of Colorado at Boulder, 45–49 (2000).
7. J. Zhang, *Chin. J. Opt. Appl. Opt.*, **2**, No. 1, 10–16 (2009) [In Chin.].
8. C. Chang, S. Wang, *IEEE T. Geosci. Remote.*, **44**, No. 6, 1575–1585 (2006).
9. R. Duda, P. Hart, *Pattern Classification and Scene Analysis*, Wiley, New York, 40–41 (1973).
10. C. Cortes, V. Vapnik, *Mach. Learn.*, **20**, 273–297 (1995).
11. M. Sarkar, *Fuzzy Set. Syst.*, **19**, No. 158, 2134–2152 (2007).
12. A. Onan, *Expert Syst. Appl.*, **42**, No. 20, 6844–6852 (2015).
13. L. Sun, C. Li, In: *2009 WRI Global Congress on Intelligent Systems*, May 19–21, Xiamen, China, 311–314 (2009).
14. J. B. Tenenbaum, V. De. Silva, J. C. Langford, *Science*, **290**, 2319–2323 (2000).