APPLY EXPLORATION OF NEURALNETWORK AND PUBLIC GREEN IN THE HANDLE OF MIXED PIXEL OF GENETIC BP PROPAGATION ALGORITHM

LV Yan [a, b,]*, ZHU Yi [a]

[a] Government GIS Research Centre, Chinese Academy of Surveying and Mapping, No.16, Beitaiping Road, Haidian District, Beijing, China
[b] School of Geomatics, Liaoning Technical University, Fuxin, Liaoning, China

KEY WORDS: Images, mixed-pixel, neural-network, Public green, genetic BP propagation algorithm

ABSTRACT:

With the rapid development of RS and more mature and the precision of number of field surveys on the basis of Quick-Bird remote sensing data through visual interpretation of urban public green vegetation diversity survey the exits of mixed. Pixel mixed is one important factor of influence accuracy of Remote sensing image classification. In this article, combine genetic algorithm with BP propagation to a new analysis module. The results show that the algorithm is effective. Remote sensing image processing detect and obtain features information by every pixel. If can identify the ratio of variety typical features compositing mixed pixel by certain means, you can solve the problem of mixed pixel and improve qualitative and quantitative accuracy of remote sensing survey.

1. INTRODUCTION

In remote sensing technologies application, in order to obtain surface information, showing the analysis result of earth observation data only by images or photographs is not enough, it must be compounded for the image analysis result and other variety information distribute of graphical coordinates and obtain accurate information with the method of field investigate and so on. Now green survey through image interpretation generally uses satellite remote sensing image as major source data with supplement of field survey and other information; Urban publics green space management and planning is based on interpreting green status apply RS data combine with ground survey green status which based on satellite remote sensing date and ground survey data as source data.

However, remote sensing image processing detect and obtain features information by every pixel. As a result of limitation of multi-spectral and hyper-spectral remote sensing image spatial resolution, a pixel often covers dozens or even hundreds square meters area, in which variety objects contain which formed mixed pixel. Mixed pixel not only affects the accuracy of identification and classification of features and it is major obstacle of remote sensing technology quantify development. If can identify the ratio of variety typical features compositing mixed pixel by certain means, you can solve the problem of mixed pixel and improve qualitative and quantitative accuracy of remote sensing survey.

The genetic algorithm modal which combined by the addressing a mixed-pixel decomposition and further classification algorithm fits analysis result super-surface based on deeply study of principle of mixed-pixel. The idea of type decomposition is identify mixed-pixel according to ratio of various types contained in mixed-pixel (type ratio), the pixel is assigned to the type which ratio is the largest. The main method are least squares, quadratic programming, etc[2]. This method shows higher performance in remote sensing classification and can solve the problem of mixed features classification.

Genetic Algorithm (GA) can overcome the limitation of BP algorithm to some extent. GA is a novel optimization algorithm which is based on the simulation of natural selection and genetic mechanism. It can make the global optimization search of simulation to the biological evolution with the adaptive probability on the computer. GA forms the search algorithm with the characteristics of "generation + test" through a comprehensive simulation of natural selection and the genetic mechanism. The operation of yaGA is a typical iterative process. And what is the basic process and structures included is as follows:

(1) To choose the coding strategy and to convert the parameter set X and the domain into the structure space of bit String.
(2) To define the f (X) which function is for the fitness.
(3) To fix on the determine the genetic strategies including n who is the size of the selected groups, the selection crossover, the variation method as well as the determination for the crossover probability, mutation probability and other genetic parameters
(4) To generate a random initialization groups P and then to calculate the value of f(x) in groups after the individual bit string decoded which can be used to reflect the fitness.
(5) To role the selection, crossover, mutation and other operators in the groups and then form the next generation groups.
(6) To determine whether the performance of the groups meets some demands or have arrived the number of iterations. If not, it will get back to (5) recently or after the modification for genetic strategies.

It is different to deal with the data directly in the multi-spectral data processing with GA for the characteristics of a large quantity and multi-spectral in the remote sensing images. Generally it is to build a model what can carry the evolution by GA. In the classification field for remote sensing data, PAL [4] and other experts has carry out to determine the combination location of hyper-plane by GA so that it can be applied in remote sensing image classification. The using process of

* Corresponding Author: Lv Yan in GIS, Tel:(010)88217807, E-mail:clsdau@126.com .

genetic and hyper-plane algorithm for classification is as follows.

First of all, give out the pattern description and matching for the training points in the training points set; then achieve the excellent search-performance by genetic algorithm and make the comparison and selection by the evolution to the variety of models and at last get the best classification program; Finally, it will achieve the purpose of pattern classification by extending the processing to the whole image.

In the hyper-plane classification model that is in the support of the optimization genetic algorithm, a genetic algorithm search-space appears after the collection for the binary strings which is from the special coding to the parameters of the hyper-plane equation. This is due to a binary string with specific median can represent all hyper-plane equation of the specific space.

The genetic algorithm can search the best ending in the search-space through the genetic manipulation operation such as choice, mutation and intercross, by the decoding process in which it can restore the binary string back to the hyper-plane parameters for the purpose of calculating the fitness of every chromosome. This is an adaptively iterative process in essence. In the training process, the difference between the training points and the mistaken points in the classification will be taken as the fitness of a series of hyper-plane. The best chromosome corresponds to the optimal classification program. That is the collection of hyper-plane with the best fitness. The way to decompose the mixed-pixel is to establish the hybrid simulation model about spectra.

That is to express the reflectance of the pixel as the function of the spectral characteristics and the comparison in types. In some cases, it is expressed as the spectral characteristics of the unit component and the function of other parameters of the ground. CharlesIchoku (1996) summarizes the mixed-pixel model as five kinds types: linear models, probability models, the geometric optics model, the random geometric model and the fuzzy analysis model, among which the linear model is widely used because of the advantages of the simple model and clear physical significance. It takes the linear model as a mixed model of pixels for simplifying the simulation process.

The basis of the linear model:
(1) Every standard feature's spectrum is a spectrum with pure pixel correspondingly.
(2) The spectrum of the mixed-pixel is the linear superposition to the spectrums of all kinds of standard features.

If the mixed pixel P contains the n type information about the ground targets, of which the multi-spectral vector is I, the vector for the type ratio of the pixel vector is R and the reflective properties of every ground targets in different bands is C, there will be the relationship of I, C, R as follows:

$$I = C \bullet R$$

Remarks:

$$I = (i_1, i_2 \ldots, i_n)'$$

$$R = (r_1, r_2 \ldots, r_n)'$$

$$C = \begin{pmatrix} c_{11} & \cdots & c_{1n} \\ \vdots & \ddots & \vdots \\ c_{m1} & \cdots & c_{mn} \end{pmatrix} \quad C_{ij} \text{ (i=1,…,m;j=1,…,n)}$$

m : the number of the bands in multi-spectral images;
n: the total number of categories

Though it is possible to measure the above parameters by instruments on the ground, because of the atmosphere, the undulating ground and other factors, the value of the reflective properties in different bands are different from what is gained from satellite photography.

The average value which is from the categories j belong to the band i of multi-spectral images in the training samples instead of cij can avoid these problems.

Premise:
If there are (1) Supposed multi-spectral image are divided into n types: T_1, $T_2 \cdots$Tn and (2) $I_j=\max\{r,i,i=1,2,\cdots,n\}$, it will make p into the categories j(that is $P \in Tj$). The classification criteria can classify the mixed-pixel into the ground target which has the largest type proportion.

2. EXPERIMENT

It chooses Ningbo in Zhejiang Province as the experimental zones. In recent years, with the spreading of reform and opening-up, the development of the city has a good development and the urban land use are rapidly expanding. As the three rivers such as Yong Jiang, Yuyao and Fenghua tntersect in Ningbo district, it can be easy to observe the erosion effect after the image processing. The model calculation and spectral analysis from the every spectrum in TM images shows that this kind of images are with the most abundant information and the synthetic image from TM2, TM4, TM7 has the realistic structure information which is favor of the city-interpretation. (such as ZHOU Cheng-hu, 2001) However, due to getting the seven bands belong to TM in several periods are difficult, it selects the kinds of data such as TM2, TM5, TM7 tri-bands in1987, 1990 and 1996 after comprehensive consideration.

Every pixel is on behalf of the area 30 × 30m in the ground, and the size of each image is 350 × 312 pixels, Figure 2 gives the TM image of Ningbo in 1990. In order to reduce the burden of programming, it makes the evaluation to the classification of the remote sensing image in support of ERDAS8.4 platform. The results is as the figure 3.

It can achieve the corrosion transformation and calculation of indicators in the environment programming of VC ++6.0 with the help of the above calculation (such as Bin He, 2001). The category image after erosion transform is as the figure 4.

The significance of that is: the larger of the city membership, the more white of the gray-value, and on the contrary the gray value will be larger.

the bigger cities in membership, the gray-value the more white, more gray value contrast of their black.

The results of the three indicators calculated in the VC after the transformation shown in table I.

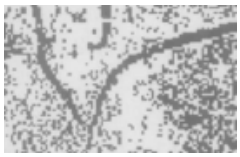

Figure I TM image of Ningbo in 1990

Figure II The classification image after Processing

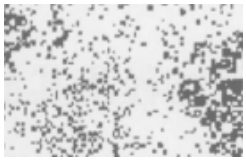

Figure III The image after erosion

Generation of training samples:

First of all, choose a standard spectral features, the choice of standards is very difficult in practice [10]. The method of making selection criteria of spectrum in this article is:

Based on the terrain and the color information in pseudo-color images it will to choose the collections of pure samples for every kind of points and to purify these collections, then to select a pure sample. The gray value of every band belong to the pure sample points constitute the standard vector of spectral features.

The concrete handling of purification is as follows: make the average gray value of all sample points in the collection of the selected region as a standard vector of spectral features. It chooses the pixels which have the largest number of the same spectral value as the standard sample. It forms the training samples in accordance with the linear combination of the standard spectrum after the choice. The model is as follows:

$$P = RC = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0.1 & 0.3 & 0 & 0.6 & 0 \\ 0.4 & 0.1 & 0.2 & 0.1 & 0.1 & 0.1 \\ 0.2 & 0.05 & 0 & 0.4 & 0.05 & 0.3 \\ 0.3 & 0.2 & 0 & 0.1 & 0.2 & 0.2 \end{pmatrix} \begin{pmatrix} C_1 \\ C_2 \\ C_3 \\ C_4 \\ C_5 \\ C_6 \end{pmatrix}$$

Remarks:

P Training sample matrix;

R The coefficient matrix of linear combination , that is the target matrix;

C the matrix of the standard spectrum

Because it is to fit the hyper-plane, the needed rank of matrix R is enough if it is equal to the dimension of the space. Of course, the more sample points the better for the result in the case of non-redundant.

Algorithm:

(1) To determine training parameters for the genetic algorithm including: crossover probability, the probability of reproduction, mutation probability, population size, number of iterations and so on.

(2) To make the linear combination of every kind of standard features as the training samples, every individuals on behalf of the corresponding linear combination coefficients.

(3) After the training, to input the gray value belong to every pixel based on the spectral angle, and the export is the decomposition results.

(4) To classify can give out the classification results according to the percentage from the decomposition.

Table 13 The results of three indicators

year	U-R	F-C	E-O
1987	0.6013	0.2949	0.7697
1990	0.6017	0.3316	0.9228
1996	0.7084	0.3166	0.8852

3. ANALYSIS AND CONCLUSIONS

This paper studies found that the GA evolution model is not only could be recived with the actual situation similar to the results, but also have the following characteristics:

(1) The evolution model is simple and also has a small amount to calculate. Because of a hyper-plane will be fitted, the necessary sample points is little, the evolution model is simple and fast training. After the model of training is over, the process of solving is known function. Given the input to seeking output of the process, the calculation amount is a very small amount. In the same computation environment, use the full-constrained least squares mixed-pixel decomposition algorithm to deal with needs 19020s, and in this paper, GA-based mixed-pixel decomposition algorithm is similar to the results (correlation coefficient is 0.99) requires only 137s, the speed increased nearly 100 times.

(2) Good adaptability. Different from the least-squares mixed-pixel decomposition algorithm, as long as the standard features spectrum is not changed, the image size change has no impact with the good evolution of GA model. Solely basis on remote sensing spectral knowledge to extract the city's complexity message is very rough, accurate extraction of urban information needs to combine other data, but the interest of this article is the calculation of three indicators and image transformation in the effects of corrosion and the effectiveness of the algorithm. As can be seen from the calculations, in 1987 and in 1990, Ningbo, the speed of the urbanization process is slow, but in 1996, larger changes have occured compare to the previous two years, and the city has developed very rapidly, U-R indicator is larger than the previous two years. F-C and E-O indicators show the large middle and little two sides, one possible explanation is Ningbo city’s urban land use structure patterns experienced the process from the simple to the complex and from complex to the simple. For example, in 1987, Ningbo City’s land usage is

relatively simple, therefore the value of F-C and E-O is relatively small. Urban land usage tends to fragmentation, so the value of F-C and E-O tend to increase before. In 1996, as a result of urban planning, urban land use combined with lot of piecemeal type, the value of F-C and E-O have reduced. Because of GA algorithm has excellent properties, it is widely used in the classification of two types model, but in the classification of mixed features is rare.

In this paper, basing on a detailed analysis of the mixed-pixel decomposition algorithm and the GA evolution principle, the mixed-pixel decomposition algorithm based on genetic algorithm is advanced, which can break down the results of fitting a hyper-plane, thereby has the result of decomposition and classification. The experiments show that GA-based mixed-pixel decomposition algorithm can receive a very similar results compared to the constrained least squares with (FCLS) mixed-pixel decomposition algorithm , but its computational complexity is greatly reduced ,the computing speed is improved and adaptability is enhanced. The author believes that in the field of mixed-pixel decomposition there are still has following questions worthy to continue to study: Mixed-pixel's classification accuracy has a great influence by its number. The further focus of the mixed-pixel classification's study is improve the classification precision based on increase the number of classification; In addition, the linear model is relatively simple, but in practical applications there are some restrictions on how to choose the actual suit situation model and improve the accuracy of decomposition is also the direction of further study.

REFERENCES

Urban Green Space Management Information, http://www.ipf.tuwien.ac.at/publications/EI_P_180_Ries.pdf.

The European Green Paper on Urban Mobility. http://repository.tudelft.nl/file/1042182/380609

Developing an Agricultural Remote Sensing Program at the University of Illinois. http://www.ilcfar.org/research/remote.pdf

Remote Sensing in Urban Areas, http://www.forum-urban-futures.net/files/KIM_gis4u_laos_23sep2008.pdf