Artificial DNA Computing-Based Spectral Encoding and Matching Algorithm for Hyperspectral Remote Sensing Data

Hongzan Jiao, Yanfei Zhong, *Member, IEEE*, and Liangpei Zhang, *Senior Member, IEEE*

***Abstract*—In this paper, a spectral encoding and matching algorithm inspired by biological deoxyribonucleic acid (DNA) computing is proposed to perform the task of spectral signature classification for hyperspectral remote sensing data. As a novel branch of computational intelligence, DNA computing has the strong computing and matching capability to discriminate the tiny differences in DNA strands by DNA encoding and matching in the molecule layer. Similar to DNA discrimination, a hyperspectral remote sensing data matching approach is used to recognize the land cover material from a spectral library or image, according to the rich spectral information. However, it is difficult to apply DNA computing to hyperspectral remote sensing data processing because traditional DNA computing often relies on biochemical reactions of DNA molecules and may result in incorrect or undesirable computations. To utilize the advantages and avoid the problems of biological DNA computing, an artificial DNA computing approach is proposed for spectral encoding and matching for hyperspectral remote sensing data. A DNA computing-based spectral matching approach is used to first transform spectral signatures into DNA codewords by capturing the key spectral features with a spectral feature encoding operation. After DNA encoding, the typical DNA database for interesting classes is constructed and saved by DNA evolutionary operating mechanisms such as crossover, mutation, and structured mutation. During the course of spectral matching, each pixel of the hyperspectral image, or each signature measured in the field, is input to the constructed DNA database. By computing the distance between an unclassified spectrum and the typical DNA codewords from the database, the class property of each pixel is set as the minimum distance class. Experiments using different hyperspectral data sets were performed to evaluate the performance of the proposed artificial DNA computing-based spectral matching algorithm by comparing it with other traditional hyperspectral classifiers, including spectral matching classifiers (binary coding, spectral angle mapper and spectral derivative feature coding (SDFC) matching methods) and a novel statistical method of machine learning termed support vector machine (SVM). Experimental results demonstrate that the proposed algorithm is distinctly superior to the three traditional hyperspectral data classification algorithms. It presents excellent processing efficiency, compared to SVM, with high-dimensional data captured by the Hyperspectral Digital Imagery Collection Experiment sensor, and hence provides an effective option for spectral matching classification of hyperspectral remote sensing data.**

***Index Terms*—Classification, deoxyribonucleic acid (DNA) computation, hyperspectral remote sensing, spectral matching.**

I. Introduction

As an integration of "imaging" and "spectroscopy," hyperspectral remote sensing provides researchers with the opportunity to perform detailed identification and mapping of constituents of the Earth's surface through the analysis of spectral features [1]. As a traditional hyperspectral analysis approach, spectral matching techniques are often used for hyperspectral signature discrimination and data classification, which concentrate on recognizing absorption band shapes in each individual spectrum [2]. Traditional spectral matching techniques, for example, spectral angle mapper (SAM) [3], often measure the similarity of the whole spectral curves of hyperspectral image spectra or field-measured spectra and spectra from spectral libraries, the unknown signature being automatically assigned to the most similar class, without requiring labeled samples in the image [4]. However, due to external factors, such as atmospheric effects, environmental radiation, shading, etc., the whole of the spectrum varies greatly compared to the laboratory measurements [5]. Therefore, the full spectral curves-based spectral matching methods need precise spectral samples and may produce significant errors. One of the major challenges in hyperspectral data exploitation is how to best utilize the spectral information provided by hyperspectral imagery to accomplish tasks such as detection, discrimination, classification, and identification, while discarding undesired information caused by unwanted interference such as noise [6].

A spectral coding-based matching technique is an available approach, which transforms spectral values into a new set of symbols in a specific manner so that a signature can be represented by the new symbols more effectively or efficiently [6]. Spectral binary coding (BC) [7], [8] is well known as a simple, effective hyperspectral analysis method with low computational cost, which encodes the data and endmember spectra into zeros and ones based on whether a band falls below or above the spectrum mean, respectively. However, the BC method may lose some important information of the spectral features [9], so that similar signatures of different classes cannot be discriminated, and it does not classify spectra with the same class, due

Manuscript received February 2, 2011; revised June 10, 2011, October 29, 2011, and February 11, 2012; accepted February 12, 2012. Date of publication April 4, 2012; date of current version September 21, 2012. This work was supported by National Basic Research Program of China (973 Program) under Grant 2009CB723905, the National Natural Science Foundation of China under Grants 40901213 and 40930532, A Foundation for the Author of National Excellent Doctoral Dissertation of P.R. China (FANEDD) under Grant 201052, Program for New Century Excellent Talents in University under Grant NECT-10-0624, the Fundamental Research Funds for the Central Universities under Grant 3103006, and academic award for excellent Ph.D. candidates funded by the Ministry of Education of China.

The authors are with the State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, Wuhan 430079, China (e-mail: zhongyanfei@lmars.whu.edu.cn).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TGRS.2012.2188856

to the heterogeneity of spectral signatures of the same class. To address these problems, some enhanced encoding approaches have been investigated for hyperspectral signature discrimination and classification. Spectral analysis manager (SPAM) [8] encodes an L-dimensional signature as a $(2L-2)$-dimensional binary codeword composed of the first L binary values used to encode the sign of the difference between a signature and its signature mean, and additional L-2 binary values to encode the sign of the difference in spectral values between a band and its adjacent band. Spectral feature-based BC (SFBC) [10], as a further extension of SPAM, uses additional $L-2$ binary values to encode a signature as a $(3L-4)$-dimensional binary codeword. The new added $L-2$ binary values are used to dictate whether the deviation of a spectral variation from the signature mean is greater than a prescribed threshold. Recently, a new signature encoding method, referred to as spectral derivative feature coding (SDFC) [11], is actually a generalization of both SFBC and SPAM. It is derived from texture features used in texture classification to dictate gradient changes between adjacent bands in characterizing spectral variations. Because SDFC is more effective than other coding methods at dictating the gradient changes in spectral variation occurring between three consecutive adjacent bands, its spectral matching performance is better than SPAM and SFBC. In addition, to improve the accuracy of BC, spectral feature probabilistic coding has also been proposed, using a criterion with memory to measure spectral similarity [12]. The traditional spectral coding-based approach, which encodes spectral signatures as binary or quaternary codewords, is conducted using the Hamming distance (HD) as the spectral similarity measure [11]. It is noteworthy that the codes for spectra are not changed during the process of matching or classification. Therefore, these methods need very precise spectral information and are still sensitive to noise.

In this paper, a new spectral matching strategy inspired by DNA computing is introduced to encode spectral curves as DNA codewords and to discriminate various spectral codewords of different classes. DNA computing is a new computational model that uses biomolecules as information storage materials and biological laboratory experiments as information processing operators [13]. This technique has been used successfully to solve the Hamiltonian path problem [14], [15]. The ability of DNA computers to perform calculations using specific biochemical reactions between different DNA strands by Watson-Crick complementary base-pairing provides a number of useful biological properties, such as massive parallelism and a huge memory capacity [16]–[18]. Recently, DNA computing has been used successfully to solve complex problems in many fields [19], such as the NP-complete problem [20], knapsack problem [21], traveling salesman problem [18], [22], weighted graph problems [23], and clustering [24]. DNA computing has also been successfully applied to signal processing fields [25], [26]. By careful designing of the mapping of signal values to DNA molecules, their annealing can be a powerful tool for the extraction of disparity information [25]. A DNA-based digital signal processing approach will consider new codeword designs to transform signal space to DNA space. A four-base (quaternary) representation provides high flexibility in choosing a codeword design scheme [26]. In addition, a DNA database of digital signal data is built to store these signals. The benefits of using DNA to store digital data, particularly signals, include, information compactness, economical database replication, efficient query mechanisms, and query time that is independent of database size [25]. DNA-based matching for digital signals is used as a search or classification mechanism by quantifying the similarity between signals [27], [28], and it is able to carry more plentiful genetic information than conventional evolution algorithms [29].

Similar to DNA-based signal matching, a hyperspectral spectral matching algorithm is used to recognize similar spectra by utilizing the rich spectral information. However, it is very difficult to utilize the advantages of DNA computing in hyperspectral spectral matching because traditional DNA computing has limitations due to its use of natural DNA and enzymes, which only recognize and act on certain nucleotide sequences [30], [31]. To avoid this problem, an artificial DNA computing model is proposed by simulating DNA computing on a conventional computer. In the artificial DNA computing model, to build the mapping relationship between DNA and hyperspectral remote sensing, an encoding method for hyperspectral data is designed. The typical spectral information of different classes is transferred from spectral space to DNA quaternary codeword space {T, C, A, G}. The four types of DNA base {T, C, A, G} stand for thymine, cytosine, adenine, and guanine, respectively. Based on the DNA spectral codeword, the artificial DNA-based spectral matching (ADSM) algorithm is proposed to obtain the optimal matching algorithm. In contrast to traditional spectral encoding and matching, such as BC and SDFC, with stationary spectral class curves, ADSM evolves the spectral DNA by its own biological properties during the process of matching and classification, such as the encoding process, selection, replication, hybridization, mutation, and structured mutation. It is important to note that the difficulty in hyperspectral remote sensing spectral matching is that spectral signals, particularly hyperspectral signals, may be characterized by spectral diversity and high dimensionality, and need large amounts of computational time. The evolution process of ADSM may avoid the errors caused by spectral variations of the same class and provide diversity tolerance for hyperspectral signature discrimination and data classification. This process is not present in other spectral matching algorithms. In addition, the three distance criteria [HD, absolute value distance (AVD), and protein value distance (PVD)] used to discriminate spectral DNA strands are introduced to obtain more integrated and precise differences between spectral DNA. Note that the PVD criterion is a new discriminating distance criterion for DNA strands, which is inspired by the biological many-to-one correspondences during the genetic information transference. To illustrate the effectiveness of the proposed design method, SAM, which is a traditional spectral matching method without an encoding process, and a novel generation learning system, SVM, are also introduced to compare results with those of the proposed spectral coding-based matching method. In the remote sensing community, SAM has been widely used as a spectral similarity measure for material identification [32]–[34], and SVM, based on recent advances in statistical learning theory, is becoming important for solving hyperspectral data classification, due to its self-adaptability and a rapid learning pace in feature space. Therefore, three data sets comprising hyperspectral spectral library data and two hyperspectral

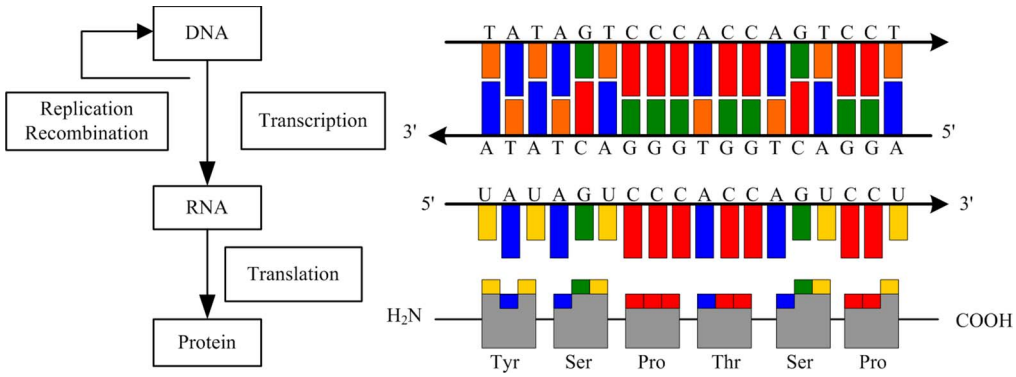


Fig. 1. Genetic rule of DNA.

images were used for experiments to carry out performance analysis of the ADSM, in comparison with BC, SDFC, SAM, and SVM.

The remainder of this paper is organized as follows. Section II provides the basic principles of DNA computing and builds the mapping relationship between DNA computing and spectral matching. Section III proposes the artificial DNA computing-based spectral encoding and matching algorithm. In Section IV, the spectral matching experiments employing a hyperspectral library and two hyperspectral images are described, and the experimental results and analyzes are provided. Section V discusses the main properties of DNA computing in theoretical and empirical terms. Finally, the conclusion and direction of future work are provided in Section VI.

II. Artificial DNA Computing

A. DNA

As we know, 99.99% of human DNA sequences are the same in every person; however, enough of deoxyribonucleic acid (DNA) is different to enable us distinguish one individual from another. The memory and expression ability of genetic information is decided by the DNA encoding and evolution mechanism based on special structures. DNA consists of four types of base: thymine (T), cytosine (C), adenine (A), and guanine (G). The DNA segments that carry this genetic information are called genes, but other DNA strands have structural purposes, or are involved in regulating the use of this genetic information. DNA strands hold the information to build and maintain an organism's cells and pass genetic traits to offspring. The genetic information will transfer to specify all proteins through transcription and translation operations, as shown in Fig. 1. Genetic information can be stored by DNA strands, and by using the DNA decoding method, the codons of DNA strands can be translated into amino acids and further into proteins. The codons are a series of three-nucleotide sequences, which specify the many-to-one correspondences during the presentation of genetic information.

DNA strands are very similar in various animals, but there are some differences at key parts, such as with chimpanzees and humans. The significantly different parts in the DNA strands determine the differences between various animals. If we can locate the position where the gene exists and determine the criterion DNA strands at this position, we can determine the category of the DNA strands.

B. Biological Mathematics of DNA Computing

A single strand of DNA can be likened to a string consisting of a combination of four different symbols: T, C, A, G. Mathematically, this means we have a four-letter alphabet $Z = \{T, C, A, G\}$ to encode information, which is more than enough, considering that an electronic computer needs only two digits, 0 and 1, for the same purpose. DNA computing research was inspired by the similarity between the way DNA works, and the operation of a theoretical device known as a Turing machine. This is naturally related to the way biological machinery works; the Turing machine processes information and stores it as a sequence, or list of symbols. Therefore, DNA computing can deal with the problems that Turing machines can solve [16]. Some of the simple operations that can be performed on DNA strands are accomplished by a number of enzymes that execute a few basic tasks. For example, restriction endonucleases will cut a DNA strand that contains the restriction site. DNA ligases will bond together, or "ligate" the end of a DNA strand to another strand. The DNA polymerases perform several functions, including replication of DNA. Exonucleases can be used to selectively destroy DNA molecules. The main biological operations of DNA computing are: synthesizing, mixing, melting, annealing, amplifying, separating, extracting, cutting, ligating, substituting, marking, destroying, and deleting. A flowchart of natural DNA computing [31] is shown in Fig. 2(a).

C. Artificial DNA Computing and Spectral Matching

DNA computing has several limitations due to the use of natural DNA and enzymes, which only recognize and act on certain nucleotide sequences, so the application of DNA computing is limited [17]. Artificial DNA computing was therefore developed to simulate natural DNA computing on a traditional computer, which can solve problems using the mechanisms of DNA encoding and optimal operating mechanisms. Artificial DNA computing encodes the information of the problem to be solved as DNA strands, recombines DNA strands to satisfy the problem constraints, and ultimately reaches the answer by an iterative optimization process. Artificial DNA computing borrows from the intelligence of biological evolution and the principle of natural DNA computing. In recent years, more biological intelligence knowledge has been introduced to enhance the evolving ability of artificial DNA computing. For example, DNA computing has been combined with the artificial immune

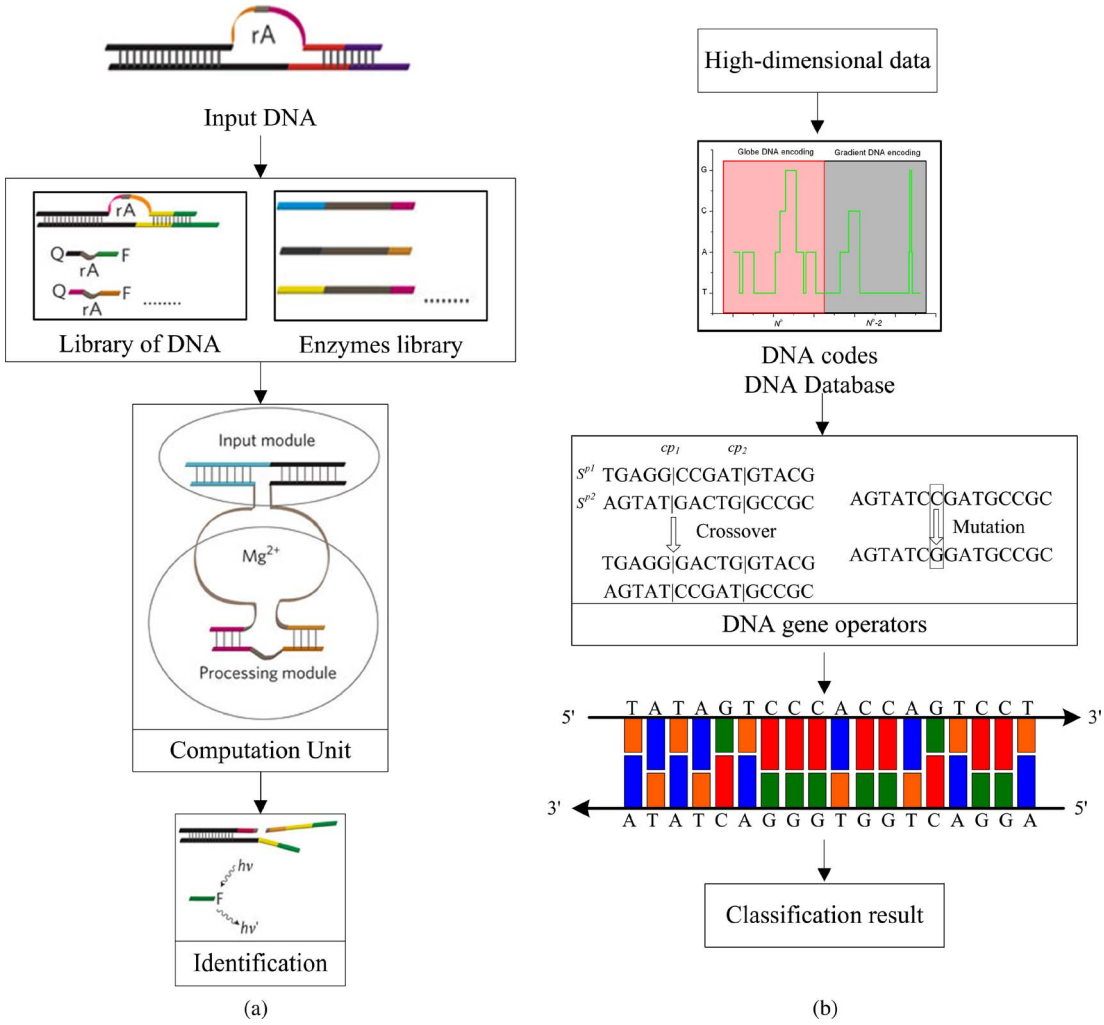


Fig. 2. Development from biological DNA to artificial DNA computing. (a) DNA computing. (b) Artificial DNA computing.

system [18], and a genetic algorithm [19]. Artificial DNA computing that simulates natural DNA computing is shown in Fig. 2(b).

Following the development of artificial DNA computing, the combination of spectral matching and DNA computing became a reality. The spectral signatures can be encoded into DNA codewords using relevant spectral feature extracting and encoding methods, and can construct the DNA database of ground objects and then accomplish DNA classification by employing a typical DNA database. By acquiring the classification result of DNA strands, the corresponding classification result of spectral matching in remote sensing can be realized in the DNA computing space. In this way, the steps in spectral matching are displaced by DNA computing procedures, and the spectral matching will be transferred from the spectral feature space to DNA computing space.

III. Artificial DNA Computing for Spectral Matching

By combining spectral matching with artificial DNA computing, the model is built up based on the DNA encoding, optimization mechanism and DNA database. Therefore, a new spectral matching mapping method is proposed, using artificial DNA computing for hyperspectral data.

In the proposed method, spectra are transferred into "DNA strands" for various ground objects, and the features of the ground objects are carried by these spectral DNA strands. With the features extracted from the whole spectral signature, or the absorption and reflectance bands, spectral matching can be performed between two signatures. However, due to the diversity of spectra, the procedure of spectral matching often needs to be optimized before performing spectral discrimination and classification.

In the procedure of DNA computing for spectral matching, spectral feature extraction is replaced by a DNA encoding procedure, and DNA recombination is regarded as a matching optimization operation. By classifying spectral DNA strands into different classes, the corresponding spectra are labeled as the same classes. Fig. 3 summarizes the relationship between DNA computing and spectral matching.

A. Signature Normalization

In order to compare a signature to a library signature, it is necessary for both signatures to have the same spectral range and resolution. We assume that atmospheric effects are accounted for by converting image radiance to surface reflectance, using appropriate atmospheric correction methods. After atmospheric correction is performed, we still must account for the

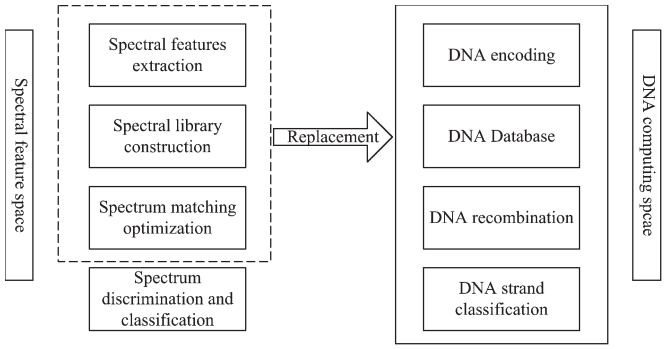


Fig. 3. Relationship between spectral matching and DNA computing.

scaling effects caused by illumination conditions. To mitigate illumination effects, we scale each cluster and library signature by its Euclidean norm. This scales signatures to unit length, while preserving spectral angles, but discards geometric albedo, which can be recovered in postprocessing by selecting from multiple same-spectrum matches.

After signature normalization, we can give the mathematical definition of the hyperspectral signature and imagery separately. Hyperspectral remote sensing data $x = (x^1, x^2, \ldots, x^{N_b})^T$ through N_b bands are observed and mapped to a finite rectangular lattice $W = ((i,j)) : 1 \le i \le N_i, 1 \le j \le N_j)$. The character T denotes the transpose of a matrix. The set $x^b = (x^b_{11}, \ldots, x^b_{N_i N_j} : b = 1, \ldots, N_b)^T$ denotes the data taken at the bth wavelength, where $x^b_{ij} \in (0, \ldots, Gl-1)$ and Gl is the number of observable gray levels. At the ijth pixel, a N_b-dimensional feature vector $x_{ij} = (x^1_{ij}, \ldots, x^{N_b}_{ij})^T$ is observed. The entire set of image data can be denoted as $x = (x_{ij} : 1 \le i \le N_i, 1 \le j \le N_j)$. Hyperspectral signatures can be represented as $y = (y^1, y^2, \ldots, y^{N_b})^T$, and a signature $y_{ij} = (y^1_{ij}, \ldots, y^{N_b}_{ij})^T,\ 1 \le i \le C,\ 1 \le j \le N$ is observed, where C is the number of signature classes, and N is the number of signatures in one class. The hyperspectral library is $l = (l^1, l^2, \ldots, l^{N_b})^T$, and a signature in the spectral library is represented by $l_i = (l^1_i, \ldots, l^{N_b}_i)^T, 1 \le i \le Cl$, where Cl is the number of classes in the spectral library. A classified image is denoted by $\omega = (\omega_{ij} : 1 \le i \le N_i, 1 \le j \le N_j)$, each pixel of which is to be assigned to one of C classes. That is $\omega_{ij} \in (1, 2, \ldots, C)$, where M is the number of classes and is assumed to be known.

B. ADSM Model Initialization

DNA population parameter setting: Before employing the ADSM model for hyperspectral data discrimination and classification, some parameters of the DNA population must be set, including the number of individuals in the population, crossover probability, mutation probability, structured mutation (deletion, insertion, and inversion) probability, the loop termination condition, adaptive global coefficient, and adaptive gradient coefficient.

C. DNA Encoding for Spectra

It is the DNA encoding method that is the primary problem with DNA computing for spectral matching. Based on the characteristics of hyperspectral data with high dimensions, the DNA code needs to satisfy three requirements:

- DNA encoding should adopt the form of a four-value codeword by a four-letter alphabet {T, C, A, G}.
- DNA encoding should capture the most relevant spectral variations in the image.
- DNA encoding should possess the capability of diversity tolerance and deal with the high dimensionality of the data.

Considering these requirements for pixel and library spectra encoding, a new encoding method named "spectral DNA encoding for hyperspectral data" is proposed. Spectral DNA encoding is composed of two parts: DNA encoding for spectral global features, and DNA encoding for spectral gradient features. The details are as follows:

1) DNA Encoding for Spectral Global Features: This part focuses on spectral global features. Using certain percentages of the reflected signatures as thresholds, the overall features of spectral absorption can be captured by DNA encoding. Due to inhomogeneous illumination effects, the value of the spectrum will be moved up or down. As a result, the thresholds need to be adaptively set by each spectrum.

For the ith band, we encode the ith band of spectral signature y^i as the global DNA codewords DNA^i_{global} by

$$DNA^i_{\text{global}} = \begin{cases} T & y^i \in [x_{\min}, T_l) \\ C & y^i \in [T_l, T_0) \\ A & y^i \in [T_0, T_r) \\ G & y^i \in [T_t, x_{\max}]. \end{cases} \tag{1}$$

Thresholds, T_0, T_l, and T_r, can be adaptively set as spectral values of each spectrum

$$T_0 = \rho \cdot \left(\sum_{i=0}^{N_b} y^i \right) / N_b \tag{2}$$

$$T_l = \frac{\left(\sum\limits_{i=0}^{N_b} \max(y^i) \right)}{k} \tag{3}$$

$$T_r = \frac{\left(\sum\limits_{i=0}^{N_b} \min(y^i) \right)}{p} \tag{4}$$

where $\max(y^i) = \begin{cases} y^i & \text{if } y^i \ge T_0 \\ 0 & \text{if } y^i < T_0 \end{cases}$, $\min(y^i) = \begin{cases} 0 & \text{if } y^i \ge T_0 \\ y^i & \text{if } y^i < T_0 \end{cases}$ and k is the number of bands whose values are greater than T_0, p is the number of bands whose value is smaller than T_0; ρ is the adaptive global coefficient of DNA encoding with domain value (0.5,1).

2) DNA Encoding for Spectral Gradient Features: This encodes a signature as a quaternary codeword, while keeping track of gradient changes in the spectral variation as spectral derivatives in three adjacent bands. DNA encoding for spectral derivative features represents interband spectral correlation. Assume that $y = (y_1, y_2, \ldots y_{N_b})^T$ is a hyperspectral signature; also, let Δ be a desired spectral value tolerance. There are four types of successive gradient changes in spectral values that can be described as follows [11]:

Type 1 $if\ |y^i - y^{i-1}| \le \Delta\ and\ |y^{i+1} - y^i| \le \Delta$
Type 2 $if\ \left(|y^i - y^{i-1}| \le \Delta\ and\ |y^{i+1} - y^i| > \Delta\right)$

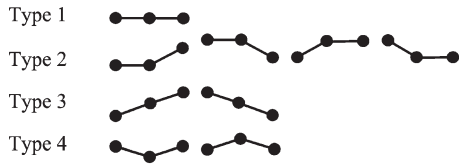


Fig. 4. Graphical representation of spectral curve characteristics.

$$\begin{aligned} &or\ \left(|y^i - y^{i-1}| > \Delta\ and\ |y^{i+1} - y^i| \le \Delta\right) \\ \text{Type 3}\ \ &if\ (y^i - y^{i-1} < -\Delta\ and\ y^{i+1} - y^i < -\Delta) \\ &or\ (y^i - y^{i-1} > \Delta\ and\ y^{i+1} - y^i > \Delta) \\ \text{Type 4}\ \ &if\ (y^i - y^{i-1} < -\Delta\ and\ y^{i+1} - y^i > \Delta) \\ &or\ (y^i - y^{i-1} > \Delta\ and\ y^{i+1} - y^i < -\Delta). \end{aligned} \quad (5)$$

Additionally, the Δ used in (6) can be set as follows:

$$\Delta = \theta \cdot \left(\frac{1}{(N_b - 1)}\right) \sum_{l=2}^{N_b} |y^i - y^{i-1}| \quad (6)$$

where θ is the adaptive gradient coefficient of DNA encoding with domain value (1) and (2).

The graphic representations of these four types of gradient change in spectral values are shown in Fig. 4 to better illustrate why these four types of gradient change can be more effective in characterizing spectral variability among three consecutive adjacent bands [11].

The DNA encoding technique is used to capture spectral texture feature changes in three consecutive adjacent bands, according to the four types of gradient changes in the spectral variation described by (7). For $1 < i < N_b$, we define spectral gradient DNA codewords DNA^i_{gradient} as:

$$DNA^i_{\text{gradient}} = \begin{cases} T & if\ y^i\ is\ type\ 1 \\ C & if\ y^i\ is\ type\ 2 \\ A & if\ y^i\ is\ type\ 3 \\ G & if\ y^i\ is\ type\ 4. \end{cases} \quad (7)$$

Two signatures, spectrum 1 and spectrum 2, which belong to the vegetation class, but are different fine categories, are represented in Fig. 5. The signatures are similar in their whole tendencies, but have small differences in the local parts, which are key points for spectral classification. The global DNA encoding method can extract the similarity of signatures that belong to the main class, while the gradient feature encoding method can find differences at the micro level. Correspondingly, the absorption codewords for the signatures are the same, such as AGGGG, but the gradient codewords are different, for example, AAT and AGG. Hence, when we need to discriminate fine categories, for example, vegetation that could be divided into tree, grass, and corn, the absorption and gradient feature encoding can be combined for that purpose.

Through DNA encoding for spectral global and gradient features, the original spectral data are transformed into spectral DNA codewords in the DNA computing space. The procedure of transformation is shown in Fig. 6. The process mentioned above can be represented in (8), and the S^{DNA} presents the DNA encoding for the spectral signature y

$$S^{DNA} = Encoding(y) \quad (8)$$

where $y = (y^1, y^2, \ldots y^{N_b})^T$ and $S^{DNA} = \{(S^{DNA\_\text{global}}, S^{DNA\_\text{gradient}})\}^T$.

It is important to note that the continuous variable of the original spectral signature is discretized into the DNA feature space. Each encoding operation can be regarded as feature extraction of a typical spectral characteristic, or should be considered as primitively classifying the spectral variable into a quaternary rough set. Discretization is a process that transforms continuous attributes into a finite number of intervals, where each interval is associated with a discrete numerical value. Discretized intervals can then be treated in a similar way to nominal values during induction and deduction. There are many advantages to using discrete values over continuous ones. The most important is that data can be reduced and simplified through discretization [36]. The discretization procedure is the foundation of artificial DNA computing.

D. Optimization Mechanism of Artificial DNA Computing

The aim of the optimization procedure of artificial DNA computing is to obtain the typical spectral DNA database, which is employed as the optimized model for spectral matching. To obtain the optimal spectral DNA database, the training samples (S) are randomly divided into two independent parts, according to different categories. One part, named $S^I \in S$, is used to initialize the spectral DNA population, and the others $(S^V \in S)$ are applied to calculate the fitness of individuals in the spectral DNA population. In this paper, some optimal strategies, such as mutation, crossover, structured mutation, and elite selection, are introduced into artificial DNA computing. The details are described as follows.

1) Initialization and Distance Discrimination Criteria: By encoding the subset S^I of the training samples, the set of training DNA strands P^{DNA} can be obtained. The initial DNA individuals are constructed by randomly selecting DNA strands according to different categories. The initial population P^I is composed of a fixed number of individuals.

Each individual of the DNA population (named as P_i^{DNA}, $i = 1 \ldots popsize$) can be regarded as a spectral DNA database (DB^{DNA}), which can be used to automatically classify the spectral signatures by the distance discrimination criteria of DNA computing, namely: HD, AVD, and PVD.

The three distance discrimination criteria are used to measure the discrepancy between the DNA strands of training samples, or pixels to be classified with the strands in the spectral DNA database, described as follows. Assume that $S^l = (s^{l\_DNA\_\text{global}}, s^{l\_DNA\_\text{gradient}}) = (s_1^l, s_2^l, \ldots, s_{2N_b-2}^l)^T$ is a spectral DNA strand belonging to the class l of spectral DNA database, and $S^p = (s^{l\_DNA\_\text{global}}, s^{l\_DNA\_\text{gradient}}) = (s_1^p, s_2^p, \ldots, s_{2N_b-2}^p)^T$ is a spectral DNA strand of a pixel to be classified.

- HD

$$HD(S^l, S^p) = \left(\frac{1}{N_b}\right) \sum_{i=1}^{N_b} \left| s_i^{l\_DNA\_\text{global}} - s_i^{p\_DNA\_\text{global}} \right| + \left(\frac{1}{(N_b - 2)}\right) \sum_{i=2}^{N_b - 1} \left| s_i^{l\_DNA\_\text{gradient}} - s_i^{p\_DNA\_\text{gradient}} \right| \quad (9)$$

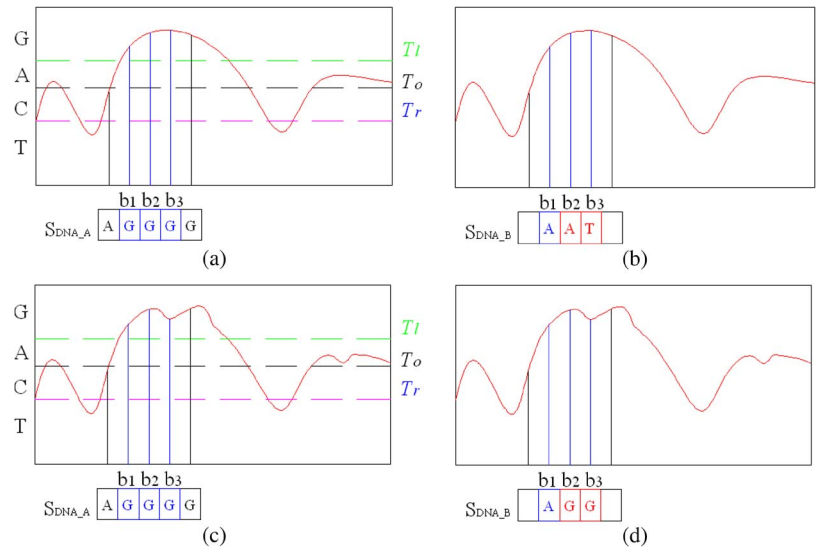


Fig. 5. Comparison of DNA encoding for different spectrum signatures. (a) Global DNA encoding for Spectrum_1. (b) Gradient DNA encoding for Spectrum_1. (c) Global DNA encoding for Spectrum_2. (d) Gradient DNA encoding for Spectrum_2.

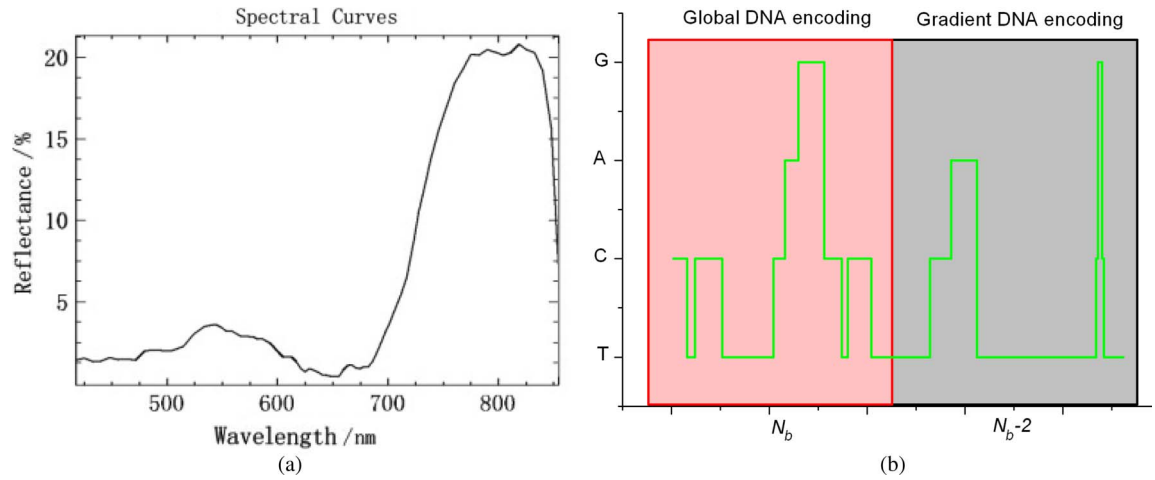


Fig. 6. DNA encoding for spectral data. (a) Original spectral data. (b) DNA codewords.

where $\{T, C, A, G\}$ should be represented by {00, 01, 10, 11}, and $|S^l - S^p|$ is the HD. For example, if $S^{l}\_^{DNA} = T \to 00$ and $S^{p}\_^{DNA} = G \to 11$, then $|S^l - S^p| = 2$.

- AVD

$$AVD(S^l,S^p)=\left(\frac{1}{N_b}\right)\sum_{i=1}^{N_b} f\left(s_i^{l\_DNA\_\text{global}},s_i^{l\_DNA\_\text{global}}\right) + \left(\frac{1}{(N_b-2)}\right)\sum_{i=2}^{N_b-1} f\left(s_i^{l\_DNA\_\text{gradient}},s_i^{l\_DNA\_\text{gradient}}\right) \quad (10)$$

where $S^l, S^p \in \{T, C, A, G\}$, and $f(S_i^l, S_i^p) = \begin{cases} 0, & if\ S_i^l = S_i^p \\ 1, & if\ S_i^l \neq S_i^p. \end{cases}$

- PVD

The PVD criterion is inspired by the many-to-one correspondences during the genetic information transference. We can translate the DNA codons into amino acids based on the corresponding rule between the genetic codons and the amino acids, as shown in Table I. Different DNA codons can be translated into the same amino acids

$$\text{Amino}(S, P) = \text{famino}(s^{DNA}, p^{DNA}) \quad (11)$$

where

$$\text{famino}(s^l, s^p) = 1/(N_b \bmod 3) \sum_{i=1}^{N_b \bmod 3} \times f\left(s_i^{l\_\text{amino\_global}}, s_i^{p\_\text{amino\_global}}\right) + 1/\left((N_b-2) \bmod 3\right) \sum_{i=1}^{(N_b-2) \bmod 3} \times f\left(s_i^{l\_\text{amino\_gradient}}, s_i^{p\_\text{amino\_gradient}}\right)$$

and $S^l and\ S^p \in \{T, C, A, G\}$,

$$f\left(s_i^{l\_\text{amino}}, s_i^{p\_\text{amino}}\right) = \begin{cases} 0, & if\ s_i^{l\_\text{amino}} = s_i^{p\_\text{amino}} \\ 1, & if\ s_i^{l\_\text{amino}} \neq s_i^{p\_\text{amino}}. \end{cases}$$

TABLE I
TRANSLATION FROM DNA CODONS INTO AMINO ACIDS

1st base	2nd base				3rd base
	T	C	A	G	
T	Phe(0)	Ser(2)	Tyr(3)	Cys(4)	T
	Phe(0)	Ser(2)	Tyr(3)	Cys(4)	C
	Leu(1)	Ser(2)	Stop(9)	Stop(9)	A
	Leu(1)	Ser(2)	Stop(9)	Try(9)	G
C	Leu(1)	Pro(5)	His(6)	Arg(8)	T
	Leu(1)	Pro(5)	His(6)	Arg(8)	C
	Leu(1)	Pro(5)	Gln(7)	Arg(8)	A
	Leu(1)	Pro(5)	Gln(7)	Arg(8)	G
A	Ile(11)	Thr(12)	Asn(13)	Ser(2)	T
	Ile(11)	Thr(12)	Asn(13)	Ser(2)	C
	Met(10)	Thr(12)	Lys(14)	Arg(8)	A
	Met(10)	Thr(12)	Lys(14)	Arg(8)	G
G	Val(15)	Ala(16)	Asp(17)	Gly(19)	T
	Val(15)	Ala(16)	Asp(17)	Gly(19)	C
	Val(15)	Ala(16)	Glu(18)	Gly(19)	A
	Val(15)	Ala(16)	Glu(18)	Gly(19)	G

The rule of DNA codons translating into amino acids can be represented by the function named "$fcodon$."

$$s_i^{\text{amino}} = fcodon\left(s_{3i-1}^{DNA}, s_{3i}^{DNA}, s_{3i+1}^{DNA}\right). \tag{12}$$

Given two pieces of DNA codewords, such as $s_1 = \{AGCACT\}$ and $s_2 = \{AGCACG\}$, the three distances can be separately calculated. The results are as follows: $HD(s_1, s_2) = 2$, $AVD(s_1, s_2) = 1$ and $PVD(s_1, s_2) = 0$. This illustrates that PVD has the best ability to tolerate the spectral diversity, HD has the best ability to discriminate the detail of signatures, and AVD considers both spectral diversity tolerance and signature discrimination ability.

2) Evaluating the Fitness of the Individuals: Considering the discretization of the DNA encoding operation, the stability features of each spectral class can be extracted for the matching classification, and the generalization and simplification ability of the classification model is taken into account due to the quaternary rough set. The classification is conducted based on the minimum spectral DNA discrimination distance criteria.

Due to the diversity of spectra, the DNA strands also present different patterns in the same category. The DNA individuals need to be optimized with the consideration of the consistency and accuracy between the classification results and the reference data set S^V. Consequently, the kappa coefficient, which is calculated on the subset S^V of the training sample set, can be chosen as the fitness function of the spectral DNA individual in the spectral DNA population

$$Fitness(P_i) = kappa(P_i). \tag{13}$$

The kappa coefficient can be defined in terms of the confusion matrix M, as follows:

$$kappa(P_i) = \frac{N\sum_{k=1}^{r} m_{kk} - \sum_{k=1}^{r}(m_{k+^*}m_{+k})}{N^2 - \sum_{k=1}^{r}(m_{k+^*}m_{+k})} \tag{14}$$

where r is the number of rows in the matrix, m_{kk} is the number of observations in row k and column k, m_{k+}, m_{+k} are the marginal totals for row k and column k, respectively, and N is the total number of observations.

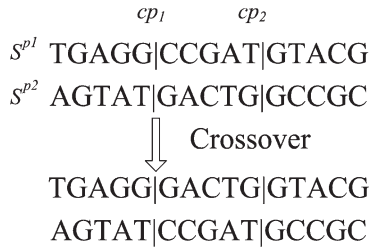


Fig. 7. Example of the two-point crossover operation.

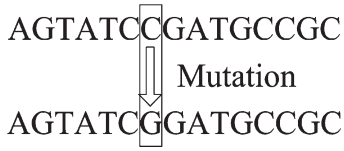


Fig. 8. Example of mutation operation ($C \rightarrow G$).

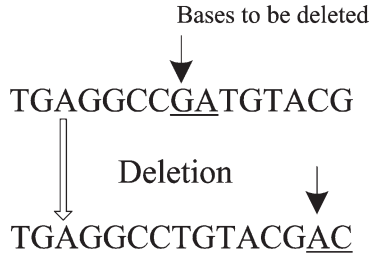


Fig. 9. Example of the deletion operation.

3) Crossover: The crossover operation is important for the entire search process. The two-point crossover operation is shown in Fig. 7. The crossover points, cp_1 and cp_2, are randomly selected, and the middle parts of the chromosomes, S^{p1} and S^{p2}, are exchanged (Fig. 8).

The different spectral DNA populations exchange their information by crossover operations, so the diversity of DNA individuals is enhanced. In this way, it is possible to find the most typical spectral DNA database. The crossover operation is the most important step in the optimization procedure.

4) Mutation: This operation is applied to a randomly selected basic nucleotide of the individual, according to the mutation probability.

The operation of mutation can generate diversity for a spectral DNA population and help genetic operations to enlarge the search range in DNA computing space.

5) Structured Mutations: The deletion, insertion, and inversion operations derived from biological DNA evolution are actually the structured mutation operations to be applied to the individuals. The purpose of adopting these operations is to try to replace those bad individuals with better ones at the start of evolution. It should be noted that the probability of these operations should be decreased, because the better individuals would have a higher risk of being disrupted by the random-based structured mutation [35].

When the deletion operation is implemented, new bases are randomly generated to make the individual abridged (Fig. 9). The superfluous bases are deleted when the insertion operation is implemented (Fig. 10). Fig. 11 shows the process of the inversion operation, with the inversion segment randomly selected. Considering the effectiveness of the structured

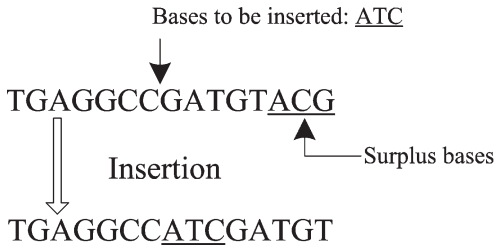


Fig. 10. Example of the insertion operation.

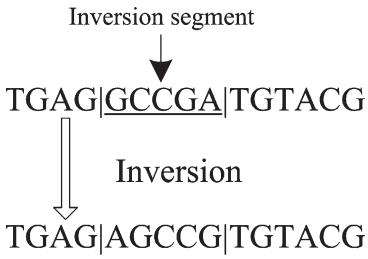


Fig. 11. Example of the inversion operation.

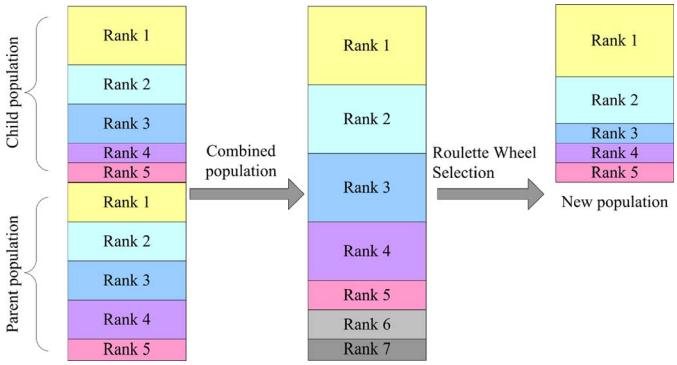


Fig. 12. Elitist selection algorithm.

mutation operations, a length of the deletion or insertion sequence containing only a few bases has sufficient ability to disrupt the individuals.

Applying structured DNA computing operators to the local parts of the spectral DNA population, the optimization procedure of artificial DNA computing will have the ability of diversity and noise tolerance. Due to the structured DNA computing operators, artificial DNA computing has more efficient and stable optimizing capabilities.

6) Elitism Selection Strategy: Population replacement follows an elitist selection strategy [37]. Sometimes, good individuals can be lost when crossover, mutation, or structured mutation results in offspring that are weaker than the parent population. An elitist selection algorithm can help prevent the loss of good solutions, once they are found, and ensure convergence. It is used to combine the parent and offspring population into a merged population, and then roulette wheel selection is carried out on the merged population (Fig. 12). The best individual in the merged population is stored as the excellent spectral DNA database P^{BDNA}.

"Roulette wheel selection" is used to determine which individual will be selected, according to the selection probability of the individuals, as calculated by the strategy of proportional assignment (15). The higher the fitness of the individual acquired, the greater the likelihood that the better individuals are favored, and will contribute with a large number of copies to the next generation.

$$R_i = \frac{Fitness_i}{\sum_{j=1}^{N} Fitness_j} \tag{15}$$

where N is the size of the population.

Due to the roulette wheel selection, the individuals with high fitness can pass into the next population, so the diversity of the new population is maintained, and premature convergence is also avoided. Due to the elitism replacement strategy, the convergence of the optimization process of the ADSM algorithm can be certified by introducing Markov chain theory [38].

7) Termination Rule: The stopping condition is different in different applications. One option is to set a fixed number of iterations as the stopping condition. Another option is to set a fixed threshold. In this paper, the optimization process will be terminated by both conditions: if the predefined maximum iteration number is reached, or the best kappa coefficient is satisfied. Finally, the classification result can be obtained using the excellent spectral DNA database P^{BDNA} with the spectral DNA similarity discrimination criteria.

The pseudocode of ADSM is given in Fig. 13(a) to illustrate the training process of ADSM in detail, and the flowchart for DNA computing in the spectral matching application is shown in Fig. 13(b).

IV. Experimental Results

The aforementioned DNA computing was coded in Visual C++ 6.0 and tested on a hyperspectral field-measured library and different hyperspectral images. Consistent comparisons between the artificial DNA computing spectral matching algorithm (ADSM) and traditional spectral matching algorithms of BC, SAM, SDFC algorithm, and support vector machine (SVM) were completed.

A. Experiment 1

The field-measured reflectance signatures were collected from the Qinghai-Tibet Plateau in August 2009 using an ASD spectroradiometer operating in the visible and infrared (325–2500 nm) region, with a spectral resolution of 1 nm. The spectral "spikes" from 1400–1535 nm are due to atmospheric absorption that was removed. A total of 2040 bands were used in this experiment 1. The signatures of seven objects were collected: grass, dry grass, wetland, rock, road, bare soil, and water. The class name, picture, signatures, and spectral DNA codewords for these classes are shown in Fig. 14, along with the mean spectral spectra. It is noted that the vegetation is composed mainly of various grasses on the Qinghai-Tibet Plateau. Due to the level of water and chlorophyll content, the signatures of different grasses can be classified into two categories: grass and dry grass.

The signatures of each class measured in the field were collected from different locations and at different times, so there are diversities in the signatures that belong to each class. When the training samples were selected, the samples from different locations were separately selected. To train and test all the algorithms, the training and test samples were randomly

ADSM methodology:

Input: Training samples (S) and hyperspectral image

Output: Best optimized DNA database (p^{BDNA})

Step 1. Randomly divide the training samples (S)

$S^I \in S$ For initialization the spectral DNA population

$S^V \in S$ For calculating the fitness of individuals

Step 2. $p^{DNA} \leftarrow Encoding(S^I)$

Step 3. Build initial DNA populations P^I

$P_i^I \leftarrow Random\{P^{DNA}\}, i=0,1,\ldots,popsize$

//Randomly selecting populations P^I

$\forall p_i^I \in P^I, f_i^I \leftarrow Fitness(p_i^I)$ //Evaluate fitness

$(p_{(i)}^I \prec p_{(j)}^I) \Leftrightarrow (f_{(i)}^I > f_{(j)}^I)$

//Sort individuals in P^I and obtain the best individual $p_{(1)}$

$p^{BDNA} \leftarrow p_{(1)}$ //Store best individual

Step 4. DNA database optimization

While not Stopping Condition **do**

p^g //Current generation, $g=1,\ldots,maxgeneration$

$p^{g-m} \leftarrow mutation(p^g)$ //mutation operation

$P^{g-m-c} \leftarrow crossover(P^{g-m})$ //crossover operation

$P^{g'} \leftarrow P^{g-m-c-sm} \leftarrow structured\ mutation(P^{g-m-c})$

//structured mutation operation

$p^{merge} \leftarrow merge(p^g, p^{g'})$

$\forall p_i^{merge} \in P^{merge}, f_i^{merge} \leftarrow Fitness(p_i^{merge})$ //Evaluate fitness

$(p_{(i)}^{merge} \prec p_{(j)}^{merge}) \Leftrightarrow (f_{(i)}^{merge} > f_{(j)}^{merge})$

//Sort individuals and obtain the best individual

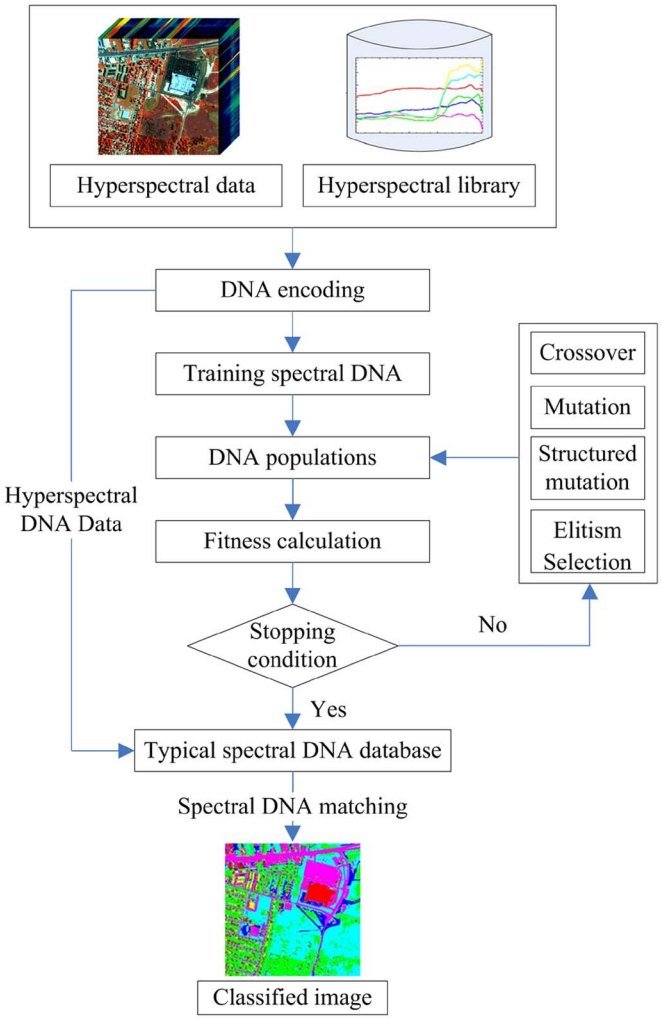
$p^{g+1} \leftarrow Roulette\ wheel\ selection(p^{merge})$

$p^{BDNA} \leftarrow p_{(1)}^{merge}$ //Store best individual

end while

Return p^{BDNA}

(a)



(b)

Fig. 13. Training and classification process of ADSM. (a) The training process of ADSM. (b) Flowchart of ADSM.

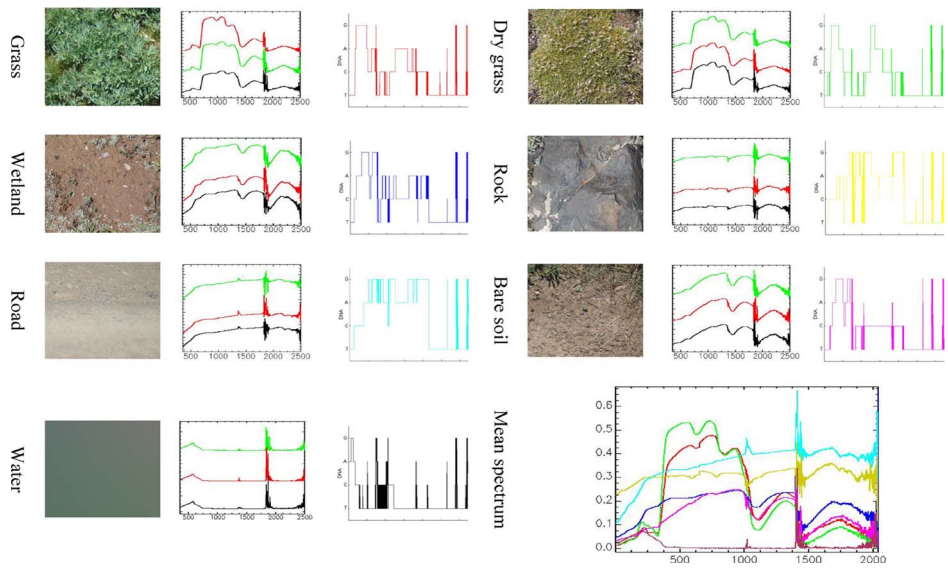


Fig. 14. Field pictures and spectral signatures.

selected from the field-measured library. The spectral classes and associated numbers of signatures in each class are shown in Table II. The main parameters of ADSM are set as follows: individual number $= 20$, maximum generation $= 200$, crossover probability $= 0.85$, mutation probability $= 0.2$, structured mutation probability $= 0.35$, termination condition $= 0.98$,

TABLE II
SPECTRAL CLASSES AND ASSOCIATED NUMBERS OF SIGNATURES IN EXPERIMENT 1

Number	Grass	Dry grass	Wetland	Rock	Bare soil	Water	Road
Total samples	36	21	9	18	12	9	15
Training samples	6	4	3	3	3	3	3
Test samples	30	17	6	15	9	6	12

TABLE III
CONFUSION MATRIX FOR CLASSIFICATION METHODS

Method	Classes	Grass	Dry grass	Wetland	Rock	Bare soil	Water	Road
BC	Grass	20	2	2	0	0	0	0
	Dry grass	6	10	0	0	1	0	0
	Wetland	0	0	4	1	2	0	0
	Rock	0	0	0	11	0	0	2
	Bare soil	4	5	0	0	6	0	2
	Water	0	0	0	1	0	6	0
	Road	0	0	0	2	0	0	8
SAM	Grass	20	1	1	0	0	0	0
	Dry grass	7	11	1	0	0	0	0
	Wetland	0	0	4	1	2	0	0
	Rock	0	0	0	9	0	0	1
	Bare soil	3	5	0	0	7	0	1
	Water	0	0	0	1	0	6	0
	Road	0	0	0	2	0	0	10
SDFC	Grass	24	1	0	0	0	0	0
	Dry grass	4	13	0	0	0	0	0
	Wetland	1	0	5	2	1	0	0
	Rock	0	0	0	10	0	0	1
	Bare soil	1	1	1	1	7	0	1
	Water	0	0	0	1	0	6	0
	Road	0	0	0	1	1	0	10
SVM	Grass	25	0	0	0	0	0	0
	Dry grass	4	17	0	0	0	0	0
	Wetland	0	0	6	0	0	0	0
	Rock	0	0	0	9	0	0	0
	Bare soil	0	0	0	6	9	0	0
	Water	0	0	0	0	0	6	0
	Road	1	0	0	0	0	0	12
ADSM	Grass	26	2	0	0	0	0	0
	Dry grass	3	15	0	0	0	0	0
	Wetland	0	0	6	0	1	0	0
	Rock	0	0	0	14	0	0	0
	Bare soil	1	0	0	0	8	0	1
	Water	0	0	0	0	0	6	0
	Road	0	0	0	1	0	0	11

adaptive global coefficient $= 0.8$ and adaptive gradient coefficient $= 1.2$. With the intention of avoiding the appearance of the "overfit" phenomenon, the training samples are divided into two uncorrelated parts. One third of the training samples from different classes are used to initialize the DNA population, and the others are considered as the validation for calculating the fitness values. With the consideration of the limited training samples, the higher mutation probability is selected to ensure the diversity of the population in the experiment.

In this experiment, the HD is taken as the distance measuring criteria for the ADSM algorithm, due to the fine spectral resolution of hyperspectral data. In executing ADSM on the hyperspectral library, the population of the typical spectral DNA database continues to be optimized. When the kappa coefficient of the individual is more than a fixed threshold, or the generation number reaches the maximum, the iterations will end. In the results shown in Table III, wetland, rock, and water can be exactly distinguished; there is, however, some confusion between grass and dry grass, and between bare soil and road.

For a more detailed verification of the results, test ground truth data were compared with the classified images to assess the accuracy of each classifier quantitatively, using both the overall accuracy measure and the kappa coefficient [39].

TABLE IV
COMPARISON OF FOUR CLASSIFIER PERFORMANCES IN EXPERIMENT 1

Method	BC	SAM	SDFC	SVM	HD-ADSM
Overall accuracy (%)	68.42	72.04	80.65	88.42	90.53
Kappa coefficient	0.6182	0.6625	0.7653	0.8591	0.8838
Total time	1.838	4.634	3.662	49.233	6.226

Tables III and IV list the results of comparisons between the original label of the test sample data and the classified label obtained by five spectral matching classifiers: BC, SAM, SDFC, SVM, and ADSM algorithms. The popular radial basis function (RBF) kernel, with the ability to solve nonlinear problems, is adopted for the SVM classifier. The parameters c and γ are set as 1000 and 0.005 by v-fold cross-validation ($v = 3$).

From Tables III and IV, it is apparent that the ADSM classifier produces better classification results than the other classifiers. The details are as follows: ADSM exhibits the best overall classification accuracy, i.e., the best percentage of correctly classified pixels among all the tested pixels considered, with gains of 22.11%, 18.49%, 9.88%, and 2.11% over the BC, SAM, SDFC, and SVM methods, respectively. The ADSM classifier improves the kappa coefficient from 0.6182 to 0.8838, an improvement of 0.2656. Because of the complexity of the ground objects and the diversity of disturbance, these conventional classification methods with bad diversity tolerance capacity perform a low-precision classification. Due to the very high-dimensional data, the SVM carries out an unsatisfactory discrimination by misclassifying between grass and dry grass, and mixing up rock and bare soil classes.

The computational time of these methods is given in Table IV. Although very high-dimensional data is used, only a small amount of time is taken to perform spectral matching classification with BC, SDFC, SAM, and ADSM. However, the classification time of SVM is sharply increased due to the statistical learning mechanism. Compared with SVM, the ADSM gives higher classification accuracy with smaller time cost. It is considered that the ADSM is capable of classifying a spectral library with field-measured reflectance signatures.

B. Experiment 2

In this experiment, the data were from an airborne imaging spectrometer (PHI). A total of 80 bands (346 $\times$ 400 pixels) were taken from the Xiaqiao test site, to be used in this experiment, and their spectral ranges were from 0.417 to 0.854 μm. Fig. 15(a) shows the experimental PHI image. The observed image was expected to fall into six classes: road, corn, bare soil, vegetable, soil, and water. The ground truth samples field map is shown in Fig. 14(b), and the reference spectral curves for six classes are presented in Fig. 14(c). There are specified training and test samples in the data set, and the list of classes and the number of training and test samples for each class is given in Table V. The partition strategy for the training process is adopted for the optimization process. One fifth of the training samples are randomly selected for the initialization of the DNA population, and the other samples are regarded as the validation for the optimal processing of artificial DNA computing.

In experiment 2, the parameters were set as shown in Table VI. The distance criteria: HD, AVD, and PVD, are introduced to evaluate the similarity of the spectra. For a convenient comparison between the ADSM algorithm and traditional

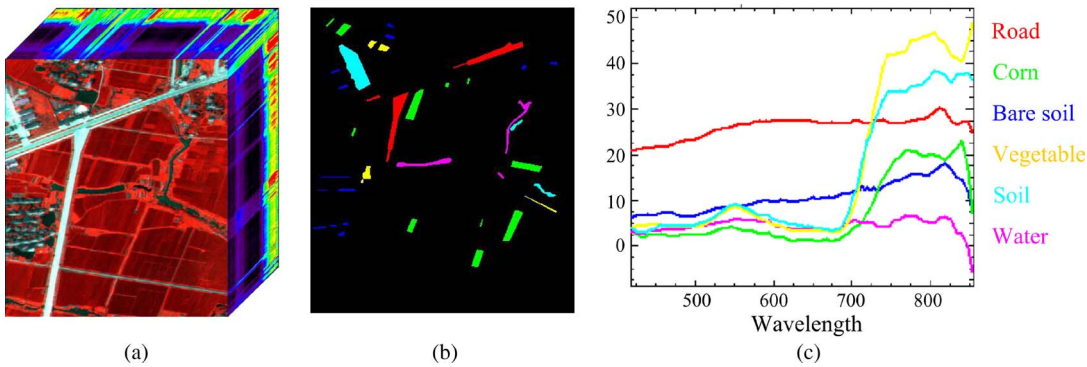


Fig. 15. Xiaqiao PHI image. (a) Xiaqiao PHI image RGB (70, 40, 10). (b) The ground truth samples. (c) Reference spectral curves for six classes.

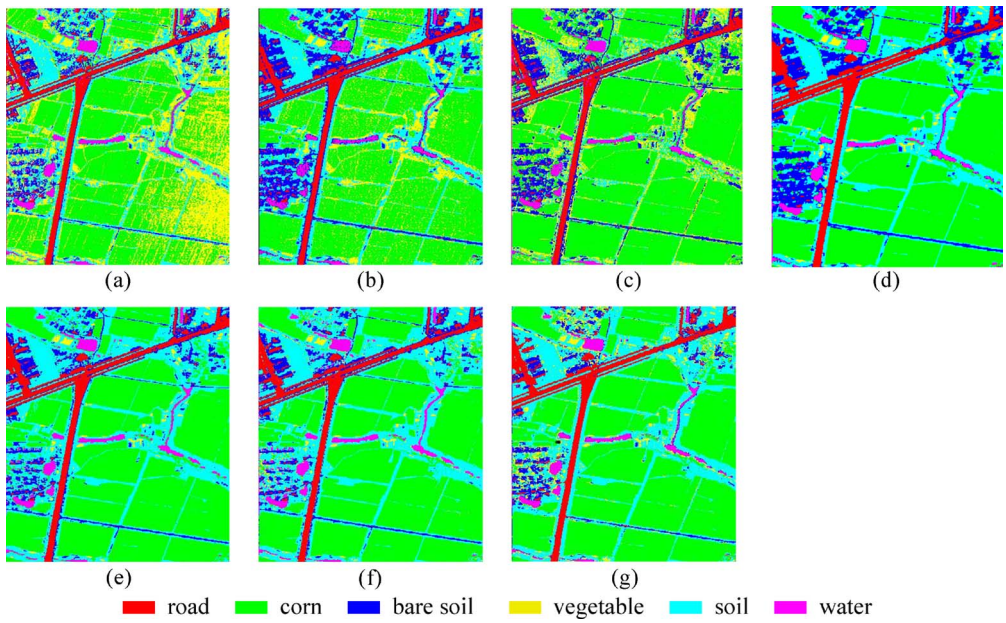


Fig. 16. Spectral matching classification results for the Xiaqiao PHI image. (a) BC. (b) SAM. (c) SDFC. (d) SVM. (e) HD-ADSM. (f) AVD-ADSM. (g) PVD-ADSM.

TABLE V
LAND COVER CLASSES AND ASSOCIATED NUMBERS OF PIXELS USED IN EXPERIMENT 2

Number	Road	Corn	Bare soil	Vegetable	Soil	Water
Total samples	1393	1749	498	452	1391	661
Training samples	138	128	146	67	130	72
Test samples	1255	1621	352	385	1261	589

TABLE VI
PARAMETER SETTINGS IN EXPERIMENT 2

ADSM parameters	
Individual number	20
Crossover probability	0.8
Mutation probability	0.2
Structured mutation probability	0.3
Maximum generation	40
Termination condition	0.99
Adaptive global coefficient	0.75
Adaptive gradient coefficient	1.5

spectral matching algorithms, and the statistical learning method SVM, Fig. 16(a)–(d) shows the classification results using BC, SAM, SDFC, and SVM classifiers, and Fig. 16(e)–(g) shows the classified results using HD-ADSM, AVD-ADSM, and PVD-ADSM, respectively. The estimation of the parameters of SVM-RBF is conducted by v-fold cross-validation on the training samples ($v = 5$). The cost coefficient and γ are set as 1000 and 0.0013, respectively. To evaluate the classification accuracy of the different classifiers, a test field map is provided based on the ground truth map in Fig. 15(b).

The visual comparisons of the seven supervised classifications in Fig. 16 show varying degrees of pixel assignment accuracy. All seven classifiers generate similar classification results in the road class. The result of the BC classifier is disappointing, with misclassification between corn and vegetable. Although SAM, with the maximum angle of 0.35 (to avoid the overfitting), can discriminate between corn and vegetable, it cannot distinguish the soil class from the corn and vegetable classes, and many unknown pixels (black pixels) exist in its classified image. The SDFC classifier can make a distinction between most classes, and the overall classification is acceptable; however, when we closely examine the inner area of each class, it is disappointing to find many misclassified points in the corn, soil, and vegetable classes. Therefore, the result of the SDFC classifier needs to be improved from the aspect of

TABLE VII
COMPARISON OF SEVEN CLASSIFICATION METHODS

Method	Class	road	corn	Bare soil	vegetable	soil	water	Total
BC	unclassified	0	0	0	0	0	0	0
	road	1151	0	100	0	0	0	1251
	corn	3	910	2	3	1	7	926
	bare soil	70	0	190	1	3	50	314
	vegetable	5	696	15	325	53	2	1096
	soil	10	15	25	56	1204	26	1336
	water	16	0	20	0	0	504	540
	Total	1255	1621	352	385	1261	589	5463
SAM	unclassified	0	6	0	1	3	36	46
	road	1211	0	112	0	0	28	1351
	corn	0	1450	0	30	5	1	1486
	bare soil	44	0	225	0	4	95	368
	vegetable	0	139	0	316	50	0	505
	soil	0	26	1	38	1199	6	1270
	water	0	0	14	0	0	423	437
	Total	1255	1621	352	385	1261	589	5463
SDFC	unclassified	0	0	0	0	0	0	0
	road	1208	0	96	0	0	0	1304
	corn	4	1471	3	31	35	5	1549
	bare soil	24	0	223	0	0	12	259
	vegetable	1	125	0	337	176	1	640
	soil	8	25	10	17	1050	24	1134
	water	10	0	20	0	0	547	577
	Total	1255	1621	352	385	1261	589	5463
SVM	unclassified	0	0	0	0	0	0	0
	road	1122	0	66	0	0	0	1188
	corn	0	1619	0	0	6	0	1625
	bare soil	133	0	286	1	3	26	449
	vegetable	0	0	0	234	3	0	237
	soil	0	2	0	150	1249	5	1406
	water	0	0	0	0	0	558	558
	Total	1255	1621	352	385	1261	589	5463
HD-ADSM	unclassified	0	0	0	0	0	0	0
	road	1197	0	94	0	0	0	1291
	corn	0	1613	4	0	2	5	1624
	bare soil	56	0	221	1	5	6	289
	vegetable	0	0	0	212	1	0	213
	soil	1	8	19	172	1253	27	1480
	water	1	0	14	0	0	551	566
	Total	1255	1621	352	385	1261	589	5463
AVD-ADSM	unclassified	0	0	0	0	0	0	0
	road	1199	0	93	0	0	0	1292
	corn	0	1607	0	1	3	0	1611
	bare soil	56	0	198	0	1	0	255
	vegetable	0	3	1	196	2	1	203
	soil	0	11	42	188	1255	24	1520
	water	0	0	18	0	0	564	582
	Total	1255	1621	352	385	1261	589	5463
PVD-ADSM	unclassified	0	0	0	0	0	0	0
	road	1200	0	85	0	0	0	1285
	corn	0	1608	0	0	3	0	1611
	bare soil	54	0	203	0	0	0	257
	vegetable	0	0	0	217	2	0	219
	soil	1	13	55	168	1256	71	1564
	water	0	0	9	0	0	518	527
	Total	1255	1621	352	385	1261	589	5463

TABLE VIII
CONFUSION MATRIX FOR CLASSIFICATION METHODS IN EXPERIMENT 2

Method	BC	SAM	SDFC	SVM	HD-ADSM	AVD-ADSM	PVD-ADSM
Overall accuracy (%)	78.42	88.30	88.52	92.77	92.39	91.87	91.56
Kappa coefficient	0.7343	0.8518	0.8548	0.9077	0.9021	0.8954	0.8913

TABLE IX
COMPUTATIONAL TIME OF THE CLASSIFICATION METHODS

Time (s)	BC	SAM	SDFC	SVM	ADSM
Training time	0.703	0.332	2.484	18.262	4.365
Classification time	0.891	1.864	1.5	6.877	2.136
Total time	1.594	2.196	3.984	25.139	6.501

stability. The SVM classifier presents the best performance on the visual interpretation, and the corn, vegetable, and soil are discriminated very well. Nevertheless, the ADSM classifiers achieve acceptable visual accuracy in the vegetation, soil, and vegetable classes and also perform satisfactorily on the road, bare soil, and water classes. It is considered that there are tiny differences in the vegetable and bare soil classes by examining the HD-ADSM, AVD-ADSM, and PVD-ADSM classification results pixel-by-pixel.

Table VII shows the accuracy comparison of the artificial DNA computing classifier with the other classifiers in terms of overall accuracy and the kappa coefficient, which is calculated based on the confusion matrix shown in Table VIII. Clearly, the artificial DNA computing classifiers, with the distances of HD, AVD, and PVD, perform better than the traditional spectral matching methods, with an overall accuracy of 92.39%, 91.87%, and 91.56%, and kappa coefficients of 0.9021, 0.8954, and 0.8913, respectively. This is followed by BC, SAM, and, lastly, the SDFC classifier. These classifiers are based on full spectral curves-based spectral matching. The SVM classifier achieves the best overall accuracy and kappa coefficient of 92.77% and 0.9077, respectively. The excellent performance of the SVM classifier benefits from the relatively low-dimensional data and much greater number of training samples. The computational complexity is given in Table IX, including the training, classification, and total time. Because the time complexities for HD, AVD, and PVD-ADSM are at the same level, their average time is considered as the computational time taken for ADSM. The training time of the ADSM classifier is much more than the other spectral matching classifier because of the optimization of the artificial DNA computing algorithm. The SVM classifier takes the longest time to train the discrimination model and perform the classification.

Tables VIII and VII list the results of accuracy evaluation between the test data and classified images, as obtained by seven classifiers: BC, SAM, SDFC, SVM, HD-ADSM, AVD-ADSM, and PVD-ADSM. From these tables, it is apparent that the HD-ADSM classifier produces better classification results than the other spectral matching classifiers when the Xiaqiao PHI data is of medium spatial resolution with 80 bands; however, the SVM classifier achieves the best result because of the better discrimination between soil and vegetable.

The details are as follows: The HD-ADSM classifier exhibits a better overall classification accuracy of 92.39%, compared to the traditional matching classifier, i.e., the percentage of correctly classified pixels among all the tested pixels, with gains of 13.97%, 4.09%, and 3.87% over the BC algorithm, SAM, and SDFC matching algorithms, respectively. However, the SVM classifier is based on structural risk minimization and exploits a margin-based criterion, achieving a higher classification accuracy for the Xiaqiao PHI data than the ADSM classifier, with a distinct discrimination between the vegetable and soil classes.

The main reason for the comparatively high accuracy achieved by the ADSM classifiers is that the BC classifier is not capable of capturing the basic feature of spectral signatures and lacks diversity tolerance; in particular, when the imagery contains heterogeneous radiation, there are many misclassified areas in the resulting image. Because the spectral angle is taken as the measure of spectral similarity, SAM is not influenced by

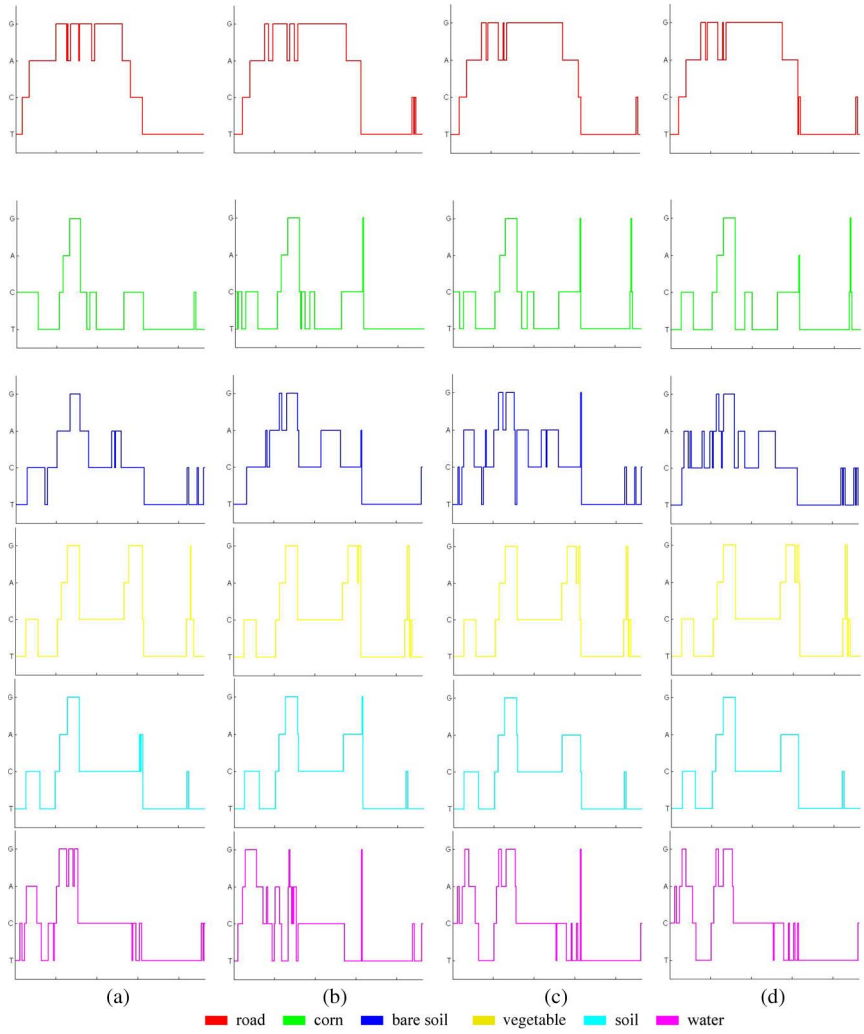


Fig. 17. Evolution process of the DNA database. (a) Initial DNA database. (b) DNA database at generation 10. (c) DNA database at generation 30. (d) Postevolved DNA database at generation 40.

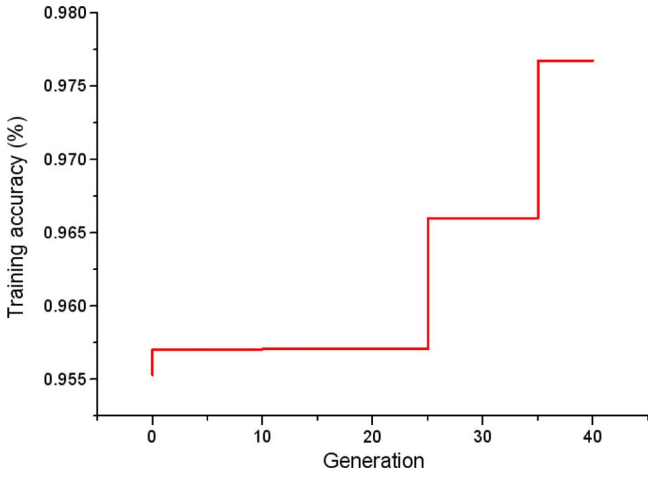


Fig. 18. Training accuracy with generation addition.

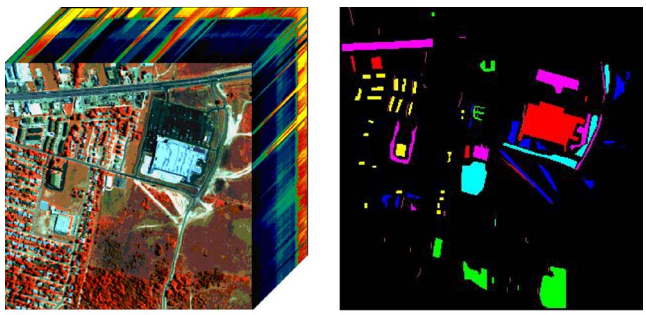


Fig. 19. Urban HYDICE data set. (a) False-color 3-D cube (R 63, G 52, B 36). (b) The ground truth for the experiment 3.

TABLE X
LIST OF CLASSES AND NUMBER OF LABELED SAMPLES IN EACH CLASS FOR EXPERIMENT 3

Number	Roof #1	Tree	Concrete road	Roof #2	Grass	Asphalt road	Shadow
Total samples	2041	1596	991	707	1332	1991	245
Training samples	36	33	10	9	13	32	5
Test samples	2005	1563	981	698	1319	1959	240

heterogeneous radiation; however, it is not able to discriminate spectral signatures with diverse properties. As a result, there are many unknown pixels (black pixels) and mislabeled pixels existing in its classified image. The SDFC classifier can only capture the mean spectral features of the training sample of one class, so its capacity for diversity tolerance is weak. As a result, there are many noisy points in the corn, soil, and vegetable classes.

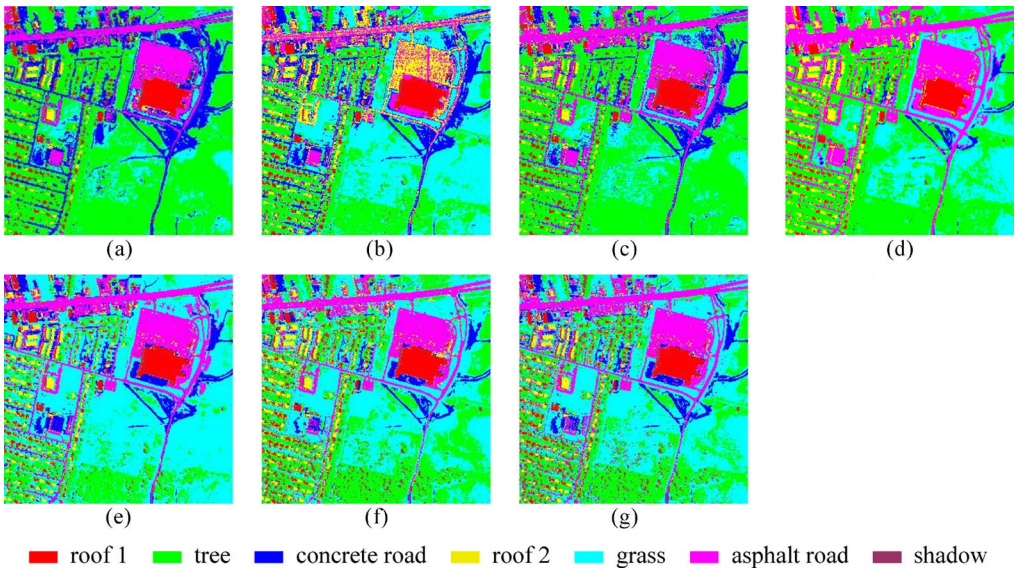


Fig. 20. Classification results by different methods. (a) BC. (b) SAM. (c) SDFC. (d) SVM. (e) HD-ADSM. (f) AVD-ADSM. (g) PVD-ADSM.

By contrast, the ADSM classifiers are data-driven and self-adaptive methods, which can adjust the DNA spectral database using the training samples. Hence, the feature capturing method of DNA computing is comprehensive and systematic, not only for the global spectral features but also for the gradient spectral features. The structured DNA computing operations enlarge the search space and make the procedure of spectral matching more robust and efficient. This enables the artificial DNA computing classifier to achieve the better accuracy.

The evolution process is presented in Fig. 17, including the initial DNA database, generation 10, generation 30, and the postevolved DNA database at generation 40. The changes in shape of the DNA database are comparatively large, in comparison with the DNA database at the initial generation and generation 10. However, in pace with the generation addition, the shape changes of the DNA database become gradually less, and the DNA database tends to stabilize from generation 30 to 40, even though there are some changes in the latter part of the evolution process because of the diversity of the class of bare soil. At the same time, the change in training accuracy (the best individual in each generation) with generation addition is given in Fig. 18. The higher training accuracy can be acquired along with the generation addition. It is considered that the optimization of the DNA system can provide a stable solution to the imagery classification problem, based on the changing trend of the DNA database, along with the generation addition.

Although the classification accuracy of ADSM is slightly lower than the statistical learning approach (SVM), we can conclude that the ADSM classifiers, based on the analysis above, are effective classifiers for PHI hyperspectral remote sensing images with medium spatial resolution and 80 bands. The reason for their advantage is that they not only obtain more details about the spectral curves, but can also avoid the interference from spectral diversity, making the classification process optimal.

C. Experiment 3

The third data set is urban data captured by the Hyperspectral Digital Imagery Collection Experiment (HYDICE) in October 1995. There are 210 bands in this data, and the spectral and spatial resolutions are 10 nm and 2 m, respectively. The image area is located at Copperas Cove, near Fort Hood, Texas, U.S., with a size of 307 $\times$ 307. Ground objects in the area include a highway near the top of the image, a shopping mall along the highway, a parking lot in front of the mall, some roads, grass, trees, and rows of houses. In addition, due to the low solar altitude, trees and houses cast long shadows on the ground [40]. This experimental data set is available from [41].

By assigning the 63rd, 52nd, and 36th bands as the R, G, and B components, respectively, a false-color 3-D cube can be obtained, as shown in Fig. 19(a). As can be seen, there are two main classes of roof: one is bright, the other, with a higher intensity of radiation absorption, resembles the shadows on the ground. Most of the roads are asphalt paved, including the highway and the parking lot, while a few are made of concrete. Trees and grasses also show some differences. Based on these analyses, we assume that there are seven classes in this data: roof#1 (bright), tree, concrete road, roof#2 (gray), grass, asphalt road, and shadow. Due to the high resolution of the spectral and spatial dimensions, the classification based on materials can be conducted by experiment [42], [43].

Before classification, bands 103–108, 139–152, and 208–210 are removed as noisy or water-absorption bands, leaving a total of 187 bands. To evaluate the classification accuracy, a test field map is provided in Fig. 19(b), based on ground truth data obtained by field sampling. Some pixels of the available samples were used as the training data set in order to train these algorithms. The list of classes and the number of training and test samples for each class are given in Table X. The parameters of the DNA computing model were set to be the same as in experiment 2 in Table VI. The training samples are divided into two parts, based on the partition strategy. One fifth of the training samples are used to initialize the DNA population, and the other training samples are applied to the optimization process of artificial DNA computing.

Fig. 20(a)–(d) shows the classification results using the BC, SAM, and SDFC matching methods, and the statistical learning method, SVM, respectively. The estimation of the parameters of

SVM-RBF is conducted by v-fold cross-validation on the training samples ($v = 5$). The cost coefficient and γ are set as 12 000 and 0.0005, respectively. Fig. 20(e)–(g) shows the classification result using HD, AVD, and PVD for the ADSM classifier. The accuracies, including total accuracy and kappa coefficient, for the several classifiers are given in Table XI, which are calculated based on the confusion matrix in Table XII.

The visual comparisons of the seven supervised classifications in Fig. 20 show varying degrees of pixel assignment accuracy. It can be seen that the classification result of the BC classifier provides a poor partition of the scene due to the misclassification of the tree and grass classes; at the same time, the concrete and asphalt road are also confused. The classified imagery of the SAM classifier, with the maximum angle of 0.4, presents a good discrimination between grass and tree, but the results of concrete and asphalt road are disappointing. The concrete road class is misclassified into asphalt road, while asphalt road and roof#2 cannot be well separated. In particular, there are many unknown pixels (black pixels) in the classified image because of the similarity measure of SAM and because the shadow under the trees is difficult to classify. The classification result of SDFC is much better than the former classifiers; however, it is sensitive to noise in the tree and asphalt areas and still has too much clutter. Furthermore, the classified result of the shadow under the trees is not acceptable. It is worth noting that the result of the SVM classifier is not very satisfactory because of misclassifying grass into the tree class, and the decrease in the concrete road class. However, the proposed classification methods are more robust to the diversity of the spectra than previous classification methods, as they exploit the global and gradient information of the spectra and optimize the matching process by DNA operators.

Table XII lists the results of the comparisons between the ground truth data and the classified images obtained by seven classifiers: BC, SAM, SDFC matching classifiers, SVM and HD-ADSM, AVD-HDSM, and PVD-ADSM methods, using overall accuracy and kappa coefficient. From the table, it is apparent that the ADSM classifiers produce better classification results than the other classifiers; furthermore, the highest accuracies are generated by HD-ADSM. The details are as follows: The HD, AVD, and PVD classifiers exhibit overall classification accuracies of 92.27%, 92.24%, and 92.42%, respectively. The HD-ADSM gives the best accuracy, with gains of 11.05%, 12.67%, 6.58%, and 4.12% over the BC, SAM, SDFC, and SVM classifier. The HD-ADSM classifier improves the kappa coefficient from 0.7622 to 0.9118, with an improvement of 0.1496. In addition, the computational complexity of these methods is compared by introducing training time, matching classification time and total time in Table XIII. Because of the reduction in the number of training samples, the training time is lower than in experiment 2, in spite of the addition of band numbers. Generally speaking, the SVM based on the statistics model consumes more time than the spectral matching classifier.

The ADSM classifier achieves a dramatically higher accuracy than the SVM classifier. A careful analysis between the Xiaqiao PHI data and the urban HYDICE data indicates that there is higher dimensionality and a smaller number of training samples in experiment 3. The accuracy of classification by a SVM can be significantly reduced by the addition of features,

TABLE XI
COMPARISON OF SEVEN CLASSIFICATION METHODS, USING OVERALL ACCURACY AND KAPPA COEFFICIENT IN EXPERIMENT 3

Method	Class	roof1	tree	concrete road	roof2	grass	asphalt road	shadow	Total
BC	unclassified	0	0	0	0	0	0	0	0
	roof1	1995	0	0	36	0	28	70	2129
	tree	0	1555	10	153	396	11	66	2191
	concrete road	0	2	652	25	349	35	0	1063
	roof2	0	0	82	470	0	14	11	577
	grass	0	6	8	1	561	3	0	579
	asphalt road	3	0	229	7	13	1842	10	2104
	shadow	7	0	0	6	0	26	83	122
	Total	2005	1563	981	698	1319	1959	240	8765
SAM	unclassified	8	14	0	58	0	75	155	310
	roof1	1991	0	0	3	0	23	8	2025
	tree	0	1511	4	0	3	1	8	1527
	concrete road	0	1	737	53	38	23	0	852
	roof2	0	0	2	566	5	935	20	1528
	grass	0	37	9	2	1273	4	0	1325
	asphalt road	6	0	229	12	0	893	4	1144
	shadow	0	0	0	4	0	5	45	54
	Total	2005	1563	981	698	1319	1959	240	8765
SDFC	unclassified	0	0	0	0	0	0	0	0
	roof1	1918	0	0	8	0	18	18	1962
	tree	0	1555	12	138	391	13	68	2177
	concrete road	0	1	756	10	69	21	0	857
	roof2	0	0	13	515	2	40	12	582
	grass	0	7	18	1	854	13	0	893
	asphalt road	3	0	182	6	3	1821	10	2025
	shadow	84	0	0	20	0	33	132	269
	Total	2005	1563	981	698	1319	1959	240	8765
SVM	unclassified	0	0	0	0	0	0	0	0
	roof1	1989	0	0	3	0	13	2	2007
	tree	0	1528	22	4	196	6	38	1794
	concrete road	0	0	485	0	0	0	0	485
	roof2	9	0	1	640	0	34	53	737
	grass	0	32	51	2	1111	1	0	1197
	asphalt road	7	0	422	26	12	1878	12	2357
	shadow	0	3	0	23	0	27	135	188
	Total	2005	1563	981	698	1319	1959	240	8765
HD-ADSM	unclassified	0	0	0	0	0	0	0	0
	roof1	1999	0	0	0	0	4	3	2006
	tree	0	1168	0	6	0	1	26	1201
	concrete road	1	0	952	2	0	1	0	956
	roof2	0	0	0	623	0	17	13	653
	grass	0	385	29	13	1302	5	2	1736
	asphalt road	1	0	0	11	17	1897	10	1936
	shadow	4	10	0	43	0	34	186	277
	Total	2005	1563	981	698	1319	1959	240	8765
AVD-ADSM	unclassified	0	0	0	0	0	0	0	0
	roof1	1928	0	1	4	0	19	3	1955
	tree	0	1355	4	1	9	2	4	1375
	concrete road	68	0	863	5	0	2	0	938
	roof2	1	2	3	603	1	53	16	679
	grass	0	105	109	7	1306	16	0	1543
	asphalt road	2	0	1	15	3	1823	10	1854
	shadow	6	101	0	63	0	44	207	421
	Total	2005	1563	981	698	1319	1959	240	8765
PVD-ADSM	unclassified	0	0	0	0	0	0	0	0
	roof1	1914	0	0	2	0	12	3	1931
	tree	0	1312	4	2	4	1	4	1327
	concrete road	81	2	927	4	0	2	0	1016
	roof2	1	1	0	586	0	28	17	633
	grass	0	146	44	10	1296	16	0	1512
	asphalt road	4	2	6	42	19	1861	11	1945
	shadow	5	100	0	52	0	39	205	401
	Total	2005	1563	981	698	1319	1959	240	8765

TABLE XII
CONFUSION MATRIX FOR CLASSIFICATION METHODS IN EXPERIMENT 3

Method	BC	SAM	SDFC	SVM	HD-ADSM	AVD-ADSM	PVD-ADSM
Overall accuracy (%)	81.67	80.05	86.14	88.60	92.72	92.24	92.42
Kappa coefficient	0.7756	0.7622	0.8312	0.8605	0.9118	0.9063	0.9084

TABLE XIII
COMPUTATIONAL TIME OF THE CLASSIFICATION METHODS

Time (s)	BC	SAM	SDFC	SVM	ADSM
Training time	0.625	0.353	2.297	14.361	3.276
Classification time	1.219	5.426	3.266	6.253	3.012
Total time	1.844	5.779	5.563	20.614	6.288

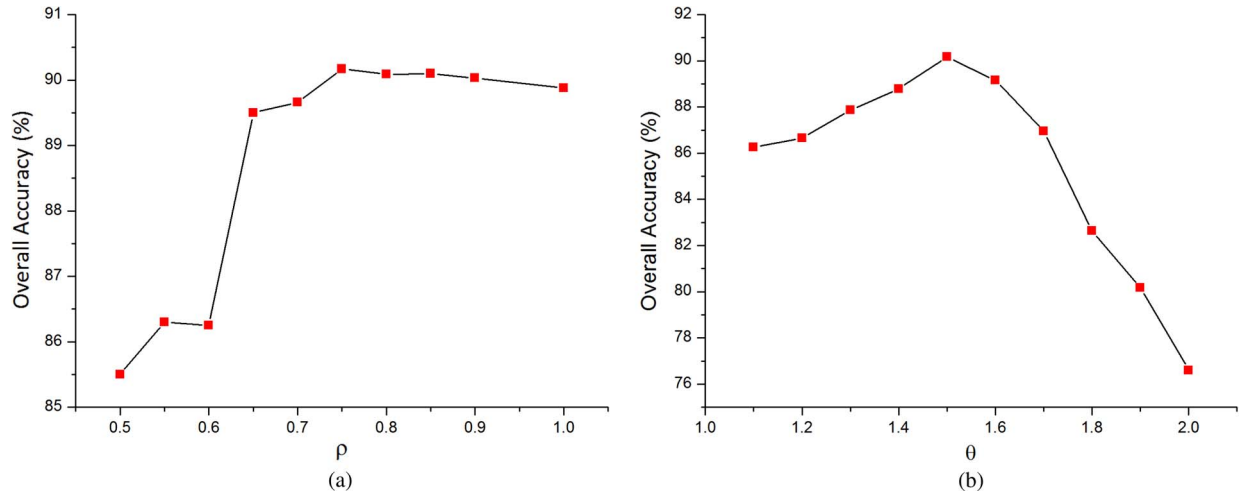


Fig. 21. OA in relation to ρ and θ. (a) OA in relation to ρ. (b) OA in relation to θ.

and this effect is most apparent with a small number of training sets because of the appearance of the Hughes phenomenon [44].

In experiment 3, the ADSM classifier achieves a better classification performance than the other classifiers because of its spectral diversity tolerance and its ability to capture global and gradient features of the spectrum. A feature of this experiment is that the spatial resolution of the urban hyperspectral data is higher than the Xiaqiao hyperspectral data, and the exact materials of the ground truth can be decided in the imagery. Therefore, the classification standard is selected based on the material class, including tree, concrete, grass, asphalt, shadow, and two roofs of different materials. Surprisingly, the result based on material classes of ADSM reaches a high accuracy classification.

V. Sensitivity Analysis

A. Sensitivity in the Adaptive Global Coefficient and the Adaptive Gradient Coefficient

The DNA encoding of the ADSM algorithm has two parameters defined by users, namely, the adaptive global coefficient ρ and the adaptive gradient coefficient θ, as the adjusting parameters for the thresholds of DNA encoding. They significantly influence the information extraction quality of DNA encoding. To analyze the effects of setting these parameters when running the ADSM algorithm, the Xiaqiao PHI image, as shown in Fig. 15, was classified using different values for the parameters. ρ was assumed to have the following values: $\rho = \{0.5, 0.55, 0.6, 0.65, 0.7, 0.75, 0.8, 0.85, 0.9, 1\}$. The relationship between the ρ value and overall accuracy with $\theta = 1.5$ is presented in Fig. 21(a). θ was assumed to have the following values: $\theta = \{1.1, 1.2, 1.3, 1.4, 1.5, 1.6, 1.7, 1.8, 1.9, 2.0\}$. The relationship between the θ value and overall accuracy with $\rho = 0.75$ is presented in Fig. 21(b).

As shown in Fig. 21(a), the higher the ρ value, the higher the OA. The reason for this is as follows: The higher ρ value makes the threshold of DNA encoding float around the middle value of the signature, and more features of reflectance signatures of different labels are extracted.

As shown in Fig. 21(b), the smaller or higher θ values bring about the smaller OA, but the higher θ values lead to a faster decrease in OA. When the θ value is set in the middle of the value range, OA is higher. The reason for this is as follows: Because the higher θ value results in a loss of the gradient variation information, the classifier does not have enough discrimination to classify the different classes; the smaller θ value leads to exceeding sensitivity to spectral diversity, so the misclassified results are acquired. This situation is similar to using different levels of low-pass filters to process the data.

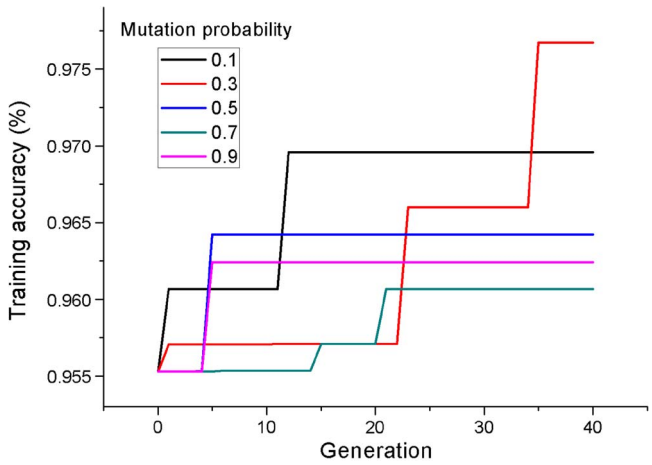


Fig. 22. Sensitivity in mutation probability (crossover probability $= 0.8$, structured mutation probability $= 0.2$).

B. Sensitivity in the DNA Optimization Operators' Parameters

The DNA optimization operators can find the optimal solution of the DNA database, which is impacted by the parameters of crossover probability, mutation probability, and structured mutation probability. To analyze the effects of setting these parameters, the ADSM algorithm for the Xiaqiao PHI image, as shown in Fig. 15, was conducted using different values for the separate parameters. The results of the training accuracy are shown in Figs. 22–24, respectively.

It can be noticed that for a value of mutation probability of 0.3, the best training accuracy can be acquired under a crossover probability of 0.8 and structured mutation probability of 0.2, when the generations reach 40 in Fig. 22. The larger the

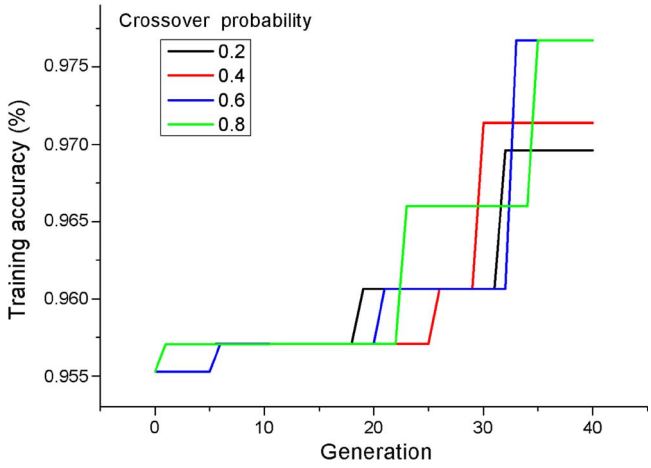


Fig. 23. Sensitivity in crossover probability (mutation probability $= 0.3$, structured mutation probability $= 0.2$).

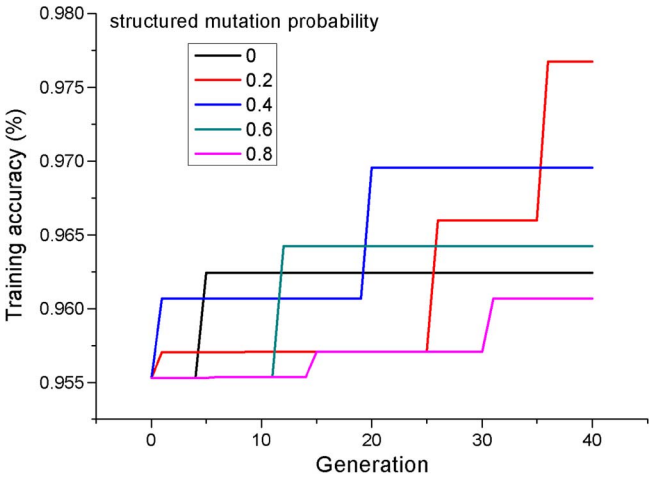


Fig. 24. Sensitivity in structured mutation probability (crossover probability $= 0.8$, mutation probability $= 0.3$).

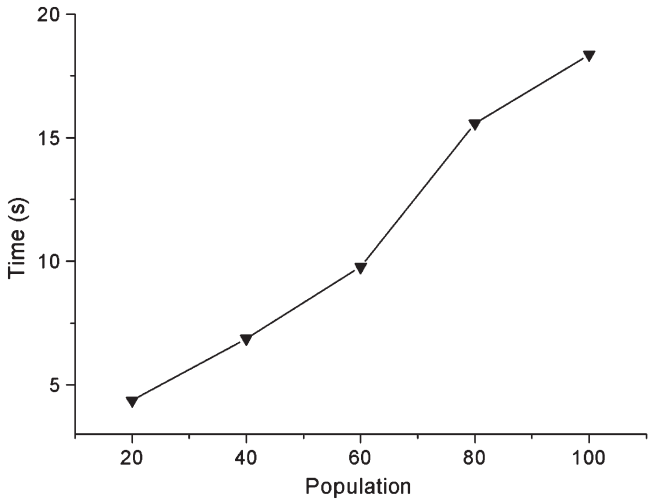


Fig. 25. Computational time with different populations.

mutation probability selected, the lower the optimal results and the slower the optimization process. The tendency is opposite when the sensitivity of crossover probability is conducted in Fig. 23. In this case, the better optimization results are found

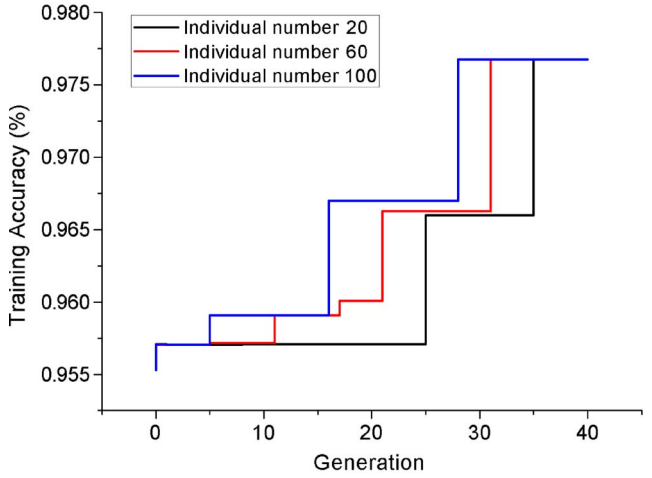


Fig. 26. Comparison of training accuracy with different initial populations.

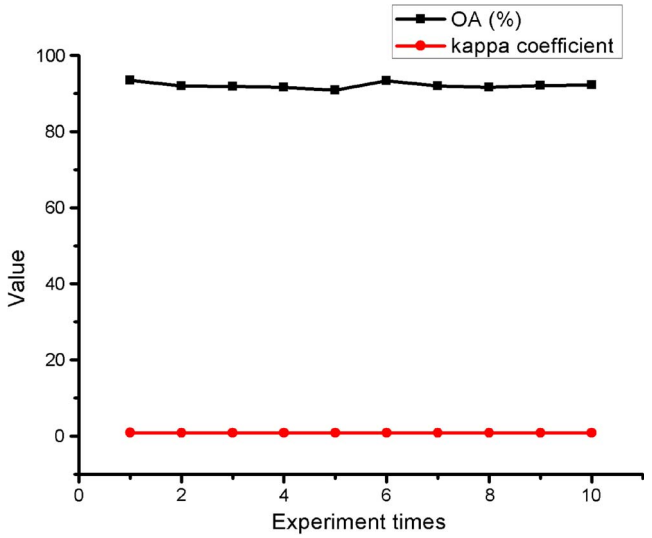


Fig. 27. Results of the cross-validation experiments.

when the crossover probability is 0.8. The sensitivity analysis of structured mutation probability is also carried out. It can be seen from the results presented in Fig. 24 that the lower probability gives better optimization. When the experiment is implemented under the structured mutation probability of 0, the results and optimal process are not satisfactory. It is confirmed that the structured mutation operator is useful for the DNA optimization, due to the more random DNA codewords.

C. Sensitivity in the Number of the DNA Population

The number of the DNA population has an effect on the range of search space of DNA computing. The computational time with different populations is given in Fig. 25. In addition, the training accuracy with different populations is shown in Fig. 26. There is an approximately linear relationship between computational time and population number. It is illustrated that a larger size of initial population brings about a wider searching space and better solutions for DNA optimization, accompanied with longer computational time.

D. Sensitivity to Training and Test Data

To validate the generalization ability and sensitivity to training samples of ADSM, a cross-validation experiment was carried out with the HYDICE urban data. The test samples, shown in Fig. 19(b), were randomly grouped into ten parts. A group of samples from the total test samples were selected as training samples, while the remaining samples were regarded as test samples. The results of ADSM with PVD distance criteria were evaluated by calculating the overall accuracy and kappa coefficient of classification, based on the remaining test samples. The classification process was repeated ten times with each group of test samples. The OA and kappa coefficient of each classification are presented in Fig. 27. It is seen that the OA and kappa coefficient keep relatively stable states.

The best OA in the cross-validation experiments is 93.43, accompanied by the best kappa coefficient of 0.9212. The worst results have an OA and kappa coefficient of 91.08 and 0.8997, respectively. The averages of OA and kappa coefficient are 92.46% and 0.9089, respectively. It is considered from the cross-validation experiments that the ADSM classifier, with fine generalization ability, is not knowledge independent of training and test samples.

VI. Conclusion

In this paper, some initial investigations into employing DNA computing for hyperspectral remote sensing data discrimination and classification were carried out, and a novel spectral matching mapping algorithm based on the artificial DNA computing model, namely, the ADSM, was proposed. The DNA encoding method can capture the rich information of spectral features and, with the DNA modulating and controlling mechanisms, the most typical DNA database can be acquired. For each pixel of a hyperspectral image, or each signature measured in the field, the distance between the measured signature and the signature in the typical DNA database is computed. The label of the minimum distance is found, based on the distance criteria of HD, AVD, and PVD, and the class property of each is set as the class with the minimum distance. The main idea of artificial DNA computing is that the spectral matching classification for hyperspectral data should be considered as the spectral DNA strands identification. Therefore, the transformation is conducted from the spectral space to DNA computing space, with the optimization by DNA evolution operations.

The objective of ADSM is the formation of DNA codewords for hyperspectral reflectance data. The classifier not only obtains more detailed information about the spectral signatures, but can avoid the interference from spectral diversity and also optimize the matching process. The hyperspectral data experiments were carried out to evaluate the performance of the proposed algorithm. It is demonstrated that the proposed algorithm is superior to the three traditional hyperspectral data classification algorithms. When the experimental data is characterized with high dimensionality and a small number of training samples, the ADSM shows higher stability and efficiency compared with the novel statistical learning method of SVM. The experiments with artificial DNA computing for ground-measured spectral discrimination and hyperspectral imagery classification have been accomplished. In our future work, in the framework of DNA computing for hyperspectral data, the ground-measured spectra will be considered to directly discriminate the hyperspectral imagery, and the gene-level-based operators will also be introduced, such as the combination of ligation, hybridization, and cleaving, which yields other cut-and-paste operations [17], [45]. In addition, we plan to enhance our algorithm, within the framework of ADSM, by utilizing the artificial immune models in high-dimensional feature space to update the evolution operators [46], [47].

Acknowledgment

The authors would like to thank the Associate Editor and the three anonymous reviewers for their helpful comments and suggestions that improved this paper.

References

[1] A. Plaza, J. A. Benediktsson, J. W. Boardman, J. Brazile, L. Bruzzone, G. Camps-Valls, J. Chanussot, M. Fauvel, P. Gamba, A. Gualtieri, M. Marconcini, J. C. Tilton, and G. Trianni, "Recent advances in techniques for hyperspectral image processing," *Remote Sens. Environ.*, vol. 113, no. Supplement 1, pp. S110–S122, Sep. 2009.

[2] A. J. Brown, "Spectral curve fitting for automatic hyperspectral data analysis," *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 6, pp. 1601–1608, Jun. 2006.

[3] F. A. Kruse, A. B. Lefkoff, J. W. Boardman, K. B. Heidebrecht, A. T. Shapiro, P. J. Barloon, and A. F. H. Goetz, "The spectral image processing system (SIPS)—Interactive visualization and analysis of imaging spectrometer data," *Remote Sens. Environ.*, vol. 44, no. 2/3, pp. 145–163, May/Jun. 1993.

[4] C.-I Chang, *Hyperspectral Imaging: Spectral Detection and Classification*. New York: Plenum, 2003.

[5] J. Chen, X. Jia, W. Yang, and B. Matsushita, "Generalization of subpixel analysis for hyperspectral data with flexibility in spectral similarity measures," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 7, pp. 2165–2171, Jul. 2009.

[6] C.-I Chang, J. Wang, C.-C. Chang, and C. Lin, "Progressive coding for hyperspectral signature characterization," *Opt. Eng.*, vol. 45, no. 9, pp. 097 002-1–097 002-15, Sep. 2006.

[7] X. Jia and J. A. Richards, "Binary coding of imaging spectrometer data for fast spectral matching and classification," *Remote Sens. Environ.*, vol. 43, pp. 47–53, 1993.

[8] A. S. Mazer, M. Martin, M. Lee, and J. E. Solomon, "Image processing software for imaging spectrometry analysis," *Remote Sens. Environ.*, vol. 24, no. 1, pp. 201–210, Feb. 1988.

[9] X. Huan, X. Tong, C. Heipke, P. Lohmann, and U. Sorgel, "Object-based binary encoding algorithm an integration of hyperspectral data and DSM," in *Proc. Urban Remote Sens. Joint Event*, 2009, pp. 1–6.

[10] S. Qian, A. B. Hollinger, D. Williams, and D. Manak, "Fast three-dimensional data compression of hyperspectral imagery using vector quantization with spectral-feature-based binary coding," *Opt. Eng.*, vol. 35, no. 11, pp. 3242–3249, 1996.

[11] C.-I Chang, S. Chakravarty, H. Chen, and Y. Ouyang, "Spectral derivative feature coding for hyperspectral signature analysis," *Pattern Recognit.*, vol. 42, no. 3, pp. 395–408, Mar. 2009.

[12] C.-I Chang, S. Chakravarty, C.-S. Lo, and C. Lin, "Spectral feature probabilistic coding for hyperspectral signatures," *IEEE Sens. J.*, vol. 10, no. 3, pp. 395–409, Mar. 2010.

[13] L. M. Adleman, "Molecular computing of solutions to combinational problems," *Science*, vol. 266, no. 5187, pp. 1021–1024, Nov. 1994.

[14] L. M. Adleman, "Computing with DNA," *Sci. Amer.*, vol. 279, pp. 54–61, Aug. 1998.

[15] R. J. Lipton, "DNA solution of hard computational problems," *Sciences*, vol. 268, no. 5210, pp. 542–545, Apr. 1995.

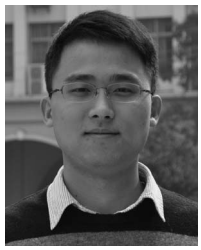
[16] C. C. Maley, "DNA computation: Theory, practice and prospects," *Evol. Comput.*, vol. 6, no. 3, pp. 201–229, Fall 1998.

[17] M. H. Garzon and R. J. Deaton, "Biomolecular computing and programming," *IEEE Trans. Evol. Comput.*, vol. 3, no. 3, pp. 236–250, Sep. 1999.

[18] S. Y. Shin, I. H. Lee, D. Kim, and B. T. Zhang, "Multi-objective evolutionary optimization of DNA sequences for reliable DNA computing," *IEEE Trans. Evol. Comput.*, vol. 9, no. 2, pp. 143–158, Apr. 2005.

[19] J. Watada and R. B. A. Barkar, "DNA computing and its applications," in *Proc. 8th Int. Conf. Intell. Syst. Des. Appl.*, Nov. 2008, pp. 288–294.

[20] Q. Liu, L. Wang, A. G. Frutos, A. E. Condon, R. M. Corn, and L. M. Smith, "DNA computing on surfaces," *Nature*, vol. 403, no. 6766, pp. 175–179, Jan. 2000.
[21] M. Darehmiraki and H. M. Nehi, "Molecular solution to the 0–1 knapsack problem based on DNA computing," *Appl. Math. Comput.*, vol. 187, no. 2, pp. 1033–1037, 2007.
[22] J. Y. Lee, S. Y. Shin, T. H. Park, and B.-T. Zhang, "Solving traveling salesman problems with DNA molecules encoding numerical value," *Biosystems*, vol. 78, no. 1–3, pp. 39–47, Dec. 2004.
[23] Z. Ibrahim, Y. Tsuboi, and O. Ono, "Hybridization-ligation versus parallel overlap assembly: An experimental comparison of initial pool generation for direct-proportional length-based DNA computing," *IEEE Trans. NanoBiosci.*, vol. 5, no. 2, pp. 103–109, Jun. 2006.
[24] R. A. B. Bakar, J. Watada, and W. Pedrycz, "DNA approach to solve clustering problem based on a mutual order," *Biosystems*, vol. 91, no. 1, pp. 1–12, Jan. 2008.
[25] S. A. Tsaftaris, A. K. Katsaggelos, T. N. Pappas, and E. T. Papoutsakis, "How can DNA computing be applied in digital signal processing?" *IEEE Signal Process. Mag.*, vol. 21, no. 6, pp. 57–61, Nov. 2004.
[26] S. A. Tsaftaris, A. K. Katsaggelos, T. N. Pappas, and E. T. Papoutsakis, "DNA computing from a signal processing viewpoint," *IEEE Signal Process. Mag.*, vol. 21, no. 5, pp. 100–106, Sep. 2004.
[27] S. A. Tsaftaris, A. K. Katsaggelos, T. N. Pappas, and E. T. Papoutsakis, "DNA-based matching of digital signals," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, 2004, vol. 5, pp. 581–584.
[28] C.-C. Tsai, H.-C. Huang, and S.-C. Lin, "FPGA-based parallel DNA algorithm for optimal configurations of an omnidirectional mobile service robot performing fire extinguishment," *IEEE Trans. Ind. Electron.*, vol. 58, no. 3, pp. 1016–1026, Mar. 2011.
[29] O. H. Kwon, K. Y. Wang, J. Y. Kim, J. Park, D. J. Chung, and C. H. Lee, "DNA inspired digital signal pattern matching algorithm," in *Proc. Frontiers Converg. Biosci. Inf. Technol.*, 2007, pp. 753–756.
[30] D. I. Liwen, "DNA computing," *Comput. Sci. Eng.*, vol. 4, no. 3, pp. 5–8, 2002.
[31] J. Elbaz, O. Lioubashevski, F. Wang, F. Remacle, R. D. Levine, and I. Willner, "DNA computing circuits using libraries of DNAzyme subunits," *Nat. Nanotechnol.*, vol. 5, no. 6, pp. 417–422, Jun. 2010.
[32] Y. Sohn and N. S. Rebello, "Supervised and unsupervised spectral angle classifiers," *Photogramm. Eng. Remote Sens.*, vol. 68, no. 12, pp. 1271–1280, Dec. 2002.
[33] S. South, J. Qi, and D. P. Lusch, "Optimal classification methods for mapping agricultural tillage practices," *Remote Sens. Environ.*, vol. 91, no. 1, pp. 90–97, May 2004.
[34] F. Zhouyu, A. Robles-Kelly, T. Caelli, and R. T. Tan, "On automatic absorption detection for imaging spectroscopy: A comparative study," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 11, pp. 3827–3844, Nov. 2007.
[35] Y. Li and J. Lei, "A feasible solution to the beam angle optimization problem in radiotherapy planning with a DNA-based genetic algorithm," *IEEE Trans. Biomed. Eng.*, vol. 57, no. 3, pp. 499–508, Mar. 2010.
[36] X. Y. Liu and H. Q. Wang, "A discretization algorithm based on a heterogeneity criterion," *IEEE Trans. Knowl. Data Eng.*, vol. 17, no. 9, pp. 1166–1173, Sep. 2005.
[37] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan, "A fast and elitist multi-objective genetic algorithm: NSGA-II," *IEEE Trans. Evol. Comput.*, vol. 6, no. 2, pp. 182–197, Apr. 2002.
[38] C. Y. Mao and Y. H. Hu, "Analysis of convergence properties of a stochastic evolution algorithm," *IEEE Trans. Comput.-Aided Des. Integr. Syst.*, vol. 15, no. 7, pp. 826–831, Jul. 1996.
[39] G. M. Foody, "Status of land cover classification accuracy assessment," *Remote Sens. Environ.*, vol. 80, no. 1, pp. 185–201, 2002.
[40] X. Liu, W. Xia, B. Wang, and L. Zhang, "An approach based on constrained nonnegative matrix factorization to unmix hyperspectral data," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 2, pp. 757–772, Feb. 2011.
[41] Hypercube. [Online]. Available: http://www.tec.army.mil/hypercube
[42] B. D. Bue, E. Merényi, and B. Csathó, "Automated labeling of materials in hyperspectral imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 11, pp. 4059–4070, Nov. 2010.
[43] Z. Fu and A. Robles-Kelly, "Discriminate absorption-feature learning for material classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 5, pp. 1536–1556, May 2011.
[44] M. Pal and G. M. Foody, "Feature selection for classification of hyperspectral data by SVM," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 5, pp. 2297–2307, May 2010.
[45] S. Shin, I. Lee, D. Kim, and B. Zhang, "Multiobjective evolutionary optimization of DNA sequences for reliable DNA computing," *IEEE Trans. Evol. Comput.*, vol. 9, no. 2, pp. 143–158, Apr. 2005.
[46] Y. Zhong, L. Zhang, B. Huang, and P. Li, "An unsupervised artificial immune classifier for multi/hyperspectral remote sensing imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 2, pp. 420–431, Feb. 2006.
[47] Y. Zhong, L. Zhang, J. Gong, and P. Li, "A supervised artificial immune classifier for remote-sensing imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 12, pp. 3957–3966, Dec. 2007.



Hongzan Jiao received the B.S. degree in land resource management from Wuhan University, Wuhan, China, in 2008, where he is currently working toward the Ph.D. degree in photogrammetry and remote sensing in the State Key Laboratory of Information Engineering in Surveying, Mapping, and Remote Sensing at Wuhan University, Wuhan.

His research interests include multi- and hyperspectral data analysis, artificial intelligence, and pattern recognition in remote sensing images.

Yanfei Zhong (M'11) received the B.S. degree in information engineering and the Ph.D. degree in photogrammetry and remote sensing from Wuhan University, Wuhan, China, in 2002 and 2007, respectively.

He has been with the State Key Laboratory of Information Engineering in Surveying, Mapping, and Remote Sensing, Wuhan University since 2007 and is currently a Professor. He has published more than ten peer-reviewed articles in international journals such as IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING and *International Journal of Remote Sensing*. His research interests include multi- and hyperspectral remote sensing image processing, artificial intelligence, and pattern recognition.

Dr. Zhong was the recipient of the National Excellent Doctoral Dissertation Award of China (2009) and New Century Excellent Talents in University of China (2009). He was a Referee of IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS PART B, IEEE JOURNAL OF SELECTED TOPICS IN APPLIED EARTH OBSERVATIONS AND REMOTE SENSING, and *Pattern Recognition*.

Liangpei Zhang (M'06–SM'08) received the B.S. degree in physics from Hunan Normal University, ChangSha, China, in 1982, the M.S. degree in optics from the Xi'an Institute of Optics and Precision Mechanics of Chinese Academy of Sciences, Xi'an, China, in 1988, and the Ph.D. degree in photogrammetry and remote sensing from Wuhan University, Wuhan, China, in 1998.

He is currently with the State Key Laboratory of Information Engineering in Surveying, Mapping, and Remote Sensing, Wuhan University, as the Head of the Remote Sensing Division. He is also a "Chang-Jiang Scholar" Chair Professor appointed by the Ministry of Education, China. He has more than 230 research papers and five patents. He is currently a Principal Scientist for the China State Key Basic Research Project (2011–2016) appointed by the Ministry of National Science and Technology of China to lead the remote sensing program in China. His research interests include hyperspectral remote sensing, high-resolution remote sensing, image processing, and artificial intelligence.

Dr. Zhang regularly serves as a Co-Chair of the series SPIE Conferences on Multispectral Image Processing and Pattern Recognition, Conference on Asia Remote Sensing, and many other conferences. He edits several conference proceedings, issues, and the geoinformatics symposiums. He also serves as an Associate Editor of the *International Journal of Ambient Computing and Intelligence*, *International Journal of Image and Graphics*, *International Journal of Digital Multimedia Broadcasting*, *Journal of Geospatial Information Science*, and the *Journal of Remote Sensing*. He is Fellow of IEE, Executive Member (Board of Governor) of the China National Committee of International Geosphere-Biosphere Programme, Executive Member for the China Society of Image and Graphics, and others.