Abstract—In this letter, a new endmember extraction algorithm based on adaptive differential evolution (DE) (ADEE) is proposed for hyperspectral remote sensing imagery. In the proposed algorithm, the endmember extraction is transformed into a combinatorial optimization problem through constructing the objective function by minimizing the root mean square error between the original image and its remixed image. DE is utilized to search for the optimal endmember combination in the feasible solution space by the DE operators, such as crossover and mutation, which have the advantage of high efficiency, rapid convergence, and strong capability for global search. In addition, to avoid the problem of parameter selection, an adaptive strategy without user-defined parameters is utilized to improve the classical DE algorithm. The proposed method was tested and evaluated using both simulated and real hyperspectral remote sensing images, and the experimental results show that ADEE can obtain a higher extraction precision than the traditional endmember extraction algorithms.

Index Terms—Differential evolution (DE), endmember extraction, hyperspectral remote sensing, spectral unmixing.

I. INTRODUCTION

HYPERSPECTRAL remote sensors have the capability of capturing hundreds of contiguous spectral bands, which can be used to improve the ability to recognize different landcover classes. However, there are often many mixed pixels in remote sensing imagery, due to the limited spatial resolution, and the mixed pixel problem not only has an effect on distinguishing ground objects but is also a barrier to the quantitative development of remote sensing technology. The spectral unmixing technique is an effective method of solving the problem of mixed pixels by endmember extraction and the estimation of the fraction based on an endmember spectral mixing model.

In spectral unmixing, endmember extraction is the key step. Many endmember extraction algorithms based on a linear spectral model have been proposed, such as the pixel purity index (PPI) [1], [2], N-FINDR [3], vertex component analysis (VCA) [4], the simplex growing algorithm [5], and the sequential maximum angle convex cone [6]. The criteria of these traditional endmember extraction algorithms are mostly based on the theory of a high-dimensional simplex structure and a linear spectral mixing model (LSMM) [7], in which they consider the endmember as the vertex of the simplex. However, the precision of the endmember extraction algorithm will be lowered when the real data cannot meet these assumptions (e.g., a simplex structure). In order to improve the accuracy, some new algorithms based on combinatorial optimization have been proposed, such as discrete particle swarm optimization (D-PSO) [8] and ant colony optimization [9]. These algorithms have improved the precision of endmember extraction. However, the parameters of these algorithms are often kept fixed throughout the optimization process and may significantly influence the optimization performance. The algorithms need to be run multiple times with different settings of the parameters to obtain the appropriate values. The time taken to find these parameters is often too long and is unacceptable for the hyperspectral remote sensing image process.

In this letter, a new endmember extraction algorithm based on adaptive differential evolution (DE) (ADEE) is proposed. The DE algorithm is a novel powerful algorithm for global optimization and has the properties of rapid convergence and simple principle. It has been widely used in many real applications; for example, it has been applied to numerical optimization [10], image processing [11], pattern recognition [12], and a lot of other mathematical and engineering fields [13]. In the proposed ADEE, the endmember extraction problem is transformed into a combinatorial optimization. The results of the spectral unmixing can be evaluated in terms of the quality of the reconstruction of the original data set using the extracted endmembers, the estimated fractional abundances, and the linear mixture model [7], [14]. The root mean square error (RMSE) between the original image and its remixed image is selected as the objective function [8], [9], and the candidate endmember solutions are set as the feasible space. The DE algorithm is used to search for an optimal combination of endmembers that minimizes the RMSE in the feasible space. In addition, an adaptive mechanism is proposed to adaptively select the appropriate value of the parameters during the process of the evolution. In order to demonstrate the performance of the proposed algorithm, ADEE was compared with the previously mentioned endmember extraction algorithms, N-FINDR, VCA, and D-PSO, using both synthetic and real hyperspectral images, and the experimental results demonstrate that ADEE performs better than the other algorithms.

The rest of this letter is organized as follows. Section II describes the background. Section III provides the detailed process of the proposed endmember extraction algorithm, ADEE.
The experimental results of the five algorithms are shown in Section IV. Finally, Section V provides the conclusions.

II. BACKGROUND

A. Endmember Extraction Problem

The LSMM, as the traditional spectral unmixing method, assumes that a linear relationship exists between the fractional abundances of the substances comprising the area being imaged and the spectra in the reflected radiation. Considering hyper-abundances of the substances comprising the area being imaged is usually used to evaluate the extraction spectral data of and the spectra in the reflected radiation. Considering hyper-abundances of the substances comprising the area being imaged is higher to a certain extent [8]. Therefore, the endmember extraction precision can be estimated easily by the endmembers and the original image using the fully constrained least squares minimization optimization problem in the continuous feature space as shown in

$$x = As + \varepsilon$$

(1)

where \(A = (a_1, a_2, \ldots, a_P) \in R^{L \times P}\) indicates the endmember spectral matrix. The last term \(\varepsilon\) is noise. \(P\) is the number of endmembers. \(S = (s_1, s_2, \ldots, s_P)^T \in R^{P \times 1}\) represents the abundance vector corresponding to the proportions of the endmembers, which satisfies two physical constraints referred to as the abundance nonnegative constraint (ANC), \(s_{ij} \geq 0, i = 1, 2, \ldots, N; j = 1, 2, \ldots, P\), and the abundance sum-to-one constraint (ASC), \(\sum_{j=1}^{P} s_j = 1\).

The abundance can be estimated easily by the endmembers and the original image using the fully constrained least squares method (FCLS), the objective of which is to find the abundance vector by minimizing the least squares error \(\|x - As\|^2\) when taking the ANC and ASC into consideration. The estimated pixel is obtained by the abundance and the endmembers according to the following:

$$\hat{x}_i = \sum_{j=1}^{P} s_{ij}a_j, \quad i = 1, 2, \ldots, N.$$  

(2)

The RMSE between pixel \(x_i\) in the original image and \(\hat{x}_i\) in the remixed image is usually used to evaluate the extraction results and is defined as follows:

$$RMSE(x_i, \hat{x}_i) = \sqrt{\frac{1}{L} \sum_{k=1}^{L} (x_i(k) - \hat{x}_i(k))^2}.$$  

(3)

Under the assumption that the number of endmembers is known, a lower RMSE value indicates that pixel \(x_i\) in the original image and pixel \(\hat{x}_i\) in the remixed image are more similar. This indicates that the precision of the endmember extraction is higher to a certain extent [8]. Therefore, the endmember extraction problem can be regarded as an optimization problem by minimizing the RMSE as follows:

$$\text{minimize } f(x_i, \hat{x}_i) = \frac{1}{N} \sum_{i=1}^{N} RMSE(x_i, \hat{x}_i).$$  

(4)

B. DE Algorithm

The DE was proposed by Storn and Price [16] to solve the minimization optimization problem in the continuous feature space as shown in

$$\min f(X_1, \ldots, X_j, \ldots, X_D) \text{ s.t. } X_j^{(L)} \leq X_j \leq X_j^{(U)},$$

\(j = 1, 2, \ldots, D\)  

(5)

where \(D\) indicates the number of the problem’s dimension and \(X_j^{(L)}\) and \(X_j^{(U)}\) indicate the minimum and maximum of the \(j\)th element of the individual vector \(X_j\), respectively. DE adopts a floating-point encoding scheme, and successive populations are generated by adding the weighted difference of two randomly selected vectors to a third randomly selected vector by the evolution operators of mutation, crossover, and selection. Nonparametric adaptation has also been considered [17]–[20] for the original DE algorithm. The process of DE can be described as the following four steps.

Step 1—Initialization: The initial population \(X\) is generated randomly. The size of the population is \(NP\).

Step 2—Mutation: DE recombines the population by adding a scaling factor \(F\) and a randomly sampled vector difference to a third vector to produce the mutant vector \(V_t^i\).

Step 3—Crossover: The trial vector \(U_t^i\) is generated by mixing the target vector \(X_t^i\) and the mutant vector \(V_t^i\) by the crossover probability \(CR\).

Step 4—Selection: DE selects individuals according to the objective function’s value of the population vector \(X_t^i\) and its corresponding trial vector \(U_t^i\). The better vector whose objective function’s value is smaller will survive to be a member of the next generation.

III. ENDMEMBER EXTRACTION BASED ON ADAPTIVE DE

ADEE is proposed in this letter. In ADEE, the endmember extraction problem is described as a minimization optimization problem of the RMSE. The original DE algorithm is highly dependent on the parameter settings, and it can take a huge amount of time to find the most suitable control parameters for a specific problem. To avoid this problem, in ADEE, an adaptive mechanism is used to automatically adapt the parameter settings during evolution. In addition, to decrease the searching feature space and improve the computational efficiency, the PPI algorithm [1], [2] is used to preprocess the original image to obtain the candidate endmembers including \(M\) pixels. The objective of the proposed algorithm is to find an optimal combination containing \(\hat{P}\) endmembers in the searching feature space by optimizing the problem (4).

A. Initialization of the Population

The first step of ADEE is to randomly initialize each individual \(X_t^i\) in the population \(X_t^i = \{X_t^1, X_t^2, \ldots, X_t^P\}\), where \(X_t^i = \{x_{ij}^1, \ldots, x_{ij}^t, \ldots, x_{ij}^P\}\) represents a possible endmember set with \(P\) endmembers in the \(t\)th generation, \(NP\) is the size of the population, and \(x_{ij}^t\) indicates the \(x_{ij}^t\)th endmember of this candidate endmember set. The feasible space of the whole endmember set can be encoded as an integer in the range of \([1, M]\), and each integer \(X_t^i\) in each individual \(X_t^i\) represents the index of the candidate endmembers and is initialized as follows:

$$x_{ij}^t = \text{rand}(1, M) i = 1, \ldots, NP, \quad j = 1, \ldots, M$$  

(6)

where the function \(\text{rand}(1, M)\) randomly produces an integer in the range of \([1, M]\) and \(M\) is the number of candidate endmembers by PPI. Fig. 1 represents an example of the initialization.

To adaptively obtain the parameters for each individual, the control parameters \(F\) and \(CR\) are encoded to an individual, and
each individual has its parameters and the value changed during the process of evolution, as shown in Fig. 1.

**B. Calculating the Objective Function**

After the initialization, an endmember set can be obtained using the element value of individual \(X_i^t\) as the index in the candidate endmember. The endmember set is applied with spectral unmixing using FCLS to obtain the abundance images of each endmember. The value of the objective function for each individual can be obtained using (3) and (4).

**C. Adaptive DE Operators**

The individuals, as candidate endmember extraction solutions, will be evolved by the DE operators of mutation, crossover, and selection. However, traditional DE needs to find suitable control parameters for different real problems [15], [21]. In this letter, an adaptive mechanism for \(F\) and \(CR\) is proposed as follows.

**Step 1—Adaptive mutation**: Each individual \(X_i^t\) is mutated with mutation rate \(p_m\) to obtain its associated mutant individual \(V_i^t\) by the strategy DE/rand/1/bin [14] as follows:

\[
V_i^t = X_i^t + F_i^t \cdot (X_{r_1}^t - X_{r_2}^t), \quad i = 1, \ldots, NP
\]

where \(F_i^t\) indicates the mutation scaling factor of the \(i\)th individual and \(r_1, r_2, r_3 \in \{1, 2, \ldots, NP\}\) are three mutually exclusive integers randomly chosen from the range \([1, NP]\), \(r_1 \neq r_2 \neq r_3 \neq i\). To adaptively determine the scaling factor \(F\), the mutation rate \(p_m\) is calculated as follows:

\[
p_m = \frac{f(X_i^t) - \min(f(X_i^t))}{\max(f(X_i^t)) - \min(f(X_i^t))}
\]

**D. Stopping Condition**

If the stopping condition is not met, go to Step C. Otherwise, the best endmember extraction solution is output. There are three stopping conditions: the value of the objective function reaches the threshold, the result converges to a fixed value, or the iteration number reaches the maximum. The flowchart of ADEE is shown in Fig. 2.

**IV. EXPERIMENTS AND ANALYSIS**

In this section, both simulated and real hyperspectral images are utilized to test the performance of the ADEE algorithm when compared with the traditional algorithms of N-FINDR, VCA, D-PSO, nonnegative matrix factorization (NMF) with a positive constraint and sum-to-one constraint [22], and endmember extraction based on traditional DE (DEE). In D-PSO, the random selection probability \(p\) is set to 0.1 and the number of particles is set to 10. In DEE, the values of \(F\) and \(CR\) are set to 0.9 and 0.2, respectively, by experience. In ADEE, the mean value of five independent runs is used as the result. The parameters in all the other methods implemented in this letter follow their original work.

**A. Experiment 1—Simulated Hyperspectral Remote Sensing Imagery**

Two simulated images were designed based on the spectra of the U.S. Geological Survey (USGS) ground-truth mineral spectra. The first image \((100 \times 100)\) (Simu-5) with 224 bands was simulated based on five spectra: alunite, buddingtonite,
calcite, kaolinite, and muscovite. The second image (160 × 160) (Simu-10) with 224 bands was simulated based on ten spectra: alunite, buddingtonite, calcite, illite, jarosite, kaolinite, muscovite, nontronite, halloysite, and pyrophyllite. The signal-to-noise ratio was set to 30. Fig. 3 shows the two simulated images, the corresponding reference spectra, and the real abundance images. The simulated data generated follow a linear distribution. The RMSE between the original image and its remixed image [RMSE(X)] [23], the RMSE between the reference abundances and the FCLS-estimated abundances associated with the extracted endmember set [RMSE(S)], spectral angle mapping (SAM), and a combination of SAM and spectral feature fitting (SFF), namely, SAM-SFF [9], are used to evaluate the precision of the algorithms. To test the performance with all the endmembers, the mean values of SAM and SAM-SFF, denoted as SAM (M) and SAM-SFF (M), respectively, are utilized to evaluate the different algorithms.

As shown in Table I, the performance of D-PSO is better than those of N-FINDR and VCA. Compared with D-PSO, DEE and ADEE can obtain better endmember extraction results than D-PSO, and the RMSE(X) of DEE and ADEE are smaller than for the other three algorithms. The RMSE(S) values of DEE and ADEE are smaller than those for the other algorithms with the data of Simu-5, but they did not perform as well as N-FINDR and VCA. In addition, the values of SAM and SAM-SFF obtained by DEE and ADEE are much closer to 1, which means that the endmembers extracted by DEE and ADEE more accurately reflect the “true” spectral signatures. Comparing ADEE with DEE, ADEE can obtain better or similar endmember extraction, without the user-defined process, than DEE. In the experiment, while the performance of NMF may not be as good as DEE and ADEE, it can simultaneously obtain the endmember signatures and the corresponding abundances.

### Table I

<table>
<thead>
<tr>
<th>No.</th>
<th>Metric</th>
<th>N-FINDR</th>
<th>VCA</th>
<th>D-PSO</th>
<th>NMF</th>
<th>DEE</th>
<th>ADEE</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>RMSE(X)</td>
<td>0.0129</td>
<td>0.1627</td>
<td>0.0018</td>
<td>0.0090</td>
<td>0.0105</td>
<td>0.0116</td>
</tr>
<tr>
<td></td>
<td>RMSE(S)</td>
<td>0.1627</td>
<td>0.1593</td>
<td>0.0214</td>
<td>0.1202</td>
<td>0.0624</td>
<td>0.0233</td>
</tr>
<tr>
<td></td>
<td>SAM (M)</td>
<td>0.7870</td>
<td>0.7874</td>
<td>0.9828</td>
<td>0.8840</td>
<td>0.9826</td>
<td>0.9830</td>
</tr>
<tr>
<td></td>
<td>SAM-SFF(M)</td>
<td>3.5620</td>
<td>0.7226</td>
<td>0.8802</td>
<td>0.7024</td>
<td>0.8814</td>
<td>0.8820</td>
</tr>
<tr>
<td>10</td>
<td>RMSE(X)</td>
<td>0.0370</td>
<td>0.0371</td>
<td>0.0360</td>
<td>0.0276</td>
<td>0.0352</td>
<td>0.0351</td>
</tr>
<tr>
<td></td>
<td>RMSE(S)</td>
<td>0.0886</td>
<td>0.0608</td>
<td>0.1847</td>
<td>0.1831</td>
<td>0.1136</td>
<td>0.0862</td>
</tr>
<tr>
<td></td>
<td>SAM (M)</td>
<td>0.9214</td>
<td>0.9125</td>
<td>0.9128</td>
<td>0.8153</td>
<td>0.9429</td>
<td>0.9385</td>
</tr>
<tr>
<td></td>
<td>SAM-SFF(M)</td>
<td>0.5646</td>
<td>0.5456</td>
<td>0.5734</td>
<td>0.5866</td>
<td>0.6131</td>
<td>0.6119</td>
</tr>
</tbody>
</table>

### B. Experiment 2—Real Aviris Image

In Experiment 2, the AVIRIS Cuprite image (400 × 350) with 50 bands from 1.9908 to 2.4790 μm is used, as shown in Fig. 4. In this area, there are many complicated mineral types, including alunite, kaolinite, chalcedony, muscovite, montmorillonite, jarosite, and calcite. Furthermore, there is also some chert, illite, and buddingtonite [24]. Because accurate estimations of the mineral types in the Cuprite mining site are difficult to confirm, the different numbers of endmembers are set to 5, 10, 15, and 20 in this experiment, respectively. In the experiment, the reference endmember spectra for the AVIRIS data are chosen from the USGS Digital Spectral Library in ENVI.

The RMSE(X), SAM, and SAM-SFF are used to evaluate the precision of the six endmember extraction algorithms. As shown in Table II, N-FINDR and VCA have better SAM (M) and SAM-SFF (M) values than DEE and ADEE when the number of endmembers is set to 5, 10, 15, and 20; however, the difference is slight. In addition, the RMSE(X) values of DEE and ADEE are lower than those of the other algorithms for all conditions, and when the number of endmembers is set 10, they can obtain better values of SAM (M) and SAM-SFF (M). Compared with DE, ADEE not only avoids the problem of parameter choice but also obtains a better result with lower RMSE.

### Table II

<table>
<thead>
<tr>
<th>No.</th>
<th>Metric</th>
<th>N-FINDR</th>
<th>VCA</th>
<th>D-PSO</th>
<th>NMF</th>
<th>DEE</th>
<th>ADEE</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>RMSE(X)</td>
<td>5.5271</td>
<td>5.8352</td>
<td>4.5321</td>
<td>4.8537</td>
<td>4.4682</td>
<td>4.2970</td>
</tr>
<tr>
<td></td>
<td>SAM (M)</td>
<td>0.9382</td>
<td>0.9296</td>
<td>0.9215</td>
<td>0.8054</td>
<td>0.9196</td>
<td>0.9338</td>
</tr>
<tr>
<td></td>
<td>SAM-SFF(M)</td>
<td>8.8662</td>
<td>0.8710</td>
<td>0.8544</td>
<td>0.7722</td>
<td>0.9660</td>
<td>0.8552</td>
</tr>
<tr>
<td>10</td>
<td>RMSE(X)</td>
<td>3.8407</td>
<td>3.9260</td>
<td>3.2797</td>
<td>3.8717</td>
<td>3.2525</td>
<td>3.1442</td>
</tr>
<tr>
<td></td>
<td>SAM (M)</td>
<td>0.8785</td>
<td>0.8675</td>
<td>0.8908</td>
<td>0.8069</td>
<td>0.8848</td>
<td>0.8906</td>
</tr>
<tr>
<td></td>
<td>SAM-SFF(M)</td>
<td>0.7970</td>
<td>0.7828</td>
<td>0.8077</td>
<td>0.7845</td>
<td>0.8302</td>
<td>0.8431</td>
</tr>
<tr>
<td>15</td>
<td>RMSE(X)</td>
<td>3.6568</td>
<td>3.8647</td>
<td>2.9444</td>
<td>3.3271</td>
<td>2.8875</td>
<td>2.8606</td>
</tr>
<tr>
<td></td>
<td>SAM (M)</td>
<td>0.9332</td>
<td>0.9302</td>
<td>0.9251</td>
<td>0.7723</td>
<td>0.9120</td>
<td>0.9230</td>
</tr>
<tr>
<td></td>
<td>SAM-SFF(M)</td>
<td>0.8850</td>
<td>0.8699</td>
<td>0.8666</td>
<td>0.6995</td>
<td>0.8234</td>
<td>0.8550</td>
</tr>
<tr>
<td>20</td>
<td>RMSE(X)</td>
<td>3.0101</td>
<td>3.2339</td>
<td>2.7488</td>
<td>3.1039</td>
<td>2.7771</td>
<td>2.6631</td>
</tr>
<tr>
<td></td>
<td>SAM (M)</td>
<td>0.9237</td>
<td>0.9280</td>
<td>0.9112</td>
<td>0.7690</td>
<td>0.9191</td>
<td>0.9151</td>
</tr>
<tr>
<td></td>
<td>SAM-SFF(M)</td>
<td>0.8586</td>
<td>0.8907</td>
<td>0.8630</td>
<td>0.7035</td>
<td>0.8570</td>
<td>0.8536</td>
</tr>
</tbody>
</table>

### C. Parameter Analysis

In this section, to test the validity of the self-adaptive strategy in ADEE, the sensitivity of the two control parameters, i.e., $F$ and $CR$, is analyzed. The values of $CR$ are set to 0.2, 0.4, 0.6, 0.8, and 1.0, respectively, and the values of $F$ are set to 0.1, 0.3, 0.5, 0.7, and 0.9, respectively. The experimental results are shown in Fig. 5, where the red asterisk indicates the result of ADEE from the evolution process, and the three subfigures represent the experiments for the Simu-5 simulated data, the
Simu-10 simulated data, and the Cuprite data when the number of endmembers is set to 10. Fig. 5 demonstrates that ADEE can obtain better or similar results, without the user-defined parameters, than DEE, and the best result of DEE is obtained when $F = 0.9$ and $CR = 0.8$.

V. CONCLUSION

In this letter, ADEE has been proposed. This approach no longer depends on the theory of convex geometry, and a self-adaptive mechanism is introduced to adaptively change the DE parameters, which makes it possible to obtain the best values without a user-defined process. The experimental results using both simulated and real hyperspectral remote sensing images show that ADEE can adaptively obtain better endmember extraction results with lower RMSE between the original image and the remixed image, a reduced RMSE between the reference and the estimated abundance, and a higher SAM and SAM-SFF accuracy. In our future work, we will consider how to automatically determine the number of endmembers, and we will also compare the endmember extraction algorithms based on evolutionary and swarm intelligence algorithms, using the nonparametric statistical tests [20].

ACKNOWLEDGMENT

The authors would like to thank the Editor, Associate Editor, and anonymous reviewers for their helpful comments and suggestions.

REFERENCES