characteristics of plants and soils (Fang et al., 2003, Kooistra et al., 2003 and Cogdill and Rippke, 2004); (iii) selecting a spectral subset of hyperspectral data for image classification (Yu et al., 2002)

Nevertheless, class information that has been used in the aforementioned studies ([Lofy and Sklansky, 2001], [Kavzoglu and Mather, 2002], [Yu et al., 2002] and [Ulfarsson et al., 2003]) for testing the performance of the GA-based band selector is broad (i.e. USGS level I or II (Anderson et al., 1976)). This means that each class possesses distinct spectral characteristics from one another, and it is relatively easy for GA to find spectral bands that maintain high spectral separability between classes. None of the aforementioned studies has tested the band selector on class information that possesses very similar spectral characteristics (e.g. species-level data). A question therefore remains if GA can deal with such complexity. The key hypothesis of this research is that the GA-based band selector can be used for selecting a meaningful subset of spectral bands that maintains spectral separability between species classes. As a result, building and testing a GA-based spectral band selection tool is the main objective of this study. The testing data in use are very high-dimensional, spectrometer records that comprise 2151 bands of leaf spectra of 16 tropical mangrove species.

2. Species-level data and methods

2.1. Species-level hyperspectral data

2.1.1. Mangrove leaf preparation

Top-level canopies of 16 tropical mangrove species (Table 1) were collected using a line-transect method in the natural mangrove forest of Ao Sawi (Sawi Bay), the province of Chumporn, the south of Thailand (10° 15′N, 99° 7′E) on February, 6, 2001. This line-transect method enabled us to collect the mangrove canopies from pioneer, intermediate, and landward zones. The canopies were only sampled from fully-grown trees (i.e. > 2.5 m tall). During the sampling campaign, species identification was carried out by the staffs of Royal Thai Forestry Department; taxonomy follows Tomlinson (1994) and Teeratanatorn (2000). At the laboratory, the leaves were then picked off the canopies for the spectral measurement.

Table 1

Thirty spectra of mangrove leaves were collected per mangrove species, using a 2151-band spectroradiometer

Mangrove species	Species code	Number of spectra
Avicennia alba	1	30
Acrostichum aureum	2	30
Bruguiera cylindrica	3	30
Bruguiera gymnorrhiza	4	30
Bruguiera parviflora	5	30
Ceriops tagal	6	30
Excoecaria agallocha	7	30
Heritiera littoralis	8	30
Lumnitzera littorea	9	30
Lumnitzera racemosa	10	30
Nypa fruticans	11	30
Pluchea indica	12	30
Rhizophora apiculata	13	30
Rhizophora mucronata	14	30
Sonneratia ovata	15	30
Xylocarpus granatum	16	30

2.1.2. Leaf spectral measurements

The leaves were randomly shuffled and separated evenly into 30 piles per mangrove species. Each pile of leaves (top side up) was placed on top of a black metal plate painted with ultra-flat black paint until the background metal plate could not be seen. Next, the spectral response of each leaf plate was recorded 20 times. Each plate was rotated 90° horizontally after every fifth record to compensate for the bi-directional reflectance distribution function (BRDF). Then, the mean of the 20 records was calculated to construct a radiance curve. Finally, the radiance was converted to a reflectance curve by using a reference panel as well as the correction of the spectrometer internal current (dark current). The steps above were followed for all other leaf plates. As a result, we have 30 reflectance curves per each mangrove species (Table 1). The spectral mean of each species is illustrated in Fig. 1.
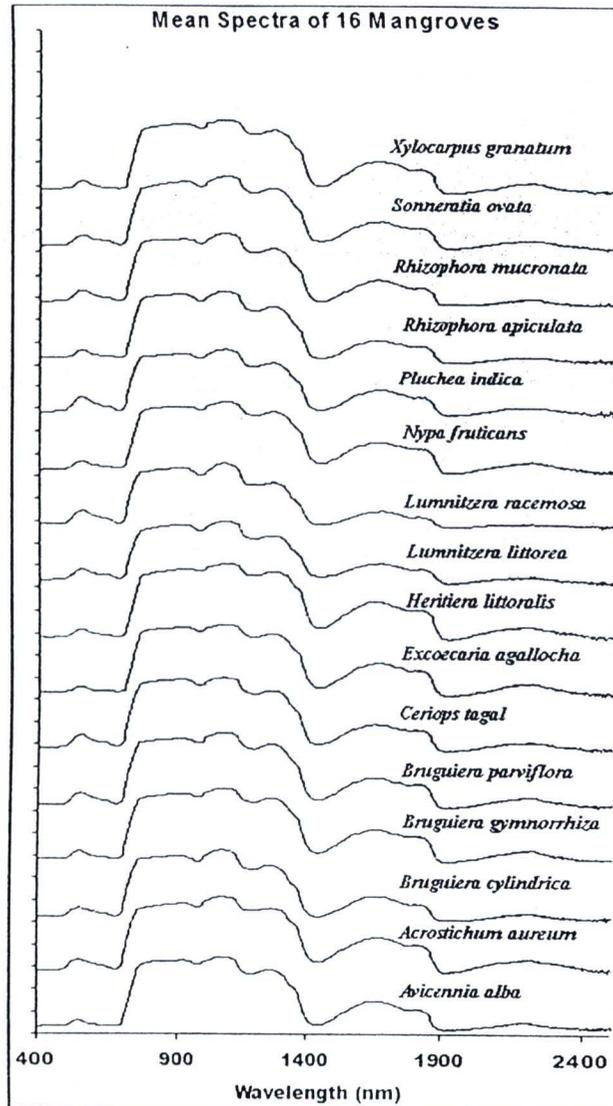
Fig. 1. The plot represents mean leaf spectra of 16 mangrove species stacked on top of one another. Each reflectance value tick mark along the y-axis is 0.2 reflectance value apart.

The spectral measurement was conducted under laboratory conditions by using a spectroradiometer (FieldSpec Pro FR, Analytical Spectral Device, Inc.). This spectroradiometer was equipped with three spectrometers (i.e. VNIR, SWIR1, and SWIR2), covering 350 nm to 2500 nm, with sampling intervals of 1.4 nm between 350 nm and 1000 nm, and 2 nm between 1000 nm and 2500 nm. The spectral resolution of the spectrometers was 3 nm for the wavelength interval 350 nm to 1000 nm, and 10 nm for the wavelength interval 1000 nm to 2500 nm. The sensor, equipped with a field of view of 25°, was mounted on a tripod and positioned 0.5 m above the leaf plate at the nadir position. A halogen lamp fixed on the tripod at the same position as the sensor of the spectrometer

was used to illuminate the sample plate. The room was conditioned to be dark with 25 °C in order to avoid unwanted external energy sources.

The reader may note that leaf stacking experiments could cause some degrees of additional uncertainty to the resultant spectra. However, we did not try to remove such an effect from our spectral data. We included them in the measured signal and assumed that they are signal noise (i.e., Measured Mangrove Signal=Mean Mangrove Signal+Mangrove Spectral Variations+Signal Noise). Moreover, we have no intention to simulate any effect of background spectra (e.g., background spectra of understory plants and soils) in our measurement. We wanted to measure "pure" mangrove spectra from the leaf plate and use them for testing the proposed algorithm.

Statistical separability of the species-level data used in this study is also plotted in Fig. 2 (after Vaiphasa et al., 2005) so as to indicate where, along the whole wavelength region, the mangroves are likely to be spectrally separable (i.e. the locations where the black line (p-value trace) are below the lower statistical threshold (α=0.01)). Full details on this statistical visualization and its limitations as well as other aspect of statistical studies can be found in the previous work (Vaiphasa et al., 2005; Vaiphasa, 2006).
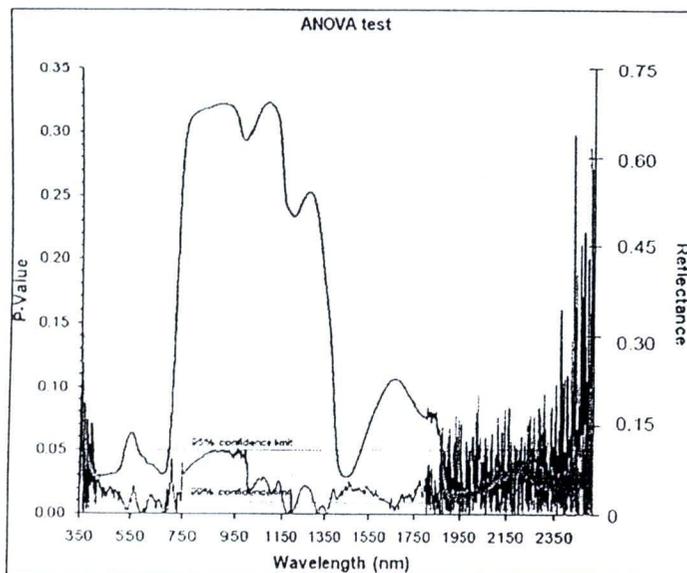

Fig. 2. After Vaiphasa et al. (2005), the plot shows p-values of the ANOVA test (black line) plotting against a laboratory reflectance of Rhizophora apiculata (gray line). The p-values indicate where, along the whole wavelength region, the mangroves are likely to be spectrally separable.

2.2. Genetic search algorithms (GA)

The theory of GA is first introduced by Holland (1975). The elaboration of its practical side including a basic computer source code can be found in Goldberg (1989). In this report, step-by-step guidelines of Goldberg (1989) are strictly followed (see Fig. 3). As a result, there is no attempt to make an exhaustive description of GA, but three major connections between the concept of GA and remote sensing applications are emphasized (i.e. gene encoding scheme, reproduction mechanism, and fitness criterion). Additionally, the code of GA used in this study has been developed in IDL language at the International Institute for Geo-Information Science and Earth Observation (ITC) (Vaiphasa, 2003).
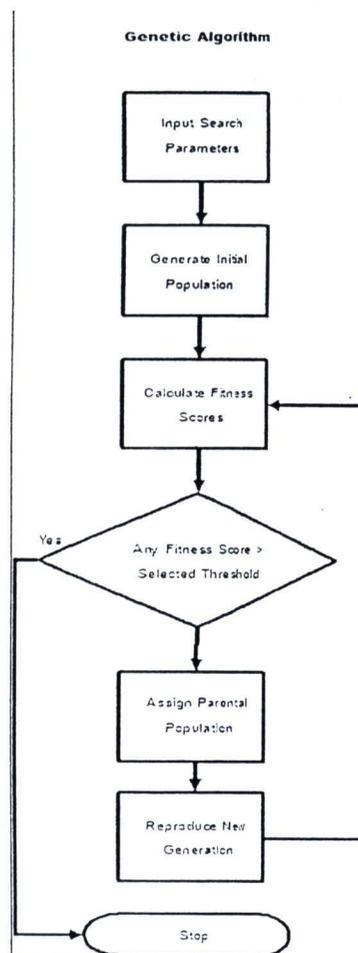
Genetic Algorithm

Input Search
Parameters

↓

Generate Initial
Population

↓

Calculate Fitness
Scores

↓

Any Fitness Score >
Selected Threshold

Assign Parental
Population

↓

Reproduce New
Generation

Stop

Fig. 3. A flowchart showing the process of the genetic algorithm used in this study.

2.2.1. Gene encoding

The gene encoding scheme in use is a direct method instead of binary encoding that is popularly used in related studies (Siedlecki and Sklansky, 1989; Kavzoglu and Mather, 2002; Yu et al., 2002). The key reasons for choosing direct encoding are that it is transparent for tracking the process of

evolution as well as straightforward for reproducing the population (i.e. crossover and mutation) (Vaiphasa, 2003).

Fig. 4a illustrates an example of two chromosomes with chromosome size (number of gene inside the chromosome) equal to 6. Following the direct encoding scheme, the 1st chromosome comprises 6 different genes flagged by the letter A to F, and the 2nd one comprises gene G to L. Each gene can be assigned with a band label. For example, the 1st chromosome, {A, B, C, D, E, F}, in Fig. 4a can be assigned with an array of band names, {Band 2, Band 8, Band 37, Band 59, Band 97, Band 99}, and, similarly, the {G, H, I, J, K, L} can be set to {Band 3, Band 5, Band 38, Band 55, Band 83, Band 100}.
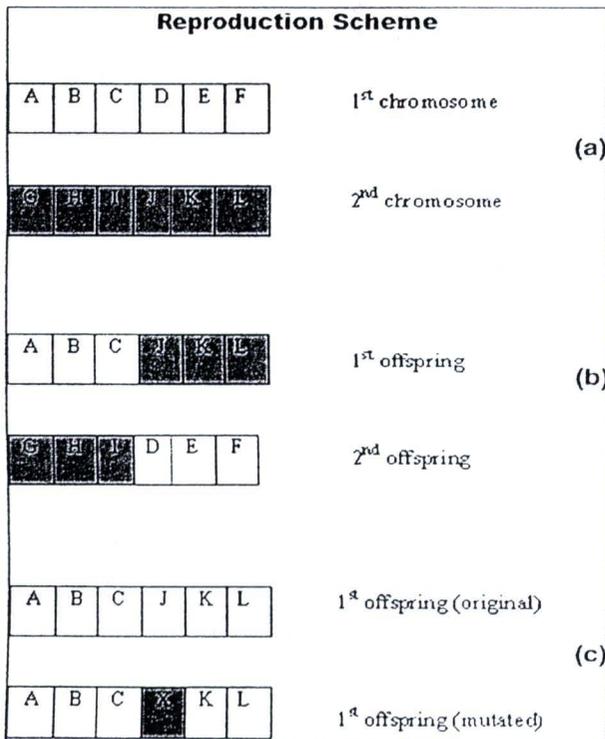

Fig. 4. The figures illustrate (a) Two parent chromosomes, (b) Two offspring chromosomes, and (c) An example of random mutation. Each block alphabet represents a spectral band label.

2.2.2. Reproduction mechanism

The mechanism of crossover and mutation is illustrated in Fig. 4. By mating the two chromosomes in Fig. 4a, the offspring that they produce share, in this example, half the characters of the first parent and the other half from the second parent. The two offspring are shown in Fig. 4b. Occasionally, some of the genes in any newly produced chromosome are randomly altered by mutation. This phenomenon causes a change in the character of the offspring, independent from the chromosome composition of the parents. The illustration of the mutation effect is shown in Fig. 4c. The "J" gene is mutated to the "X" gene through random mutation. In remote sensing context, this is equal to a random flip of a band label inside a chromosome.