A Hybrid Approach for Supervised Spectral Band Selection in Hyperspectral Images Classification

Seyyid Ahmed Medjahed[1], Mohammed Ouali[2,3]

[1] Centre Universitaire Ahmed Zabana Relizane,
Algérie

[2] Thales Canada Inc.,
Canada

[3] University of Sherbrooke,
Computer Science Department,
Canada

seyyidahmed.medjahed@cu-relizane.dz, mohammed.ouali@usherbrooke.ca

Abstract. Recently, hyperspectral imagery has been very active research field in many applications of remote sensing. Unfortunately, the large number of bands reduces the classification accuracy and computational complexity which causes the Hugh phenomenon. In this paper, a new hybrid approach for band selection based is proposed. This approach combines the advantage of filter and wrapper method. The proposed approach is composed of two phases: the first phase consists to reduce the number of bands by merging the highly correlated bands, and, the second phase uses a wrapper approach based on Sin Cosine Algorithm to select the optimal band subset that provides a high classification accuracy. In addition, a new binary version of Sin Cosine Algorithm is proposed to adapt it to the band selection problem. The performance evaluation of the proposed approach is tested on three publicly available benchmark hyperspectral images. The analysis of the results demonstrates the efficiency and performance of the proposed approach.

Keywords. Spectral band selection, hyperspectral image, classification, sin cosine algorithm, optimization.

1 Introduction

Hugh phenomenon is one of the major challenge problem in hyperspectral image classification [9, 4, 13]. Because the large number of bands and highly correlated bands, several problems are present such as low quality of classification map, difficulty of processing, etc. To avoid the curse dimensional problem and reduce computational time, the band selection methods are used. The aims of band selection are to select an optimal band subset by removing the highly correlated band, irrelevant and non-informative bands [6, 9, 5].

Generally, feature selection approaches are subdivided into two categories. The first category is known as filter approach [7]. This category uses a statistical measure to compute the importance of a feature and it is independent of any classification method. Therefore, the features are selected based on their statistical score. Among the filter approaches, we cite: Relief, independent component analysis (ICA), principal component analysis (PCA), minimum redundancy maximum relevance (mRmR), mutual information maximization (MIM), factor analysis (FA), conditional informax feature extraction (CIFE), mutual information feature selection (MIFS), fast correlation-based filter (FCBF), etc. [12, 1].

The second category of feature selection approaches is wrapper approach. In this category,

the problem of feature selection is regarding as a search problem and it uses generally machine learning methods to measure the importance of a subset of candidate features [7, 3, 11]. Among these approaches, we cite, BPSO-feature selection [8], Simulated Annealing-Support Vector Machine [7], Grey Wolf Optimizer for Band Selection [9], Binary Bat Algorithm for Feature Selection [11], etc.

The difference between filter and wrapper approaches can be summarized in these points: filter approach measures the importance of feature individually, the wrapper approach measures the relevance of a feature subset using training model. Also, wrapper approach provides in many cases the best subset of features but the risk of overfitting is high.

In this paper, we propose a new hybrid band selection approach to reduce the highly correlated bands which are neighbors and the irrelevant bands. This approach is based on two steps, firstly, we merge the adjacent bands which are highly correlated and secondly, we select a subset of bands that provides a high classification accuracy. Three hyperspectral images widely used in the literature will be considered to test the performance of the proposed approach. The rest of paper is organized as follows: in section 2, we detailed the proposed approach, in section 3, the experimental results are presented and conclusions are given in section 4.

2 The Proposed Band Selection Approach

The proposed approach is essentially based on two important steps. Figure 1 illustrates the proposed approach for band selection.

2.1 First Step

The first step consists to decrease the number of bands by merging the adjacent bands which are considered as highly correlated. We compute the Pearson correlation coefficient between adjacent bands. Considering the set of bands $B = \{b_1, ..., b_N\}$. The correlation coefficient is defined as follows:
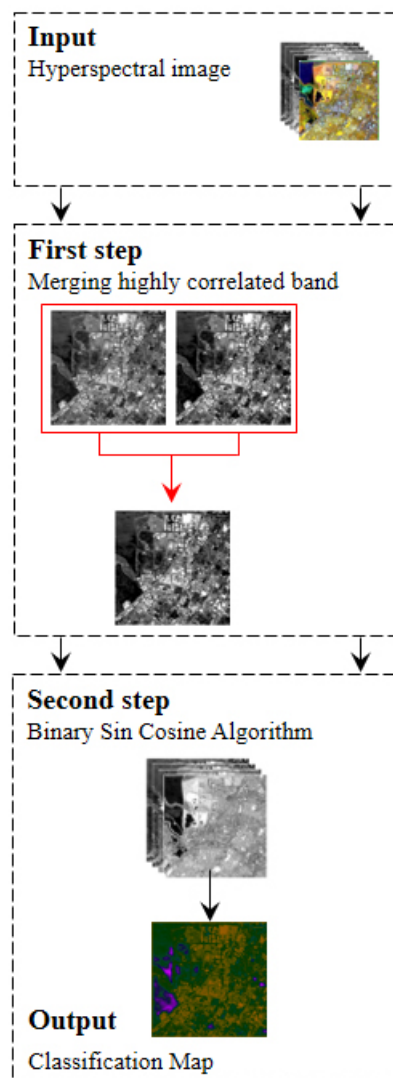

Fig. 1. The architecture of the proposed approach for spectral band selection

$$\rho(b_i, b_j) = \frac{\sum_{b_i, b_j}}{\sqrt{\sigma_{b_i} \sigma_{b_j}}}, \qquad (1)$$

where ρ is the Pearson correlation between the bands b_i and b_j. σ is the variance and $\sum_{b_i, b_j}$ is the covariance between the bands b_i and b_j.

Two bands b_i and b_j are considered as highly correlated if the Pearson correlation coefficient is
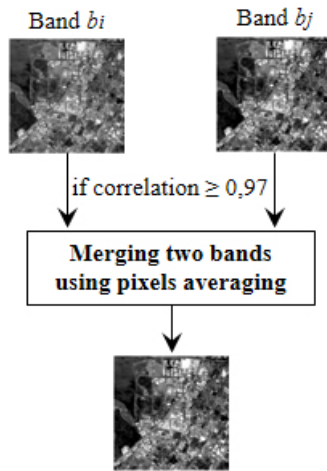
Fig. 2. The first step using to merge highly correlated bands

close to 1. This indicates that the adjacent bands b_i and b_j are strongly linear dependency.

For merging, we propose to take the average value between each pixels the two images. Figure 2 summarizes the first process.

2.2 Second Step

The second step consists to select from the subset of bands provided by the first step the optimal band subset. For this process, we propose to use a wrapper approach based on Sin Cosine Algorithm. The basic idea is using SCA to optimize the classification accuracy [10].

Sine Cosine Algorithm is an optimization algorithm developed by Seyedali Mirjalili [10]. SCA generates a set of random solutions called search agents, which are evaluated using the fitness function.

This algorithm uses a mathematical schema based on sin and cosine functions to designed the position updating [10]. This model is defined as follows:

$$X_i^{t+1} = \begin{cases} X_i^t + r_1 \times \sin(r_2) \times \left| r3P_i^t - X_i^t \right|, & r_4 < 0.5, \\ X_i^t + r_1 \times \cos(r_2) \times \left| r3P_i^t - X_i^t \right|, & r_4 \geq 0.5, \end{cases} \quad (2)$$

where r_1, r_2, r_3 and r_4 are random numbers.

r_1 is used to perform an exploration or exploitation and it is defined as follows:

$$r_1 = a - t\frac{a}{T}, \quad (3)$$

where T is the total number of iteration, t is the current iteration and a is a constant, r_2 defines the direction of the movement, r_3 provides a random weight, r_4 is a random variable between 0 and 1 using to choose between sin and cosine function [10].

A binary version of SCA algorithm is proposed. This binary version uses the transfer function defined on [2]. The transfer function allows to mapping $\mathbb{R}$ solution in $[0,1]$ solution and it given as follows:

$$g(X_i^{t+1}) = \left| \frac{2}{\pi} \arctan\left(\frac{2}{\pi} X_i^{t+1} \right) \right|, \quad (4)$$

$$v_i = \begin{cases} 1 \text{ if } g(X_i^{t+1}) > \beta, \\ 0 \text{ otherwise}, \end{cases} \quad (5)$$

where β is a random variable between 0 and 1. In this case, if $v_i = 1$ then the band b_i is used, otherwise, the band b_i is removed.

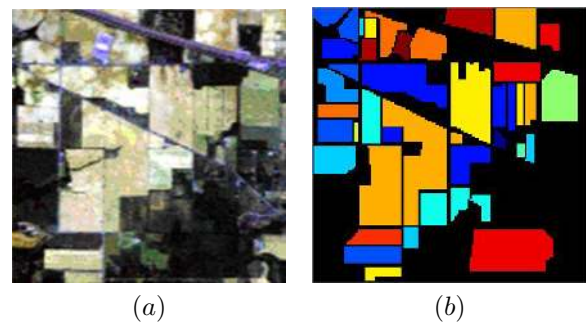The pseudo code of binary CSA algorithm is described as follows:

Fig. 3. Indian Pine (a) RGB image (b) Ground truth

The fitness function used in this study is the classification accuracy rate provided by k nearest neighbor (k-NN).
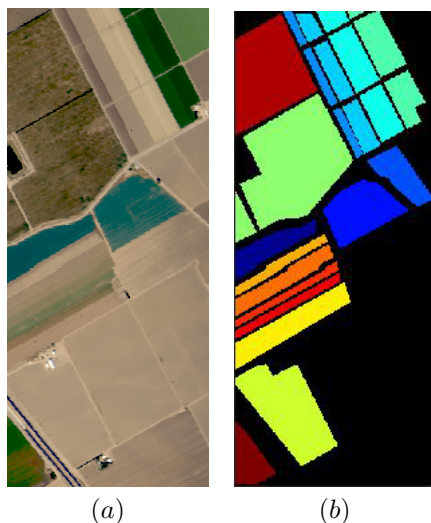
(a) $\qquad$ (b)

Fig. 4. Salinas (a) RGB image, (b) Ground truth

Algorithm 1 Binary Sin Cosine Algorithm

1: Initialize search agents position $X = X_1, ..., X_N$
2: Initialize the total number of iteration T
3: **for** $t = 1$ to T do **do**
4:   **for** $i = 1$ to $number\ of\ search\ agents$ do **do**
5:     Compute the fitness function
6:     Store the best solution
7:     Update parameters r_1, r_2, r_3 and r_4
8:     Update the position of each search agents using (2)
9:     Compute the binarization using (4 and 4)
10:   **end for**
11: **end for**
12: Return the best solution

---

3 Experimental Results

To assess the performance of the proposed approach, we conduct the experimentation on three publicly available benchmark hyperspectral images called: Indian Pines, Salinas and Pavia University.

Indian Pines is a hyperspectral image acquired by AVIRIS sensor over the Indian Pines, Northwestern, Indiana, USA. It is 145×145 pixels and 224 bands taken in the wavelength range $0, 4 - 2, 5\mu m$. The classes of this images are: Alfalfa, Corn-notill, Corn-mintill, Corn, Grass-pasture, Grass-trees, Grass-pasture-mowed, Hay-windrowed, Oats, Soybean-notill, Soybean-mintill, Soybean-clean, Wheat, Woods, Buildings-Grass-Trees-Drives, and Stone-Steel-Towers.

Salinas is hyperspectral image taken by AVIRIS sensor over Salinas Valley, California, USA. It is 512×217 pixels and composed of 224 bands taken in the wavelength range $0, 4 - 2, 5\mu m$.

It composed of 16 classes: Broccoli-green-weeds-1, Broccoli-green-weeds-2, Fallow, Fallow-rough-plow, Fallow-smooth, Stubble, Celery, Grapes-untrained, Soil-vinyard-develop, Corn-senesced-green-weeds, Lettuce-romaine-4wk, Lettuce-romaine-5wk, Lettuce-romaine-6wk, Lettuce-romaine-7wk, Vineyard-untrained and Vineyard-vertical-trellis.

Pavia University is acquired over University of Pavia, Italy. The size of this image is 610×340 pixels and it composed of 103 bands taken in the wavelength range $0, 43 - 0, 86\mu m$. It composed of 9 classes: Asphalt, Meadows, Gravel, Trees, Painted Metal Sheets, Bare Soil, Bitumen, Self-Bloking Bricks, and Shadows.
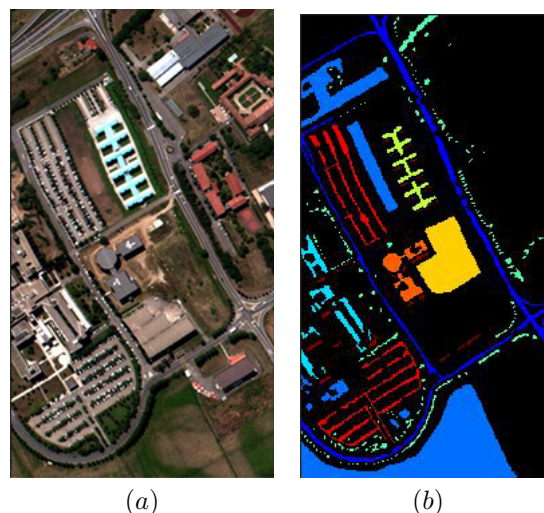

(a) $\qquad$ (b)

Fig. 5. Pavia University (a) RGB image (b) Ground truth

3.1 Parameters Settings

The parameters of the proposed approach are adjusted as follows: 10% of pixels are used for training. 40% of pixels are used for testing and the rest 50% of pixels are used for validation phase.

For the first step, we consider two adjacent bands highly correlated if the Pearson correlation coefficient is greater than $0,97$.

For the second phase, the binary CSA algorithm is set as follows: we use 30 search agents and 500 iterations. The fitness function is the classification accuracy rate provided by k-NN classifier using the Euclidean distance and $k = 7$.

3.2 Results and Discussions

The assess of the proposed approach is conducted in term of overall accuracy (OA), average accuracy (AA), individual class accuracy (ICA) and the number of selected bands. The results are described in tables 1, 2 and 3.

Table 1. OA (%), AA (%) and ICA (%) obtained by the proposed approach applying on Indian Pine

Class	Classification all bands	This approach Step 1	Step 2
Alfalfa	3,57	0,00	3,57
Corn-on till	57,41	58,23	70,25
Corn-min till	44,98	42,77	53,61
Corn	29,37	19,58	25,17
Grass/pasture	78,97	79,66	86,55
Grass/tree	96,80	96,12	97,72
Grass/pasture-mowed	11,76	5,88	35,29
Hay-windrowed	99,30	98,95	99,30
Oats	0,00	0,00	8,33
Soybeans-no till	56,68	56,16	71,58
Soybeans-min till	75,97	73,52	78,12
Soybeans-clean till	28,65	26,12	36,80
Wheat	93,50	95,12	97,56
Woods	95,39	95,13	94,47
Bldg-grass-tree-drives	18,97	12,50	25,00
Stone-steel towers	80,36	87,50	83,93
AA	54,48	52,95	60,46
OA	67,92	66,60	73,46
Number of bands	224	133	114

Tables 1, 2 and 3 present the classification accuracy obtained by our approach for Indian pines, Salinas and Pavia University. The first

Table 2. OA (%), AA (%) and ICA (%) obtained by the proposed approach applying on Salinas

Class	Classification all bands	This approach Step 1	Step 2
Brocoli_green_weeds_1	98,18	97,68	99,59
Brocoli_green_weeds_2	99,82	99,78	99,82
Fallow	99,58	94,94	99,66
Fallow_rough_plow	99,64	98,69	99,52
Fallow_smooth	96,76	96,20	97,95
Stubble	99,87	99,87	99,79
Celery	99,39	99,53	99,86
Grapes_untrained	81,09	77,85	82,82
Soil_vinyard_develop	99,73	99,38	99,65
Corn_green_weeds	92,22	91,61	94,61
Lettuce_romaine_4wk	94,85	90,80	97,97
Lettuce_romaine_5wk	99,39	97,93	99,74
Lettuce_romaine_6wk	98,36	97,82	98,55
Lettuce_romaine_7wk	92,21	89,56	95,79
Vinyard_untrained	56,87	54,41	69,34
Vinyard_vertical_trellis	98,43	96,41	98,43
AA	94,15	92,65	95,82
OA	89,10	87,52	91,55
Number of bands	224	118	103

column is the class number. The second column represents the results produced by using only k-NN without band selection approach, all the bands are used for the train. The third column is the results obtained by the first step of the proposed approach and the last column is the results obtained by the second step of the proposed approach.

The rows are the individual class accuracy and the two rows befor the last are the average accuracy and overall accuracy respectively. The last row is the number of selected bands.

As seen in tables 1, 2 and 3, the average accuracy and the overall accuracy produced by proposed approach are very satisfactory. The proposed approach outperforms and leads to a higher classification accuracy.

Moreover, we observe that the classification accuracy increases significantly when using the two steps of the approach compared to the classification without band selection. For Indian pines image, we record a $73,46\%$ of classification accuracy when using the proposed approach and $67,92\%$ when using classification without band selection.

Table 3. OA (%), AA (%) and ICA (%) obtained by the proposed approach applying on Pavia University

Class	Classification all bands	This approach Step 1	This approach Step 2
Asphalt	88,77	87,96	89,80
Meadows	97,85	97,93	98,30
Gravel	70,24	75,87	78,49
Trees	86,57	88,47	90,43
Painted Metal Sheets	99,26	99,38	99,26
Bare Soil	60,37	68,22	76,84
Bitumen	82,33	87,72	91,73
Self-Bloking Bricks	85,57	86,88	87,69
Shadows	100	99,65	100
AA	85,66	88,01	90,28
OA	88,42	89,95	91,87
Number of bands	103	51	42

The same observation is remarked for Salinas image with $91,55\%$ of classification accuracy and $89,10\%$ when we do not use band selection. Also, for Pavia university, the proposed approach reached $91,87\%$ of accuracy which is better compared to classification without band selection $88,42\%$.

In addition, the number of bands reduces significantly without changing the classification accuracy rate. $49,10\%$ of bands are removed in Indian Pines and $54,01\%$ of bands are removed from Salinas. For Pavia University $59,22\%$ of bands are removed considering as irrelavant and highly correlated bands.

The classification maps provided by our approach are illustrated in figure 6.

The classification maps illustrated in figure 6 are good and represents clearly delination of complex regions. The classes are homogenous and separable.

4 Conclusions

In this paper, a novel supervised band selection approach is proposed. The basic idea of this approach is to develop a hybrid approach based on two steps. The first step attempts to merge the adjacent bands highly correlated. The second step consists to optimize the classification accuracy rate using Sin Cosine Algorithm. A new binary version

Indian Pines   Salinas

Pavia University
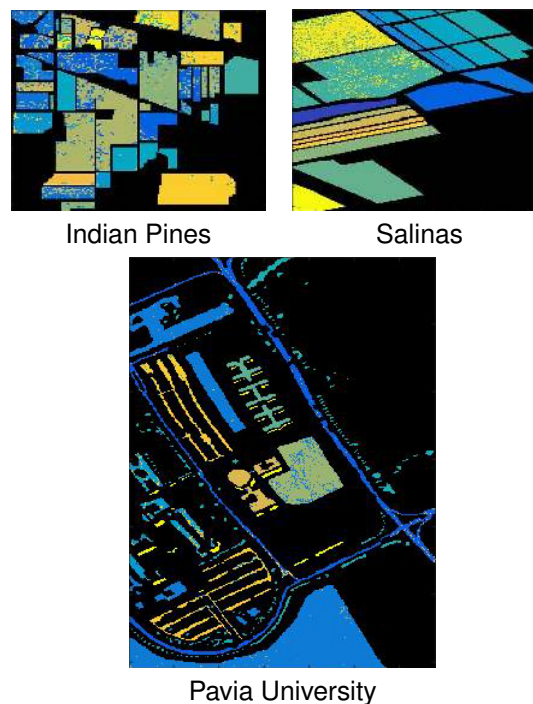
Fig. 6. The classification maps obtained by the proposed approach

of SCA algorithm is proposed based on transfer function. The fitness function used in this study is the classification accuracy rate produced by k-NN algorithm. The experimentation is conducted on three hyperspectral images namely: Indian Pines, Salinas and Pavia University. The approach is evaluated in term of individual class accuracy, overall accuracy and average accuracy. The experimental results demonstrate the effectiveness of the proposed approach and show the capacity to challenge the problem of band selection. Future work will concern the improvement of the fitness function by using other criteria.

References

1. **Brown, G., Pocock, A., Zhao, M. J., & Luj, M. (2012).** Conditional likelihood maximisation: A unifying framework for information theoretic feature selection. *The Journal of Machine Learning Research*, Vol. 13, pp. 27–66.

2. **Crawford, B., Soto, R., Astorga, G., García, J., Castro, C., & Paredes, F. (2017).** Putting continuous metaheuristics to work in binary search spaces. *Complexity*, Vol. 17, pp. 1–19.

3. **Lewis, D. D. (1992).** Feature selection and feature extraction for text categorization. *In Proceedings of Speech and Natural Language Workshop*, Morgan Kaufmann, pp. 212–217.

4. **Liu, H., Yang, S., Gou, S., Liu, S., & Jiao, L. (2018).** Terrain classification based on spatial multi-attribute graph using polarimetric sar data. *Applied Soft Computing*, Vol. 68, pp. 24–38.

5. **Medjahed, S. A. & Ouali, M. (2018).** Band selection based on optimization approach for hyperspectral image classification. *The Egyptian Journal of Remote Sensing and Space Science*.

6. **Medjahed, S. A. & Ouali, M. (2018).** Svm-rfe-ed: A novel svm-rfe based on energy distance for gene selection and cancer diagnosis. *Computacion y Sistemas*, Vol. 22, No. 2, pp. 675–683.

7. **Medjahed, S. A., Ouali, M., Saadi, T. A., & Benyettou, A. (2015).** An optimization-based framework for feature selection and parameters determination of svms. *International Journal of Information Technology and Computer Science (IJITCS)*, Vol. 7, No. 5, pp. 1–9.

8. **Medjahed, S. A., Saadi, T. A., Benyettou, A., & Ouali, M. (2015).** Binary cuckoo search algorithm for band selection in hyperspectral image classification. *IAENG International Journal of Computer Science*, Vol. 42, No. 3, pp. 183–191.

9. **Medjahed, S. A., Saadi, T. A., Benyettou, A., & Ouali, M. (2016).** Gray wolf optimizer for hyperspectral band selection. *Applied Soft Computing*, Vol. 40, pp. 178–186.

10. **Mirjalili, S. (2016).** SCA: A sine cosine algorithm for solving optimization problems. *Knowledge-Based Systems*, Vol. 96, pp. 120–133.

11. **Nakamura, R. Y. M., Pereira, L. A. M., Costa, K. A., Rodrigues, D., & Papa, J. P. (2012).** Bba: A binary bat algorithm for feature selection. *Conference on Graphics, Patterns and Images (SIBGRAPI)*.

12. **Peng, H., Long, F., & Ding, C. (2005).** Feature selection based on mutual information: criteria of max-dependency, max-relevance, and min-redundancy. *IEEE Transactions on pattern analysis and machine intelligence*, pp. 1226–1238.

13. **Zhang, Q., Yuan, T., Yang, Y., & Chunhong, P. (2015).** Automatic spatial–spectral feature selection for hyperspectral image via discriminative sparse multimodal learning. *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 53, No. 1.