Journal of Information Hiding and Multimedia Signal Processing ©2015 ISSN 2073-4212
Ubiquitous International Volume **6**, Number **1**, January **2015**

Hyperspectral Image Feature Reduction Based on Tabu Search Algorithm

Ying Cui[1,2,a], Jia Qi Wang[1], Shu Bin Liu[2], and Li Guo Wang[1]

[1]College of Information and Communication Engineering
Harbin Engineering University
145 Nantong Avenue, Nangang District, Harbin, Heilongjiang
[a]cuiying@hrbeu.edu.cn

[2]Remote Sensing Technology Center
Heilongjiang Academy of agricultural science
368 Xuefu Avenue, Nangang District, Harbin, Heilongjiang

Received May, 2014; revised April, 2014

Abstract. *High dimension coupled with abundant spectral information are among the greatest challenges for traditional image processing measures. These problems can often be handled effectively by dimensionality reduction which has been studied in recent years. Monte Carlo feature reduction method(MCFR) can potentially calculate optimal feature reduction number, but it costs a long time by sampling large numbers of random samples. In this paper, a new algorithm based on tabu search optimization technique and Compactness-Separation Coefficient(CS Coefficient) was developed to perform dimensionality reduction and calculate optimal feature reduction number. The advantages of this algorithm can be further exploited for hyperspectral data classification by introducing classifiers such as Support Vector Machine(SVM) and Relevance Vector Machine(RVM). Experimental results obtained from new algorithm are superior to those of MCFR, with less optimization time and higher classification accuracy.*
Keywords: Feature reduction method, Tabu search, Support vector machine, Relevance vector machine

1. **Introduction.** Hyperspectral remote sensing is the multidimensional feature information retrieval technology, including target detection technology and spectral imaging technology. For classification applications, hyperspectral data have provided huge opportunities for its abundant spatial and spectral information. However, high spectral dimension coupled with spectral resolution also put forward great challenges for traditional image classification algorithms. The classification processing may come across dimension disaster when training samples are limited. One approach that is frequently employed to mitigate this problem involves dimensionality reduction. In recent years, many scholars study dimension reduction processing[1-3], but all of them merely reduce the features redundant and can't provide optimal feature reduction number. ZHAO Chun-hui and QI Bin proposed Monte Carlo feature reduction method (MCFR) to solve this problem[4], but the MCFR costs long optimization time by sampling a large number of random samples to get optimal feature reduction number.Tabu search (TS) is a simulation of human intelligence process proposed by F.Glover, YANG Zhe-hai raised TS to perform dimensionality reduction by using Optimal Index Factor (OIF) as fitness function[15]. ZHU Yan combined TS with band correlation coefficient to reduce hyperspectral data's dimension[16].

The two methods above can reduce hyperspectral data's dimension in a short time, but the fitness functions can't provide different bands contribution for terrain classification.

Compactness-Separation Coefficient in MCFR can provide both compactness within class and separability between classes in each band. TS can obtain optimum solution in a short time. Therefore in this paper, a new model is developed which incorporates the use of TS with Compactness-Separation Coefficient for dimensionality reduction. Experimental results are demonstrated on two classifiers acquired by Support Vector Machine(SVM) and Relevance Vector Machine (RVM), both of which have the same functional form. The performance measures used are running time, the number of feature reduction bands and overall classification accuracy.

The remaining of this paper is organized as follows. Section 2 presents the basic concept. In Section 3, we explain the feature reduction modules and algorithm. We give the process of TSFR. In Section 4, by means of simulation examples, we evaluate the performance of proposed algorithm. Finally, concluding remarks are given in Section 5.

2. The related concept introduction.

2.1. **Dimensionality reduction methods.** The increasing availability of data from hyperspectral sensors makes features distinguishing more accurate. However, abundant spectral information cause Hughes, high calculation amount and high data redundancy, and make data processing difficult. Consequently, there is a clear need of dimensionality reduction before data processing. Dimensionality reduction methods can be grouped into four categories: band selection, data resource division, feature detection and data fusion. This paper focuses on band selection. Most of band selection methods are based on information content and separability between classes. Methods based on information content commonly use variance and image entropy as criterion. Criteria used in methods based on separability are mainly standard distance, dispersion, Jeffreys-Matusita distance and so on[17]. This paper introduces CS Coefficient to select optimal bands. CS Coefficient reflects both compactness within class and separability between classes.

2.2. **Tabu Search Algorithm.** TS is a metaheuristic method for combinatorial optimization problems. Its popularity has grown due to its strong local search ability. It has achieved great success in the field of combinatorial optimization, machine learning, circuit design and neural network. Conceptually, TS starts with an incumbent solution, establishes a set of candidate solutions in the neighborhood, which can be obtained by $2-opt$, $k-opt$ and so on. The tabu list can help avoid unwanted cycling and escape from local optimal solutions. The aspiration criterion allows the solution whose tabu length is not zero to move out of tabu list. The stopping criterions are mainly based on maximum iterations, limited frequency, deviation and so on.

TS parameters involve neighborhood, tabu list, tabu length, candidate and aspiration criterion. The tabu list and aspiration criterion reflect diffusion and concentration strategies respectively. The concentration strategy means starting with an incumbent solution and searching for local optimal solution in the neighborhood. The diffusion strategy is to jump out of local optimal solution.

3. Feature reduction modules and TSFR.

3.1. **The CS Coefficient calculating module.** Assuming that X=x_i is all the objects, these samples are classified to M kinds $Z = \left\{z^1, z^2, ..., z^M\right\}$, where z_i stands for all the objects belonging to the $i-th$ kind. The numbers of each sample are $\{\beta_i\}$, $i = 1, 2, ..., M$. The $k-th$ feature band compactness coefficient is defined as:

$$Compactness(k) = \frac{1}{M}\sum_{i=1}^{M}\frac{1}{\beta_i(\beta_i-1)}\sum_{u=1}^{\beta_i}\left|z_{u,k}^{i}-z_{v,k}^{i}\right| \tag{1}$$

The $k-th$ feature band separation coefficient is defined as:

$$Separation(k) = \frac{1}{M(M-1)}\sum_{i=1}^{M}\sum_{j\neq i}\frac{1}{\beta_i\beta_j}\sum_{u=1}^{\beta_i}\sum_{v=1}^{\beta_j}\left|z_{u,k}^{i}-z_{v,k}^{j}\right| \tag{2}$$

The $k-th$ feature band CS Coefficient is

$$CS(k) = \alpha Compactness(k) - (1-\alpha)Separation(k) \tag{3}$$

where α is compactness and separation regulatory factor. CS Coefficient is decided by separation coefficient when α=0, otherwise being decided by compactness coefficient when α=1. Document[4] shows that classification accuracy is highest when α=0.6 Therefore, α is 0.6 in this paper.

3.2. **The neighborhood generating module.** The proper type of neighborhood provides guidance for choosing proper candidate solutions. This paper uses $k-opt$ method to generate neighborhood solutions. Suppose k is 6 and there is a set of solutions as follows:

$$S1 : 110010101110$$

Generate 6 random numbers between 1 and 12 and change the value in the corresponding position such as 1 to 0 or 0 to 1. Then the neighboring solution of S_1 is

$$P1 : 011000001011$$

The $k-opt$ method can traverse more solutions than $2-opt$ method.

3.3. **Tabu Search feature reduction (TSFR).** In this paper, TS is implemented within feature reduction so that an optimal feature reduction number can be calculated. The main idea of TSFR is to use CS Coefficient as fitness function, calculating CS Coefficient of each band and finding the optimal combination of bands based on TS. The flowchart of TSFR is shown as Fig.1. The implementation of TSFR is as follows:
Step 1 Initialization-Setting the initial parameters.
Step 2 Calculate CS Coefficient of each band using(3).
Step 3 Generate the initial solution randomly.
Step 4 Generate neighboring solution using $k-opt$ and choose the optimal solution.
Step 5 Determine whether in tabu list, if not, replace 'best-so-far' solution, modify tabu list and run to Step 7; if so, continue.
Step 6 Determine whether aspiration criterion is satisfied, if so, replace 'best-so-far' solution, modify tabu list and run to Step 7; if not, choose non-tabu optimal solution, replace 'best-so-far' solution, modify tabu list and continue.
Step 7 Determine whether stopping criterion is satisfied, if not, return to Step 4.
Step 8 Output optimal feature reduction number.

3.4. **The Classification Methods.**

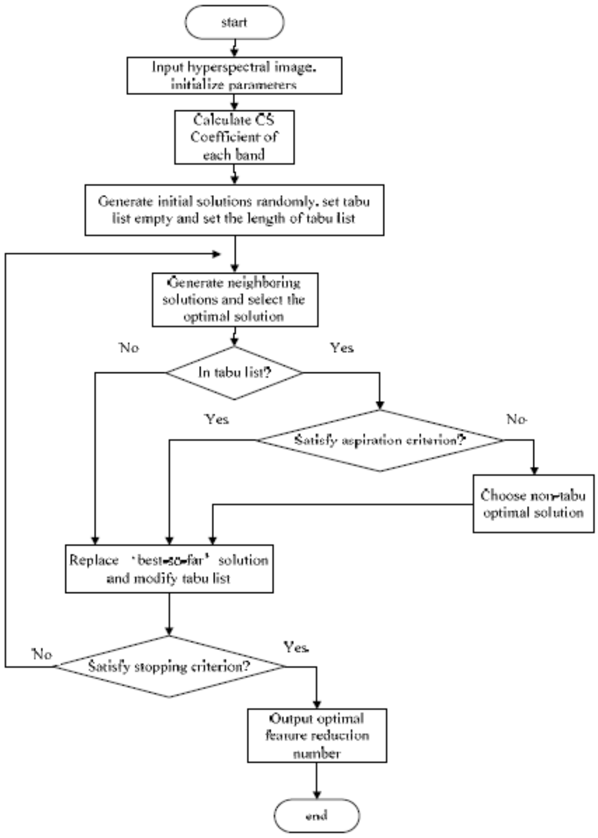

FIGURE 1. Flowchart of TSFR

3.4.1. *Support Vector Machine.* SVM mechanism is to find an optimized decision boundary which satisfies classification requirements. Take classification of data with two types as example, the training sample set is given as $(xi, yi), i = 1, 2, ..., l$, $x \in R^n$,$y \in \{\pm 1\}$, the hyperplane is denoted by $(\omega \cdot x) + b = 0$.The hyperplane can classify all samples correctly and posses class interval when it satisfies the following restriction:

$$y_i[(\omega \cdot x_i) + b] \geq 1, i = 1, 2, ..., l \tag{4}$$

The construction of optimal hyperplane can be transformed into constrained formula as follows:

$$\min \Phi(\omega) = \frac{1}{2}\|\omega\|^2 = \frac{1}{2}(\omega' \cdot \omega) \tag{5}$$

By introducing Lagrange function, the optimal classification function can be obtained as follows:

$$f(x) = \mathrm{sgn}\{(\omega^* \cdot x) + b^*\} = \mathrm{sgn}\{(\sum_{j=1}^{l} a_j^* y_j (x_j \cdot x_i) + b^*\}, x \in R^n \tag{6}$$

where

$$\begin{cases} \omega^* = \sum_{i=1}^{n} \alpha_i x_i y_i, 0 \leq \alpha_i \leq C \\ b^* = -\frac{1}{2}[\min_{y_i=1}(\omega^* \cdot x_i) + \max_{y_i=-1}(\omega^* \cdot x_i)] \end{cases}$$

where C is regularization coefficient, and C is 8 in this paper.

3.4.2. *Relevance Vector Machine.* Relevance Vector Machine is based on Bayesian model. For binary classification, the label of training samples $\{t_i\}_{i=1}^{N}$ can only be 0 or 1. Judge category using S function as follows:

$$P(t_i = 1 \,|\omega\,) = \sigma[y(x_i;\omega)] = \frac{1}{1+e^{-y(x_i;\omega)}} \tag{7}$$

If each observation is an independent event, then the possibility of observed result being t is

$$P(t\,|\omega\,) = \prod_{i=1}^{N} \sigma[y(x_i;\omega)]^{t_i}\{1-\sigma[y(x_i;\omega)]\}^{1-t_i} \tag{8}$$

The solution of weight w using method of maximum likelihood is

$$\omega_{MP} = \arg\max p(\omega\,|t,\alpha) = \arg\max \log\{p(t\,|\omega\,)p(\omega\,|\alpha\,)\} \tag{9}$$

$$\log\{p(t\,|\omega\,)p(\omega\,|\alpha\,)\} = \sum_{i=1}^{N}[t_i \log y_i + (1-t_i)\log(1-y_i)] - \frac{1}{2}\omega^T A\omega \tag{10}$$

4. Simulation results and discussions.

4.1. **Hyperspectral Data.** The hyperspectral image used in this paper is original AVIRIS image, which was a part of remote sensing research area in Northwest Indian, Indiana, USA. It was shot in June, 1992. It contains a mixture of crops and forest vegetation area. The characteristics of the AVIRIS data are shown in Table1. The simulated image of band 50, 27 and 17 is shown in Fig.2. In this paper, we choose three kinds of beans from this image to conduct feature reduction, the corresponding ground truth image is shown in Fig.3. In Fig.3, each kind of bean has 500 pixels. The test samples are composed of 750 pixels by choosing 250 pixels from each kind randomly. The training samples are composed of 30 or 60 pixels by choosing 10 or 20 pixels from each kind randomly in the remaining 750 pixels.

TABLE 1. Characteristics of the AVIRIS Data used in the Experiment

Characteristics	Parameters
Spatial resolution	20m*20m
Image size	145*145 pixels
Pixel depth	16 bit
Bands number	220
Wavelength range	400 2500nm
Spectral resolution	10nm approximately

4.2. **Experiment parameters selection.** The method proposed in this paper (TSFR) is compared with MCFR. SVM and RVM are used to verify the accuracy of classification and test the universality of TSFR. The kernel function of SVM and RVM is Radial Basis Function (RBF), and the formula is shown as follows:

$$k(x,x') = \exp(-\eta\|x-x'\|^2) \tag{11}$$

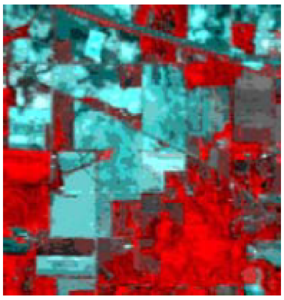

FIGURE 2. Simulated image of band 50, 27 and 17

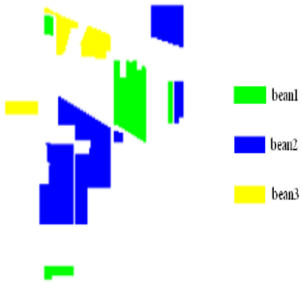

FIGURE 3. The bean ground truth image

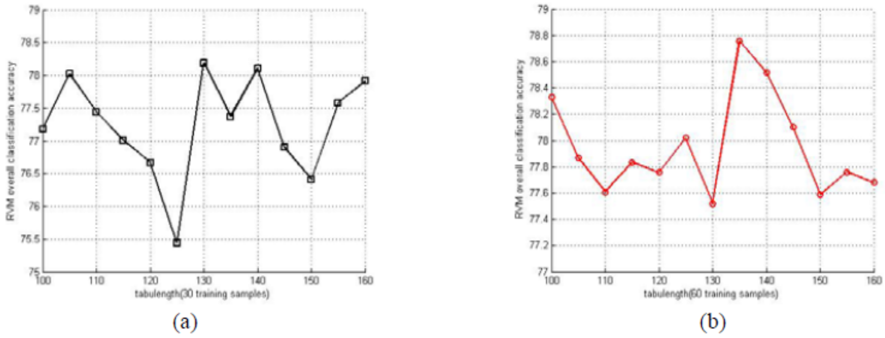

FIGURE 4. The effects of tabu length to RVM classification accuracy (a) The effect of tabu length under 30 training samples(b) The effect of tabu length under 60 training samples

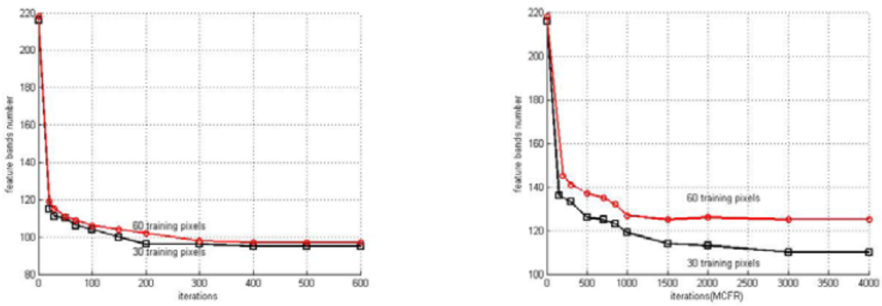

FIGURE 5. Feature bands number of different iterations (a) The TSFR feature bands number curve (b) The MCFR feature bands number curve

where η is kernel function parameter. In this paper η is equal to 0.5. The experimental results in this paper are the average of 20 experiment results.

Tabu list contains tabu length and object function. Tabu length is a key parameter, which is the maximal time that tabu objects are not selected. Commonly, the length of tabu list should not be lower than half scale of the question to be solved. The effects of tabu length are shown in Fig.4. From Fig.4 we can determine that tabu length of 30 training samples is 105 and tabu length of 60 training samples is 135.

The curves of feature bands number under different iterations based on the MCFR and TSFR are shown in Fig.5. From Fig.5, we can observe that the TSFR feature bands number is stable when iteration is over 300. However, the MCFR feature bands number is stable when iteration is over 2000.

Experimental results are demonstrated on three aspects including running time, number of feature reduction bands and overall classification accuracy. The overall classification accuracy is defined as:

$$OA = \frac{1}{n}\sum_{i=1}^{N} m_{ii} \tag{12}$$

where n is the total number of samples and $m_i i$ stands for the number of correctly classified samples for class i.

4.3. **Results and analysis.** The feature reduction number and running time of TSFR and MCFR are listed in Table2. The comparison of overall classification accuracy between SVM, MCSVM and TSSVM with different training samples numbers are shown in Table3. Similarly, the comparison of overall classification accuracy between RVM, MCRVM and TSRVM for different training samples numbers are shown in Table4.

Table 2. Comparison of feature reduction number and running time with MCFR and TSFR

Number of pixels in training samples	Feature Reduction Method	Number of feature reduction bands	Running Time(s)
30	MCFR	113	42.87471
30	TSFR	95	6.003483
60	MCFR	126	69.127787
60	TSFR	97	6.010688

Table 3. Comparison of overall classification accuracy with SVM, MCSVM and TSSVM

Number of pixels in training samples	Classification Method	Number of feature reduction bands	Overall Classification
30	SVM	–	83.73%
30	MCSVM	113	95.57%
30	TSSVM	95	96.98%
60	SVM	–	81.60%
60	MCSVM	126	96.69%
60	TSSVM	97	98.29%

From Table2, we can conclude that TSFR can obtain less feature bands in shorter time than MCFR and TSFR can save a great deal of time. Table3 and Table4 show that the overall accuracy with feature reduction data by TSFR has significant improvement compared with those of original data. Among them, the overall accuracy of SVM increases 13.25% for thirty training samples and 16.69% for sixty training samples. While the overall accuracy of RVM increases 18.02% for thirty training samples and 17.03% for sixty training samples. We can conclude that the overall accuracy of TSSVM and TSRVM are superior to those of MCSVM and MCRVM.

Table 4. Comparison of overall classification accuracy with RVM, MCRVM and TSRVM

Number of pixels in training samples	Classification Method	Number of feature reduction bands	Overall Classification
30	RVM	–	60.00%
30	MCRVM	113	75.72%
30	TSRVM	95	78.02%
60	RVM	–	61.73%
60	MCRVM	126	80.13%
60	TSRVM	97	78.76%

5. **Conclusions.** The TSFR is proposed and implemented in this paper, in which Tabu Search and CS Coefficient are used to build dimensionality reduction module. The proposed method could successfully get optimal feature reduction number with less running time than MCFR. Furthermore, the universality of TSFR is tested by SVM and RVM classifiers. The experimental results show that the overall classification accuracy with feature reduction data of TSFR is better than the accuracy with feature reduction data of MCFR.

Acknowledgement. This work was supported by the National Natural Science Fund of China (No. 61275010), the doctoral Fund of Ministry of Education of China (No.20132304110007), Heilongjiang postdoctoral special fund (LBH-TZ0420) and theFundamental Research Funds for the Central Universities (HEUCF140813). The authors would like to thank the associate editor and reviewers for their comments that help in improving this paper.

REFERENCES

[1] J.Y. Shi, X.B. Zou, J.W. Zhao, and H.P. Mao, Selection of wavelength for strawberry NIR spectroscopy based on BIPLS combined with SAA, *Journal of Infrared and Millimeter Waves*, vol. 30, no. 5, pp. 458-462, 2011.

[2] Y.H. Wang, S.F. Tian, and H.k. Huang, Feature Weighted Support Vector Machine, *Journal of Electronics & Information Technology*, vol. 3, no. 3, pp. 514-518, 2009.

[3] J. Li, X.B. Gao, and L.C. Jiao, A new feature weighted fuzzy clustering algorithm, *Acta Electronica Sinica*, vol. 34, no. 1, pp. 89-92, 2006.

[4] C.H.Zhao, B. Qi, and Y. Eunseog, Hyperspectral image classification based on Monte Carlo feature reduction method, *Journal of Infrared and Millimeter Waves*, vol. 32, no. 1, pp. 62-67, 2013.

[5] W. Xia, B. Wang, and L.M. Zhang, Blind unmixing based on independent component analysis for hyperspectral imagery, *Journal of Infrared and Millimeter Waves*, vol. 30, no. 2, pp. 131-136, 2011.

[6] F.A. Mianji, and Y. Zhang, Robust hyperspectral classification using relevance vector machine, *IEEE Trans. on Geoscience and Remote Sensing* vol. 49, no. 6, pp. 2100-2112, 2011.

[7] B. Qi, C.H. Zhao, Y.L. Wang, Hyperspectral imagery classification based on SVM and RVM, *Journal of Jilin University(Engineering and Technology Edition)*, vol. 42, pp.143-147, 2013.

[8] Y. Gao, X.S. Wang, Y.H. Cheng, Dimensionality Reduction of Hyperspectral Data Using Nonnegative Sparsity Graph, *Journal of Electronics & Information Technology*, vol. 35, no. 5, pp.1177-1184, 2013.

[9] C.H. Zhao, B. Qi, Y. Zhang, Hyperspectral Image Classification Based On Variational Relevance Vector Machine, *Acta Optica Sinica*, vol. 32, no. 8, 2012.

[10] H.Q. Zhao, P.Y. Wang, L. Li, The Parameter Selection and Convergence Analysis of Tabu Search Optimization Algorithm, *Communication and Information Processing*, vol. 32, no. 2, pp. 28-33, 2013.

[11] F. Samadzadegan, H. Hasani, T. Schenk, Simultaneous feature selection and SVM parameter determination in classification of hyperspectral imagery using Ant Colony Optimization, *Canadian Journal of Remote Sensing*, vol. 38, no. 2, pp. 139-156, 2012.

[12] J. Qian, K.Z. Deng, H.D. Fan, Dimensionality Reduction and Classification for Hyperspectral Remote Sensing Imagery Based on Laplacian Eigenmap, *Remote Sensing Information*, vol. 27, no. 5, pp.3-7, 2012.

[13] C.H. Zhao, B. Qi, Y.L.Wang, An Improved N-FINDR Hyperspectral Endmember Extraction Algorithm, *Journal of Electronics & Information Technology*, vol. 34, no. 2, pp. 499-503, 2012.

[14] F. Glovera, S. Hanafib, Tabu Search and Finite Convergence, *Discrete Applied Mathematics*, vol. 119, no. 1-2, pp. 3-36. 2002.

[15] Z.H. Yang, Y.Z. Zhang, D.P. Gong, Z.X. Li, J.F. Han, The Feature Selection In Hyperspectral Images Based on Tabu Search, *Hydrographic Surveying And Charting*, vol.26, no.4, pp.11-14, 2006.

[16] Y. Zhu, X.L. Liu, Z.H. Yang, Dimensionality Reduction in Hyperspectral Classification and the Application of Tabu Search Algorithm, *Journal of Zhengzhou Institute of Surveying and Mapping*, vol. 24, no. 1, pp. 22-29, 2007.

[17] Y. Tian, Research on Dimension Reduction Method of Hyperspectral Remote Sensing Images, Harbin Engineering University, pp. 2-11, 2008.