IJCTA, 9(10), 2016, pp. 4603-4609

Hyperspectral Image Classification Improved with ELRMA Denoising

***Gouri Swamynadhan **Nidhin Prabhakar TV, K.P Soman**

Abstract : In this paper, scope of improvement of hyperspectral images with a preprocessing technique is studied using various classification methods. Hyperspectral images are of great scope in exploration as it provides wider and precise information. Reflectance of a hyperspectral image contains spectral information of pixels as well as spatial information. Hyperspectral images have wide range of applications in diverse fields of remote sensing such as geology, oil spill detection, land cover classification, mineral detection, bio mass detection, urban planning and forest study. Since hyperspectral images are subjected to noise, denoising using enhanced low rank matrix approximation(ELRMA) is applied as a preprocessing technique. Low rank matrix approximation(LRMA) is enhanced using a non-convex regularization treating it as a convex optimization problem and it is applied to hyperspectral images. Using ELRMA technique denoising of hyperspectral images are done effectively and the improvement is analyzed using subspace pursuit algorithm, GURLS and random forest classification methods.

Keywords : Hyperspectral Images; ELRMA; Denoising; Classification; SP; RandomForest; GURLS

1. INTRODUCTION

Hyperspectral images (HSI) are of high spectral and spatial information. HSI images are captured in hundreds of contiguous bands [1]. It is stacked one on top of the other and thus forms a cube of data. Hyperspectral sensors has the potential to capture images of scene in hundreds of narrow bands over visible and infrared regions. It provides the wavelength or spectral information in bands and spatial information as 2D pixels. It has wide range of application like hydrology, environmental monitoring, pollution detection, study of vegetation and lot more [2]. Since hyperspectral images possess high spectral information and data abundancy, linear and nonlinear noises like Gaussian noise, impulse and lot more will affect hyperspectral images both spatially as well as spectrally [3]. Because of this denoising of hyperspectral images has wide scope and study on this topic is done profusely. Denoising is done as a preprocessing step to remove noise and to enhance the quality of hyperspectral images. Denoising can be done by various means such as wavelet, Total Variation denoising and many more. Another method used for HSI denoising is based on the concept of low rank matrix approximation. Because of high correlation of pixels in hyperspectral images it exhibits low rank property and hence LRMA method can be effective in HSI denoising.

In [4], Wei He et al., proposed a method based on low rank matrix approximation in HIS in which LRMA method is applied in each band since noise in each band has variations with the other one and so a noise adjusted iterative LRMA is proposed. A regularized SVD is used to solve the LRMA. In [5] hyperspectral image classification and denoising of data as a preprocessing technique using AB filter is discussed. Classification is done using optimization

* Center for Computational Engineering and Networking (CEN), Amrita School of Engineering, Coimbatore Amrita Vishwa Vidyapeetham, Amrita University, India gouri.swaminadhan@gmail.com

** Center for Computational Engineering and Networking (CEN), Amrita School of Engineering, Coimbatore Amrita Vishwa Vidyapeetham, Amrita University, India kp_soman@amrita.edu

based basis pursuit algorithm solved via ADMM. In [6], sparse based HIS classification by spatial preprocessing is studied. Classification accuracy is improved by this spatial preprocessing step. In [7], BaassouBelkacem et al., describes an iterative support vector machine for HSI classification which coordinate both spatial and spectral data. Spatial parameters are included using a majority voting process which utilizes the spatial information of the neighborhood classes.

Recently Ankit Parekh and Ivan W. Selesnick suggested an Enhanced LRMA (ELRMA) method [8], in which low rank matrix approximation is framed using a convex optimization with non-convex regularization. Non-convex penalty functions are used for the computation of non-zero singular values. The method is used for denoising of grey scale images.

In this paper, the method described in [8] is mapped for HSI denoising. The hyperspectral images of AVIRIS sensor is used for the experimentation. After the preprocessing, the classification is done using sparsity based subspace pursuit (SP) algorithm, GURLS and random forest. The effectiveness of each classifier is analyzed in terms of overall accuracies obtained.

2. LRMA

Matrix approximation is a much used term in machine learning and it has wide range of application in machine learning, image processing and signal processing. In usual denoising method a low rank matrix is estimated from the noisy image Y by adding additional noise W, problem is formulated as

$$\mathrm{Y} = \mathrm{X} + \mathrm{W},\ \mathrm{X}, \mathrm{Y}, \mathrm{W} \in \mathrm{R} \tag{1}$$

where W stands for zero-mean additive white Gaussian noise (AWGN).Since hyperspectral images are noisy additional noise is not added. The LRMA problem can be formulated as,

$$\{\psi(\mathrm{X})\} := 1/2\,\|\mathrm{Y} - \mathrm{X}\|^2 + \lambda \sum_{i=1}^{k} \phi(\sigma_i(\mathrm{X}), a)\} \tag{2}$$

where $k = \min(m, n)$, ϕ is the non-convex regularizer and it induces sparsity. In equation (2), the low rank matrix approximation is framed using a convex optimization with non-convex regularization. The non-convex penalty function is used in such a way that the convexity of objective function is not disturbed. Nuclear norm minimization (NNM) is a peculiar case of LRMA in which $\phi(\mathrm{X}) = |\mathrm{X}|$. It is convex and its solution is obtained by soft thresholding [8].

For the precise calculation of non-zero singular values and to ensure the condition of sparsity, non-convex penalty functions are used. The penalty function is parameterized by $a > 0$ [8].The non-convex penalty function $\phi = \mathrm{R} \rightarrow \mathrm{R}$satis?es the following assumptions

- It is continuous on R.
- Differentiable twice and should be greater than zero.
- It should satisfy the condition of symmetry.
- $\phi'(\mathrm{X}\,;0) > 0, \forall\ \mathrm{X} > 0$
- $\phi''(\mathrm{X}\,;0) \le 0, \forall\ \mathrm{X} > 0$
- $\phi'(0+\,; a) = 1$
- $\inf_{\mathrm{X} \neq 0} \phi''(\mathrm{X}\,; a) = \phi''(0^+\,; a) = -a$
- $\phi(\mathrm{X}\ ;0) = |\mathrm{X}|$

Rational penalty function which satisfies the above assumption is given as,

$$\phi(\mathrm{X}\,; a) = \frac{|\mathrm{X}|}{1 + a\,|\mathrm{X}|/2},\ a \ge 0$$

Non-convex penalty functions ensure sparsity better than previous methods like NNM. While adding the non-convex penalty functions, the convex nature of the objective function get disturbed. Hence the convexity of the objective function can be ensured by the following lemmas [8].

Lemma 1 : Let $\phi = R \to R$, non-convex penalty satisfies assumptions defined above, the function $S = R \to R$, is given as,

$S\,(X\,;a)\,;\phi(X\,;a) - |X|$ is differentiable twice and satisfies $-a \leq S''\,(X\,;a) \leq 0$.

Lemma 2 : Let, $\phi = R \to R$, be the non-convex penalty function $f: R \to$ the function given as

$$f(t) := \frac{1}{2}t^2 + \lambda\phi\,(t, a) \quad \text{For } \lambda > 0$$

Is strictly convex if $0 \leq a < \frac{1}{\lambda}$

The ELRMA problem is solved by the following theorem.

Theorem 1 : Let, $Y = U\Sigma V^T$ be the SVD of Y and $\phi = R \to R$ be the non-convex penalty function. If $0 \leq a < 1/\lambda$, then global minimizer of (2) is obtained as,

$$\overline{X} = U.\,\Theta\,(\Sigma; \lambda, a\,).\,V^T,$$

Where Θ represents threshold function associated with ϕ [8]

For unitary matrix U and V, $\phi(X) = \phi(UXV)$, by making use the unitary property of Frobenius norm and SVD [8]

$$\overline{X} = \underset{x}{\arg\min}\left\{\frac{1}{2}\|\,Y - X\,\|_X^F + \lambda\phi(X)\right\}$$

$$\overline{X} = \underset{x}{\arg\min}\left\{\frac{1}{2}\|\,\Sigma - U^T XV\,\|_F^2 + \lambda\phi(U^T XV)\right\}$$

$$= U.\underset{x}{\arg\min}\left\{\frac{1}{2}\|\,\Sigma - X\,\|_F^2 + \lambda\phi(X)\right\}V^T$$

Solving this we get,

$$\Theta(\Sigma) := \underset{x}{\arg\min}\left\{\frac{1}{2}\|\,\Sigma - X\,\|_F^2 + \lambda\phi(X)\right\}$$

3. PROPOSED SCHEME

Classification of images can be done using various means. Non-linear classification and feature detection of images can be done using support vector machines, thresholding technique and many other ways [9]. In this paper HIS images are denoised using ELRMA and denoised images are classified using different classification methods like SP, GURLS and random forest. Classification is analyzed using the overall accuracies obtained and Kappa coefficient. Overall accuracies is obtained by adding the number of correctly classified pixels and dividing by the total number of pixels. Kappa coefficient is calculated using

$$k = \frac{N\sum_{c=1}^{n} m_{c,c} - \sum_{c=1}^{n}(B_c S_c)}{N^2 - \sum_{c=1}^{n} B_c S_c}$$

- Where is the class number
- N represents total number of classified pixels

- $m_{c,c}$ is the number of pixels belonging to ground truth class c also belonging to the same class
- S_c is the total number of pixels classified to class c
- B_c is the total number of ground truth pixels in class c

Overall procedure is shown in Fig. (1).Various classification techniques are used to find out the accuracy of hyperspectral image denoising and are explained below.

A. Subspace pursuit(SP)

Subspace pursuit is a sparsity based classification algorithm. In Subspace pursuit (SP) sparse representation of testing sample is estimated using a dictionary matrix. Sparse representation of training samples are identified in accordance with the training sample from the dictionary matrix [3].

Algorithm for subspace pursuit

- Dictionary matrix $A = [A_1, A_2, A_3 ... A_T\}]$ of size $B \times T$ is taken as the training sample with K sparse levels
- y is the test pixel vector
- the problem is formulated as $x = \min \|y - Ax\|_2 ; \leq$ K, where x is the output variable.
- For the first iteration residue r is taken as y.
- Index $= \underset{i=1,2,\ldots T}{\arg\max} < y, A_i >$ is evaluated and first K max values are taken. Let the index values chosen be $[i_1, i_2, \ldots i_K]$.
- $\widetilde{A} = [A_{i_1}, A_{i_2}, \ldots A_{i_k},]$
- $x = (\widetilde{A}^t \ \widetilde{A})^{-1} \ \widetilde{A}^t \ y$
- After each iteration dictionary matrix is updated with the maximum values and corresponding x is calculated.
- Residual vector is calculated using $Ax - y$ Class labels are determined.

B. Gurls

The GURLS toolbox, which is an abbreviation for Grand Unified Regularized Least Squares, has been specialized to solve multiclass problem, it is based on regularized least square (RLS). It does not give sparse solutions. Different regularization parameters are used to calculate the classifier but it will not increase the complexity. Training and testing of classifiers can be done using linear as well as Gaussian kernels. Kernel which is used here is Gaussian. [10]

C. Random Forest

Random Forest is a cluster of trees, in which each classification tree corresponds to independently sampled random vector. When a data is given for classification, each tree does classification and forest chooses the one with maximum vote [11]. Generalized error of random forest is reduced when more number of trees are used. It mainly depends on strength and correlation of individual tress. It is very robust and efficient [12].Its major benefits are robustness against noise and outliers and it is very easy to use also, it can handle unbalanced data efficiently using balancing methods like Synthetic Minority Oversampling (SMOT), oversampling of training data and under sampling [13]. In SMOT oversampling of minority class is done by creating examples rather than replacing. Minority class is taken and synthetic examples are added from the nearest neighbors [14].

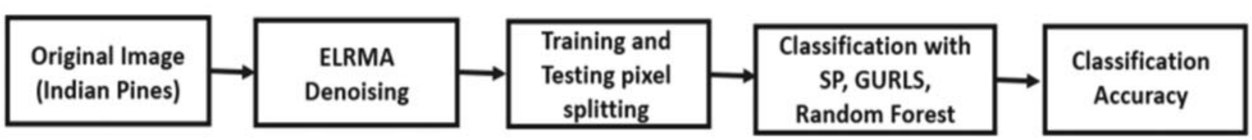

Fig. 1. Block diagram of overall procedure.

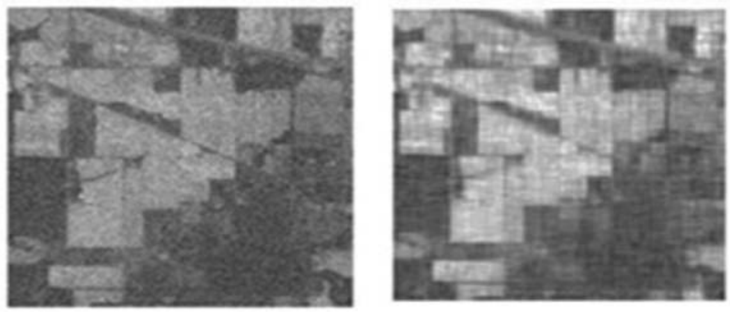

Fig. 2. 165th band (*a*).Without Denoising (*b*).With Denoising

Table 1. Classification Accuracies of Different Classifiers.

		ACCURACY (%)					
		SUBSPACE PURSUIT		***GURLS***		***RANDOM FOREST***	
Class	***Class name***	***Without denoising***	***With denoising***	***Without denoising***	***With denoising***	***Without denoising***	***With denoising***
1	Alfalfa	95.65	86.96	76.09	56.52	64.6	56.3
2	Corn-notil	58.33	64.57	79.48	83.33	64.2	68.1
3	Corn-mintill	30.96	45.18	67.71	66.02	48.8	45.2
4	Corn	46.84	48.12	49.37	60.76	27.6	79.2
5	Grass-pasture	74.53	82.61	86.54	89.86	79.2	85
6	Grass-trees	88.77	90	98.9	98.08	96.5	99.5
7	Grass-pasture-mowed	89.29	82.11	82.14	71.43	16	38.8
8	Hay-windrowed	88.7	95.04	94.35	98.74	94.49	99.5
9	Oats	100	100	65	30	11.1	21.8
10	Soybean-notill	39.81	60	73.46	74.18	70.9	70.2
11	Soybean-mintill	66.03	71.1	85.5	88.64	89.5	84.9
12	Soybean-clean	51.26	59.36	74.7	70.83	58.7	38.1
13	Wheat	99.51	98.07	98.05	99.02	97.7	89.9
14	Woods	93.04	95.86	94.55	97.87	95.1	95.8
15	Buildings-Grass-Trees-Drives	54.4	66.58	65.8	65.28	25.6	33.5
16	Stone-Steel-Towers	93.55	90.37	53.76	68.82	86.4	95.3
	Overall accuracy	65.49	71.9	82.28	84.2	76.01	80.11
	Kappa coefficient	0.6033	0.6843	0.7968	0.8187	0.7248	0.7749

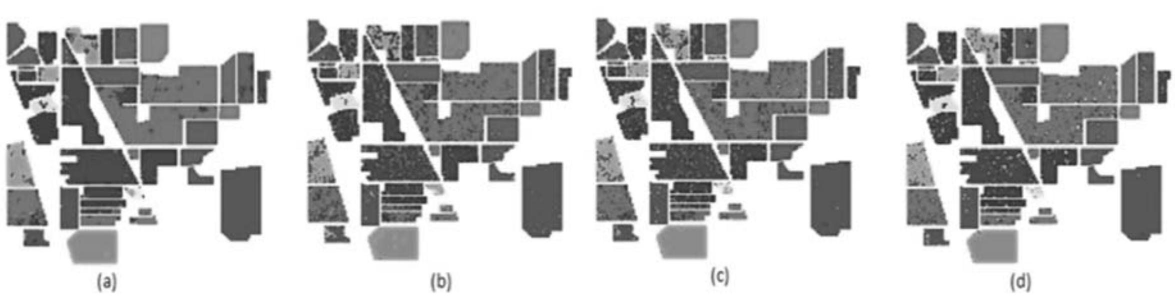

Fig. 3. Classification map of SP with denoising corresponding to training sets (*a*) 10% with accuracy 71.17 % (*b*) 20%with accuracy78.76% (*c*) 30%with accuracy 82.78% and (*d*) 40% with accuracies 86.78%

4. RESULT AND DISCUSSION

Denoising of hyperspectral images are done using Enhanced LRMA method which uses non-convex penalty function with convex objective function. Denoising of images are done using ELRMA. It is done multiple times to fix the regularization parameters. Hyperspectral images of AVIRIS sensor over Indian pines is used for experiment. Airborne Visible/Infrared Imaging Spectrometer (AVIRIS) is designed and developed by NASA Jet Propulsion Laboratory (JPL) for earth remote sensing. Scanning technique which is used is whiskbroom scanning [15]. It possess 220 bands and ground truth contain 16 classes. Data set has a dimension of 145 × 145 × 224.. Spatial accuracy is 20m/pixel. Dataset is classified using different methods of classification before preprocessing and after preprocessing. Sparsity based classifier subspace pursuit, kernel based classifier GURLS and random forest are the different classifiers used. Classification is done for data with and without preprocessing. Results obtained shows an increase in overall accuracy and kappa coefficient for data which is denoised than without any preprocessing. For the experiment 10, 20, 30 and 40 percentage of training pixels are taken randomly and whole sample is given as testing and classification is done.

Enhanced low rank matrix approximation is applied to data with a regularization parameter of λ ranges from 100 to 200 and between 0.001 and 0.01. After many iterations parameter is fixes as $\lambda = 130$ and $a = 0.005$ for best denoising results. Fig.2 shows the denoising effect in 165th band. After applying denoising, classification is done. Over all accuracy, class wise accuracy and kappa coefficient for all the classification methods used are given in Table 1. From the table it is clear that for 10 percent training data there is an accuracy of 71.17 percent and for without processing overall accuracy is 65.49 percent. There is an increase in accuracy of about 5%. Similarly the process is repeated for 20, 30 and 40% and there is considerable increase in accuracies.

Another classification method used is GURLS. It showed considerable increase in accuracy when compared to sparsity based classification. For 10% training data accuracy obtained after preprocessing is 82 % and without denoising is 84%. Similar process is repeated for 20, 30 and 40%. For 40% over all accuracy obtained is 94.5% after denoising.

Random forest is a decision tree based classifier. Classification is done using WEKA tool which is developed by University of Waikato, New Zealand. Our dataset is multiclass and unbalanced. In order to handle unbalanced data Weka provides an algorithm called SMOT. It is applied as a preprocessing step before classification and classification is done using random forest. Without applying SMOT the accuracies which is obtained with denoising shows an improvement when compared to without denoising but from the class accuracies it is clear that minority classes have very small class accuracy and so there is misclassification. In order to tackle this issue SMOT is applied. Results obtained for with denoising shows considerable increase in overall accuracy as well as class accuracies for minority classes when compared with that of without denoising from 76% for without denoising to 80% for with denoising.

5. CONLUSION

ELRMA denoising with non-convex penalty functions in hyperspectral data is an effective denoising technique for hyperspectral images. Since hyperspectral images have high noises and redundancy it is necessary to apply denoising in order to retrieve the useful data. ELRMA does this role of retrieving information by avoiding noises in an effective way with a convex objective function. Denoising is applied to each bands and classification is done using subspace pursuit algorithm, GURLS and random forest. All classification methods shows considerable increase in accuracies and kappa coefficient for data after denoising than compared to before denoising. Among the various classification methods used maximum accuracy of 84% for 10 % training data is given by GURLS. From this observation it is concluded that ELRMA denoising is a good denoising technique for hyperspectral images.

6. REFERENCES

1. G. Chen and S. Qian, “Denoising of hyperspectral imagery using principal component analysis and wavelet shrinkage,” IEEE Trans. Geosci.Remote Sens., vol. 49, no. 3, pp. 973-980, Mar. 2011.

2. J. Bi oucas-Dias, A. Plaza, G. Camps-Valls, P. Scheunders, N. Nasrabadi, and J. Chanussot, “Hyperspectral remote sensing data analysis and future challenges,” Geoscience and Remote Sensing Magazine, IEEE, vol. 1, no. 2, pp. 6-36, 2013.
3. C. Aswathy, Sowmya, V., and Soman, K. P., "ADMM based hyperspectral image classification improved by denoising using legendre fenchel transformation", Indian Journal of Science and Technology, vol. 8, no. 24, 2015.
4. He, Wei, et al. “Hyperspectral image denoising via noise-adjusted iterative low-rank matrix approximation.” IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 8.6 (2015): 3050-3061.
5. Haridas, Nikhila, et al. “Hyperspectral Image Denoising Using Legendre-Fenchel Transform for Improved Sparsity Based Classification.” Intelligent Systems Technologies and Applications. Springer International Publishing, 2016. 521-528.
6. Kavitha Balakrishnan, Sowmya V., and K. P. Soman. “Spatial preprocessing for improved sparsity based hyperspectral image classification.“ International Journal of Engineering Research and Technology. Vol. 1. No. 5 (July-2012). ESRSA Publications, 2012.
7. Baassou, Belkacem, et al. “Hyperspectral image classification based on iterative Support Vector Machine by integrating spatial-spectral information.” 2013 IEEE International Geoscience and Remote Sensing Symposium-IGARSS. IEEE, 2013
8. Parekh, Ankit, and Ivan W. Selesnick. “Enhanced low-rank matrix approximation.” IEEE Signal Processing Letters 23, no. 4 (2016): 493-497.
9. D.Suganya , K.Krishnaveni, Brain Image Segmentation Methods using Image Processing Techniques to Analysis ADHD . IIR Journal of Computing Paradigms Volume: 01 Issue: 01, June 2016, Pages: 39-42
10. Tacchetti, Andrea, et al. “GURLS: a least squares library for supervised learning.” Journal of Machine Learning Research 14.1 (2013): 3201-3205.
11. Yi, Zhenping, and Jingchang Pan. “Application of random forest to stellar spectral classification.” Image and Signal Processing (CISP), 2010 3rd International Congress on. Vol. 7. IEEE, 2010.
12. Breiman, Leo, and Adele Cutler. “Random forests-classification description.” Department of Statistics, Berkeley 2 (2007).
13. Dittman, David J., Taghi M. Khoshgoftaar, and Amri Napolitano. “The effect of data sampling when using random forest on imbalanced bioinformatics data.” Information Reuse and Integration (IRI), 2015 IEEE International Conference on. IEEE, 2015.
14. Chawla, Nitesh V., et al. “SMOTE: synthetic minority over-sampling technique.” Journal of artificial intelligence research 16 (2002): 321-357
15. kgowey, “Airborne visible / infrared imaging spectrometer,”. [Online]. Available: http://aviris.jpl.nasa.gov/. Accessed: Jun. 28, 2016.