International Journal of Emerging Technology and Advanced Engineering
Website: www.ijetae.com (ISSN 2250-2459, ISO 9001:2008 Certified Journal, Volume 6, Issue 12, December 2016)

Threshold Selection to Minimize the Redundancy in Hyperspectral Band Selection

M. Kalamuthumari[1], R.Vidhya[2]

[1]*Research Scholar, Institute of Remote Sensing, Anna University, Chennai, India*
[2]*Professor, Institute of Remote Sensing, Anna University, Chennai, India*

***Abstract*— Hyperspectral band selection is the efficient way to improve the classification accuracy. Unsupervised rank based band selection methods are preferred due to its less computational complexity. The bands are ranked based on their information content and redundancy among the top ranked are removed using divergence measure. The divergence value vary depends on the band image size and content. A procedural way of selection is adapted in the study. nMax divergence value of band images are used for analysis and the fraction of Avg, Max and Max-Avg divergence values are also tested for the threshold parameter identification. Two study area site with different spatial and spectral size with different number of feature classes are used for the study. The experimental results proved that (Max-Avg) divergence value is the suitable parameter for threshold selection to reduce the redundancy among the prioritized bands.**

***Keywords*— band selection, divergence measure, entropy, hyperspectral image, redundancy**

I. INTRODUCTION

Hyperspectral imaging collects data to generate a "data cube" that can reveal objects and information which conventional multispectral scanners cannot pick up. The hyperspectral leads to an improvement in the performance of detection and classification process due to its enhanced ability to identify materials on the basis of their spectral signatures. However, using hyperspectral data is much more complex than multispectral data (Thenkabail et al., 2004). Large volume of hyperspectral data is leading to numerous complex technical issues such as data storage volume, transmission bandwidth, atmospheric corrections, computing bottlenecks in data analysis, and new algorithms for data utilization.

Due to high correlation between adjacent bands, the data redundancy exists in hyperspectral data (Amarsaikhan et al., 1999). The information increases greatly with the increase of band number. Because of the existence of band correlation and data redundancy, the number of image channels and the number of information dimensions are not equal (Xiaoguang et al., 2002).

It is therefore advantageous to remove bands that convey little or no discriminatory information. Generally, the band selection methods consider the following rules (Shaw et al., 2003). *1)* The amount of information of selected bands should be large; *2)* The selected bands should have maximum class separability; *3)* The correlation among selected bands should be small.

Band Subset Selection methods select the group of spectral bands that maximize the class separability. The ranking methods select band subset from the hyperspectral dataset to represent the data based on optimality criterion such as variance (Chang et al., 1999), mutual information (Guo et al., 2006), non-Gaussianity (Chang et al., 2006 and Chang 2007) and signal-to-noise ratio (Sun et al., 2014). Many band selection algorithms are proposed to give priority score to spectral band depending on their information content (Bajcsy 2004; keshava 2004; Sarhrouni et al., 2012; and Sun et al., 2014). The uniqueness of spectral band is characterized based on the criterion. The main advantage of the ranking-based unsupervised band selection methods is the algorithm need to be executed only once to find band subset. These methods are stable and low computational cost is involved. The above discussed methods were not considered the correlation between the bands during the sorting procedure.

There are many criterias that can be used to measure the distance between two distributions, like the correlation coefficient, the Bhattacharya distance, and the Kullback-Leibler Divergence (KLD). The correlation coefficient assumes Gaussian distribution, where as the other two measures can be applied to other types of distributions (Zhouyu et al., 2006). Redundant bands has be identified and removed even if they have high relevant information for better accuracy.

The earlier work by Chang et al., (1999) and by Du et al., (2007) used divergence for band selection and the criteria for the threshold divergence value are left to users. The selection of optimized threshold becomes crucial without a priori knowledge. If the threshold is too low, many redundant bands will still be included, reduces efficiency.

On the other hand, if the threshold is too high, some important features may be lost. Thus, an optimized threshold for redundancy reduction should be selected so as to properly balance feature loss with redundancy, can be efficiently done by considering the overall feature information and the redundancy information at the same time.

In this approach, both the amount of information and band decorrelation are considered. The high informative bands are ranked using entropy measure and to keep bands with low correlation amongst themselves, divergence metric is used. If the divergence between two bands is below a prescribed threshold, the band with lower priority is removed. Through classification experiments, it is concluded that the classification precision is significantly improved using this method.

Instead of directly measuring the redundancy between individual bands, entropy is used to measure the overall information carried by the selected spectral bands and divergence is used to detect redundancy contained within the selected bands. In this paper a simplified method to select the threshold divergence from the image itself is discussed and proved that the 50% of selected bands gives moreover the same accuracy as that of whole data set.

II. Previous Work

Effective use of features of hyperspectral data and the selection of suitable bands are especially significant for improving classification accuracy (Rui et al., 2005). Some algorithms compute the correlations between the spectral bands and reserve the bands with least correlations or reject some highly correlated bands to reduce the redundancy (Rui et al., 2005 and Sotoca et al., 2007). In most cases, high spectral correlation of hyperspectral image data enables us to reduce the feature dimensionality without substantial loss of classification accuracy (Martinez-Uso et al., 2007). The bands that have the abundance information, minimized relativity, more discrepancy and separability are considered to be the best bands (Xiaoguang et al., 2002). In the past, many criteria have been proposed for band selection i.e., to find bands that are crucial and significant in terms of information conservation.

For instance, distance measures such as Bhattacharyya distance, Jeffreys–Matusita distance (Simin et al., 2009 and Ifarraguerri et al., 2004) , angle measures (Keshava 2001), Correlation measures like Spectral Correlation Mapper (Rui et al., 2005), information-theoretic approaches such as divergence, transformed divergence, mutual information (Chang et al., 1999), eigenanalysis (Bajcsy et al., 2004), information-entropy-based band selection (koosanit et al., 2012) and class-separability based (Xiaoguang et al., 2002) have been applied to Hyperspectral images for optimal band selection. Band selection based on criteria such as high-order moments (Chang et al., 1999), information theory (Wang et al., 2004), novel entropy-based (Wang et al., 2004) and Mutual Information (Guo et al., 2006 and Martinez-Uso et al., 2007) select bands by using the information content in each band. Du (2003) used high-order moments for band ranking and divergence for band decorrelation. The divergence takes the dissimilarity that exists among various selected bands, and it is a simple and efficient measurement of statistical class separability used in pattern recognition (Swain et al., 1978). The divergence takes into account the correlation that exists among the various selected bands and influences the classification capabilities of the spectral bands that are selected (Du 2003). The concept of entropy has been widely used in data compression to measure information content of a source, using the uncertainty as a measure to describe the information contained in a source. In image analysis, the entropy-based threshold considers the image as an information source with a probability vector described by its grey-level image histogram (Chang et al., 2006 and Barbieri et al., 2011).

III. Methodology

A. Band Selection Method

Hyperspectral data is a three dimensional array with the width and length corresponding to spatial dimensions and the spectral bands as the third dimension, which are denoted by I,J and L in sequence. R is the image cube and $R_{ij} \in R_L$ is an observation vector of a single pixel. The information quantity available in a band is an imperative parameter used in evaluation of band.

International Journal of Emerging Technology and Advanced Engineering
Website: www.ijetae.com (ISSN 2250-2459, ISO 9001:2008 Certified Journal, Volume 6, Issue 12, December 2016)

The entropy is often used to show information quantity. Here, an algorithm based on information content in bands is proposed. Entropy is used to access the information content and prioritise the bands. The redundancy is removed using an adaptive threshold divergence which is based on the image content.

1) Information Entropy: According to Shannon (1948), entropy is the only function which satisfactorily measures the confidence of information. The information entropy measure is defined below.

$$H(X) = -\sum_{i=1}^{n} p(x_i)\log_b p(x_i) \qquad (1)$$

Where $p(x_i)$ is the probability mass function of outcome x_i.

H is the entropy measure, p is the probability density function of reflectance values in a hyperspectral band and n is the number of distinct reflectance values. Some methods [13] directly use entropy as a criterion for band selection. In these methods, the entropy is calculated to estimate the level of information contained in each individual band or wavelength interval. Generally, if the entropy value H is high then the amount of information in the data is large. Thus, the bands are ranked in the ascending order from the band with the highest entropy value to the band with the smallest entropy value i.e. large amount of information to small amount of information (Yin et al., 2010 and chang et al., 2011).

2) Divergence: The widely used metrics to select the distinctive bands similarity metric such as distance, correlation, divergence etc. and the measurement is taken on each pair of bands (Yin et al., 2010). According to Chang et al., (1999), divergence performs very well in terms of capturing similarity and dissimilarity between two images and can be used for band subsetting purpose. The band image with larger entropy contains target information, and should be preserved. Since the band prioritization does not consider the spectral correlation, divergence is used to decorrelate prioritized bands.

The number of bands selected depends upon the threshold value. If the divergence between two bands is below a prescribed threshold, the band with lower priority is removed from the list of bands. In information theory (Cover et al., 1991), a criterion called "divergence," is used to measure the discrepancy between any two probability distributions.

The Kullback Leibler divergence (KL-divergence) is a natural distance function from the "true" probability distribution, p, to the "target" probability distribution, q.

Let p=$\{p_k\}_{k=1}^{K}$ and q= $\{q_k\}_{k=1}^{K}$ be two image arrays which are generated by stacking one row after another, where K is the number of pixels in each image. Their divergence, denoted by D(p,q) is defined as

$$D(p,q)=L(p;q)+L(p;q) \qquad (2)$$

Where L(p;q) is the divergence of p with respect to q defined by

$$L(p;q)=\sum_{k=1}^{K} \log\left(p_k/q_k\right) \qquad (3)$$

and L(q;p) is the divergence of q with respect to p defined by

$$L(q;p)=\sum_{k=1}^{K} \log\left(q_k/p_k\right) \qquad (4)$$

Since L(p;q)≠ L(q;p) both are used (Cover et al., 1991).

The divergence criterion is applied to the prioritized set of bands to select the significant bands and thereby identify optimal band set. In other words, the redundant bands are removed and optimal bands are selected without a priori knowledge. Thus, a good threshold for redundancy reduction is selected so as to optimally balance feature loss with redundancy.

B. Threshold Selection Scheme

Correlation coefficient between bands shows the correlation of different bands. Lower the correlation coefficient, lesser the redundancy of information is included. Unlike correction coefficient value which lies between -1 to +1, the divergence value vary depends on the band image. In most of the cases 0.7 or 0.8 correlation coefficient value is taken and highly correlated bands are removed. The reason for divergence measure instead of correlation coefficient is correlation between adjacent bands is significant under many circumstances; it affects the results optimality (Oh et al., 2004).

Guo et al., (2006) proposed an algorithm to eliminate redundancy with this rule "If a band decreases the error probability, it will be retained even if it contains redundant information". To reduce the redundancy and at the same time to retain high informative bands, the bands are sorted based on entropy. It is observed that some adjacent bands have the same entropy value. Only one band is selected from both bands having the same value. Divergence is not applied for adjacent bands where as applied on prioritized bands to remove the redundancy.

International Journal of Emerging Technology and Advanced Engineering
Website: www.ijetae.com (ISSN 2250-2459, ISO 9001:2008 Certified Journal, Volume 6, Issue 12, December 2016)

In this band selection method, without a priori knowledge the bands are selected. The thumb rule is that number of bands should be high enough to get high accuracy. With pair wise performance nearly 50% of bands can be selected. The threshold selection is generally based on experiments only. The selection of threshold varies with the input image size. The thresholds are selected based on the divergence among the bands. The divergence between the prioritized bands is sorted in descending order and the nth maximum value is used as a threshold. Atleast one band is selected among the two adjacent prioritized bands which is having divergence value above the threshold or otherwise both the bands are selected.

C. Classification

In order to evaluate the amount of information and class separability in the selected bands, a supervised classification algorithm, called Support Vector Machine can be applied. SVM classification is much more effective than other conventional nonparametric classifiers in terms of classification accuracy, computational time, and stability to parameter setting. A detailed description about SVM is given by Burges (1998) and an overview in the context of remote sensing is given by Huang et al., (2002).Since SVMs are said to exhibit low sensitivity to the Hughes phenomenon, this method is adopted for the study. A classification scheme is defined based on the most important land use types present in the study area. For each land-use class, ground truths are manually collected with reference to in situ knowledge and Imagery from Google Earth is used as a support tool.

When the pixel-level ground truth is unavailable, the classification maps from all the original bands can be considered as ground truth, and those from the selected bands are compared with them (Du et al., 2008). This is under the assumption that, using all the original spectral bands (after bad band removal), the best or at least satisfying classification performance can be provided. For classes with similar but separable spectra, is a reasonable assumption (Platt et al., 2004). Such a method based on image similarity provides quantitative evaluation even in an unsupervised situation or in the lack of pixel-level ground truth.

Accuracy assessment is performed based on Overall Accuracy (OA) and Kappa analysis. An estimation of the remotely sensed classification agrees with the reference data classification performance is assessed at class level.

IV. EXPERIMENT

In the EO-1 Hyperion data of about 242 spectral bands, the bands with high sensor noise, duplicated bands in the NIR spectrum due to the detector materials change, and heavy atmospheric water absorption bands are removed. After initial radiometric and atmospheric correction, the data are geo referenced and converted to surface reflectance. Two sets of data set are considered in the experiment:

(i) *Subset #1.*The data set #1 with size of 232 by 190 pixels is a subset of Hyperion data of Muthupet lagoon, Tamil Nadu, India acquired in September 2012. It contains 106 spectral bands after removal of noisy bands. Seven ground truth classes considered in experiment are Aquaculture, Intertidal, Mangroves very dense, Mangroves dense, Mangroves sparse, Mangroves plantation and Mud flat vegetation

(ii) *Subset #2.*The data set #2 with size of 128 by 210 pixels which is a subset of Hyperion data of Pichavaram, Tamil Nadu, India acquired in January 2013. It contains 91 spectral bands after removal of noisy bands. Nine ground truth classes considered in experiment are Aquaculture, Water, Sand, Mangroves dense, Mangroves sparse, Agriculture, Plantation, Fallow land and Mud flat vegetation

The proposed band selection method is applied to the test images. The criterion for calculating the divergence threshold is based on sorted divergence values from the images. The value of n^{th} maximum is decided by the user and nine threshold values are selected for experiment. The 5^{th}, 7^{th}, 10^{th} and15^{th} max threshold values are used to find the bands. The experiment is repeated for 20^{th}, 30^{th}, 40^{th} and 50^{th} Maximum divergence. The criteria, Divergence and the number of bands selected are given in table 1.

TABLE I
Max Criteria Threshold and Number of Bands

Criteria	Muthupet		Pichavaram	
	Threshold Divergence	Number of Bands	Threshold Divergence	Number of Bands
5th Max	29469400	54	15374000	47
7th Max	28715500	54	13757100	47
10thMax	25935900	55	9087770	48
15thMax	22890200	59	6930510	50
20thMax	20843700	60	4028250	53
30thMax	13119700	64	2414330	57
40thMax	9646940	66	1704660	60

EXPLORING RESEARCH AND INNOVATIONS

International Journal of Emerging Technology and Advanced Engineering
Website: www.ijetae.com (ISSN 2250-2459, ISO 9001:2008 Certified Journal, Volume 6, Issue 12, December 2016)

It may be useful for the user to avoid restricting the threshold selection. Keeping that in mind the average divergence value and ratio of its value are used as threshold value.

Similarly (maximum-Avg) value is also tested. The criteria, Divergence and the number of bands selected are given in table 2.

TABLE II
Adaptive Threshold and Number of Bands

Criteria	Muthupet		Pichavaram	
	Threshold Divergence	Number of Bands	Threshold Divergence	Number of Bands
AvgDiv	9950040	61	17285000	45
AvgDiv/2	4975020	74	8642520	48
AvgDiv/3	3316680	78	5761680	50
(maximum-avgdiv)	34984900	52	3485300	54
(maximum-avgdiv)/2	17492500	61	1742650	60
(maximum-avgdiv)/3	11661600	64	1161770	66

From the result it is clear that the threshold either avg based or Max-Avg can be used for band selection in this method. The given band selection method does not require reference data as that of mutual information based method. It is a simple way of selecting bands even though pair wise is applied, it is not on the adjacent bands therefore the advantage of hyperspectral is utilised where in alternate band selection method selects bands without considering the band information.

As mentioned previously, entropy based band prioritization does not take care of spectral correlation, the divergence measure is used to remove the redundancy. It is observed that the adjacent bands are also selected based on the information content and the divergence value. The selection method is not considering the band adjacency in the spectrum and the selected bands are heighted in Figure 1

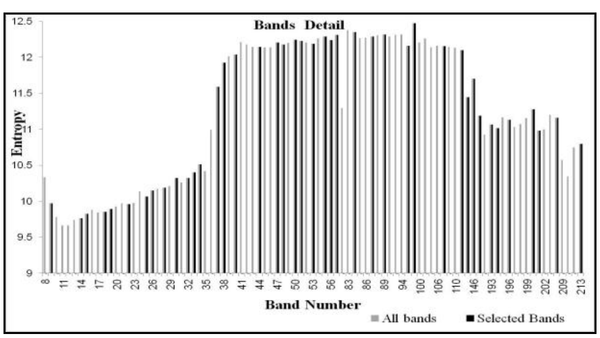

Fig. 1 Selected bands Vs All bands

Since this method is based on the image size and the distribution of features, the band selection method is tested on two sets of data. The threshold selection is adaptive in nature. One of the significant weaknesses of the mutual information based method is that it relies heavily on the availability of a given reference map. Considering the limitation of its availability, it becomes prohibitive to apply this technique in practice. In terms of the first difference, the proposed method directly estimates the utility of each band to classification rather than the correlation between two band images. This avoids the problem of removing bands where two spectral bands are adjacent but are also highly valuable to classification. For the second difference, the proposed method does not rely on the availability of a given reference map, and is therefore suitable for more applications. The classification accuracy for the two subset images are given in figure 2 and figure3.

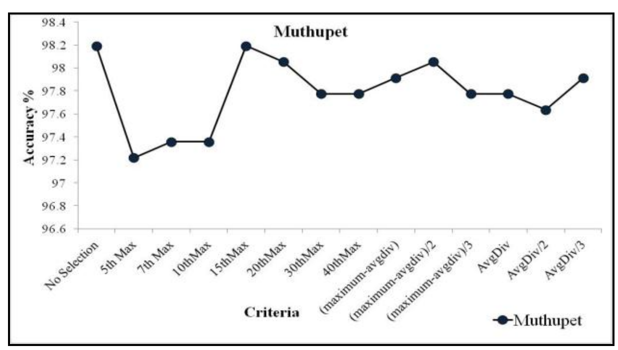

Fig. 2 Muthupet subset image Classification accuracy for different Threshold values

International Journal of Emerging Technology and Advanced Engineering
Website: www.ijetae.com (ISSN 2250-2459, ISO 9001:2008 Certified Journal, Volume 6, Issue 12, December 2016)

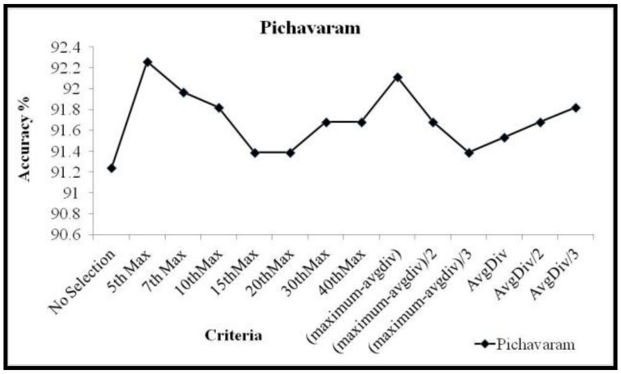

Fig. 3 Pichavaram subset image Classification accuracy for different Threshold values

TABLE III
Number of Bands and the change in Accuracy

Criteria	Muthupet			Pichavaram		
	Number of Bands	Overall Accuracy %	Change in Accuracy %	Number of Bands	Overall Accuracy %	Change in Accuracy %
No Selection	106	98.1919	0	91	91.2409	0
5th Max	54	97.2184	-0.9735	47	92.2628	1.0219
7th Max	54	97.3574	-0.8345	47	91.9708	0.7299
10thMax	55	97.3574	-0.8345	48	91.8248	0.5839
15thMax	59	98.1919	0	50	91.3869	0.146
20thMax	60	98.0529	-0.139	53	91.3869	0.146
30thMax	64	97.7747	-0.4172	57	91.6788	0.4379
40thMax	66	97.7747	-0.4172	60	91.6788	0.4379
(maximum-avgdiv)	52	97.9138	-0.2781	45	92.1168	0.8759
(maximum-avgdiv)/2	61	98.0529	-0.139	48	91.6788	0.4379
(maximum-avgdiv)/3	64	97.7747	-0.4172	50	91.3869	0.146
AvgDiv	61	97.7747	-0.4172	54	91.5328	0.2919
AvgDiv/2	74	97.6356	-0.5563	60	91.6788	0.4379
AvgDiv/3	78	97.9138	-0.2781	66	91.8248	0.5839

The accuracy curve can be divided in to two segments; S1 is based on the maximum divergence which is controlled by the user and S2 is adaptive threshold. In both the segments, at the beginning of the curve, a slow rise is observed and then decrease in accuracy where as the number of bands selected for the criteria is increased. (Refer table1). When the selected number of bands are in the range of 50-60% of the original set gives almost the same accuracy as that of whole data bands.

The threshold analysis reveals that the selection should be adaptive in nature since it varies with the size of the image. The main advantage of adaptive threshold is that without a prior knowledge the bands can be selected. The detailed status of number of bands selected and the change in classification accuracy with respect to the criteria are given in table 3. The classification accuracy gained using all bands is used as a reference for the change calculation.

International Journal of Emerging Technology and Advanced Engineering
Website: www.ijetae.com (ISSN 2250-2459, ISO 9001:2008 Certified Journal, Volume 6, Issue 12, December 2016)

The minimum number of bands selected by the Maximum-Average divergence criterion and the change in classification accuracy is -0.2781% in case of subset #1 and 0.8759% for subset#2.

V. CONCLUSIONS

This paper presented an adaptive threshold for band divergence based on image content. Since most information-based band selection approaches use only the entropy or the mutual information with a given reference map.

The proposed method revises them by devising an adaptive threshold measurement for band decorrelation. This scheme proved that it is more suitable for applications where the ground reference is not available. As shown by hyperion dataset, the proposed method could effectively reduce the redundant bands with a minor classification accuracy loss .i.e. almost 50-45% bands selection performed with less than 1% accuracy loss. At the same time standing classification accuracy is preserved, by reducing the processing time, storage space and communication bandwidth of the subset data.

REFERENCES

[1] Amarsaikhan, D., and Ganzorig, M., 1999; Different approaches in feature extraction for hyperspectral image classification. Proceedings of the 20th Asian Conference on Remote Sensing, Hong Kong, 22–25 November (Hong Kong: Joint Laboratory for GeoInformation Science of the Chinese Academy of Sciences and the Chinese University of Hong Kong), pp. 434–438

[2] Bajcsy.P and Groves.P,2004; Methodology for Hyperspectral Band Selection, Journal of Photogrammetric Engineering and Remote Sensing, vol. 70, pp. 793-802

[3] Barbieri, A. L., de Arruda, G. F., Rodrigues, F. A., Brunoa, O. M., Costa, L. da F. 2011; An entropy based approach to automatic image segmentation of satellite images, Physica A 390, 512–518

[4] Bruce. L.M., Koger. C.H. and J. LI.J, 2002 Dimensionality Reduction of Hyperspectral Data Using Discrete Wavelet Transform Feature Extraction. IEEE Transactions on Geoscience and Remote Sensing, Vol .40, No. 10, pp .2331-2338.

[5] Burges. C. J. C.,1998; A tutorial on support vector machines for pattern recognition , Data Mining Knowledge Discover., vol. 2, no. 2, pp. 121–167.

[6] Chang, C.-I, Du, Y., Wang, J., Guo, S.-M., Thouin, P.D. 2006; Survey and comparative analysis of entropy and relative entropy thresholding techniques, IEE Proceedings of Vision, Image and Signal Processing, 153 (6), pp. 837-850

[7] Chang. C and Du,Q. 1999; A Joint Band Prioritization And Band-Decorrelation Approach To Band Selection For Hyperspectral Image Classification, IEEE Transactions on Geoscience and Remote Sensing, vol. 37, no. 6, pp. 2631–2641.

[8] Chang. C.-I. and Wang. S., 2006; Constrained band selection for hyperspectral imagery," IEEE Trans. Geosci. Remote Sens., vol. 44, no. 6, pp. 1575– 1585.

[9] Chang.C and Safavi.H , 2011 Progressive dimensionality reduction by transform for hyperspectral imagery , Pattern Recognition 44 (2011) pp 2760–2773

[10] Chang.C.I,2013; Hyperspectral Data Processing: Algorithm Design and Analysis. New York, NY, USA: Wiley, 2013.

[11] Cover.T.M. and Thomas. J. A., *Elements of Information Theory*. New York: Wiley, 1991

[12] Du. Z, Jeong. M.K, and Kong. S.G, 2007; Band Selection of Hyperspectral Images for Automatic Detection of Poultry Skin Tumors, IEEE Transactions on Automation Science and Engineering, vol. 4, no. 3.

[13] Du.Q, 2003; Band selection and its impact on target detection and classification in hyperspectral image analysis, Adv. Tech. Anal. Remotely Sensed Data, pp. 374–377, 2003.

[14] Du.Q, and Yang.H, 2008; Similarity-Based Unsupervised Band Selection for Hyperspectral Image Analysis, IEEE Geoscience And Remote Sensing Letters, vol. 5, no. 4, pp 564-568

[15] Fu. Z, Caelli. T, Liu.N, and Robles-Kelly. A, 2006; Boosted Band Ratio Feature Selection for Hyperspectral Image Classification, Proceedings of the 18th International Conference on Pattern Recognition (ICPR'06) 2006

[16] Guo.B, Gunn. S.R., Damper. R.I., and Nelson. J.D.B,2006; Band Selection for Hyperspectral Image Classification Using Mutual Information, Geoscience and Remote Sensing Letters, IEEE Volume 3, Issue 4, pp:522 – 526

[17] Hongzhi.W and Angelopoulou,2004; Average Normalized Information: a Novel Entropy-Based Band Selection Method" Stevens Institute of Technology Technical Report, CS, 13, 2004

[18] Huang.C, Davis. L. S., and Townshend. J. R. 2002; An assessment of support vector machines for land cover classification, International Journal of Remote Sensing .,vol. 23, no. 4, pp. 725–749.

[19] Ifarraguerri.A, Prairie. M.W.,2004; Visual method for spectral band selection, Geoscience and Remote Sensing Letters, IEEE Vol 1, no 2, Page(s):101 – 106

[20] Jia.X, Kuo.B, and Crawford.M. 2013; Feature mining for hyperspectral image classification, Proc. IEEE, vol. 101, no. 3, pp. 676–697, Mar. 2013.

[21] Keshava.N 2004; Distance metrics and band selection in hyperspectral processing with applications to material identification and spectral libraries, IEEE Trans. Geosci. Remote Sens., vol. 42, no. 7, pp. 1552–1565.

[22] Keshava.N, 2001; Best Bands Selection for Detection in Hyperspectral Processing, IEEE International Conference on, 2001 Volume 5, 7-11 May 2001 Page(s):3149 - 3152 vol.5

[23] Koonsanit.K, Jaruskulchai.C, and Eiumnoh.A,2012; Band Selection for Dimension Reduction in Hyper Spectral Image Using Integrated Information Gain and Principal Components Analysis Technique, International Journal of Machine Learning and Computing, Vol. 2, No. 3.

[24] Liu, Yi; Li, Jun; Du, Peijun; 2016 Class-Oriented Spectral Partitioning for Remotely Sensed Hyperspectral Image Classification IEEE Journal Of Selected Topics In Applied Earth Observations And Remote Sensing pp1-21 2016

[25] Martınez-Uso.A, Pla.F, Sotoca.J.M, and Garcıa-Sevilla.P 2007; Comparison of Unsupervised Band Selection Methods for Hyperspectral Imaging., Part I, LNCS 4477, pp. 30–38, Springer-Verlag Berlin Heidelberg

[26] Oh, I.-S., Lee, J.-S., Moon, B.-R.2004; Hybrid genetic algorithms for feature selection. IEEE Trans. Pattern Analysis and Machine Intelligence 26, 2004:1424 -1437

[27] Platt R. V. and Goetz. A. F. H. 2004; A comparison of AVIRIS and Landsat for land use classification at the urban fringe, Photogrammetric Engineering and Remote Sensing., vol. 70, no. 7, pp. 813–819.

[28] Rui. H. and Mingyi. H. 2005; Band Selection Based On Feature Weighting For Classification of Hyperspectral Data IEEE Geoscience and Remote Sensing Letters, 2(2):156–159.

[29] Sarhrouni.E, Hammouch.A, and Aboutajdine.D 2012; Band selection and classification of hyperspectral images using mutual information: An algorithm based on minimizing the error probability using the inequality of fano, in Proc. IEEE Int. Conf. Multimedia Comput. Syst., pp. 155–159.

[30] Shannon. C. E., 1948; A Mathematical Theory of Communication, Bell System Technical Journal, Vol. 27, pp. 379-423 and pp. 623-656, 1948.

[31] Shaw. G.A and Burke.H, 2003; Spectral imaging for remote sensing, Lincoln Lab. J., vol. 14, no. 1, pp. 3–28, 2003.

[32] Shen. S. S. and Bassett. E. M., Information theory based band selection and utility evaluation for reflective spectray systems, Proceedings of SPIE, vol. 4725, 2002.

[33] Simin.C, Rongqun.Z, Wenling.C, and Hui.Y2009; Band Selection of Hyperspectral Images Based on Bhattacharyya Distance Wseas Transactions On Information Science And Applications , no 7, Vol 6, pp. 1165-1175 .

[34] Sotoca.J.M, Pla.F, and Sanchez. J. S.2007; Band Selection in Multispectral Images by Minimization of Dependent Information, IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews, 37(2):258–267.

[35] Sun.K, Geng.X, Ji.L, and Lu.Y. 2014; A new band selection method for hyperspectral image based on data quality, IEEE J. Sel. Topics Appl. Earth Observations Remote Sens., vol. 7, no. 6, pp. 2697–2703.

[36] Swain.P and Davis S. M., Remote Sensing: The Quantitative Approach. New York: McGraw-Hill, 1978

[37] Thenkabail, P.S., Enclona,E.A., Ashton,M.S., and Van Der Meer. B 2004; Accuracy assessments of hyperspectral waveband performance for vegetation analysis applications, Remote Sensing of Environment vol.91(no.3-4)

[38] Xiaoguang.J, Wang. L, and Changyao.2002; Spectral Characteristics and Feature Selection of Hyperspectral Remote Sensing Data, Remote Sensing Remote Sensing Technology and Application, vol. 17, 2002, pp. 59-65

[39] Yin.J, Wang.Y and Zhao.Z, 2010; Optimal Band Selection for Hyperspectral Image Classification Based on Inter-class Separability, IEEE Symposium on Photonics and Optoelectronic 2010