Journal of Basic and Applied Engineering Research
Print ISSN: 2350-0077; Online ISSN: 2350-0255; Volume 2, Number 16; April-June, 2015 pp. 1405-1410

http://www.krishisanskriti.org/jbaer.html

Kernel based Fuzzy Classification–SMIC Tool

Prashansa Agarwal[1] and Anil Kumar[2]

[1]Meerut Institute of Engineering and Technology, Meerut
[2]Indian Institute of Remote Sensing, ISRO, Dehradun
E-mail: [1]prashu.201290@gmail.com, [2]anil@iirs.gov.in

Abstract—*A pixel containing more than one class known as mixed pixel are not handled by the hard classifiers. This requires the emergence of soft classifiers like fuzzy c-means (FCM) .It requires less parameter and less training data. One drawback of FCM is that it cannot handle the non-linearly separable classes. So the concept of Kernels is applied which maps the training samples to higher dimensional space. Thus the various kernels were incorporated in FCM to form kernel based fuzzy classifier (KFCM). This paper presents the nine single kernels as well the composite kernels which is the combination of two weighted kernels. KFCM is more robust and performs better than FCM by reducing the computational complexity. However, composite kernels gave more effective results as compared to KFCM using single kernel as it has suppressed the noise pixels within the water body.*

Keywords*: Fuzzy c-means, Kernel Fuzzy c-means (KFCM), Land Cover/Land Use (LULC).*

1. INTRODUCTION

Remote Sensing is defined as an art and science of gathering useful information about the earth's surface from a distance [8]. It is an important data source for providing effective Land Use/ Land Cover information which is retrieved by digital image classification. In general, classification algorithms are statistical in nature and assign single Land Use/ Land Cover class to each pixel which is known as Hard Classification. However, a pixel may contain more than one class which is known as mixed pixel .Mixed pixels can't be handled by hard classifier. Coarser is the spatial resolution, the problem of mixed pixel increases causing erroneous classification. For retrieving the classes from mixed pixel information soft classification approach may be required.

Also, sometimes the classes are non-linearly separable. To overcome this problem, the concept of Kernels is used which maps the training samples to higher dimensional space where they are considered to be linearly separable.

To handle the mixed pixels, where the object definition is not crisp and knowledge about the object is vague, fuzzy set theory comes into account. There are various classifiers like Fuzzy *c*-means, Possibilistic *c*-means, Noise Clustering with Entropy, Noise Clustering without Entropy and FCM with entropy. One of the most popular fuzzy clustering techniques is the Fuzzy *c*-means (FCM) which assigns the sample data points to multiple clusters thus overcoming the drawback of hard classifiers. The FCM classifier is effective in the presence of spherical and non-overlapping data clusters. For non-spherical overlapping data clusters and non-linearly separable classes, Kernel based Fuzzy *c*-means (KFCM) were introduced. In this paper various kernels used in work are characterized under the local kernels, global kernels, spectral kernel and hyper-tangent kernel.

Proportions of the Land Use Land Cover within mixed pixels can be extracted using soft classifier algorithm. The fraction images represent the proportions of each land cover within pixel [7]. This Land Cover information is useful for resource management. With the help of fraction images, the discrete classification maps of any type can be produced out of the continuous land cover information.

The algorithms for soft classification can be grouped under statistical, fuzzy set theory and neural network based. Statistical classifiers have multivariate normal distribution for each class. It includes Maximum Likelihood Classification that allocates the pixel of a class having the highest probability density and performs the hard classification of the remote sensing data. This classifier has been modified suitably for producing soft classification [1]. Another statistical method is the Linear mixture model (LMM) used in the geological and geophysical studies and performs soft classification [11]. Artificial Neural Network (ANN) is a learning algorithm and the upcoming machine learning is the Support vector machine (SVM). SVM is a powerful methodology for solving problems in nonlinear classification, function estimation and density estimation which has led to many recent developments in the Kernel based learning methods [6].

Some of the commercial software like ERDAS Imagine, ENVI etc revealed that some classifiers have been implemented. Minimum Distance classifier, Mahalonobis Distance, Parallelepiped Classifier, Maximum Likelihood Classifier are available in ERDAS Imagine yielding only hard output. All these classifiers along with Linear Mixture Model, Decision Tree, Artificial Neural Network and Support Vector Machine are available in ENVI (Environment for Visualizing Images) producing both hard and soft outputs. Maximum Likelihood Classifier, Box Classifier, Cluster analysis, Mahalonobis

Distance are available in ILWIS (Integrated Land and Water Information System). Fuzzy *c*-means is available in PCA Geomatica. However, these tools have limited soft classifiers and do not have fuzzy classifiers. This paper presents the capabilities of "Kernel Based Fuzzy Classification" methods of SMIC (Sub Pixel Multi Spectral Image Classifier) tool developed using Java Programming language for classifying multi-spectral data from remote sensing satellites.

2. REVIEW OF SOFT CLASSIFIERS

Land Cover information has been identified as one of the crucial data components for many aspects of global change studies and environmental applications [10]. Derivation of the land cover information from the remote sensed image is known as classification. The various algorithms developed were maximum likelihood classifier (MLC), neural network classifier and decision tree classifier. The maximum likelihood classifier is a parametric classifier based on the statistical theory. It assumes the normal distribution of class which is one of its limitations. Linear Mixture model (LMM) is based on the assumption that spectral response of a pixel is the linear sum of the spectral response of classes (within the pixel) weighted by their corresponding proportional area, the classifier can be used to produce soft classification, which sum to one for a pixel [1]. Due to the limitations of MLC and LMM, fuzzy set theory and neural network were used. Neural networks avoid some of the problems of the maximum likelihood classifier by adopting a non parametric approach. Artificial neural network (ANN) has a complex structure for processing high dimensional data. In case of linear mixture model (LMM), the number of classes to be extracted from the data should be less than or equal to dimensionality of the data plus one.

Decision tree classifier is a non parametric classifier which breaks a complex problem into multiple stages of simple decision making processes [9]. Decision trees can be univariate or multivariate depending upon the number of variables used at each stage. Multivariate decision trees are more accurate than the univariate decision trees. Support vector machine (SVM) works well with small training data set yielding high classification accuracy.

The structure of SVM is less complex in comparison to neural network. Unlike decision tree classifier, SVM do not require the generation of rules that depend heavily on the knowledge from experts. In practice, the SVM has been applied to optical character recognition, handwritten digit recognition and text categorization [4, 12]. SVM works well for high dimensional data such as classification of hyper spectral data.

The FCM approach has been adopted as it requires less parameter and less training data. FCM has an advantage over SVM as it does not require learning. However, it is not able to deal the nonlinear data for which kernels are incorporated in FCM. Thus, a java based tool implementing different kernels under various soft classifiers for land cover mapping known as SMIC has been developed.

3. SMIC TOOL COMPONENTS

SMIC tool has been developed to handle the multi-spectral images. In this tool, Kernel Fuzzy *c*-Means (KFCM) have been incorporated in a supervised mode for soft classification.

The various components of SMIC tool are shown in Fig. 1 and their description is as follows:

File: It performs file read, display and saving operations. When the file is read, the height, width of the image and the number of bands can also be read simultaneously.

Tools: It allows zooming in and out and the creation of false color composite (FCC). It also provides enhancement facilities while reference information is generated.

Signature Files: It allows the user to create the training data/ reference data/signature data which is pure in nature.

Classifiers: It contains the kernel fuzzy *c*-means classifier and the 9 kernels which can classify independently or can be used in the combination.

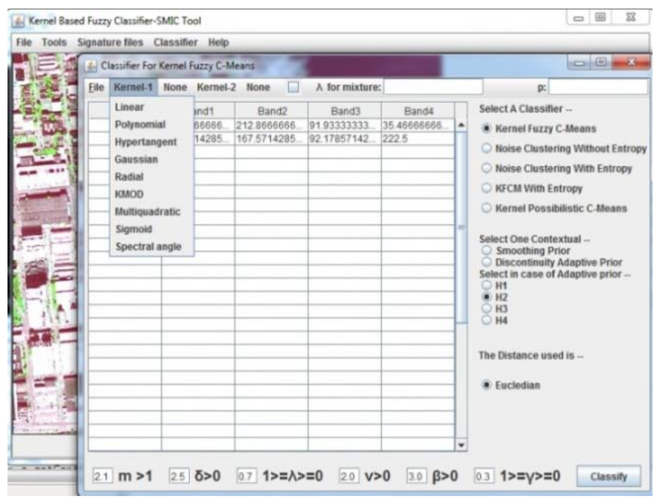

Fig. 1: Components of SMIC tool

Soft classification requires number of input parameters like weighting exponent (m) to be defined before as the classifiers need to be trained before the data is classified. In this tool, the kernel fuzzy *c*-means classifier was tested in supervised mode using the Euclidean norm only. The squared distance that is used in various kernels is given in equation (1) as follows:

$$d_{ij}^2 = (x_i - v_j)^t A\,(x_i - v_j) \quad (1)$$

Here, x_i is the unknown vector denoting spectral response of the pixel. v_j is the mean vector or the vector of cluster centre. A is the weight matrix. For Euclidean norm A=I (Identity matrix).

Supervised classification requires the reference data which is the training data for classifying the image. This reference data is the pure reference data i.e. the pure pixel for each class is considered while preparing the reference data. The various kernels used were linear, polynomial, Gaussian, radial basis, inverse multi-quadratic, hyper tangent, spectral angle, sigmoid, KMOD. The tool incorporates the composite kernels as well which is obtained by the combination of two kernels by choosing their weighting component.

4. KERNELIZED FUZZY CLASSIFIER

The various kernels have been incorporated in FCM that forms the kernelized fuzzy classifier. This section deals with the objective functions of FCM, KFCM and various implemented kernels.

4.1. Fuzzy *c*-Means

The Fuzzy *c*-means (FCM) is a partitioning algorithm that partitions the feature space and form clusters. FCM uses the fuzzy partition to make each data point with the membership grade of value given in interval of [0, 1] to determine its degree of belonging to the various groups [12]. FCM is based on the minimization of the objective function is given in equation (2) as follows [3]:

$$J_m(U,V) = \sum_{j=1}^{n}\sum_{i=1}^{c} \mu_{ij}^{m} \left\|X_j - V_i\right\|^2, 1 \leq m < \infty \quad (2)$$

where, n is the total number of pixels, c is the number of classes, μ_{ij} is the fuzzy membership value of the i^{th} pixel for class j, m is the weighing component, X_j is the vector pixel value and V_i is the mean vector of cluster j.

The fuzzy membership value is calculated through an iterative optimization of the above equation with update of membership μ_{ij} and cluster centers V_i are given in equation (3) and (4) respectively [3].

$$\mu_{ij} = \frac{1}{\sum_{k=1}^{c}\left(\frac{d_{ij}}{d_{ik}}\right)^{\frac{2}{m-1}}} \quad (3)$$

$$V_i = \frac{\sum_{j=1}^{n} {\mu_{ij}}^{m} X_j}{\sum_{j=1}^{n} {\mu_{ij}}^{m}} \quad (4)$$

The FCM algorithm is described as follows [14]:

i. Initializing each clustering center V_i (i=1,2,---,c).
ii. Using (3) determine the membership matrix U.
iii. Calculating the objective function with (4), if it is less than the determined threshold value, or the change relatively in the last objective function is less than a certain threshold value, and then stops the algorithm.
iv. According to (4), the clustering centers can be amended. Back to step (i).

4.2. Kernel Fuzzy *c*-Means

FCM is effective only in clustering the crisp, spherical and non-overlapping data. When dealing with non-spherical shape and much overlapped data, such as the Ring dataset FCM cannot always work well. Therefore we use the kernel method to construct the nonlinear version of FCM, and construct a kernel-based fuzzy c -means clustering algorithm (KFCM). Their common feature is that the data processing methods are applied in kernel mapping [13].

The basic idea of the kernel method is mapping the input pattern space R^p into a high-dimensional feature space R^q with the non-linear mapping$\emptyset(.)$, and doing the corresponding linear operation in high dimensional feature space [14]. Fig. 2 shows the mapping of input space to higher dimension space. $\emptyset(x_i)$ is used to represent the image of the pattern vector x_i (i=1,2,-------,n) in high dimensional feature space. If we choose the kernel function K(x, y), then in the feature spaceR^q, Euclid distance between the pattern vector x_i and x_j is given in equation (5) [2]:

$${d_{ij}}^2 = K(x_i, x_i) + K(x_j, x_j) - 2K(x_i, x_j) \quad (5)$$

The objective function is defined as in equation (6):

$$J_m(U,v) = \sum_{i=1}^{c}\sum_{k=1}^{n} {\mu_{ik}}^{m} \left\|\emptyset(x_k) - \emptyset(v_i)\right\|^2 \quad (6)$$

where v_i (i=1,2,----,c) is the clustering centre of the input space, unfold and do the Kernel substitution, there is equation (7)

$$\left\|\emptyset(x_k) - \emptyset(v_i)\right\|^2 = K(x_k, x_k) + K(v_i, v_i) - 2K(x_k, v_i) \quad (7)$$

$d(x,y)$ can be defined as in equation (8):

$$d(x,y) = \left\|\emptyset(x_k) - \emptyset(v_i)\right\| \quad (8)$$

$d(x,y)$ is the Euclidean distance of the feature space. The membership function μ_{ij} and the cluster centre v_i can be given as in equation (9) and (10) respectively:

$$\mu_{ij} = \frac{\left(1/{d_{ki}}^2\right)^{\frac{1}{m-1}}}{\sum_{j=1}^{c}\left(1/{d_{kj}}^2\right)^{\frac{1}{m-1}}} \quad (9)$$

$$v_i = \frac{\sum_{k=1}^{n} {\mu_{ik}}^{m} K(x_k, v_i) x_k}{\sum_{k=1}^{n} {\mu_{ik}}^{m} K(x_k, v_i)} \quad (10)$$

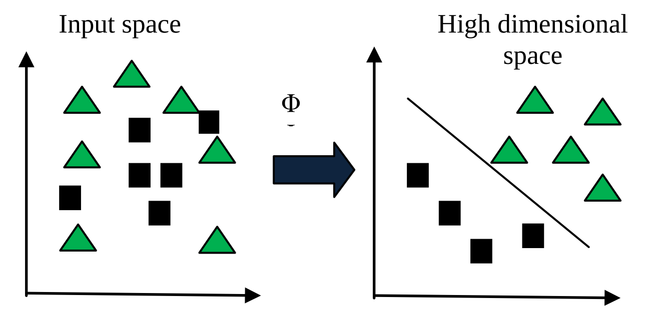

Fig. 2: Higher dimension feature space using kernel

The KFCM algorithm is described as follows [14]:

i. Setting the parameter of kernel σ, the number of clusters c, the fuzzy index m, the convergence precision ε, the maximum number of iterations k.
ii. Initializing each clustering centre v_i (i=1,2,--,c).
iii. Repeating the following operations, if it is less than a determined threshold value, or the change relatively in the last objective function is less than a certain threshold value, then stop the algorithm.
    a. According to (9), updating the membership degree with the current clustering centers.
    b. According to (10), updating each clustering centre with the current clustering centers and the membership degrees.

4.3. Description of Kernels

The kernels that are used in work are known as local kernels, global kernels, spectral kernels and hyper tangent kernel [6]. These kernels are described below from equation (11) to equation (19).

4.3.1. Local Kernels. They are based on the evaluation of the quadratic distance between any two training samples. All kernels that are based on the distance function are Local kernels.

Gaussian

$$K(x, x_i) = exp(-0.5(x - x_i)A^{-1}(x - x_i)^t) \quad (11)$$

Here, A=I, Identity matrix i.e. the Euclidean norm.

Radial Basis

$$K(x, x_i) = exp(-\|x - x_i\|^2) \quad (12)$$

KMOD

$$k(x, x_i) = exp(1/(1 + \|x - x_i\|^2)) - 1 \quad (13)$$

Inverse Multiquadratic

$$K(x, x_i) = \frac{1}{\sqrt{(\|x - x_i\|^2 + 1)}} \quad (14)$$

4.3.2. Global Kernels. Those samples that are far away from each other have an influence on the kernel value. All the kernels which are based on the dot-product are global.

Linear

$$K(x, x_i) = x.x_i \quad (15)$$

Polynomial

$$K(x, x_i) = (x.x_i + 1)^p \quad (16)$$

Sigmoid

$$K(x, x_i) = tanh(x.x_i + 1) \quad (17)$$

4.3.3. Spectral Angle Kernel. In order to fit hyper spectral point of view, it is of interest to consider new criteria that take into consideration spectral signature concept. Spectral angle (SA) $\alpha(x, x_i)$ is defined in order to measure the spectral difference between x and x_i while being robust to differences of the overall energy (e.g. illumination, shadows etc).

$$\alpha(x, x_i) = arccos((x.x_i)/(\|x\| \, \|x_i\|)) \quad (18)$$

4.3.4. Hyper Tangent Kernel. The hyper tangent kernel is given as follows [5]:

$$K(x, x_i) = 1 - tanh\left(-\frac{\|x_i - x_j\|^2}{\sigma^2}\right), \sigma > 0 \quad (19)$$

5. OUTPUTS AND RESULTS

The outputs from this tool generate the fraction images. Fig. 3 shows the fraction images that were generated from FCM classifier. Fig. 4 shows the fraction images that were generated from the SMIC tool using Gaussian kernel. The noisy pixels in the water body are seen. Fig. 5 shows the fraction images that were generated from the SMIC tool using linear kernel. Here the noisy pixels are less compared to Gaussian kernel. Fig. 6 depicts the fraction images generated from KFCM classifier using composite kernel (linear with gaussian). All the outputs were generated keeping the weighted exponent (m=2.1). The output pixel values were represented between 0 and 1 membership values.

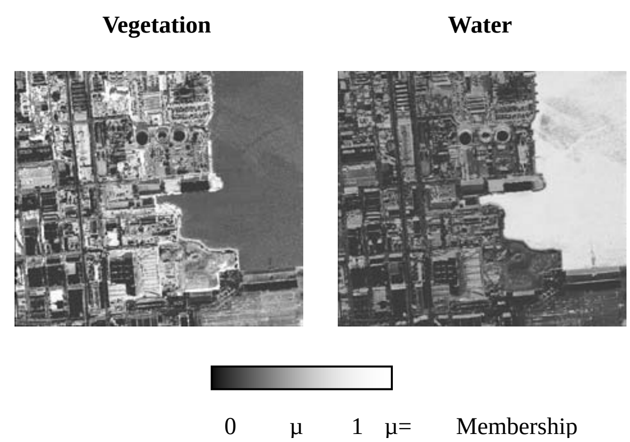

Fig. 3: Fraction images generated from FCM classifier

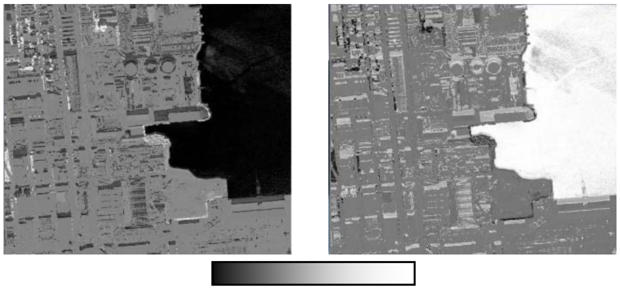

Fig. 4: Fraction images generated from KFCM classifier using single kernel (gaussian)

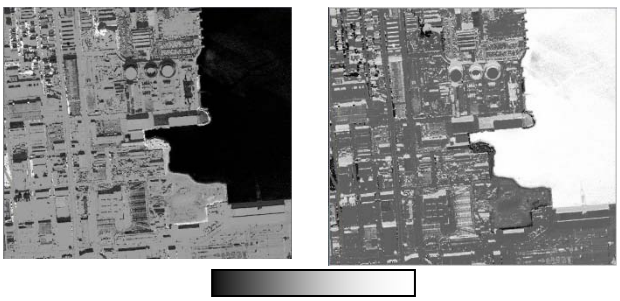

Fig. 5: Enhanced fraction images generated from KFCM classifier using single kernel (linear)

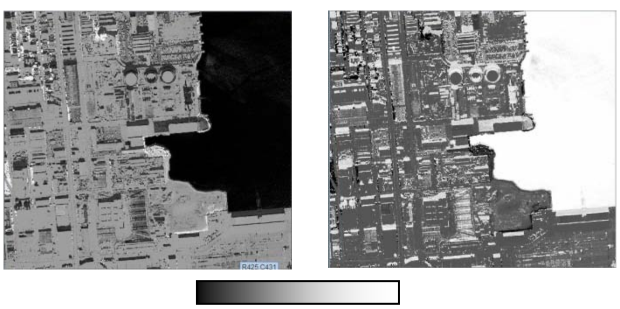

Fig. 6: Best fraction images generated from KFCM classifier using composite kernel (linear with gaussian)

The noisy pixels in the water body were suppressed producing best results. The two classes namely, vegetation and water were taken and the signature data for each land cover class were generated. Using these signature data, fraction images for each land cover class were generated.

6. CONCLUSIONS

SMIC tool has been developed in JAVA programming language that provides land cover information useful for resource management. This tool has been developed in graphical user friendly environment so that this tool can be used by any resource management professionals. The tool is platform independent as JAVA has been used for programming.

FCM has shown good results for linear classes. In case if the classes are non-linear, kernel based fuzzy *c*-means classifier (KFCM) has shown better results compared to Fuzzy *c*-means (FCM). Kernel approach maps the training samples to higher dimensional space. Compared to FCM algorithm, the KFCM algorithm can cluster more efficiently using kernel functions.

The results showed that kernel method is an effective approach to construct a robust image clustering algorithm. Thereby, KFCM enhances the effects of image classification. Thus the provision of 9 kernels along with the composite kernels performs classification of remote sensed image. While observing the homogeneous class water which means it has less variance. However, due to some man-made or natural activity, there is turbulence in some portion of water body indicating high variance and that is shown by the gray patch. That gray patch has been suppressed in case of composite kernel as compared to linear and Gaussian kernel.

7. ACKNOWLEDGEMENTS

The first author is thankful to IIRS/ISRO for permitting to accomplish the M.Tech project to study Kernel based fuzzy classifiers for which the tool has been developed and for providing the guidance in the research work and paper presentation.

REFERENCES

[1] Aziz, M.A., *Evaluation of Soft Classifiers For Remote Sensing Data*, Indian Institute of Technology Roorkee, Roorkee, India, unpublished Doctoral thesis, 2004.

[2] Dulyakarn, P., and Rangsanseri, Y., "Fuzzy *c*-means clustering using spatial information with application to remote sensing", in *paper presented at the 22nd Asian Conference on Remote Sensing*, Singapore., 5, November, 2001, pp.5-9.

[3] Dwivedi, R.K., *Fuzzy Hybrid Approaches For Soft Classification of Satellite Images*, Indian Institute of Technology Roorkee, Roorkee, India, unpublished Doctoral Thesis, 2014.

[4] Joachims, T., "Text categorization with support vector machines—learning with many relevant feature", in *Proceedings of European Conference on Machine Learning*, Chemnitz, Germany, (Berlin: Springer), April 10, 1998, pp. 137–142.

[5] Kaur, P., Gupta, P., and Sharma, P., "Review and Comparison of Kernel Based Fuzzy Image Segmentation Techniques", *I.J. Intelligent Systems and Applications*, 7, 2012, pp.50-60.

[6] Kumar, A., Ghosh, S.K., and Dadhwal, V.K., "Study of Mixed Kernel Effect on Classification Accuracy Using Density Estimation", in *Proceedings of Mid-term synposium ISPRS*, ITC- The Netherlands, 8-11th May, 2006.

[7] Kumar A., Ghosh S.K., and Dadhwal V.K., "Sub-Pixel Land Cover Mapping: SMIC System", in *Proceedings of ISPRS International Symposium on "Geospatial Databases for Sustainable Development*" Goa, India, September 27-30, 2006.

[8] Lillesand, T.M., and Kiefer, R.W., "Remote Sensing and Image Interpretation", Book News Inc., Portland, Or., USA, 2003.

[9] Safavian, S. R., and Landgrebe, D., “A survey of decision tree classifier methodology”, *IEEE Transactions on Systems, Man, and Cybernetics*, 21, 1991, pp.660–674.

[10] Sellers, P.J., Meeson, B.W., Hall, F.G., Asrar, G., Murphy, R.E., Schiffer, R.A., Bretherton, F.P., Dickinson, R.E., Ellingson ,R.G., Field, C.B., Huemmrich, K.F., Justice, C.O., Melack, J.M., Roulet, N.T., Schimel, D.S. and Try, P.D., “Remote sensing of the land surface for studies of global change: models-algorithms-experiments”, Remote Sensing of Enviro., 51, 1995, pp. 3-26.

[11] Smith, M.O., Johnson, P.E., and Adams, J.B., “Quantitative determination of mineral types and abundances from reflectance spectra using principal component analysis”, *Journal of Geophysical Research*, 90, 1995, pp. 797-804.

[12] Vapnik, V. N., “*The Nature of Statistical Learning Theory*” (New York: Springer-Verlag), 1995.

[13] Xinbo, G. and Weixin, X., “Development of fuzzy clustering theory and its research progress of application”, *Chinese Science Bulletin*, 44(21), 1999, pp.2241-2251.

[14] Yung-song, S., and Yu-feng, S., “Remote Sensing Image Classification and Recognition Based on KFCM”, *The 5th International Conference on Computer Science & Education*, Hefei, China, August 24-27, 2010.