

SUB-PIXEL LAND COVER MAPPING: SMIC SYSTEM

Anil Kumar[@], S. K. Ghosh^{*}, V. K. Dhadhwal

Indian Institute of Remote Sensing, Dehradun,

^{*} Department of Civil Engineering, Indian Institute of Technology, Roorkee,

[@] Corresponding Author: anil@iirs.gov.in

ABSTRACT:

Keywords: Sub-Pixel, Fuzzy Logic, Possibilistic c-Means, Support Vector Machine, Membership Value, Supervised Classification.

Statistical based algorithms, like Linear Mixture Model (LMM); Fuzzy Set Theory based algorithms, like Fuzzy c-means (FCM), and Artificial Neural Network (ANN) based classification algorithms have been incorporated in different commercially available digital image processing softwares; like neural network and fuzzy c-mean in PCI Geomatica, Linear Mixture Model (LMM) in ERDAS, ENVI etc. But there are hardly any digital image processing software's available having Support Vector Machines (SVM) based algorithms for multi-spectral remote sensing data sub-pixel classification. This paper presents the capabilities of *SMIC (Sub-Pixel Multi-Spectral Image Classifier)* System developed at Indian Institute of Remote Sensing, Dehradun, India, using JAVA programming language for land cover mapping using multi-spectral remote sensing satellite data at sub-pixel level using advance classifiers like, Support Vector Machine (SVM). This system can handle multi-spectral image of any bands. In this system four different classification algorithms (like; Fuzzy c-Means, Possibilistic c-Means, Maximum Likelihood Classifier, Support Vector Machine) have been incorporated in supervised mode for sub-pixel classification.

Reference data can be generated from this system in two modes i.e. pure reference data as well as mixed reference data. There is also provision in this system for saving the membership values generated using different classifiers for sub-pixel land cover mapping. As this system is capable for generating hard as well as sub-pixel level outputs, so for both types of outputs, different accuracy assessment methods have been incorporated. For hard classification output error matrix as well as K_{hat} coefficient have been used for assessing the accuracy of classified output. For sub-pixel classification, when reference as well as output is soft, Fuzzy Error Matrix (FERM) has been incorporated in this system for assessing the accuracy of soft classified output. As this system is developed in JAVA language, so this system is platform independent.

1. INTRODUCTION:

Preparation of the land cover maps from remote sensing images is viewed as a classification problem. In general, classification algorithms are statistical in nature and assign a unique value to each pixel. However, a pixel may contain more than one class, and such pixels are known as mixed pixels. As the remote sensing images become coarser, the problems of mixed pixel increases, leading to erroneous classification. To account for information with a mixed pixel, the analyst has to make an assessment regarding the proportion of classes within a mixed pixel. Normal statistical classifiers are not able to provide such information and hence different types of classifications are required. These classifiers are known as soft classifiers.

There are large numbers of soft classifier algorithms in digital image processing for capturing the respective proportions of land cover classes within mixed pixels and characterize land cover more accurately. A fraction images represents the proportions of each land cover within pixel. This land cover information is very important for land resource management. One important use of the fraction image information is that discrete classification maps of any type can be produced out of the continuous land cover information if desired. Soft classifier algorithms typically require training pixels, often called "image endmembers," which can be either pure pixels or mixed pixels with known class proportions.

The algorithms for sub pixel classification can be grouped under statistical, fuzzy set theory and some neural network based. Statistical based sub-pixel classifiers are dependent on multivariate normal distribution, while in the

case of Linear Mixture Model (LMM) number of classes to be extracted from the data should be less than or equal to dimensionality of the data plus one. Artificial Neural Network is a learning algorithm but does not always generalize well. Upcoming machine learning Support Vector Machines (SVM) is a group of theoretically superior machine learning algorithms. It is found to be competitive with the best available machine learning algorithms in classifying high-dimensional data sets (*Huang et al. 2002*).

SVM can work well with a small training data set yielding high classification accuracy. It is well suited for high dimensionality data such as classification of hyperspectral data, where data is available in hundreds of bands. Further, SVM is robust to the overfitting problem as it relies on margin maximization rather than finding a decision boundary from the training samples. The structure of an SVM is less complex in comparison to Neural Network, which has a complex structure for processing high dimensional data. Further, when compared to some of nonparametric classifier, namely Decision Tree Classifier, SVM does not require the generation of rules that heavily depend on the knowledge from experts. This is crucial for achieving high classification accuracy.

A review of some of the commercial software reveals that only one or two soft classifiers are implemented, e.g. Linear Mixture Model is available in ERDAS Imagine, Fuzzy C- Means is available in PCI Geomatica. SVM based algorithms are not available in any of the software, which is fast becoming popular amongst remote sensing analyst. This paper presents the capabilities of "Sub-pixel Multispectral Image Classifier" SMIC system developed at Indian Institute of Remote Sensing, Dehradun, India, using JAVA

programming language for classifying multi-spectral data from remote sensing satellites at sub-pixel level using advance classifiers like, Support Vector Machine (SVM).

2. SOFT CLASSIFICATION – CLASSIFIERS AND METHODS:

Land cover information has been identified as one of the crucial data components for many aspects of global change studies and environmental applications (Sellers *et al.* 1995). The derivation of such information increasingly relies on remote sensing technology due to its ability to acquire measurements of land surfaces at various spatial and temporal scales. One of the major approaches to deriving land cover information from remotely sensed images is classification. Numerous classification algorithms have been developed since the first Landsat image was acquired in early 1970s (Townshend 1992, Hall *et al.* 1995). Among the most popular are the maximum likelihood classifier (MLC), neural network classifiers and decision tree classifiers. The MLC is a parametric classifier based on statistical theory. Despite limitations due to its assumption of normal distribution of class signature (e.g. Swain and Davis 1978), it is perhaps one of the most widely used classifiers (e.g. Wang 1990, Hansen *et al.* 1996). Neural networks avoid some of the problems of the MLC by adopting a non-parametric approach. Their potential discriminating power has attracted a great deal of research effort. As a result, many types of neural networks have been developed (Lippman 1987); the most widely used in the classification of remotely sensed images is a group of networks called a multi-layer perceptron (MLP) (e.g. Paola and Schowengerdt 1995, Atkinson and Tatnall 1997).

A decision tree classifier takes a different approach to land cover classification. It breaks a complex classification problem into multiple stages of simple decision-making processes (Safavian and Landgrebe, 1991). Depending on the number of variables used at each stage, there are univariate and multivariate decision trees (Friedl and Brodley, 1997). Univariate decision trees have been used to develop land cover classifications at a global scale (DeFries *et al.* 1998, Hansen *et al.* 2000). Though multivariate decision trees are often more compact and can be more accurate than univariate decision trees (Brodley and Utgof 1995), they involve more complex algorithms and, as a result, are affected by a suite of algorithm-related factors (Friedl and Brodley, 1997).

The support vector machine (SVM) represents a group of theoretically superior machine learning algorithms. SVM employs optimization algorithms to locate the optimal boundaries between classes. Statistically, the optimal boundaries should be generalized to unseen samples with least errors among all possible boundaries separating the classes, therefore minimizing the confusion between classes. In practice, the SVM has been applied to optical character recognition, handwritten digit recognition and text categorization (Vapnik 1995, Joachims 1998b). These experiments reveals that SVM to be competitive with the best available classification methods, including neural networks and decision tree classifiers. The superior performance of the SVM has been demonstrated in classifying hyperspectral images acquired from the Airborne

Visible/Infrared Imaging Spectrometer (AVIRIS) (Gualtieri and Crompt, 1998). While hundreds of variables were used as the input in the experiments mentioned above, there are far fewer variables in data acquired from operational sensor systems such as Landsat, the Advanced Very High Resolution Radiometer (AVHRR) and the Moderate Resolution Imaging Spectro radiometer (MODIS). As these are some of the major sensor systems from which land cover information is derived, an evaluation of the performance of the SVM using images from such sensor systems should have practical implications for land cover classification. A Java based software implementing different types of soft classifiers for land cover mapping, known as SMIC (Sub-pixel Multispectral Image Classifier) has been developed. The details of this classifier are given ahead.

3. COMPONENTS OF SMIC:

This system has been designed to handle multi-spectral image of any sensor. In this system, four different classification algorithms (Fuzzy c-Means, Possibilistic c-Means, Maximum Likelihood Classifier, Support Vector Machine) have been incorporated in a supervised mode for sub-pixel classification. Figure 1 shows the different components of SMIC and is discussed below.

File: This option is for performing the file read, display and storing operations. One of the unique aspect incorporated regarding reading of a file is that any number of bands in file can be read simultaneously. Thus high dimensionality data sets can be read and displayed on the screen.

Utilities: This option allows for creation of a False Color Composite, and also allows for zooming in and out. Further simple linear enhancement facilities can be carried out which helps while generating reference information.

Signature files: The option allows the user to create of reference data, both pure and mixed in nature. The reference data can then be used for any of the four soft classifiers.

Classifiers: In this option four classifiers had been incorporated. There is also option for saving the membership values generated through sub-pixel classifier option using any four classifiers algorithms incorporated in the SMIC system.

Accuracy Assessment: This option allows for accuracy assessment, which is a critical part for evaluating the output. Here the hard classifier output is evaluated through error matrix as well as through kappa coefficient. The accuracy of sub-pixel classifier is checked using fuzzy error matrix.

When carrying out soft classification, it requires a number of input parameters to be defined or assessed beforehand as the classifiers have to be trained before the data can be classified. Some of these parameters are weight matrix 'A' known as the norm, a weighting exponent m etc.

In this system provision has been kept that the user can select weight matrix 'A' according to type of norm, while calculating squared distance in fuzzy c-means as well as in possibilistic c-means. As mentioned in following equation;

$$\text{Squared distance } d_{ij}^2 = (x_i - v_j)^T A (x_i - v_j) \quad (1)$$

Where x_i is the vector denoting spectral response of a pixel i (i.e. a vector of spectral response of a pixel in various bands of a multi-spectral remote sensing image), v_j is the vector of cluster center, A is the weight matrix, used as three

namely Euclidean, Diagonal and Mahalonobis norm. The formulations of each norm are given as (Bezdek, 1981),

$$\begin{aligned} A &= I && \text{Euclidean norm} \\ A &= D_j^{-1} && \text{Diagonal norm} \\ A &= C_j^{-1} && \text{Mahalonobis norm} \end{aligned}$$

Where, I is the identity matrix, D_j is the diagonal matrix having diagonal elements as the eigen values of the variance covariance matrix, C_j given by;

$$C_j = \sum_{i=1}^N (x_i - c_j)(x_i - c_j)^T \quad (2)$$

$$\text{where } c_j = \sum_{i=1}^N \frac{x_i}{N}$$

While incorporating maximum likelihood classifier algorithm, provision had been kept that in the Gaussian expression the weight matrix can be any of three norms as mentioned in equation (1). The probability density function $P(x/j)$ can be written as (Polubinskas *et al.* 1995),

$$P(x/j) = \frac{1}{(2\pi)^{n/2} |A_j|^{1/2}} e^{(-1/2(x-v'_j)^T A_j (x-v'_j))} \quad (3)$$

where n is the dimension of data;

While incorporating Support Vector Machine (SVM) algorithm, having constraint on their application in remote sensing due to its binary nature, requires multiclass classifications to be based upon a large number of binary analyses. In this algorithm density estimation has been done, which is based on the support vector machines (SVM) approach and it uses the Mean Field (MF) theory for developing an easy and efficient learning procedure for the SVM. It is well know that if data is non-separable training sets, than kernel functions can be used, which behaves like an inner product in high dimensions, to build a maximal margin hyperplane there, for separating out different classes. Different kernels used in SVM algorithm incorporated in this system are; Gaussian Kernel, Radial Basis, KMOD, Inverse Multiquadric, Linear Kernel, Polynomial Kernel, Sigmoid Kernel, Spectral Kernel. Also there is the provision kept in this system that mixed kernel can be generated will using any of the two kernel in combination.

As it is well known that in supervised classification reference data is required. This reference data can be generated from this system in two modes i.e. pure reference data as well as mixed reference data. Reference data in pure or in mixed form can be used to all the classification algorithms incorporated in this system for generating fraction images. There is also provision in this system for saving the membership values generated using different classifiers for sub-pixel classification.

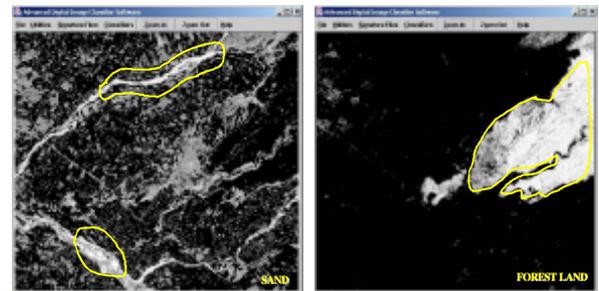


Figure 1. Main Window of the SMIC System

Accuracy assessment is very critical component in classification for assessing the accuracy of fraction images generated using sub-pixel classifier. As this system is capable for generating hard as well as sub-pixel level outputs, so for both types of outputs, different accuracy assessment methods have been incorporated. For hard classification output error matrix as well as K_{hat} coefficient have been used for assessing the accuracy of classified output. For sub-pixel classification, when reference as well as output is soft, Fuzzy Error Matrix (FERM) has been incorporated in this system for assessing the accuracy of soft classified output.

4. GENERATION OF OUTPUT:

The outputs from this system can be generated in hard as well as in soft classification (fraction images). For an example fraction images generated using possibilistic Fuzzy c-means algorithm had been shown in figure 2. In this total six land cover classes were taken and signature data for each land cover class were generated. Using these signature data fraction images for each land cover class were generated and output pixel value were represented between 0 to 1 membership values.



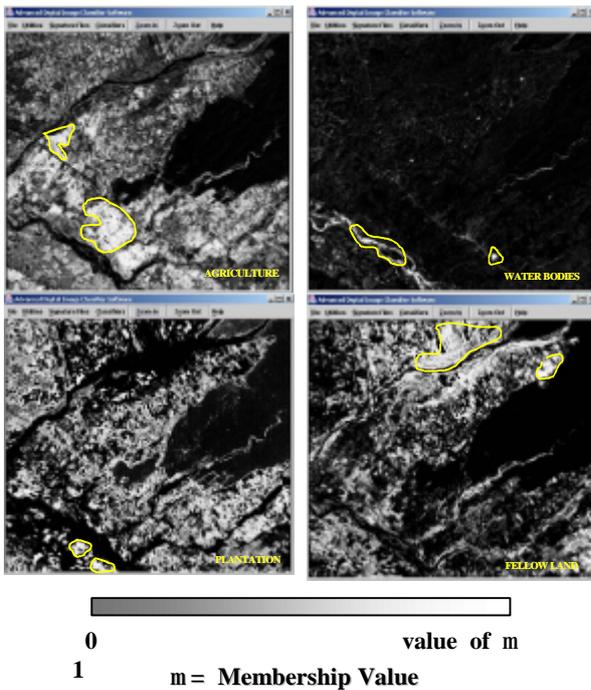


Figure No. 2: Fraction of Images generated from SMIC System

5. CONCLUSIONS AND FUTURE SCOPE:

The main objective of developing *SMIC* (Sub-Pixel Multi-Spectral Image Classifier) in JAVA programming language is to provide resources management professionals the confidence to learn and use the sub-pixel classification approach. At present, there is hardly any commercial digital image processing systems dedicated to advanced sub-pixel classifier are available and if available, they are than it is very costly. This system has been developed in a very graphical user-friendly environment, so that any resources management professionals can easily use this system and learn the supervised fuzzy sub-pixel classifier approach. As this system is developed in JAVA environment, so this system is platform independent. This software is a basic version presently and is being updated and refined to incorporate many other aspects of image processing.

Reference:

1. Atkinson, P. M., and Tatnall, A. R. L., 1997, Neural networks in remote sensing. *International Journal of Remote Sensing*, 18, 699–709.
2. Brodley, C. E., and Utgoff, P. E., 1995, Multivariate decision trees. *Machine Learning*, 19, 45–77.
3. C. Huang, L. S. Davis, J. R. G. Townshend, 2002, "An assessment of support vector machines for land cover classification" *IJRS*, vol. 23, no. 4, 725–749.
4. DeFries, R. S., Hansen, M., Townshend, J. R. G., and Sohlberg, R., 1998, Global land cover classifications at 8km spatial resolution: the use of training data derived from Landsat imagery in decision tree classifiers. *International Journal of Remote Sensing*, 19, 3141–3168.
5. Friedl, M. A., and Brodley, C. E., 1997, Decision tree classification of land cover from remotely sensed data. *Remote Sensing of Environment*, 61, 399–409.
6. Gualtieri, J. A., and Crompton, R. F., 1998, Support vector machines for hyperspectral remote sensing classification. In *Proceedings of the 27th AIPR Workshop: Advances in Computer Assisted Recognition, Washington, DC, Oct. 27, 1998* (Washington, DC: SPIE), pp. 221–232.
7. Hall, F. G., Townshend, J. R., and Engman, E. T., 1995, Status of remote sensing algorithms for estimation of land surface state parameters. *Remote Sensing of Environment*, 51, 138–156.
8. Hansen, M., DeFries, R. S., Townshend, J. R. G., and Sohlberg, R., 2000, Global land cover classification at 1 km spatial resolution using a classification tree approach. *International Journal of Remote Sensing*, 21, 1331–1364.
9. Hansen, M., Dubayah, R., and DeFries, R., 1996, Classification trees: an alternative to traditional land cover classifiers. *International Journal of Remote Sensing*, 17, 1075–1081.
10. Joachims, T., 1998a, Making large scale SVM learning practical. In *Advances in Kernel Methods—Support Vector Learning*, edited by B. Scholkopf, C. Burges and A. Smola (New York: MIT Press).
11. Joachims, T., 1998b, Text categorization with support vector machines—learning with many relevant features. In *Proceedings of European Conference on Machine Learning, Chemnitz, Germany, April 10, 1998* (Berlin: Springer), pp. 137–142.
12. Lippman, R. P., 1987, An introduction to computing with neural nets. *IEEE ASSP Magazine*, 4, 2–22.
13. Paola, J. D., and Schowengerdt, R. A., 1995, A review and analysis of back propagation neural networks for classification of remotely sensed multi-spectral imagery. *International Journal of Remote Sensing*, 16, 3033–3058.
14. Paola, J. D., and Schowengerdt, R. A., 1997, The effect of neural network structure on a multispectral land-use/land cover classification. *Photogrammetric Engineering and Remote Sensing*, 63, 535–544.
15. Safavian, S. R., and Landgrebe, D., 1991, A survey of decision tree classifier methodology. *IEEE Transactions on Systems, Man, and Cybernetics*, 21, 660–674.
16. Sellers, P. J., Meeson, B. W., Hall, F. G., Asrar, G., Murphy, R. E., Schiffer, R. A., Bretherton, F. P., et al., 1995, Remote sensing of the land surface for studies of global change: models—algorithms—experiments. *Remote Sensing of Environment*, 51, 3–26.
17. Swain, P. H., and Davis, S. M. (editors), 1978, *Remote Sensing: the Quantitative Approach* (New York: McGraw-Hill).
18. Townshend, J. R. G., 1984, Agricultural land-cover discrimination using Thematic Mapper spectral bands. *International Journal of Remote Sensing*, 5, 681–698.

19. Townshend, J. R. G., 1992, Land cover. *International Journal of Remote Sensing*, 13, 1319–1328.
20. Vapnik, V. N., 1995, *The Nature of Statistical Learning Theory* (New York: Springer-Verlag).
21. Vapnik, V. N., 1998, *Statistical Learning Theory* (New York: Wiley).
22. Wang, F., 1990, Fuzzy supervised classification of remote sensing images. *IEEE Transactions on Geosciences and Remote Sensing*, 28, 194–201.