

ANALYSIS OF THE EFFECTIVENESS OF SPECTRAL MIXTURE ANALYSIS AND MARKOV RANDOM FIELD BASED SUPER RESOLUTION MAPPING OVER AN URBAN ENVIRONMENT

D. R. Welikanna ^{a*}, V.Tolpekin ^b, Yogesh Kant ^c

^bDept of Earth System Analysis, ITC The Netherlands – valentyn@itc.nl

^cIndian Institute of Remote Sensing (NRSA), 04 Kalidas Road, Dehradun – yogesh@iirs.gov.in

^aSabaragamuwa University of Sri Lanka, Faculty of Geomatics - drw@sab.ac.lk

Commission WgS VII/4(1)

KEY WORDS: Hyperspectral Images, Spectral Mixture Analysis, Markov Random Field, Simulated annealing, Super Resolution Mapping

ABSTRACT:

The information in a pixel of satellite data within the instantaneous Field of View (IFOV) of the sensor is a mixture of different land cover types, and the individual land cover components can be estimated using soft classification techniques. However these techniques do not account for the spatial distribution of the class proportions, the information itself has a great relevance. In this study Markov Random Field (MRF) based Super Resolution Mapping (SRM) with certain modifications have been analysed for its performance with respect to the linear unmixing technique applied on hyperspectral data. This testing was carried out over a heterogeneous urban environment which was defined by the Vegetation, Impervious surface and Soil (V-I-S) model that has been used as an accepted alternative in characterising the urban land cover components. Linear unmixing technique with a hyperspectral remote sensing image (Hyperion) has been used to generate fractions according to the spectral variability of the V-I-S classes. Modified MRF based SRM technique was applied on IKONOS, ASTER MSS and Landsat ETM+ images with markedly different spatial and spectral resolutions. Reference map for the validation were created from the IKONOS MSS image using hard Maximum likelihood classification. The super resolution maps which contain the spatial information were again turned in to fractions representing each class (V-I-S). Next the results of MRF based SRM technique and the linear unmixing technique were validated using three measures of accuracy with respect to the reference fractions of the IKONOS image, the Area error proportion (AEP), Root Mean Square Error (RMSE) and the Correlation Coefficient (CC). The accuracy statistics for the Optimized Super Resolution Map (OSRM) fractions and the reference fractions showed a higher CC value in the range of 0.7 with respect to the linear unmixing fractions which lies in the range of 0.5. Results were justified by the overall RMSE and AEP values dropping from 0.7 to 0.5 and 1.8 to 1.4 respectively. In the case of ASTER visible to infrared (VNIR) image the correlation of the OSRM with the reference again showed a higher value in the range of 0.7 than the linear unmixing results which showed a correlation in the range of 0.5. Here it has been seen the overall RMSE and the AEP values were dropped for the SRM than the linear unmixing results from 0.7 to 0.6 and 1.8 to 0.6 respectively. In addition the results for the ASTER short wave infrared (SWIR) image and the Landsat image also followed a same trend which finally envisaged the improved subpixel representation of these land cover classes with the use of MRF based SRM techniques than the linear unmixing technique. The contextual refinement brought in by the MRF based SRM technique can produce accurate land cover components at a sub-pixel level even in a heterogeneous urban environment.

1. INTRODUCTION

For sustainable urban development and to understand and manage the urban environment, accurate information about the urban land cover classes are important. But the identification of these classes in a heterogeneous urban environment is a very difficult task. Under such circumstances remote sensing act as a highly productive tool in monitoring and understanding the urban morphology and its environment. In remote sensing several models and highly advance techniques have been developed for the study of urban land cover. One of the models developed to understand and to standardise the urban environment is described with the Vegetation- Impervious surface-soil (V-I-S) model (Ridd 1995). This model is becoming an accepted alternative to parameterize biophysical composition of urban environments (Wu & Murray, 2002; Wu, 2004). The spatial composition of a segment of urban landscape may be classified by the Vegetation-Impervious Surface-Soil

(V-I-S) model (for more details see Ridd, 1995). But the productivity of the remote sensing techniques depends on the ability of the technique to correctly categorize these land cover components from remote sensing images for there spatial and spectral arrangement. Studying these techniques for their accuracies and productivity is also very important for the future of urban land cover studies.

Hyperspectral images which correspond hundreds of wavelength channels for the same area on the surface of the earth have the ability to detect and identify individual materials or land cover classes in a highly efficient manner. With its high spectral resolution (narrow and continuous bands), hyperspectral images have an advantage over multispectral images to address the spectral variation of each V-I-S components in a heterogeneous urban environment. This is the prerequisite of its use in urban composition studies. But the

* Corresponding author. Email: drw@sab.ac.lk

problem is the mixed pixels affecting the images. Mixed pixels are the pixels in an image which represents a reflectance value corresponding to a combination of reflectances of several individual materials within the IFOV of the sensor. When a pixel observed by a sensor covering a large (low spatial resolution images) heterogeneous area on earth surface it tends to become a heavily mixed pixel.

Modelling V-I-S classes with the use of end-members is very difficult due to the complex spectral variations of these classes. Conventional classification techniques assign a single pixel into one land cover class, in the case of a coarse spatial resolution image this will directly affect the information extracted by the remote sensing image. Spectral Mixture Analysis (SMA) is utilized to calculate the land cover fraction within a pixel and model a mixed spectrum as a combination of spectra for pure land cover types, called endmembers. It does not show the spatial distribution of class proportions (spatial contextual information) within a pixel. Modelling the contextual information in land cover classification can improve the accuracy of the classification significantly (Tso & Mather, 2001). In this regard super resolution mapping (SRM) splits the pixel into sub pixels and assigns the classes to the sub pixels depending on their spatial and spectral information, generating an image with higher resolution than the original image. Markov random field model is another method (Hailu, 2006; Kasetkasem & Arora, 2003) of super resolution mapping and it describes the spatial dependencies (spatially closer pixels tend to be more alike than the more distant ones) quite accurately. In the current research MRF based SRM technique has been attempted over a heterogeneous urban environment of Dehradun city, India defined by V-I-S model and analysed with respect to the linear unmixing technique. The study has been developed to understand the effectiveness of handling the mixed pixel problem in coarse resolution satellite imagery with the integration of the contextual information by MRF based SRM technique. Hyperion hyperspectral image was used for the linear unmixing technique to identify the V-I-S classes.

2. MRF BASED SRM TECHNIQUE

MRF have been used by statistical physicists to explain various phenomena occurring among neighbouring particles because of their ability to describe local interaction between them. MRF model has also been used to explain why intensity values of adjacent pixels of an image are more likely to be same than different values. In the interpretation of a scene contextual information is very important to classify a pixel in context with respect to the other measurements. This will lead to more complete information with the minimization of possible ambiguities, recovery of missing information and correction of errors. Context can be defined in three dimensions; the spectral, spatial and temporal dimensions (Anne *et al.*, 1996). The spectral dimension can be defined as the different bands of the electromagnetic spectra while spatial context refers to the correlation between the class labels of neighbouring pixels.

2.1 The Maximum a posterior probability in image analysis

In statistical or Bayesian rules, the decision making is based on the concept of the maximum a posterior (MAP) probability (Christopher & Bishop 2006). The Bayes theorem which converts the prior probability to a posterior probability takes a form as below.

$$p(c|y, M_w) = \frac{p(c)p(y|M_w)}{p(y)} \quad (1)$$

Where $p(c)$ is the prior probability that the given pattern belongs to class c , $y = \{y_1, y_2, \dots, y_I\}$ is a set of observations, I is the number of observations, $p(y|M_w)$ is the conditional probability of the observation set y for a given model M_w , $p(y)$ is the probability of the observation set y and $p(c|y, M_w)$ is the probability that the pattern belongs to class c , for the observation set and the model which is the posterior probability. Since $p(y)$ is not related to any class, we can use the eq (1) for making the classification decisions.

$$c' = \underset{c}{\operatorname{arg\,max}} \{p(c)p(y|M_w)\} \quad (2)$$

Where c' is the estimated pattern class. If $p(c)$ is unknown, it is set to constant and the MAP estimation becomes maximum likelihood (ML) estimation.

2.2 Mathematical concepts in MRF based SRM technique

Let y be a coarse resolution multispectral remote sensing image containing k spectral bands. The pixel locations are denoted as $b_i \in B$, where B is the set of pixels with size $M \times N$. Spatial resolution of the image is denoted with the pixel size R . The resulting super resolution map is a classified map which is denoted by c and defined on a set of pixels A . The area the pixels in A covers is as the same as B , but bears a high spatial resolution which is denoted by a lower pixel size of r . The original image with the pixel size of R will be referred to as coarse resolution image and the super resolution map with pixel size r will be referred to as the fine resolution image. The scale factor of the super resolution map is denoted as S which is the ratio between the coarse and fine resolution pixel sizes. This means each pixel b_i will contain S^2 fine resolution pixels of $a_{j/i}$ or a_j of a fine resolution multispectral image x . Furthermore it is assumed that every pixel in image x can be assigned to a unique class $c(a_j) = \alpha$, where $\alpha \in \{1, 2, \dots, L\}$. For convenience the class of pixel a_j will be referred to as c_j . The relation established between images y and x can be described by:

$$y(b_i) = \frac{1}{S^2} \sum_j x(a_{j/i}) \quad (3)$$

For each pixel a_j in A , a symmetric neighbourhood $N_j = N(a_j)$ can be defined. Where N_j a set of pixels neighbouring pixel a_j . A set of pixels inside a square window, excluding the pixel a_j at the centre of the window and with the window size W being the length of one side, defines the complete neighbourhood. Here neighbourhoods are defined with their order, and first and second order neighbourhood systems are used commonly. A clique is defined as a subset of this in which all pair of pixels are mutual neighbours. The size of the window also determines the amount of cliques which can fit inside as neighbourhood. The contextual constraints introduced from the neighbouring pixels will be modelled by the means of prior energy. In this study the minimum neighbourhood order considered is second

order with respect to the scale factors used. To implement this growth of the neighbourhood size in relation to the scale factor a relationship between S and the window size W has been developed as $W=S \cdot I$. Under the given window size the number of maximal pixel neighbours N_n to be considered is also realized by $N_n = ((W \times 2) + 1)^2 - 1$.

2.3 Initial SRM Generation and the optimization

In the first step the fraction image produced by Singular Value Decomposition (SVD) is the input for the initial SRM generation (Geladi & Hans 1996). Each pixel was divided by S to get the sub-pixels, and then each pixel was randomly labelled with the corresponding classes from the coarse fraction image. The classes were defined by their mean vectors and covariance matrices. Prior to that the separability of these classes defined on the coarse spatial resolution image was assessed using the Transform Divergence (TD) and Jeffries-Matusita distance (JM). The parameters related to Gibbs potential function were also estimated. The Gibbs potential function will have low values in the case of high spatial resolution images, due to more number of pure pixels leading to neighbouring pixels occupying the same class. The number of subpixels allocated to the given class for the initial SRM can be expressed as $n = \theta_{bi} \times S^2$, where θ_{bi} is the proportion of a given class in pixel b_i of the coarse resolution image. Thus; n number of sub-pixels will be randomly labelled within the original pixel y_i with respect to this class proportions. The mean and covariance matrices were defined by carefully selecting pure training pixels for the desired land cover classes from the coarse resolution image. And fractions for each land cover class in pixel b_i is defined by SVD. The output of this initial step is an initial SRM with many isolated pixels. This should be optimized considering the spatial dependence between the pixels, to get the optimized SRM with more contextually smoothed classes. The optimization process is based on the Simulated Annealing (SA) algorithm.

2.4 MRF, Gibbs distribution and SA algorithm

The observed coarse spatial resolution multi spectral image y , is represented in vector form as $y(b_j) \in R^k$ for a pixel b_j where R denotes the set of real numbers representing the intensity values and K for the number of spectral bands. Let $A^j = \{a_1^j, \dots, a_s^j\}$ represent a set of pixels in the set A , which corresponds to the same area as the pixel b_j in set B . As a resulting SRM pixel is pure and can represent only by one land cover class its value c_j can only take an integer value that corresponds to a particular land cover class at pixel a_j in the actual scene. The properties inherent to MRF as a random field with respect to its neighbourhood are as follows

Positivity where $p(c_j) > 0$, this states that the joint probability $p(c_j)$ of any random field is uniquely determined by its local conditional probabilities. *Homogeneity*, this specifies the conditional probability for the label of a pixel, given the label of the neighbourhood pixel, regardless of the relative location of the pixel. *Markovianity*, this shows that labelling of the pixel is only dependent on its neighbouring pixels. That is, the conditional probability of the value of a pixel given pixel values of the entire image excluding the pixel of interest is equal to the conditional probability of a value of a pixel given values of its neighbouring pixels. This can be expressed in mathematical terms as below.

$$p(c_j | c_{c-j}) = p(c_j | c_{N_j}) \quad (4)$$

Where c_{c-j} refers to all the classes except the class at j , and c_{N_j} refers to the classes neighbouring the class at j . In the context of land cover classification this property implies that same land cover class is more likely to occur in connected regions than isolated pixels (Cutler *et al.*, 1997). The MRF model characterizes the spatial dependence among neighbouring sites; however a direct implementation of equation (4) is not simple because the probabilities can take up any values. As a result we introduce GRF. With the above properties we equalize the MRF with the Gibbs Random Field (GRF). Therefore the posterior probability for each class C given the observed image y can be specified by the means of a posterior energy function.

$$p(c | y) = \frac{1}{Z} \exp\left(\frac{-U(c | y)}{T}\right) \quad (5)$$

Where

$$Z = \sum \exp\left(\frac{-U(c | y)}{T}\right) \quad (6)$$

Where Z is the normalizing constant and T is a constant termed temperature. Here we can define an energy function $U(c)$.

$$U(c) = \sum_{j,i} \sum_{l \in N_j} V_2(c_j, c_l) \quad (7)$$

Hence $V_2(c_j, c_l)$, is the Gibbs potential function with respect to pairwise clique. And the prior probability can be expressed as:

$$p(c) = \frac{1}{Z} \exp\left(-\frac{U(c)}{T}\right) \quad (8)$$

In the study to maximize the posterior energy function, when a pixel belong to the same class as the class of the pixel in focus it was given a value 0 and 1 otherwise, and they were weighted with respect to the Euclidian distance (Kasetkasem *et al.*, 2005). According to the equation (5) to allocate pixel a_j to class α , the posterior probability has to be maximized. As described in equation (2) a MAP solution can be incorporated for this.

$$c_j = \text{arg max} \{ p(c | y) \} \quad (9)$$

Using the Bayes rule, the likelihood probability, which is the conditional distribution of the observed data y given the class c assumed as a Gaussian distribution and can be defined as:

$$p(y|c) = \frac{1}{\sqrt{2\pi^p} \sqrt{|\sum a|}} \exp(-U(y|c)) \quad (10)$$

Where ρ is the number of bands, Σ_α is the class-conditional covariance matrix for class α . The likelihood energy can be defined by using the mean vector μ_α of class α and the covariance matrix, by the following equation:

$$U(c|y) = \sum_i \frac{1}{2} (y(b_i) - \mu_\alpha)^T \Sigma_\alpha^{-1} (y(b_i) - \mu_\alpha) + \frac{1}{2} \ln \det \Sigma_\alpha \quad (11)$$

The mean vector and covariance matrices were given by,

$$\Sigma_i = \Sigma (y(b_i)) = \sum_{\alpha=1}^L \theta_\alpha (b_i) \Sigma_\alpha \quad (12)$$

$$\mu_i = \mu (y(b_i)) = \sum_{\alpha=1}^L \theta_\alpha (b_i) \mu_\alpha \quad (13)$$

Where θ_α is the proportion of land cover class α in pixel $y(b_j)$. And a class can be assigned to a pixel when prior $p(c)$ and posterior $P(c|y)$ probabilities are maximized. This is equivalent to minimizing the energy functions:

$$c_j = \arg \min \{ U(c) + U(c|y) \} \quad (14)$$

The likelihood energy in equation (11) depends on the Mahalanobis distance of the pixels and the prior energy in equation (7) depends on the pairwise clique's V_2 . For the pairwise clique potential function following equation has been used

$$V_2(c_j, c_l) = \frac{1}{\sum_{l \in N_j} \omega(a_l)} \cdot \omega(a_l) \delta(c_j, c_l) \quad (15)$$

Where $\omega(a_l)$ is the weight function. The weight function is made to stay inversely proportional to the square of the distance $d(a_j, a_l)$ between pixel a_j and a_l and described in mathematical form as:

$$\omega(a_l) = \frac{1}{N n_j \cdot d(a_j, a_l)^2} \quad (16)$$

Where $N n_j$ is the amount of pixels in the neighbourhood for the pixel j . For the equation (14) a special parameter λ , the smoothness parameter has been used. This controls the balance between the two energy functions. In previous studies this parameter was only used to control the prior energy, which result in optimal λ in between 0 and ∞ . Here we normalize this parameter between 0 and 1:

$$c_j = \arg \min \{ \lambda U(c) + (1 - \lambda) U(c|y) \} \quad (17)$$

This equation should be minimized to find the most reasonable class for the pixel a_j . Once the global energy is constructed and the smoothness parameters have been determined, as the next step we start the pixel labelling. We use the MAP estimation for this. To find the MAP solution we adopt the SA algorithm in this study. SA algorithm allows the randomness (T), to decrease in an iterative way that the best solution for equation (17) can be made. The temperature will be decreased according to the criterion called cooling schedule. The process is repeated until the system becomes frozen ($T \rightarrow 0$), which means pixels stop updating. The cooling schedule associated with this study can be expressed as:

$$T(t) = T_0 \times T_{upd} \quad (18)$$

Where $T(t)$ is any next temperature value depending on the iterations. According to Hailu (Hailu, 2006), the value for the updating schedule T_{upd} typically lies in the range of 0.8 to 0.99. So for this research we have incorporated a value of 0.9 for T_{upd} . The initial temperature T_0 use to control the randomness of the optimization algorithm ($T_0 \propto \text{Randomness}$). So a high temperature can increase the probability of a pixel being replaced by a new class label though the energy of a new class is higher. Here we incorporate a value of 3.0 for the initial temperature. This finding justified the findings of the earlier studies of Kasetkasem (Kasetkasem & Arora, 2005). The updating process for the new temperature value is controlled by two constraints, if the number of pixels that have a different value after one updating is less than a predefined value, minimum activation threshold (*min_acc_thr*) or by keeping the counter with in a limit defined by the variable called, number of iterations (*Niter*). The pixel updating was performed by a row wise visiting scheme, which determines the order of the pixels in the SRM for the updating. The SRM will gradually change reducing the number of isolated pixels. The Gibbs potential function will force the SA algorithm to iteratively generate a SRM which is close to the solution of the MAP criterion in equation (2).

2.5 Measures of Accuracy

First the generated SRM was tested for the fitness of use with respect to the reference image with the use of a confusion matrix. Two measures of accuracy the overall accuracy (*OA*) and the Kappa coefficient (*K*) derived by the confusion matrix has been used for the assessment of the SRM in this study. Three other measures of accuracy the CC, RMSE and AEP, has been incorporated to evaluate the agreement between the fraction Images generated from the linear unmixing and SRM technique, with the reference fraction image. Take a set of known proportions as Y and the set of estimated proportions as A , let n be the total number of pixels.

$$C \ C = \frac{C_{Y_j \cdot A_j}}{S_{Y_j} \cdot S_{A_j}} \quad (19)$$

$$R \ M \ S \ E = \sqrt{\frac{\sum_{i=1}^n (Y_{ij} - A_{ij})^2}{n}} \quad (20)$$

$$AEP_j = \frac{\sum_{i=1}^n (y_{ij} - A_{ij})}{n} \quad (21)$$

Where $C_{y_iA_j}$ is the covariance between Y and A for class j and S_{y_j} and S_{A_j} are the standard deviations of Y and A for class j . These measures play a major role in this study.

3. THE RESULTS AND DISCUSSION

3.1 Linear unmixing results from Hyperion image

Figure 1 shows the linear unmixing results on Hyperion data. Due to the broader class definitions involved in this study, the representative endmembers for these classes (V-I-S) may not be the most pure pixels in the image. Highly positive and negative values correspond to the areas having high RMSE values (Figure 1(d)). Table 1 gives; the mean abundance of the endmembers derived from the Hyperion data shows a positive value. And the low standard deviation shows the acceptable deviation from the mean. The minimum and maximum abundances for vegetation and soil are more likely to be the same.

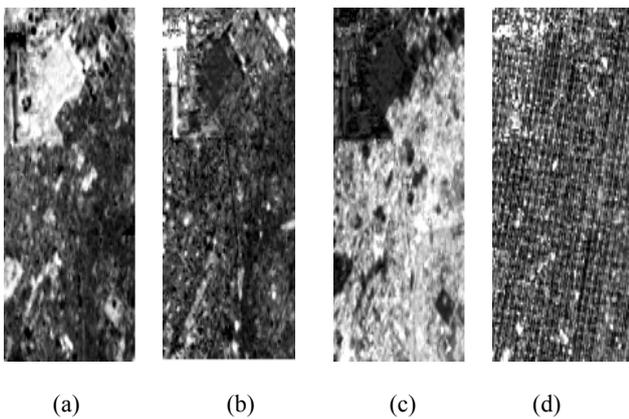


Figure 1. Endmember abundance images (a) Vegetation (b) Soil (c) Impervious (d) Error image showing RMSE

lass	min	Mean	Max	<0%	>1%	Stdv
V	-0.913	0.500	1.362	38.190	0.305	0.283
I	-0.990	0.176	1.365	12.020	0.532	0.193
S	-0.524	0.466	2.200	13.540	3.093	0.346
RMSE	0.000	0.141	1.000	0.045	0.007	0.071

Table 1. Statistics of the linear unmixing results

3.2 Test experimental results on IKONOS degraded image

To understand the efficiency of the MRF based SRM technique first it was applied on a degraded IKONOS image. Here the degraded IKONOS image pixel is made to have a value, which is the mean of 4x4 pixels in the fine resolution image (IKONOS 4m multispectral image). This makes the degraded image high in mixed pixels making it one of the best inputs for the MRF based SRM technique. The scale factor was set to $S=4$. This S value defines a neighbourhood by a 7x7 square window. Good class separability leads to a higher spectral distance between the class means which can make the prior constraint provide more

information for the pixels to be classified correctly. The λ value controls the smoothness and it assures the pixels to be labelled according to the majority of the pixels in the neighbourhood. These effects were tested by comparing the SRM results with a maximum likelihood classification results on a IKONOS image. The respective kappa accuracies and the OSRM for each of the optimum λ values are shown below in table 2.

λ	Initial Kappa	Kappa
0.1	0.2813	0.4610
0.25		0.5041
0.3		0.5075
0.35		0.5114
0.4		0.5084
0.45		0.3491
0.85		0.0002

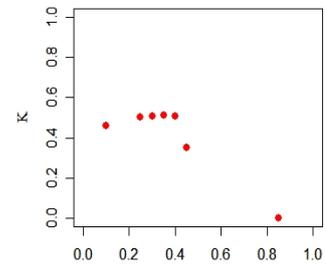


Table 2. Kappa agreement for different λ values and the resulting plot for IKONOS degraded image

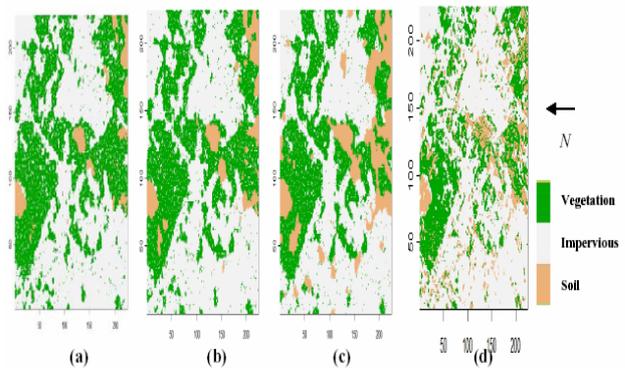


Figure 2. Effect of smoothing parameter on SRM, $\lambda=$ (a) 0.3, (b) 0.35, (c) 0.40 and (d) Reference image

The graph shows the effect of $\lambda(\Lambda)$ on the quality of the SRM. Resulting K values increases with the increase of the λ values and it reach a unique optimal value of 0.35 where $K=0.5114$. An abrupt drop in the K values can be seen for the value of λ from 0.4 to 0.45. Beyond this the K values started dropping steadily. This drop shows that at the range starting from 0.45 to 1 for the λ values it starts giving lower weight to the likelihood energy than the prior energy. Therefore 0.35 value was considered as the optimal smoothing parameter.

The visual interpretation of figure 2 shows that the vegetation and impervious classes are more closely classified to the reference image. But the soil class, especially in the upper left corner of the image has been misclassified for the vegetation class. The soil class in most of the parts of the image represents a very small distribution. As a result it represents a very small number of pixels with respect to the other two classes. So the contextual information provided by this class will not be enough to satisfy the Markovian property. The OSRM were turned in to fractions to compare it with the linear unmixing results on Hyperion image. Initially the OSRM with 4m spatial resolution has to be turned into fractions of 30m spatial resolution to compare with Hyperion fractions of 30m. This aggregation was possible mainly because the OSRM represent only the pure pixels. By analyzing the two images in a pixel by

pixel basis, it has been found out that 7x7 pixels of the OSRM corresponds to a 1 pixel of the Hyperion image. Therefore 49 pixels from OSRM were used to generate a class proportion measure in a single 30m pixel. The same procedure was carried out to convert the reference image of 4m spatial resolution in to 30m fractions. The results of these comparisons are shown in the below tables.

Class	CC	RMSE	AEP
V	0.6432	0.2749	1.6385
S	0.2460	0.1870	-0.0823
I	0.7196	0.2688	0.2910
Total		0.7307	1.8472

(a)

Class	CC	RMSE	AEP
V	0.8521	0.2271	-0.3296
S	0.6846	0.1692	1.7299
I	0.9200	0.1584	0.0579
Total		0.5547	1.4582

(b)

Table 3. Accuracy statistics for IKONOS degraded image by CC, RMSE and AEP (a) between fractions of Hyperion linear unmixing & Reference fractions (b) between OSRM fractions and the reference fractions

The CC and the RMSE in tables 3 (a) and (b), respectively shows the significant improvement in the classification accuracy with respect to the maximum likelihood results. The soil class which initially shows a less correlation with 0.2460 values, has been improved up to 0.6846 with the SRM, this is confirmed by the RMSE values improving from 0.1870 to 0.1692. The AEP also follows the same trend and overall values for RMSE and AEP shows the improvement in the accuracy with the values dropping from 0.7307 to 0.5547 and 1.8472 to 1.4582 respectively. This drop in overall RMSE shows the improvement in the prediction (the biasness and variance), while the AEP shows the success rate in maintaining the known proportions during prediction (Tatem *et al.* 2001)

3.3 The experimental results on ASTER (VNIR) image

The ASTER image was processed in way that it prevents original pixels getting resampled. In other words the three bands of the ASTER (VNIR) region were kept to the original grid size of 16m. This leaves the pixels getting resample by any interpolation method. The classes defined possessed a high separability with the TD values ranging from 1.9 to 2.0. A scale factor of 4 which generates SRM with spatial resolution of 4m was experimented here and the optimal smoothness parameter for the best SRM was tested. The experiments show that for the optimum results λ can take a value in the range of 0.4 to 0.5. Here the highest value of $K=0.4018$ is reached for a λ value of 0.45, with a slightly higher accuracy than the values of 0.5 and 0.4 respectively. The results are shown below in table 4 and figure 3 shows the resulting OSRM.

	Initial Kappa	Kappa
0.3	0.2134	0.3754
0.4		0.3962
0.45		0.4018
0.5		0.4016
0.55		0.3750
0.6		0.2561
0.9		0.0003

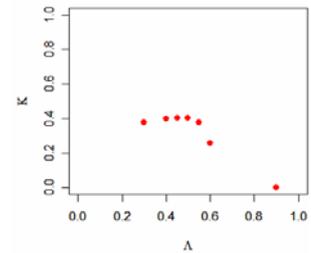


Table 4. Kappa agreement for different λ values and the resulting plot for ASTER (VNIR) image

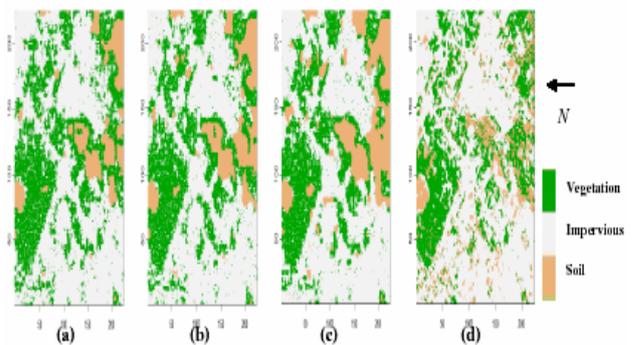


Figure 3. Effect of smoothing parameter on SRM, $\lambda=$ (a) 0.4, (b) 0.45, (c) 0.5 and (d) Reference image

These results were analysed with respect to the fractions generated from the MLC results on the IKONOS image. Below table 5 shows the comparative results.

Class	CC	RMSE	AEP
V	0.6432	0.2749	1.6358
S	0.2460	0.1870	-0.0823
I	0.7196	0.2688	0.2910
Total		0.7307	1.8472

(a)

Class	CC	RMSE	AEP
V	0.8208	0.2199	-0.2875
S	0.6267	0.1868	0.9729
I	0.8453	0.2024	-0.0539
Total		0.6091	0.6315

(b)

Table 5. Accuracy statistics for ASTER (VNIR) image by CC, RMSE and AEP (a) between fractions of Hyperion linear unmixing & Reference fractions (b) between OSRM fractions and the reference fractions

The correlation between the OSRM fractions and the reference fractions have been improved to an average in the range of 0.7, which is a good correlation between the known and the prediction. The total RMSE shows the biasness in the prediction with a high value of 0.6091 and AEP shows the over estimation of the area with a value of 0.6315. This may be explained by the errors inherited from the co-registration between the ASTER image and the reference IKONOS image, which leads to certain

mismatch between the set of pixels considered for these accuracy measures.

3.4 Experimental results on ASTER (SWIR) image

The test were further carried out in finding the effects of the MRF based SRM technique on ASTER (SWIR) image. For this experiment, certain differences were made in the preparation of the reference data set using IKONOS image. This image was resampled to grid size of 5m before applying the MLC classification to the image using the classes defined on the ASTER (SWIR) image. This makes a scale factor of $S=6$ to be defined between the two images, making the maximal pixel neighbours of 120 to be considered.

In contrast to the previous results the optimal λ values shows a different range in achieving the optimal results. Looking at the results (Table 6), the optimal values for λ seems to stay in the range of 0.1 to 0.3. And the highest accuracy levels with a K value of 0.3682 has been achieved for a λ value of 0.25. One of the reasons for this range may be the spatial resolution of the image, which makes the scene contains more homogeneous representation of objects. This means when the pixels are divided to sub-pixels the spatial dependency between the sub-pixels will not be increased significantly, still leaving the sub-pixels as mixed pixels.

λ	Initial Kappa	Kappa
0.1	0.2472	0.3417
0.2		0.3570
0.25		0.3682
0.3		0.2882
0.4		0.0081
0.45		0.0605
0.7		0.0000

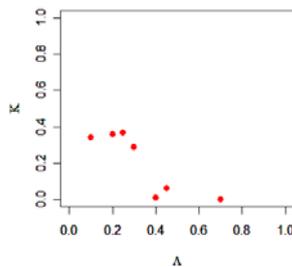


Table 6. Kappa agreement for different λ values and the resulting plot for ASTER (SWIR) image

The resulting super resolution maps for these range is shown by figure 4.

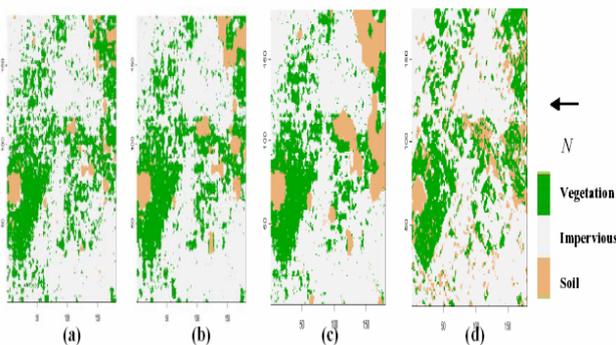


Figure 4. Effect of smoothing parameter on SRM, $\lambda =$ (a) 0.1, (b) 0.2, (c) 0.25 and (d) Reference image

The cross validation with the reference image showed the results in table 7.

Class	CC	RMSE	AEP
V	0.6419	0.2795	0.6178

S	0.3800	0.1746	-0.2194
I	0.6638	0.3032	0.2464
Total		0.7573	0.6448

(a)

Class	CC	RMSE	AEP
V	0.8109	0.1949	-0.1390
S	0.5425	0.2004	0.6791
I	0.8127	0.2133	-0.0169
Total		0.6086	0.5232

(b)

Table 7. Accuracy statistics for ASTER (SWIR) image by CC, RMSE and AEP (a) between fractions of Hyperion linear unmixing & Reference fractions (b) between OSRM fractions and the reference fractions

According to the table 7(a) Hyperion fractions from linear unmixing shows a marginal association with reference fractions from IKONOS image, with the CC being 0.6419 and 0.6638 for the vegetation and impervious classes respectively. The poor classification results for the soil class has been repeated again with the CC of 0.3800. Finally with respect to the tables 7(a) and (b) the significant improvements in the classification results can be seen, with the CC improving in to the range of 0.8 for the vegetation and impervious classes and for the soil class the correlation value of 0.38 in the Hyperion linear unmixing has been improved to value of 0.5425 in the SRM results. These results shown by the CC have been mirrored by the total AEP and RMSE values. The biasness in the predictions using SRM technique is less than what it is in the linear unmixing, with the total RMSE falling down from 0.7573 to 0.6086 and the total AEP is falling from 0.6448 to 0.5232. The overall results indicate that higher classification accuracies can be obtained using the SRM with respect to the linear unmixing techniques.

3.5 Experimental results on Landsat ETM+ image

Table 8 shows the results for the Landsat ETM+ image. The optimal λ value was determined as 0.25 with highest K value of 0.4224. The optimal range for smoothing parameter lies between 0.1 to 0.3.

λ	Initial Kappa	Final Kappa
0.1	0.2459	0.3753
0.2		0.3985
0.25		0.4224
0.3		0.4112
0.45		0.2226
0.7		0.0270

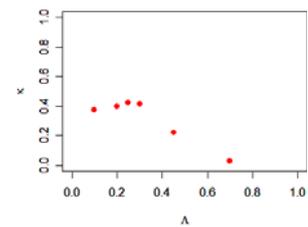


Table 8. Kappa agreement for different λ values and the resulting plot for Landsat ETM+ image

This shows a similar behaviour of the MRF model parameters in generation of the SRM for the different image characteristics of the Landsat image and the ASTER (SWIR) image. The OSRM produced with optimum λ value of 0.25 according to figure 5(b) shows a high agreement with the reference image than the other two images. The resulting images shows an significant smoothness for each class, this makes the OSRM more visually appealing, but it degrade the quality of the image

due to the missing of the details. One reason for this is the high scale factor $S=6$ used, which invariably grow the neighbourhood size. This significantly improves the effect of prior energy, with many sub pixel configurations with equivalent contextual energy, making the OSRM to become smoother even with a low smoothing parameter value λ . When the likelihood energy is controlled by a higher smoothing value $(1-\lambda)$ the more homogeneous objects (eg. Vegetation patches) with more number of pure pixels tend to be classified properly than the scattered object (eg. soil).

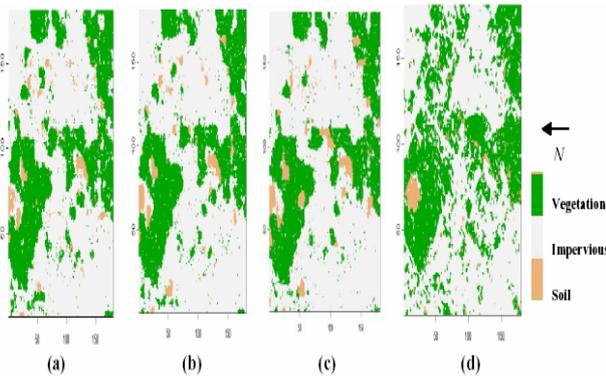


Figure 5. Effect of smoothing parameter on SRM, $\lambda=$ (a) 0.1, (b) 0.2, (c) 0.25 and (d) Reference image

The results of the OSRM and the Hyperion linear unmixing were comparatively analysed as shown in table 9. According to the experiments the linear unmixing results have a marginally high correlation with respect to the OSRM results with the reference fractions generated from the IKONOS MLC classified image. With the overall RMSE dropping slightly from 0.7429 to 0.7079 MRF base SRM results shows its advantage in correctly predicting the fractions. And this has been justified by an identical AEP measure of 0.0134 for the OSRM results

Class	CC	RMSE	AEP
V	0.8230	0.3247	0.7307
S	0.5708	0.1549	-2.6871
I	0.7682	0.2633	0.2222
Total		0.7429	-1.7342

(a)

Class	CC	RMSE	AEP
V	0.6883	0.2882	0.2759
S	0.3251	0.1440	-0.1624
I	0.7290	0.2757	0.1001
Total		0.7079	0.0134

(b)

Table 9. Accuracy statistics for Landsat ETM+ image by CC, RMSE and AEP (a) between fractions of Hyperion linear unmixing & Reference fractions (b) between OSRM fractions and the reference fractions

The overall results suggest that the MRF base SRM technique on Landsat image has a marginally higher advantage than the linear unmixing technique with the hyper spectral remote sensing image.

4. CONCLUSION

Parameter determination is an important step for optimizing the performance of the super resolution mapping. This study applies the MRF based super resolution mapping to a

significant number of real satellite images with different characteristics for the first time.

The results obtained on applying the MRF based SRM technique using IKONOS image shows that to model the contextual information needed for an accurate land cover classification, the global energy needs to be modelled with a higher λ value (smoothness parameter) in the range of 0.3 to 0.5. This smoothness parameter range was justified with the similar results obtained from ASTER (VNIR) image. The λ range experienced in case of low resolution satellite images like ASTER (SWIR) and Landsat image was different where the global energy has a tendency to depend on likelihood energy more than the prior energy, claiming the λ values to stay in the range of 0.1 to 0.3 and also the range for the optimal λ values getting narrowed down to a lower range. In other words, it is observed that the class labelling depends more on the spectral information. The overall results suggests that with the different input image characteristics (spectral and spatial resolutions), the input parameters should be tuned to obtain optimal results. Moreover, these parameters tend to change, depending on the image characteristics, especially the spatial resolution. In high spatial resolution images with the Gibbs potential function taking low values the pixels corresponding to low energy state have a high probability in generating a class. This means that the prior energy brings in the additional information to classify the image pixels correctly. Hence this can lead it taking a higher weight (λ) than in the case of a lower spatial resolution image. Though the likelihood energy poses these variations due to the radiometric range of the dataset, result from IKONOS and ASTER (VNIR) images suggest that there is no significant difference in treating the likelihood energy for the optimal results. The observation on ASTER (SWIR) and Landsat images also justifies the above conclusions. Hence, it can be concluded that the spectral range of these MSS images may not have an effective influence over the optimal SRM results.

The comparative analysis between the Hyperion linear unmixing results and the MRF based SRM results based on RMSE, AEP and CC shows that the MRF models produce super resolution maps with accurate fractional representation for each class at the subpixel level than the linear unmixing results.

REFERENCES

Ridd.M.K., 1995. Exploring a V-I-S (Vegetation-Impervious surface-Soil) model for urban ecosystem analysis through remote sensing: comparative anatomy for cities. *International Journal of Remote Sensing*, 16, pp. 2165-2185.

Wu,C., and Murray,A.T., 2002. Estimating impervious surface distribution by spectral mixture analysis. *Remote Sensing of Environment*, 84, pp. 493-505.

Wu.C., 2004. Normalized spectral mixture analysis for monitoring urban composition using ETM+ Imagery. *Remote Sensing of Environment* 93, pp. 480-492.

B.Tso and Mather, P., M.,C., 2001. *Classification methods for remotely sensed data*. London etc, Taylor & Francis.

Hailu, R.K., 2006. Suitability of Markov Random Field-based Method for Super-Resolution Land Cover Mapping, *MSc Thesis*, ITC, Enschede.

Kasetkasem,T.,Arora,M.K., and Varshney,P.K.,2003. Sub-pixel Land cover Mapping Based on Markov Random Field Models. *IEEE Trans.on Geosci.and Remote Sensing*, pp. 3456-3458.

Anne H Schistad Solberg,T. T., and K Jain,A., 1996. A Markov Random Field Model for Classification of Multisource Satellite Imagery. *IEEE Trans.on Geosci.and Remote Sensing*, 34(1), pp. 100-113.

Christopher,M.Bishop., 2006. *Pattern recognition and machine learning*. Springer, New York.

Geladi,P., and Hans, G., 1996. *Multivariate Image Analysis*. John Wiley & Sons Ltd, West Sussex, England.

Cutler,M.E, Lewis,H.G and Atkinson,P.M., 1997. Mapping sub-pixel propotinal land cover with AVHRR imagery. *International Journal of Remote Sensing*, 18(4), pp. 917-935

Kasetkasem,T.,Arora,M.K., and Varshney,P.K.,2005. Super-resolution land cover mapping using a Markov random field based approach. *Remote Sensing of Environment*, 96, pp302-314.

Tatem,A.J., Lewis,H.G., Atkinson,P.M., and Nixon,M.S., 2001. Super-resolution target identification from remotely sensed images using a Hopfield neural network. *IEEE Trans.on Geosci.and Remote Sensing*, pp. 781-796

