CONTEXTUAL CLASSIFICATION OF REMOTELY SENSED DATA USING MAP APPROACH AND MRF

R. Khedam[a*], A. Belhadj-Aissa[a]

[a] Image Processing Laboratory, Electronic and Computer Science Faculty, Technology and Sciences University (USTHB), B.P. 32, El Alia, Bab Ezzouar, 1611, Algiers, Algeria.
radjakhedam@lycos.com, h.belhadj@lycos.com

Commission PS WG VII/1

KEY WORDS: Remote sensing, classification, supervised, Bayes, contextual, Markov, optimisation.

ABSTRACT:

Classification of land cover is one of the most important tasks and one of the primary objectives in the analysis of remotely sensed data. Recall that the aim of the classification process is to assign each pixel from the analysed scene to a particular class of interest, such as urban area, forest, water, roads, etc. The image resulting from the labelling of all pixels is henceforth referred to as "a thematic map". Such maps are very useful in many remote sensing applications especially those concerned with agricultural production monitoring, land change cover and environmental protection. Conventional classification methods commonly named "punctual methods", classify each pixel independently by considering only its observed intensity vector. The result of such methods has often "a salt and pepper appearance" which is a main characteristic of misclassification. In particular of remotely sensed satellite imagery, adjacent pixels are related or correlated, both because imaging sensors acquire significant portions of energy from adjacent pixels and because ground cover types generally occur over a region that is large compared with the size of a pixel. It seems clear that information from neighbouring pixels should increase the discrimination capabilities of the pixel-based measured data, and thus, improve the classification accuracy and the interpretation efficiency. This information is referred to as the spatial contextual information. In recent years, many researchers have proven that the best methodological framework which allows integrating spatial contextual information in images classification is Markov Random Fields (MRF). In this paper, we shall present a contextual classification method based on a *maximum a posterior* (MAP) approach and MRF. An optimisation problem arises and it will solved by using an optimisation algorithm such as Iterated Conditional Modes (ICM) which occurs the definition and the control of some critical parameters : neighbouring size, regularisation parameter value and criterion convergence. Test data available is SPOT image of "Blida" region sited at 50km on the south west of Algiers (Algeria). This image acquired on February 1986, contains seven main classes. The result of our contextual classification process is an interpretable and more easily exploitable thematic map.

1. INTRODUCTION

In the current decade, global environmental change has reached beyond the research domain and become a major national and international policy issue. The project "Analysis of multitemporal remotely sensed data ; multispectral and interferometric SAR imagery for land cover change in northern Algeria" was established in our laboratory on January 2004. The project has the objective of analysing the spatial characteristics, temporal dynamics, and environmental consequences of land-use and land-cover changes which have occurred in northern Algeria over the period of 1980 and 2004 as a result of a range of socio-economic , biogeophysical and natural driving forces. Especially, over the last three years, northern Algeria has known two natural catastrophes ; flood and mudflows of Bab El Oued city happened on November 10, 2001 and a strong earthquake which struck Boumerdes city on May 21, 2003. These two events have caused land cover changes, land degradation and serious materials damage. The data analysis is used to project plausible future changes in land use and land cover under different assumptions of future natural, demographic, economic, technological, social and political development. Given the current techniques available, remote sensing is recognized as an efficient tool for earth watching and land monitoring and provides the most feasible approach to land surface change detection at regional, continental or global scales. Remote sensing is a collective name for several techniques which study at distance the ground surface or the atmosphere. Sensors installed on satellites or airplanes receive and/or send radiation to the earth. The variation in amount and wavelength of the reflected energy between studied objects or phenomena gives the object its spectral signature and makes it possible to distinguish between different types of land use, vegetation, soils etc. Remotely sensed data are being and will continue to be used to retrieve information on a land cover map which hold an important place at each step of a territory planning project. For a better characterization of land cover mapping, data classification approaches are generally proposed to obtain also a robust objects or classes identification. Conventional automatic classification techniques called also "punctual classifications" classify each pixel independently without tacking into account information given by its context which a very helpful information because the response and class of two or more spatially neighboring pixels are highly related. Different approaches have been taken to incorporate context in classification of remote sensing data and have named "contextual classification" (Kartikeyan *et al.*, 1994). We find approaches based on clustering pixels of the image according to the similarity of their response (Amadamn *et al.*, 1988, Kettig *et al.*, 1976), relaxation techniques where probabilities of neighboring pixels are used iteratively to update the probability of a given pixel (Richards *et al.*, 1981) and approaches known as “sequential compound decision theory” and which attempts to decide the label for one pixel based on the observation at all other pixels in the image (khazenie *et al.*, 1990). More recently, another type of approaches called “global approaches” have been evolved (Braathen *et al.*, 1993, Pieczynski, 2000). These new approaches are two types ; MAP (maximum *a posterior* probability) and MPM (maximum of posterior marginal probability) and both present a classic case of an ill-posed problem (Maroquin *et al.* 1987) and their solution must be given through stochastic or deterministic optimization algorithms. In this paper, we shall solve a MAP problem according to a Markov Random Field (MRF) model which provides a methodological framework to avoid combinatorial problem and effectively incorporate contextual information through ICM (Iterated Conditional Modes) which is a deterministic optimization algorithm. The supervised contextual classifier is applied on data set of SPOT multispectral image acquired on February 1986 covering initially, an agricultural region sited in northern of Algeria and contains the famous city of Blida. Our objective is to obtain an exploitable land cover map to use it later for change detection process.

2. NOTATION

We assume that a classified image X and observed data Y are realisations of stochastic processes $\boldsymbol{X}$ and $\boldsymbol{Y}$, respectively. $Y = \left\{Y^1, Y^2, \ldots, Y^K\right\}$ are multispectral data observed through K spectral bands and are supposed to be acquired on a finite rectangular lattice $W = \left\{s = (i,j) : 1 \le s \le S\right\}$, s is the site of the *ijth* pixel and S is lattice's area. The set $Y^k = \left\{y_s^k, \ldots, y_S^k\right\}$ where $k = 1, 2, \ldots K$**,** denotes the data taken at the *kth* wavelength, where $y_s^k \in \{1, 2, \ldots, NG\}$ and NG is the number of observable grey levels. It is also possible to describe the multispectral data with $Y = \left\{y_s, 1 \le s \le S\right\}$ where $y_s = \left\{y_s^1, y_s^2, \ldots, y_s^K\right\}$ is a feature vector observed on the site s called also a spectral signature on site s.

3. PUNCTUAL CLASSIFICATION APPROACH

Image classification can be done visually, by visual interpretation of the data, or digitally where numerical procedures, usually statistically based decision rules, automate the classification process. While a visual classification is superior in the interpretation of spatial information (textural and contextual information), computers can handle the spectral information more efficiently. Conventional digital classifiers, called also punctual classifiers, are entirely based on spectral pattern recognition. Indeed, in punctual classification, the spectral signature y_s which represents the observed intensity vector is the only aspect used to classify a pixel on site s. The parameters of the distribution are learnt from training samples in a supervised classification approach, and from test image pixels by suitable clustering method in an unsupervised approach. The pixels of the image are then classified by calculating, from their observed response, the likelihood that they have come from different classes. By this procedure, it can be seen that the decision taken for a pixel is based solely on the response to that pixel. For this reason, techniques based on this approach have been called “punctual or blind approaches” (Braathen *et al.*, 1993). These approaches have been widely used for classification and have given fairly good results for a wide variety of images (Desachy, 1991). The most used supervised punctual method is a maximum likelihood method where the analyst supervises the classification by identifying representative areas, so called training zones. These zones are then described numerically and presented to the computer algorithm which classifies the pixels of the entire scene into the respective spectral class that appears to be most alike. In a maximum likelihood classification, the distribution of the response pattern of each class is assumed to be normal (gaussian). It means that the feature vector observed y_s is drawn from a *“gaussian distribution”*. So, the likelihood probability to assign a pixel y_s to the class x_s is given as fellows:

$$P\left(x_s / y_s\right) = -\frac{1}{2}\left(y_s - \mu_{xs}\right)^T . \sum_{xs}^{-1} . \left(y_s - \mu_{xs}\right) - \frac{1}{2} ln\left\{ \left| \sum xs \right| \right\} \qquad (1)$$

Where m_{xs} and $\sum_{x_s}$ are statistic parameters of class xs estimated during training step process. The decision to assign one pixel from the analysed scene to a particular class is then given as follows:

$$y_s \in x_i \ \ if \ \ P(y_s/x_i) \rangle P(y_s/x_j) \ \ for\ each\ j \neq i \qquad (2)$$

The accuracy of such methods is very much affected by a “salt and pepper” appearance characterizing misclassification of some pixels. It means that intensity vector is insufficient and then leads to incorrect classification of pixels. In particular of remotely sensed data, adjacent pixels are related or correlated, both because imaging sensors acquire significant portions of energy from adjacent and because ground cover types generally occur over a region that is large compared with the size of a pixel. Using coherent contextual information for classification efficiency and accuracy in remote sensing has long been desired. Contextual information is important for the interpretation of a scene. When a pixel is considered in isolation, it may provide incomplete information about the desired characteristics. However, the consideration of the pixel in its context, more complete information might be derived. We can define three kinds of context: 1) spectral context, 2) spatial context and 3) temporal context (Khedam *et al.*; 2001). The basic philosophy in non punctual approaches is that the response and class of two spatially adjacent pixels are highly related. For example, if *(i, j)* and *(m, n)* are two neighbouring pixels and if *(i, j)* belongs to class k, then there is a high possibility that pixel *(m, n)* also belongs to the same class k. Therefore, the decision for a pixel is taken based not only on the observation at *(i, j)* but also on all observations at *(m, n)* where *(m, n)* is neighbour of *(i, j)*. Non punctual approaches can be contextual or global (MAP and MPM) approaches. We are interested in this paper on MAP approach.

4. MAP CLASSIFICATION APPROACH

In term of global approach where the class assigned to a site depends not only on the spectral feature of the site itself, but also on the spectral feature of all pixels in the image, our goal is to find the optimal classified image $X^* = \left\{x_s, \ldots, x_S\right\}$ based on the observed data Y. Each site of the segmented image is to assigned into one of M classes; that is, $x_s = \{1, 2, ..., M\}$ where M is the number of classes assumed to be known in supervised classification process. This optimisation is executed from the view point of the maximum *a posterior* (MAP) estimation as follows:

$$X^* = X_{MAP} = \underset{X \in \Omega}{argmax}\{P(X/Y)\} \tag{3}$$

Where Ω is labelled configurations set. Following Bayes theorem, equation (3) becomes:

$$X_{MAP} = \underset{X \in \Omega}{argmax} \left\{ \frac{P(Y/X)P(X)}{P(Y)} \right\} \tag{4}$$

The modelling of both class conditional distribution *P(Y/X)* and prior distribution *P(X)* becomes an essential task (Khedam *et al.*, 2001). *P(Y)* is the probability distribution of the observed data and doesn't depend on the labelling *X*. Note that the estimate (4) becomes the pixel–wise non–contextual (punctual) classifier if the prior probability doesn't have any consequence in formulating (4). *P(Y/X)* is the conditional probability distribution of the observation *Y* given the labelling *X*. A commonly used model for *P(Y/X)* is that the feature vector observed Y_s is drawn from a "*gaussian distribution*". For a Markov Random Field (MRF) of field labelled *X* and so, according to the Hammerslay-Clifford theory, *P(X)* can be expressed as a Gibbs distribution with "*Potts model*" as energy function model. The global MAP estimate given by equation (3) is equivalent to the minimisation of the followed *a posterior* global energy function (khedam *et al.*, 2002):

$$X_{MAP} = \underset{X \in \text{Ù}}{\arg\min}\{U(X/Y)\} \tag{5}$$

Where:

$$U(X/Y) = \left\{ \sum_{s \in S} \left\{ \frac{1}{2}(y_s - \text{ì}_{xs})^T . \Sigma_{xs}^{-1} . (y_s - \text{ì}_{xs}) + \frac{1}{2} ln\left\{ \left| \sum xs \right| \right\} \right\} + \right.$$
$$\left. + \sum_{q \in Q} \text{â} \sum_{r \in Vs} (1 - \text{ä}(x_s, x_r)) \right\} \tag{6}$$

Where m_{xs} and $\sum_{x_S}$ are statistic parameters of class xs estimated during training step process and β is parameter regularisation and is frequently user specified. $\ddot{a}$ is Kroeneker function calculated on the neighbourhood *Vs* of site *s* on all clique *q* from set clique *Q* (figure 1). Neighbourhood system *Vs* can be 4-connexivity given by equation 7 or 8-connexivity given by equation 8.

$$V_{(i,j)}^4 = \{ (k,l) \in S, \ 0 < (i-k)^2 + (j-l)^2 \leq 1 \} \tag{7}$$

$$V_{(i,j)}^8 = \{ (k,l) \in S, \ 0 < (i-k)^2 + (j-l)^2 \leq 2 \} \tag{8}$$

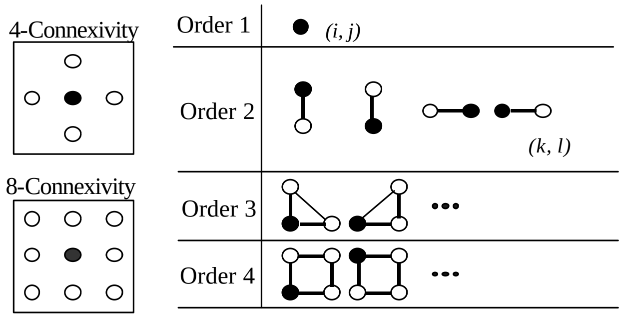


Figure 1. Neighbourhood and cliques

Once MAP classification problem is formulated as an energy minimisation problem, it can be solved by an optimisation algorithm. Among the most effective algorithms for optimisation in the framework of image MRF modelling are Simulated Annealing (SA) (Geman *et al.*, 1984) whose the computational demands are well known and Iterated Conditional Modes (ICM) (Besag, 1986) which is a computationally feasible alternative of the SA with a local minimum convergence of the energy function. To use ICM algorithm, global minimisation energy function of equation (6) must be transformed on the followed local minimisation energy function under the assumption of the independence of conditional probabilities:

$$U(x_s/y_s) = \left\{ \left\{ \frac{1}{2}(y_s - \text{ì}_{xs})^T . \Sigma_{xs}^{-1} . (y_s - \text{ì}_{xs}) + \frac{1}{2} ln\left\{ \left| \sum xs \right| \right\} \right\} \right.$$
$$\left. + \text{â} \sum_{r \in V_S} (1 - \text{ä}(x_s, x_r)) \right\} \tag{9}$$

The first term on the right hand of (9) called data attach term, can be consider as the potential function of one-order clique and is often used to provide an initial configuration for the contextual classification process. The second term called regularisation term, defines pair-cliques potential function which explicitly describes local spatial interactions in neighbourhood *Vs* . ICM flow chart is presented on figure 2.It can be resumed on five steps as follows:
Step 1: Estimate statistic parameters set (m_{xs} , Σ_{xs}) from the training samples of each class from classes set *A*.
Step 2: Based on (m_{xs} , Σ_{xs}), estimate an initial classification using the non-contextual pixel-wise maximum likelihood decision rule. We use first term of equation (9).
Step 3: Choose an appropriate value of β, an appropriate shape and size of neighbourhood system *Vs* and an appropriate convergence criterion.
Step 4: Perform the local minimisation defined by equation (9) at each pixel in specified order: update y_s by the class x_s that minimises equation (9).
Step 5: Repeat step (3) until convergence. Iterative algorithms often pose convergence problem. Convergence criterion which we have adopted in this study is a zero number of pixels changing classes between two consecutive iterations. This number of pixels is calculated on the whole image and thus for all classes. We

have thought of a local criterion convergence which can be regarded as a zero number of pixels which change state on each class, other classes being masked. This procedure can be seen as the decomposition of ICM process on a number of under-processes. Each under-process relates to one class and is slow or fast according to the heterogeneity of this class (Khedam *et al.*, 2002).

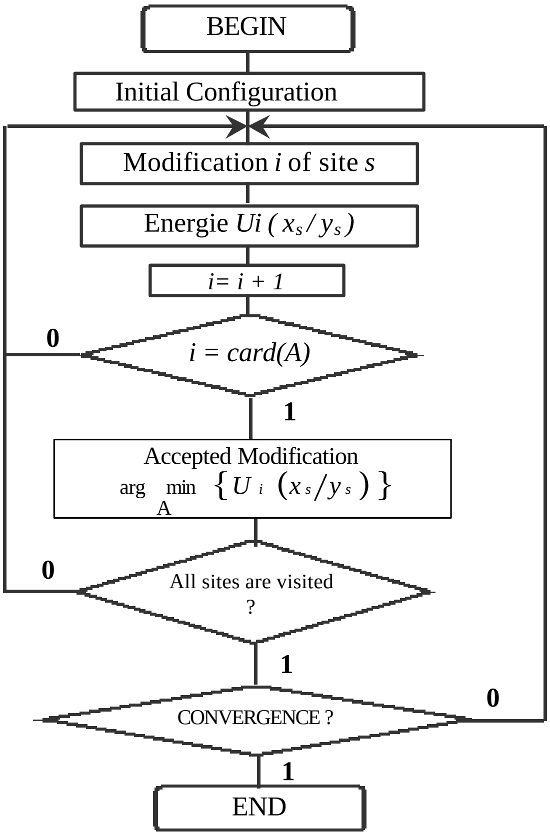


Figure 2. ICM flow chart

ICM algorithm is looked as a regularisation process of an initial labelled configuration. The regularisation is operated through Potts model which is a function of regularisation parameter β and a neighbourhood topology adopted in the image (4-connexivity or 8-connexivity). ICM Development consists to sweep the whole of sites image (initial configuration) and to choose for each site the class which minimises the energy function given by expression (9). This operation must be repeated a number of times to reach a stationary state flowing the selected convergence criterion. This relaxation technique is fast, but strongly depends on the initial configuration and regularisation parameters. A stochastic algorithm like a simulated annealing or genetic algorithm (Khedam *et al.*, 2003) can be operated in the same way but using a random initial configuration and allowing local energy increasing. Optimal convergence is obtained after a great number of iterations.

5. EXPERIMENTAL RESULTS

We have tested the presented classification process on SPOT image acquired on February 23, 1986. It contains three spectral bands covering Blida region located at 50km in the south-west of capital Algiers (north of Algeria) as shown on figure 3. Our data set of size 256x256 is presented on figure 4. A composite color have be done on this set. The aim of this pre-processing is to have a better visual interpretation of the scene and to be able to identify representative areas which will constitute a training base for the supervised process. Recall that prior to supervised classifications an unsupervised cluster classification can be applied to uncover the major land cover classes that exist in the image, without prior knowledge of what they might be. Seven discriminating classes have been defined and presented on table 1. According to these classes, we define a training samples image (figure 5.(a)) which will be used for classification and ground truth image (figure 5.(b)) which will be used for assessing classification accuracy. The training stage is important since its characteristics determine the outcome of the classification. In theory, a statistically based algorithm requires a minimum of n+1 pixels for training in each class (Brogaard *et al.*, 1998), where n is the number of wavelength bands. However, in practice, the use of a minimum of 10n to 100n is advised by Lillesand and Kiefer (1994). Numbers of pixels used for classification and those used for assessing classification are given on table 2. To applied ICM algorithm, a good initial configuration is required. For our study, we first execute the gaussian maximum likelihood algorithm to product a punctual classification (figure 6. (a)) which may be non-optimal and need so to be improved. Secondly, we execute a regularisation process with taking a punctual classification result as the initial configuration. Generally, regularisation parameter β is selected in an empirical way. In the present study, β is taken 0,8 and a 8-connexivity is adopted. ICM classification result is presented on figure 6.(b). A statistical classification assessing is carried out by means of confusion matrix established between truth ground and obtained classifications. From this matrix, is calculated the statistical parameter "kappa" (Congalton, 1991) which is a global indicator of classification accuracy. Let be Xij the confusion matrix elements, $Xi+$ the total sum of elements in lines, $X+i$ the total sum of elements in columns, Xii the diagonal elements, N the total number of the pixels of the matrix and M the number of considered classes "Kappa" is given by the following expression:

$$\hat{K} = \frac{N\sum_{i=1}^{M} X_{ii} - \sum_{i=1}^{M}\left(X_{i+}\times X_{j+}\right)}{N^{2} - \sum_{i=1}^{M}\left(X_{i+}\times X_{j+}\right)} \qquad (10)$$

Note that if a confusion matrix is established between the classified image and a truth ground representing only some homogeneous pieces of the scene, then expression (10) represents statistical parameter called "Khat". There is another more significant criterion introduced recently by Shabah (Shabah *et al.*, 2001). It is a local kappa calculated for each class i and given by the following expression:

$$\hat{K}_i = \frac{NX_{ii} - \left(X_{i+}\times X_{j+}\right)}{NX_{i+} - \left(X_{i+}\times X_{j+}\right)} \qquad (11)$$

The next section deals on the discussion of the experimental results.

Figure 3. The position of the study area in the northwestern part of Algeria,

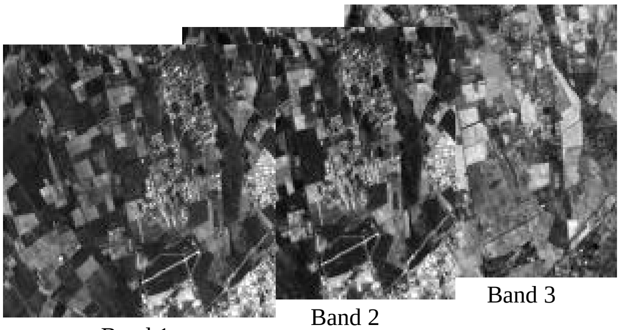
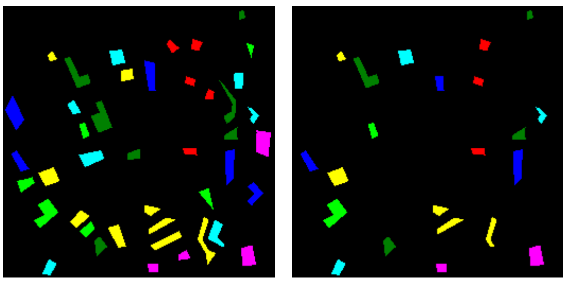


Figure 4. Data set study

CLASS	Object
1	Less dense urban
2	Less dense vegetation
3	Blida airport
4	Non cultivate fields
5	Dense urban (Blida city)
6	Cultivate fields
7	Dense vegetation

Table 1. Classes of study zone

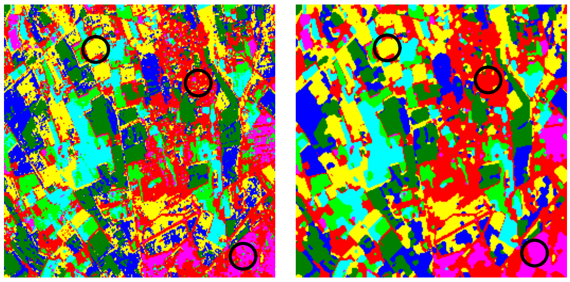


(a) (b)

Figure 5. (a) Training samples image – (b) Truth ground image

Classes	Training pixels	Ground truth pixels
1	357	220
2	1052	483
3	1348	620
4	773	378
5	643	290
6	984	349
7	1100	519

Table 2. Numbers of training pixels and ground truth pixels



(a) (b)

Figure 6. (a) Punctual classification (initial configuration) – (b) ICM classification (β=0,8 and 8-connexivity)

6. DISCUSSION

We have presented the optimisation algorithm we used to obtain Maximum *a posterior* (MAP) classification of remotely sensed data. This iterative algorithm is based on Markov Random Field (MRF) and exploits spatial class dependencies between neighbouring pixels in an image. It is a simpler and faster version of Geman's algorithm (Geman *et al.*, 1984). Applied on the data set of size 256x256, ICM convergence is reached after 13 iterations only. For this reason, ICM classification algorithm is selected to keep the computational complexity of MAP approach at an acceptable level. Performance of the obtained classifications is evaluated by calculating kappa parameters derived from confusion matrix and given by equation 10 and 11. The resulted classified imagery using context is find to reveal globally and locally more patch-like and meaningful patterns. This visual interpretation is confirmed by statistics given on table 3 and by graph of figure 7. It is shown that the incorporation of contextual information leads to impressively improved results, up to 84% of global accuracy is achieved in comparison with the out put derived from traditional punctual maximum likelihood (MLLH) classifier where only around 70% of global accuracy is obtained. Also, the classification accuracy is improved for each class.

Approach	Kappa (%)
Punctual (MLLH)	72.6
MAP (ICM)	84.23

Table 3. Global classification accuracy

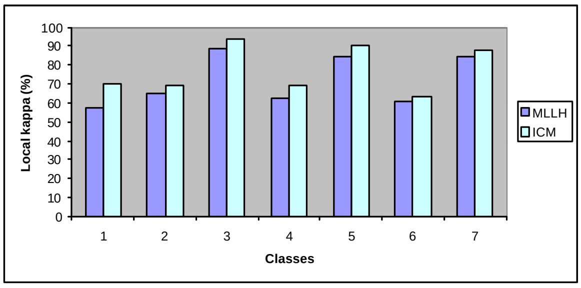


Figure 7. Local classification accuracy

7. CONCLUSION

The purpose of this work is to design robust algorithm for classification of remotely sensed images. Our experience confirms that context information plays an important role in the task of scene interpretation. At the pixel level, context information provides neighbourhood information around a pixel, and helps to increase the reliability of each detect object. Discrete random fields, especially the Gibbs Random Fields (GRF) and Markov Random Fields (MRF) provide a methodological framework which allows the integration of context information in satellite data classification. A powerful of these models is that the prior probability density function modelled by the use of the contextual information and the class conditional probability density function modelled by the use of the observed data from one or more sensors, can be easily combined through the use of suitable energy function. Once the posterior energy model and the associated parameters have been defined, pixel labelling is found out by using the MAP estimate which is equivalent to a minimum energy function in terms of GRF-MRF modelling. For a non-convex energy function, the solution space may contain several local minimum. To find a global minimum which is a truly MAP estimate, the solution is to use an optimisation algorithm among which ICM is the most know and used. The ICM algorithm is sub-optimal and converges only to a local minimum of the energy function. However, classification result of such algorithm is acceptable and shows that the incorporation of contextual information successfully improves classifier performances by more than 10% in terms of global accuracy. However, algorithms and methods to construct more complex models and to efficiently integrate context (context at object level which is useful for obtaining a coherent interpretation of the whole scene) in order to achieve higher classification accuracy, are still significant issues worthy of further investigation.

References

Amadamn, M., and King, R. A., 1988. Low level segmentation of multispectral images via agglomerative clustering of uniform neighborhoods. *Patterns recognition*, 21, pp. 261-268.

Besag, J., 1986. On the Statistical Analysis of Dirty Pictures. *Journal Royal of Statistics: Soc.* B, 48, 3, pp. 259-302.

Braathen, B., Pieczynski, W., and Masson, P., 1993. Global and local methods of unsupervised Bayesian segmentation of images. *Machine Graphics and vision*, vol. 2, no. 1, pp. 39-52.

Brogaard, S., Prieler, S., 1998. Land cover in the Horqin Grasslands, North China. Detecting changes between 1975 and 1990 by means of remote sensing. *Interim report on work of IIASA*, IR-98-044/July.

Congalton, R. G., 1991. A Review of Assessing the Accuracy of Classific ations of Remotely Sensed *Data*. *Remote Sensing of Environment*, no. 37, pp. 35-46.

Desachy, J., 1991. Interprétation automatique d'images satellitaires : le système ICARE. Thèse de Doctorat en informatique, Université Paul Sabatier, Toulouse, France.

Geman, S. and Geman, D. 1984. Stochastic Relaxation, Gibbs Distribution, and the Bayesian Restoration of Images. *IEEE Trans. Pattern Analysis and Machine Intelligence, PAMI-6,* pp. 721-741.

Kartikeyan, B., Gopalakrishna, B., Kalubarme, M. H., and Majumder, K. L., 1994. Contextual techniques for classification of high and low resolution remote sensing data. *IJRS*, 1994, vol. 15, No. 5, pp. 1037-1051.

Kettig, R. L., d Landgrebe, D. A., 1987. Classification of multispectral image data by extraction and classification of homogenous objects. *IEEE. Transactions on Geoscience and Remote Sensing,* GE-14, pp. 17-26.

Khedam, R., Belhadj-Aissa, A., 2001. General Multisource Contextual Classification Model of Remotely Sensed Imagery based on MRF. IEEE / ISPRS Workshop on Remote Sensing and Data Fusion Over Urban Areas, Rome, Italy, November 8-9th 2001. *IEEE Catalog Number 01EX482C.* ISBN 0-7803-7060-0.

Khedam, R., Belhadj-Aissa, A. and Ranchin, T., 2002. Study of ICM parameters influence on imges satellite contextual classification. *In proceedings of the 22nd symposium of the European association of remote sensing laboratories, Prague, Czech,* 4-6 june, pp. 79-85.

Khedam, R., Belhadj-Aissa, A., 2003. Contextual fusion by genetic approach applied to the classification of satellite images. *Remote sensing in transition, Goossens (ed.), 2004 Millpresse, Rotterdam,* ISBN 90 5966 007 2.

Lillesand, T. M., and Kiefer, R. W., 1987. Remote sensing and image interpretation. *John Wiley and sons: New York.*

Marroquin, J., Mitter, S., Poggio, T., 1987. Probabilistic solution of ill-posed problems in computational vision. *Journal of the American Statistical Association,* no. 82, pp. 76-89.

Pieczynski, W., 2000. Segmentation Statistiques d'Images. Notes de cours, Institut National des Télécommunications, France, 95 p.

Shabah, M. A., Dikshit, O., 2001. Improvement of classification in urban areas by the use of textural features: the case study of Lucknow city, Uttar Pradesh. *IJRS*, 22, 4, pp. 565-593.

Richards, J. A., Landgrebe, D. A., and Swain, P. H., 1981. Pixel labelling by supervised probabilistic relaxation. *IEEE. Transactions on Pattern Analysis and Machine Intelligence, PAMI-3,* pp. 188-191.

Toussaint, G. C., 1978. The use of context in pattern recognition. *Pattern Recognition,* 10, pp. 189-204.