Journal of Geomatics Vol 10 No. 1 April 2016

Object based classification techniques for citrus orchards

Kadapala Bharath Kumar Reddy[1], J. Mohammed Ahamed[2], R. Hebbar[a], Uday Raj[2] and Rama Rao[1]
[1]Indian Institute of Space Science and Technology, Thiruvananthapuram, Kerala, India
[2]RRSC – South, National Remote Sensing Centre, Bengaluru, Karnataka, India
Email: bharathkadapala@gmail.com, (mohammedahamed_j, hebbar_kr, uday_raj)@nrsc.gov.in

(Received: Mar 02, 2016; in final form: Apr 13, 2016)

Abstract: Cultivation of citrus orchards is characterized by small field size, sparse distribution with different age groups and coexisting spectrally similar crops. Thus, classification of citrus crop using satellite data is quite challenging. The present study was carried out to evaluate object based classification techniques for mapping citrus orchards located at Indi hobli of Bijapur district, Karnataka state. Two date LISS-IV and single date Cartosat-1 data were used for classification. Spectral signature of young orchards less than 5 to 6 years and poorly managed orchards were mixed with pomegranate, sugarcane and grape orchards. Two approaches of segmentation techniques namely, threshold & clump and lambda schedule, were tested. The results revealed that the single date satellite data showed classification accuracy (around 75%) using both threshold and clump and lambda schedule segmentation approach. Inclusion of second date data along with vegetation indices significantly improved the mapping accuracy (around 85%) by eliminating short duration crops from evergreen citrus orchards. The study explored the potential use of high resolution data for inventory of citrus orchards and the methodology could be refined for operational application using textural and contextual information.

Keywords: Object based image analysis, Full lambda schedule Segmentation, Threshold and clumping, Citrus, Brovey transform

1. Introduction

Image classification is perhaps the most important part of digital image analysis. It is very nice to have a "pretty picture" or an image, showing a magnitude of colors illustrating various features of the underlying terrain, but it is quite useless unless it is known what the colors mean (Korgaonkar, 2012). The data obtained through remote sensing satellites has huge applications in the fields of agriculture, urban modeling, disaster management etc. For all these applications a well classified data is required i.e., image classification has to be carried out in order to group all the pixels into several land cover classes and this further can be used according to user needs.

1.1 Classification methods

Classification methods can be mainly categorized into two types namely:

1. Pixel based classification
2. Object based classification

1.1.1 Pixel based classification: The pixel based methods use the digital number associated with the pixel in order to assign it to a specific class *i.e.*, the spectral information of the pixel is used for the classification. There are mainly two methods in pixel based classification methods, Supervised and Unsupervised classification.

1.1.2 Object based classification: The object based image analysis (OBIA) or specifically Geographic Object Based Image Analysis (GEOBIA) when it comes to satellite images delineate readily usable objects from imagery while at the same time combining image processing and GIS functionalities in order to utilize spectral, texture and contextual information in an integrative way(Blaschke, 2010). Dissatisfaction of using pixels solely in the classification has been mentioned long back (Cracknell, 1998). Spectrally similar but compositionally different land cover may be misclassified. Similarly, the spectral heterogeneity of the land cover can lead to rogue pixels appearing within classes creating a 'salt and pepper' effect. In addition to this, the increased application of higher resolution imagery is problematic as it is difficult to classify accurately using traditional pixel-based methods. The increased amount of spatial information often leads to an inconsistent classification of pixels(Whiteside, 2005). In GEOBIA the spatial information in the neighborhood is also considered which allows increasing the dimensionality of the feature space of a pixel when compared to the traditional pixel based methods where, only spectral values of the pixel are used in the feature space and hence giving us the more reliable results than those methods. Hence, in this study the object based classification methods were evaluated to identify the best object based method in mapping citrus orchards.

2. Data used

Multi-spectral images of LISS IV camera and CARTOSAT data were used. Indi taluk of Bijapur District, Karnataka was identified for the study. This taluk is located with the following boundaries of coordinates: Upper bound ULX: 75° 43'36.13", ULY: 17° 0' 20.49" and Lower bound LRX: 76° 03'3.16" LRY: 16° 56'17.04". The dominant crop is Citrus with small amount of grapes, sugarcane and other crops grown in the region.

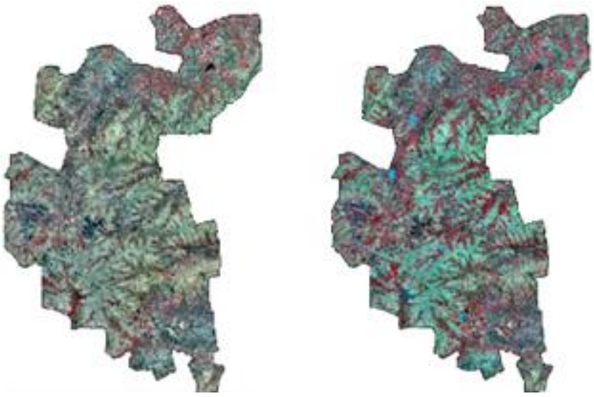


Figure: 1a) LISS IV May Figure: 1b) LISS IV November

Classification was done classification for the data obtained in the month of May (figure 1a) and also using the stacked data of May and November. Summer crops get harvested in the month of May and monsoon crops get harvested in the month of November (figure 1b). Therefore, these two months data was used to exclude seasonal crops as much as possible in the images.

LISS IV sensors have three bands with a spatial resolution of 5.8 m. CARTOSAT-1 panchromatic has one broad band with very good spatial resolution—2.5 m. Combining these two images to yield a three band data set with 2.5 m resolution provides the best characteristics of both sensors (NRSC, 2004).

Brovey Transform was used to merge the two data sets. The Brovey Transform was developed to visually increase contrast in the low and high ends of an image's histogram (*i.e.*, to provide contrast in shadows, water and high reflectance areas such as urban features). Consequently, the Brovey Transform should not be used if preserving the original scene radiometry is important. However, it is good for producing RGB images with a higher degree of contrast at the tails of the image histogram.

Our objective was to perform a binary class classification for the identification of Citrus orchards in the given area. Object based analysis of the image was carried out to identify the citrus orchards and mark all the other regions as the background.

3. Methodology and work flow

In ERDAS imagine, Objective Workstation was used. The first step was pixel level binary classification of the image and assigns probability for each and every pixel depending on how close it is to Citrus class. For this step sixtyfour training samples were given from both Citrus and Non Citrus equally covering all the variations. The next step *i.e.*, objects formation step which is an important step in the analysis. Two different methods were followed to create objects in this step. One was segmentation using full lambda schedule and the second was threshold and Clump.

3.1 Segmentation using full lambda schedule

This method iteratively merges adjacent segments based on a combination of spectral and spatial information. Merging occurs when the algorithm finds a pair of adjacent regions, *i* and *j*, such that the merging cost $t_{i,j}$ is less than a defined threshold lambda value:

$$t_{i,j} = \frac{\frac{|O_i|.|O_j|}{|O_i| + |O_j|}.\left\|u_i - u_j\right\|^2}{length(\partial(O_i, O_j))} \quad 1$$

where,

O_i is the area of region *i*.

Oj is the area of region *j*.

u_i is the average value in region *i*.

u_j is the average value in region *j*.

$\|u_i\text{-}u_j\|$ is the Euclidean distance between the spectral values of regions *i* and *j*.

length($\partial(O_i,O_j)$) is the length of the common boundary of O_i and Oj (Robinson, 2002).

While not necessarily derived from the Pixel Probability Layer, the raster object segments will have the zonal mean pixel probabilities as attributes. Output from the Probability Pixels to Raster Objects Operator contains pixels that are grouped as raster objects which have associated probability attributes (Erdas, 2013).

3.2 Threshold and clumping

This operator performs a threshold operation on a pixel probability layer which keeps only pixels with a probability greater than or equal to the threshold value specified. It converts the pixels to binary values (0's and 1's), then performs a contiguity operation (clump) on the binary values of 1(Imagine Objective, 2010).

Now the objects formed were edited from the above two methods by applying filters over the objects to refine the results. Focal filter was used in the beginning to make the object edges sharp and remove very narrow strips from the objects formed. This filter replaces the most repeating value in the window of specific size around the pixel. The optimum value for this was found to be 3. The next filter used was probability filter which removes the objects of lower probability and gives us more reliable objects after filtering. Later size filter was used to filter out very small objects. Finally, re-clump filter was used on objects to form bigger objects. Raster to vector conversion was made using polygon tracing and later two more filter are applied on the vector objects. The smoothening filter which eliminates the sharp edges of objects and the island filter takes out island like gaps in the formed objects were used in post processing operation. The work flow is pictorially represented in a flowchart in the Appendix-1.

4. Results and discussions

The main aim of the entire process was to classify the mature citrus orchards. It can be seen how different parameter sets (Table 1) had performed in citrus dominant and non-citrus dominant regions from images shown (Figure 2a, 2b).

It was also observed that few parameter sets had classified the citrus very well in its dominant region but there was a lot of misclassification in the non-citrus dominant region. Two different assessment methods were used to calculate accuracy. The first method is from area estimation based on the actual data obtained from the government statistics. This method might not give precise idea about which method had the best performance because if some non- citrus dominant regions are classified as citrus and vice-versa, it may match the reported area even though spatially the classification was incorrect. The area accuracy results of this method are mentioned in table 1. To overcome the above mentioned problem the second method, i.e., Point based assessment was carried out. Here, random points were generated with 30samples per class. Later analysis was carried out to identify correctly classified samples using ground based GPS locations. The results of this assessment method are described in the form of confusion matrix (Table 2). It can be seen that highest accuracy of 85.07 per cent was observed for the set 11 in which NDVI was a stacked along with multi-data data using lambda schedule segmentation.

Table 1: Object creation methods, parameters used and respective area statistics and accuracy obtained in identifying citrus orchards

Single Date/ Multi Date	Object Creation Method	No.	Spectral	Texture	Size	Shape	Probability Threshold	Estimated area (ha)	Mapping Accuracy (%)
Single Date	Segmentation- Lambda Schedule	Set1	0.8	1	0.1	0.05	NA	2490	61.81
		Set2	0.6	1	0.15	0.05	NA	1146	63.59
		Set3	0.6	1	0.15	0.05	NA	1140	63.27
	Threshold/ Clump	Set4	NA	NA	NA	NA	0.95	1888	95.24
		Set5	NA	NA	NA	NA	0.9	1152	63.91
		Set6	NA	NA	NA	NA	0.95	1557	86.40
Multi Date	Segmentation- Lambda Schedule	Set7	0.5	1	0.02	0.02	NA	843	46.79
		Set8	0.6	1	0.15	0.05	NA	1402	77.79
		Set9	0.8	1	0.15	0.05	NA	1411	78.30
	Threshold/ Clump	Set10	NA	NA	NA	NA	0.9	1460	81.03
Multi Date + NDVI	Segmentation- Lambda Schedule	Set11	0.7	1	0.15	0.05	NA	1440	79.90

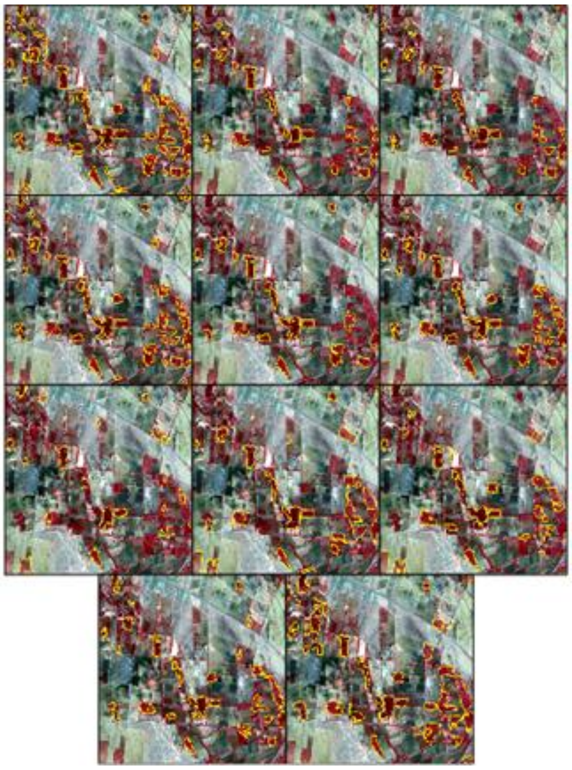
4.1 Lambda schedule segmentation (single date)

In raster object operation (ROO) process the probability filter values were tested iteratively from 0.45 to 0.8. The optimized value for probability filter was found to be 0.6 for lambda schedule and 0.7 for the Threshold and Clump process. For Lambda Schedule it was observed that by giving 0.7 as probability filter value area under citrus drastically changed resulting in under classification and decreasing it to 0.5, resulted in misclassifying non-citrus regions. Hence the optimum value for this parameter was found to be 0.6. In threshold and clump method the value was increased to 0.7 because in this step, only high probable pixels were taken into account in the raster object creation (ROC) step itself which precedes ROO step in segmentation. Hence, higher values were needed to be kept as cut-off for probability filter.

The set 1 estimated the area as 2490 hectares and overall accuracy was 73.13 per cent. But it was observed that a lot of non-citrus regions were classified as citrus orchards. Hence, changes were made in probability filter in ROO step and spectral value in ROC step by increasing the probability filter value to 0.6 and reducing the spectral value to 0.6, which gave very less area estimate 1146 hectares as compared with 1800 hectares of reported area. However, misclassification of non-citrus regions was reduced.

4.2 Threshold and clump segmentation (single date)

In threshold and clump the threshold probability in ROC was initially set to be 0.95 (table 2) and probability value of 0.6 in ROO, which resulted in smaller object size and misclassification of non-citrus segments. Even though the area estimated looked very precise (95.24 per cent) spatially there was significant misclassification. Hence, the probability filter value in ROO step was increased to 0.8 which reduced the misclassification. It also had smaller segments of objects existing adjacently which could join to form a bigger object if threshold in the object formation step is reduced. To achieve this, there was a need to decrease probability value in ROC. Hence, two changes were made i.e., the threshold probability was set to 0.9 and probability filter in ROO was set to 0.7.



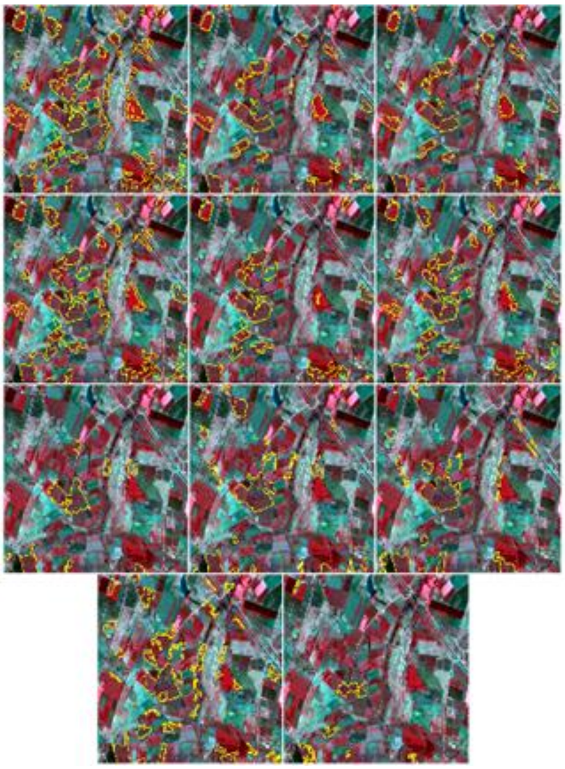


Figure 2a Figure 2b

Figure 2: From top left to bottom right in each of the figures show the classification results for 11 parameter sets from (a) citrus dominant region; and (b) non-citrus dominant region

Table 2: Table of confusion matrices for selected sets using point based assessment method

Lambda Schedule (Set 1)				Threshold and Clump(Set 5)			
	Citrus	Non citrus	Total Classified		Citrus	Non citrus	Total classified
Citrus	31	15	46	Citrus	27	11	38
Non citrus	3	18	21	Non Citrus	7	22	29
Total samples taken	34	33		Total samples taken	34	33	
Accuracy	91.17 %	54.55%	**73.13%**	Accuracy	79.41 %	66.67%	**73.13%**
Lambda Schedule (Set 2)				Lambda Schedule (Set 9)			
	Citrus	Non citrus	Total classified		Citrus	Non citrus	Total classified
Citrus	28	10	38	Citrus	29	9	38
Non citrus	6	23	29	Non Citrus	5	24	29
Total samples taken	34	33		Total samples taken	34	33	
Accuracy	82.35 %	69.70%	**76.12%**	Accuracy	85.29 %	72.72%	**79.10%**
Threshold and Clump (Set 10)				Lambda Schedule (Set 11)			
	Citrus	Non citrus	Total classified		Citrus	Non citrus	Total classified
Citrus	30	12	42	Citrus	31	6	37
Non citrus	4	21	25	Non Citrus	4	26	30
Total samples taken	34	33		Total samples taken	34	33	
Accuracy	88.23 %	63.63%	**76.12%**	Accuracy	91.14 %	78.90%	**85.07%**

The best methods in the two types of algorithms using single date data had accuracies around 76.12 per cent for lambda schedule segmentation (set 2 of table 2) and 73.13 per cent (set 5 of table 2) for threshold and clump method of segmentation. Hence there is a need to increase the feature space to capture the seasonal variability of crops and improve the classification. The data for the month of November was added which gave us better results (set 9 and 10 of table 2) compared to above mentioned single date classification methods.

4.3 Multi-date and multi-date with thematic layer included

Classification carried out using multi-date data, set 9 and set 10 (table 2) reduced the misclassification. Overall accuracy was 79.1 and 76.12 per cent respectively. The results of both the methods were observed and found that threshold clump method of segmentation had more misclassification occurring in non-citrus regions. This was observed clearly in the second accuracy assessment (Table 2). Also it was observed that threshold and clump method had unusual shaped objects. However, lambda schedule segmentation method reduced the misclassification significantly and the boundary shape of the objects created matched the field boundaries. Therefore, by considering the above factors into account lambda schedule segmentation method was used to include one more additional layer (Set 11 of table 1) of Normalized difference vegetation index (NDVI) to further attempt in improving the classification accuracy.

Since NDVI layer describes additional spectral information, iteratively a value of 0.7 was selected as spectral weightage instead of 0.8 which performed better by reducing misclassification with multi-date data. A lot of misclassification were removed which is desired. It was observed that temporal data with additional NDVI as a layer included in classification using lambda schedule segmentation significantly reduced the misclassification and also increased the mapping accuracy 79.9 per cent (table 1).

5. Conclusion and limitations

From the above results it was observed that lambda schedule segmentation algorithm worked well as a raster object creation method with multi-date data and when the data was combined with NDVI which adds the additional spectral information the results were enhanced. Although through the first assessment method where only total area is taken into consideration threshold and clumping method shows an accuracy of over 95.24% for set 4, however when checked with ground based GPS points, misclassification was clearly observed (table 2). Similar reason can also be concluded for other methods using different parameter sets. Even in the set 11 where NDVI was used along with multi-date data, it was observed that trees along the road side and clusters of trees were classified as citrus.

This method is applicable only to the mature Citrus orchards at this resolution limits. A similar attempt was made on young citrus plants too but unconvincing results were observed because of the resolution of this data as young citrus plants get mixed-up with other orchards. The young plants were not identifiable in the data; they looked similar to other fallow lands which made the method specific for mature orchards at this level of resolution.

Further this work can be validated by using this technique in dominant citrus growing regions in India. Use of higher resolution data products may enhance the classification accuracy. Reported statistics from state departments include both mature and young citrus orchards. Since, the classification technique targeted on mature citrus orchards, area estimated by this technique might be on par with the reported data.

References

Blaschke, T. (2010). Object based image analysis for remote sensing. ISPRS Journal of Photogrammetry and Remote Sensing, 1-6.

Blaschke T. and J. Stroble (2001). What’s wrong with pixels? Some recent developments interfacing remote sensing and GIS, *GIS-Zeitschrift für Geoinformationsysteme* 6:12-17.

Cracknell, A. (1998). Synergy in remote sensing — what’s in a pixel? International Journal of Remote Sensing, 2025–2047.

Crisp, D.J. P. Perry and N.J. Redding (2002). Fast segmentation of large images. Edenberg, South Australia: Defence Research and Surveillance & Reconnaisanse Division,*DSTO Electronics and Surveillance Research Laboratory*.

Erdas (2010). IMAGINE Objective. Norcross, USA: ERDAS Inc.

Erdas (2013). Erdas Field Guide.

Korgaonkar, G.S., R.R.Sedamkar and K. Bhandari (2012). Hyperspectral image classification on decision level fusion. International Journal of Computer Applications (IJCA), 1-2.

NRSC (2004). IRS-P6 data user’s Manual. Hyderabad.

Robinson, D. J. (2002). Implementation of a fast algorithm for segmenting SAR imagery. Edinburgh, South Australia, Australia: DSTO Electronics and Surveillance Research Laboratory.

Whiteside, T. and W. Ahamad (2005). A comparison of object-oriented and pixel-based methods in northern

Australia. Proceedings of SSC2005 Spatial intelligence, innovation and praxis: The national biennial Conference of the Spatial Sciences Institute, September 2005. Melbourne: Spatial Sciences Institute.

APPENDIX – 1

(Flowchart describing the methodology adopted in identification of citrus orchards)