MONITORING LAND COVER IN SOUTHERN FRANCE: A PROJECT FOR TEACHING REMOTE SENSING CLASSIFICATION TECHNIQUES

drs. R. Knippers (International Institute for Aerospace Survey and Earth Sciences ITC)
drs. B. Köbben (Cartography Section, Faculty of Geographical Sciences, Utrecht University)
PO Box 80115, 3508 TC Utrecht, The Netherlands

Abstract
This paper describes a project on monitoring the land cover of the Calavon Valley in the Departement de Vaucluse in Southern France, using remote sensing imagery and classification techniques. Its main purpose is the teaching of these techniques to students of Cartography at Utrecht University and the International Institute for Aerospace Survey and Earth Sciences (ITC) in the Netherlands. It is also used to get an insight into the possibilities and problems of methods of ground-truth gathering and various classification techniques.

1. Introduction

The Calavon Valley is an agricultural area, with a wide variety of crops. The traditional vineyards have in a great extent been removed in favour of orchards and annual fruit crops (hence its nickname 'Fruit garden of France') and irrigated vegetables.
The Terrain is very hilly and parcels are irregular and very small. The average size of 432 parcels in a test area (around the village of Bonnieux) was determined and was found to be little over 1,5 ha. The agricultural area is mainly located on and near the valley floor, while the steeper slopes and hilltops are covered with woodlands and the so-called 'Garrigue', a thorny thicket.

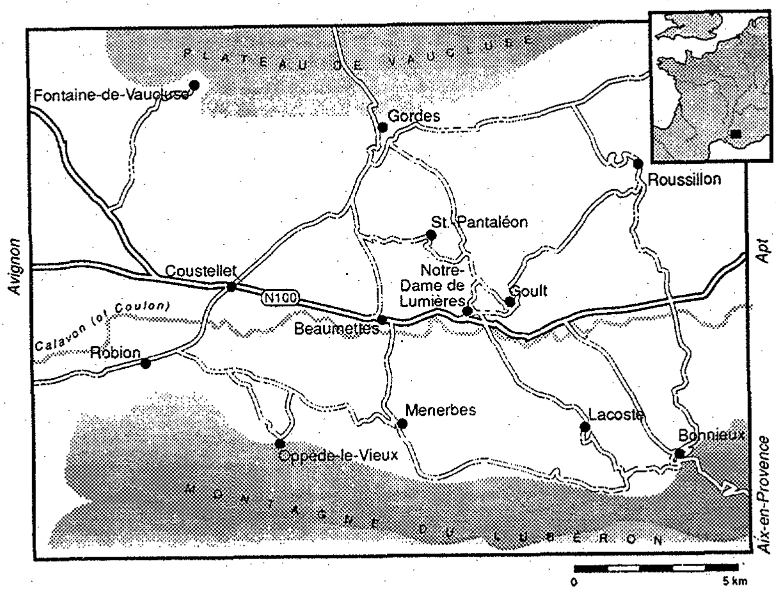

Figure 1: The fieldwork area

This valley has for many years now been the destination for a group of students majoring in Cartography of both Utrecht University and ITC. Two weeks of fieldwork offer them the opportunity to offer various cartographic skills[1]. From 1989 on, remote sensing has been one of them and in recent years an exercise has been developed which lets the students complete a landcover classification of the valley, using SPOT-XS imagery.

The next paragraphs describe and evaluate the results of this exercise. Some attention will be given to the pre-processing of the imagery and the methods the students use to collect ground-truth. Comparisons will be made between the classification schemes used and the accuracy of the resulting landcover maps.

2. Outline of the exercise

The objectives of the land cover classification exercise are

- to experience the difficulties in choosing classes and training areas
- to be confronted with the problems in locating training areas in the field and determining the landcover class of pixel groups
- to gain some insight in the relationship between the accuracy and the amount of ground-truth collected and the accuracy of the resulting supervised classification
- to gain some experience in using standard image processing software to classify satellite data

The evening before the ground-truth gathering the students and staff jointly define a preliminary list of land cover classes. Many of the students tend to mix up land use and land cover. For the 1994 fieldwork they agreed on the following list of land cover classes:
1: deciduous forest; 2: coniferous forest; 3: garrigue (thorny thicket); 4: old vineyard; 5: new vineyard; 6: orchard; 7: grains; 8: bare ground/rock outcrops; 9: built-up area; 10: lavender; 11: grass; 12: water.

The first day in the field nine teams of two or three students have to locate training areas (class samples) in their sub-area (see figure 2; each area is approx. 3 km^2).

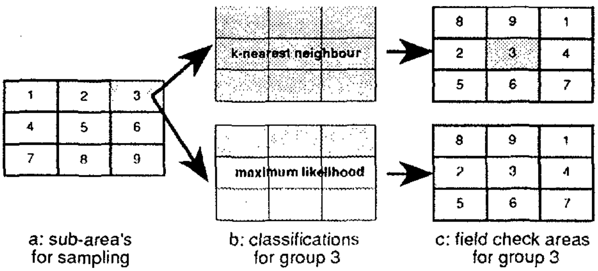

Figure 2: Sub-area's of the fieldwork area for 9 groups.

Each team has the opportunity to use an ILWIS[2] system for one evening to enter their class samples and classify the satellite imagery. Hardcopies at scale 1:15,000 of the classification are produced to be able to check the result of the classification.

[1] An overview of all fieldwork activities can be found at the ITC contribution in the poster session s.
[2] Integrated Land and Water Information System. A low-cost PC-bases GIS system with good image processing capabilities. ILWIS is programmed and marketed by the ITC.

The second day in the field each team checks the result of the classification. Based on this field check, an error or reliability matrix is produced. From this matrix the Percentage Correctly Classified has to be calculated.

The required result of this two and a half day's exercise are training data for a supervised classification, a classified satellite image and an error-matrix as a result of the field check.

3. Pre–processing of the satellite imagery

Imagery of the SPOT–XS instrument (multispectral data in three bands: green, red and near infrared with a resolution of 20x20 meters) is used for this exercise. Because of financial limits the images can be up to 3 years old.
Relief-displacement in the satellite imagery was corrected by means of a Digital Elevation Model[1]. The DEM was obtained through a linear interpolation of the 50m contours. These contours were densified where necessary and spot heights were included. Comparing the image and map co-ordinates of 20 check points distributed over the area showed an overall planimetric image accuracy of 23.18 m, which is very good considering the spatial resolution of the imagery of 20 m.
The spectral information within the imagery is ordered according to a vegetation reflection model. Such a model is particularly suitable for a land cover classification. The result of applying this model is a Leaf Area Index (LAI) image with information about the green leaf area, and an intensity image with information about the terrain slopes and the surface roughness of the land cover classes. The DEM was again used, now to separate the terrain slope information from the information on the surface roughness[2]. This corrected intensity image (IntCor) and the leaf area index image (LAI) were used for the classifications described in paragraph 5.

4. Methods of locating ground–truth

The main problem in locating ground-truth in the field is locating yourself on the satellite imagery: On what pixel am I actually standing now? This, we found, is a considerable problem for many students. We found the main tool necessary is a geometrically corrected data with a topographic reference.
Therefore we are using a false-colour coded hardcopy of the SPOT-XS data scaled to 1:15,000 with the main roads and streams on it, digitised from the topographic map, in combination with a transparent overlay of the 1:25,000 topographic map, enlarged to1:15,000.

Students are advised to first devise a route through there sub–area, in which they have to try to cover as big an area as possible and also as many different forms of landcover as possible. The experience with the surroundings they got in earlier exercises during this fieldwork, helps them with this.
The training areas should be groups of pixels with similar spectral characteristics of which you can determine the land-cover and which you can locate on the SPOT-image. Each team tries to find several samples of every land-cover class.

After returning from the field each team had to enter the samples in the ILWIS system. This appeared to be time consuming and error prone. Each team took about 3 hours to enter the samples in the ILWIS system. This appeared to be mainly because of difficulties in finding back on the screen–image, the locations of the samples they noted down on the hardcopy image. This is due to considerable colour differences between the hardcopy and the screen.
Using Global Positioning Systems (GPS) to assist ground-truth location could overcome these problems. Differential GPS with simple code receivers form an excellent data collec-

tion tool for locating ground-truth in this small scale area. Besides capturing the location of the samples, it is possible to add attribute data for each sample, like land cover class, parcel size, quality parameters, etc. During the 1995 fieldwork, tests will be made on a procedure to import the GPS-data into the ILWIS system and perform the classification. The findings hereof will be presented at the plenary session.

5. Classification methods

After the ground-truth is used to create sample sets for the sub-area's of every group, the ILWIS system is used to create land-cover classifications using two methods. For every group, these two classifications will be performed on the whole fieldwork area, as seen in figure 2b.
Using two different classification methods lets the students discover the considerable differences between the two resulting land-cover classifications (discussed in detail in the next paragraph) and thereby letting them appreciate the impact of choosing a classification method and its parameters.

The two methods used are among the most used in image processing practice: *k-Nearest Neighbour* and *Maximum Likelihood*.

Using the k-Nearest Neighbour classification, for every pixel to be classified (**o** in figure 3a) the nearest neighbours in the feature space are determined. Only neighbours within a certain search radius, as chosen by the user, are considered. If there is more than one neighbour, the class with the most neighbours (predominant class) will be selected. This would be class A in figure 3a. If no neighbours are found within the search radius the pixel is labelled 'not classified' [3,4].

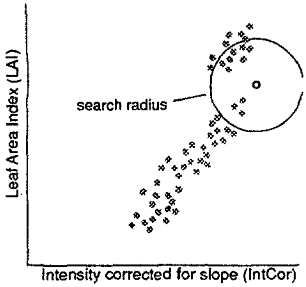

Figure 3a: k-Nearest Neighbour

Ideally the students would be allowed to experiment with the parameters to the classification, thus being able to compare the results and finally use the one that appears to give the best results. But this would make it impossible to compare the results of the various groups. Furthermore, the tight schedule does not permit much experimenting. Therefore a fixed set of parameters is determined by the staff and used by all the groups.

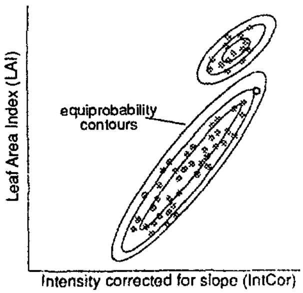

Figure 3b: Maximum Likelihood

The Maximum Likelihood classification works with likelihood (or equiprobability) contours which are constructed for all the classes in the sample sets. This results in areas of diminishing probability around the mean of every class sample, whose shape and size is determined by the statistical description of the sample classes. The pixel to be classified (**o** in figure 3b) is attributed to the class with the highest likelihood (probability) for that pixel. In figure 3b, this would be class B. If the likelihood of the pixel belonging to any class is less than a user-defined threshold value, the pixel is labelled 'not classified' [3,4].

All groups will perform these two classification methods and because every group uses its own sample sets, the results will differ considerably. At this stage the students will learn that:

- The area in which the samples are taken has great impact on the classification result. They will find that their particular area might not be a good representative for the total fieldwork area. In most instances they will have failed to find samples for one or more classes, therefore these will not be in their sample sets. This will result in a large proportion of unclassified pixels in the area where this class is an important part of the landcover. The groups that sample in the central river valley are a good example for this. They will seldom find good samples for 'garrigue/thorny thicket'. Therefore, the classifications of these groups will have a great deal of unclassified pixels on the Luberon hills, where this landcover is mainly found.
- The k-Nearest Neighbour classification usually yields more 'unclassified' pixels compared with the Maximum Likelihood classification. Because of this, students might at this stage consider the latter method the 'best'. This is of course much too simple a view, which we will try to rectify at the next stage (checking the results – see paragraph 6).

Afterwards the classifications for each group are printed. The prints covers only that part of the area in which the group will perform its classification check (see figure 2c).

6. Checking classification results

As Lillesand and Kiefer stated "a classification is not complete until its accuracy is assessed" [4, p.612]. Student groups go back into the field and check the results of their two classifications on a sub-area, other then their original sampling area (figure 2c).
Although taking random samples is the method most suited for a good classification assessment, this method has some practical disadvantages:

- Determining random pixels is easy enough, but reaching them in the terrain takes too much time, if they can be located at all. Here also, GPS could be of great use in the future, as we hope to demonstrate at the plenary session.
- random pixels can easily be border pixels or mixels (pixels in whose 20x20 m. area more than one type of landcover can be found).

To overcome this, students are advised to follow the same procedure as they did for taking ground-truth samples, namely traversing the area and noting down all those pixels they can not only confidently locate on the image but also determine the land cover of. These are all added for the whole area and a confusion matrix is constructed for both of the classifications (table 1).

When looking at these results, students should make several observations :

- The overall classification results, expressed as Percentage Correctly Classified (PCC) is low: 64,5% for the k-Nearest Neighbour classification and 59,2% for the Maximum Likelihood classification. Considering the inexperience of the students in gathering ground-truth and the difficulty of the small and irregular parcels, this was to be expected.
- The k-Nearest Neighbour classification yields significantly better results than the Maximum Likelihood classification. This should rectify the initial impression of the latter to be the 'best' method, because it resulted in less pixels labelled 'unclassified'. It's the quality that counts, not the quantity!
- There's much confusing between the two types of forest in both classifications.
- The same goes for the two types of vineyard, together with orchard and, in a lesser extent, bare ground.

FOUND IN FIELD

CLASSIFIED

Class	dec. for.	con. for.	garri-gue	old vin.	yng. vin.	orch.	grain	bare grnd	built –up	lav.	gra.	wat.	oth.	commis-sion
decid. forest	237	106	14		39		1						15	175
conif. forest	103	149											16	119
garri-gue	2		230	5	2		3	5					3	20
old vineyd.	1		18	339	23	10	1			1			15	69
young vineyd.				26	512	9	20						30	85
orchard			15	86	5	145	59						21	186
grains			6	3	25	4	260	2						40
bare ground	11		66	19		7	4	10					4	111
built-up	1	1		16	1				23				4	23
laven-der		2	8	4	2		2			3				18
grass										7	60			7
water												5	1	1
unclas-sified	23	15	16	88	45	33	76	2	10		1	2	146	311
omis-sion	141	124	143	247	142	63	166	9	10	8	1	2	109	

PCC (Percentage correctly classified) = $\frac{\text{\# of correct pixels}}{\text{total \# of pixels}} = \frac{2119}{3284} \times 100\% =$ 64,5 %

CLASSIFIED

Class	dec. for.	con. for.	garri-gue	old vin.	yng. vin.	orch.	grain	bare grnd	built–up	lav.	gra.	wat.	oth.	commis-sion
decid. forest	289	89												89
conif. forest	109	327				5								114
garri-gue	16		145		2	7								25
old vineyd.	16	2		201	89	9								116
young vineyd.	2			53	421	9	5		4					73
orchard				55	1	143	83	25					25	189
grains	1		72		15	15	187	2	7				8	120
bare ground								0					20	20
built-up					1				24					1
laven-der	1			5		5				12				11
grass											5			0
water												3	2	2
unclas-sified	22	50	44	108	242	46	53	12				3	187	580
omis-sion	167	141	116	221	350	96	141	39	11	0	0	3	55	

PCC (Percentage correctly classified) = $\frac{\text{\# of correct pixels}}{\text{total \# of pixels}} = \frac{1944}{3284} \times 100\% =$ 59,2 %

Table 1: Error matrices of some of the groups for the k–Nearest Neighbour (top) and Maximum Likelihood (below) classifications of June 1994. Values are *number of pixels,*.

- Some classes (built–up, lavender, water and grass) have very little samples attributed to them, either rightly or wrongly. When we get students to look back to their original training sets, these classes are usually also badly represented. Because of this, no valid conclusions can be drawn about these landcover classes.

To improve classification results, several things could be done. Firstly, classes with much confusion between them could be merged (eg. the two types of forest). Secondly, the classes with insufficient ground–truth data (eg. water) should be deleted all together. But the main improvement would come from gathering new ground–truth, using the things learned from this first–time experience on making a landcover classification. It's one of those things that only comes with experience!

7. Conclusion

This short, two–day exercise gives cartography students a good introduction to the problems and pitfalls of classifying satellite images. They get to experience all the stages, setting up the landcover classes, planning and executing ground–truth gathering, feeding the data to the image processing system, setting up and executing two types of classification and checking the results.

Thus the objectives for the exercise are met:

- The students get to experience the difficulties in choosing classes and training areas and see the considerable influence this has on results.
- They are confronted with the problems in locating training areas in the field and determining the landcover class of pixel groups during the first stage.
- They gain some insight in the relationship between the accuracy and the amount of ground-truth collected and the accuracy of the resulting supervised classification by being able to compare ground–truth collected with classification results and the real–life situation.
- They gain some experience in using standard image processing software to classify satellite data by using the ILWIS package on the fieldwork to do all image processing.

Literature

[1] Bargagli, A., Geometric aspects and DTM requirements related to feature extraction from SPOT images. ITC Msc–thesis, Enschede (1990).
[2] Pickering, R.P., Digital image analysis of SPOT multi–spectral data for topographic mapping. ITC Msc–thesis, Enschede (1990).
[3] ITC, ILWIS 1.41 User's Manual. ITC (1993).
[4] Lillesand, T.M. & R.W. Kiefer, Remote sensing and image interpretation. Wiley & Sons (1994).