THEMATIC MAPPING OF THE NAIVASHA-REGION (KENYA) FROM LANDSAT IMAGES

Árpád BARSI

Department of Photogrammetry
Technical University of Budapest
H-1521 Budapest, Hungary
Tel: 463 1187

Abstract

At the University of Karlsruhe as a graduating student I could join to a project created by the European Community for supporting the developing countries. In my thesis work I classified two images with relatively few ground data. I used the excellent possibilities of the DIDIX image processing package and passing-through with ARC/INFO. The described classification will be followed in the future by careful analysis and suggestion for cultivation.

Keywords: thematic classification, LANDSAT.

Some Words about the Working Area and the Data

From the very dry desert to the afro-alpine swamp six ecologic zones can be differentiated in Kenya. Most of the area is savannah. The meaning of the indian-spanish word is all but a forest. The flora is classified into two categories: high value rangeland and low value rangeland.

Vegetation types of the first category are bushland, woodland, grassland. All these types have subcategories depending on the species, height and swamping rate. There are much mixed classes (e.g. bushed grassland, wooded grassland, etc.), what makes the uniquee classification very difficult. Mostly the classification errors could be traced back to this fact. Vegetation categories of the low value rangeland are the forest, swamp and barren land.

Lake Naivasha can be found in the Central Kenya Rift, to the Northwest of Nairobi. The region has an area of 900 km^2(30 × 30 km), that means 1000 × 1000 pixels in LANDSAT TM images. The average height above the sea level is 2010 m, the highest point of the region is 2634 m. I have used two 'multitemporal' images: the first was made on 10 July 1987, the second on 1 March 1989. In order to get the best results I used all bands except the 6th one, the thermal infrared band.

Fig. 1. The image taken in March

Using the thematic classification I had to determine 38 categories. These categories were described in the previous phase of the project. I would like to show as an example the most analysed category:
Bush- and shrubland

b1   8   rift bush
b2   9   lava bush
b3   10   mountain bush
b4   11   high mountain bush
b5   12   clear mountain bush
b6   13   clear dry cedar
b7   14   clear moisture cedar
b8   15   clear bamboo

The first column contains class codes, the second one shows the identifiers for the later ARC/INFO analysis, and the last one includes class names. As this example shows, the classes were differentiated not only by species but also by geographic placement and cultivation.

Digital Image Processing

In the Institute for Photogrammetry and Remote Sensing of the University Karlsruhe an image processing system was developed for remote sensing purposes running on UNIX. DIDIX (Digital Image Data Processing in UNIX) is a modular package mirroring the philosophy of the operation system: it contains special modules for displaying, classification, Fourier-transformation, etc. and the user should select the ones that are necessary for his job.

After rectification and registration of the bands of the satellite images I have selected the training areas with my project adviser who knew this region well. During the selection it was very helpful that I could display both images simultaneously so the selections of the agricultural areas were simplified a bit.

The selection of the training areas was followed by the thematic classification. The mathematical formulas of the classification could be used only in that case when the distribution of the selected pixels by each training area was normal. Because of the inadequate ground information (obsolete topographic maps, quickly changing vegetation) I had to make an effort for using the training areas maximally but reliably.

The normality analysis module not only decides with a very clever algorithm whether the pixels of a given training area show normal distribution, but makes recursive subclass building, too. By using some connecting modules statistical values can be derived (Kernel hyperconfidence ellipsoid, Jeffrey–Matusita distance, transformed divergence) which are the basic data of the statistical analysis. Applying the values written in table form it can be examined to what extent the subclasses overlap, how many classification errors (a priori errors) can be expected, and whether new training areas should be created or some of them must be deleted. Moreover, the locations of thematic classes in the intensity space are displayed as well.

After the correction and modification of the training areas the new situation must be analysed again and these steps must be iterated until the estimated values of the classification errors or the related overlap values become acceptable.

g1 with 8 subclasses

scale : 1cm ~ 10.60 grayvalues

C1

103

67

31   52   31   52   31   52   C2

3

2

1

cl1:	26 pixels
cl2:	104 pixels
cl3:	126 pixels
cl4:	125 pixels
cl5:	310 pixels
cl6:	30 pixels
cl7:	39 pixels
cl8:	57 pixels

cl1:	0.0	0 it
cl2:	-1.0	1 it
cl3:	-1.9	1 it
cl4:	-0.4	1 it
cl5:	-1.2	1 it
cl6:	0.0	0 it
cl7:	0.0	0 it
cl8:	-1.3	10 it

C3

69

39

52   76   52   76   52   76   C4

3

2

8   5

6   7 4

C5

146

94

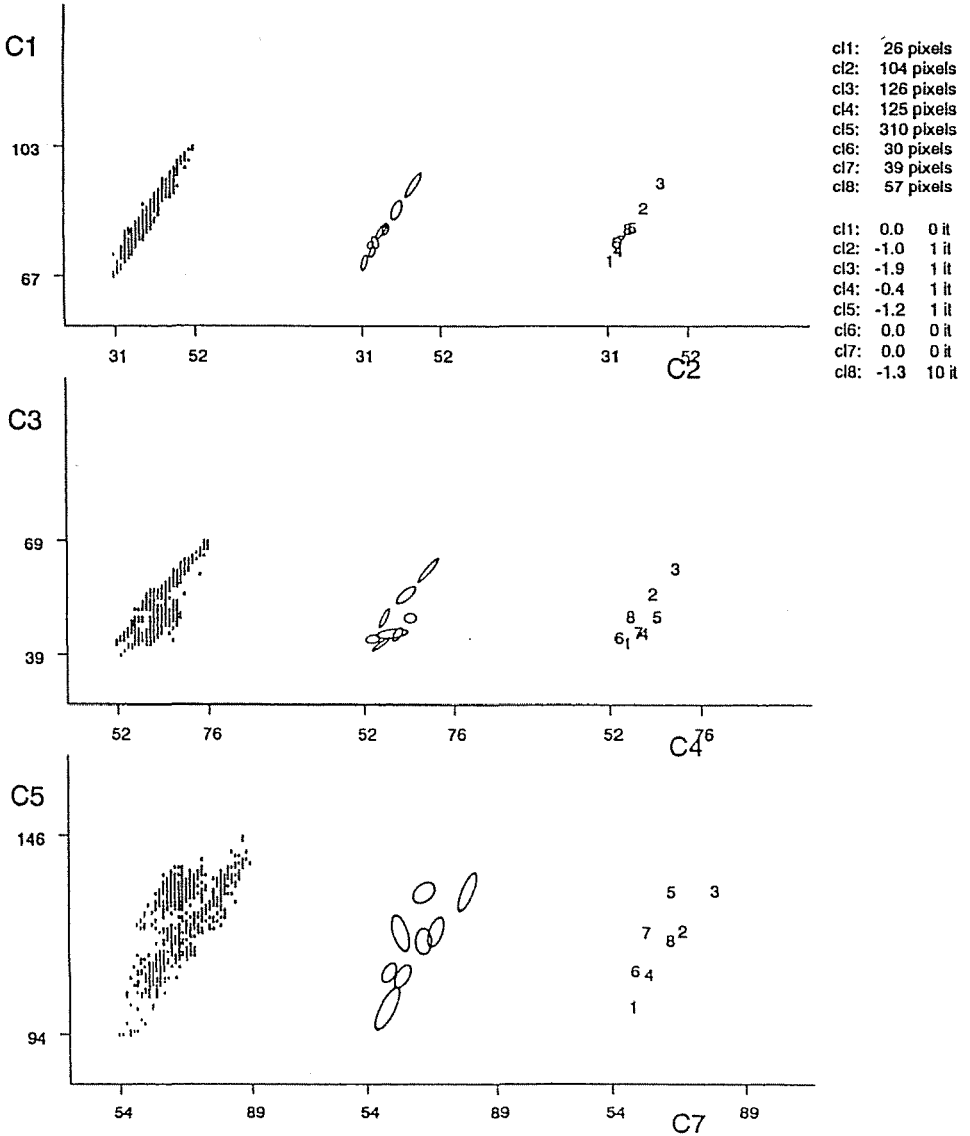54   89   54   89   54   89   C7

5   3

7   8 2

6   4

1

Fig. 2. A SYNCLASS detail

Now I can complete the classification using the so optimized training areas.

The classifier program differs from the conventional image processing packages. We can choose from several mathematical methods (e.g. maximum likelihood, Mahalanobis), but the most remarkable is that the program calculates the second or, when needed, the n-th probable classification. This makes the creation of a more accurate mathematical model of the thematic mapping possible.

Fig. 3. Last classification of the July image

In the result image we can find rejected pixels that would be classified with very low probability in spite of the carefully selected training areas. In consequence of this fact we need to select further training areas for

repeating the analysis and the classification. For example, after the first classification I got 14.0% of rejected pixels from the image taken in March and 9.4% of the image from July. After the last (fourth) classification these data are the following: 6.1% and 2.3%. (From these 6% of the March image those pixels should be subtracted, on which there were no data originally!)

Image analysis in ARC/INFO

After the digital image processing I submitted the result images to further analysis by applying ARC/Info's GRID module. I simplified the technics of this analysis using the Map Algebra Language so I could concentrate on the theoretical solution.
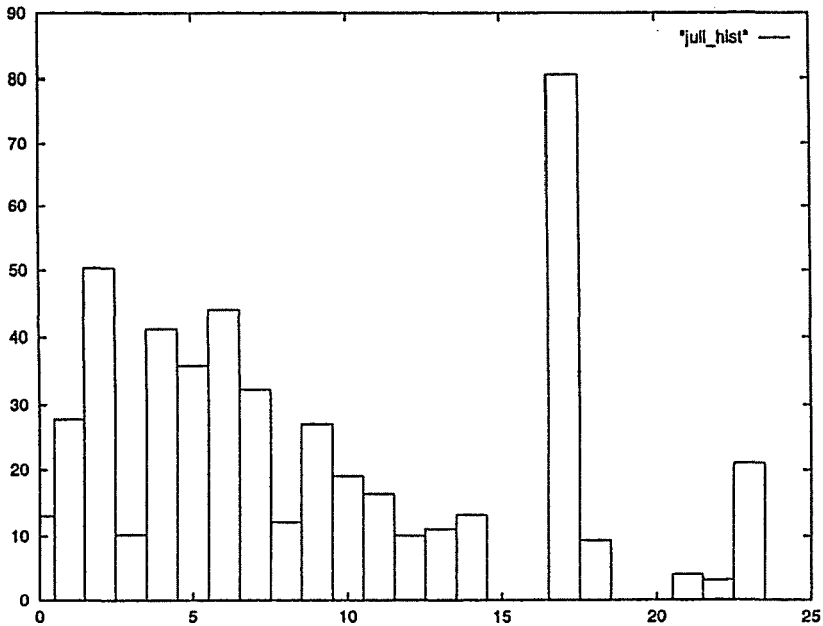

Fig. 4. The main diagonal of the combination matrix

After format conversion I produced all possible combinations of the pixels of both images with the combinatorial procedure called COMBINE. It was evident to collect these values into a matrix. Instead of absolute pixel values (frequencies) I used percentages. The meaning of a matrix element is: x percent of all pixels of a class (row) of the July image was classified as a class (column) in the March image. The elements of the main diagonal

are very interesting. They include the pixels which were classified into the same category in both images. A peak can be seen in the class 11 (lake deep) with the value of 80.5%, that means this group is stationary and can be classified reliably, whereas the agricultural categories have about 10% because of the changes.

After the general analysis I examined the cultivated areas. Using an AML-routine I could distinguish the change of the cover of agricultural areas and the possible, long-term change of cultivation. By introducing new identifiers these new thematics became selectable and displayable.

As endproduct I finished three images. On the first one I put the pixels of agricultural cultivated areas (all possible combinations) on the satellite image of March. Thus, the distribution of agriculture is visible. – By producing other images I used a crosscombination to avoid disturbing effects. It means that instead of clouds, shadows and rejected pixels I inserted the corresponding pixels of the other image and in this way I got two filtered, aesthetic images. (Using this method from the nearly 61 thousand rejected pixels of the March image only 3016 remain!)

References

1. ARC/INFO User's Guide, Cell-based Modelling with GRID Analysis, Display and Management ESRI 1991.
2. DIDIX On-line Manual and Help, IPF Karlsruhe, 1994.
3. JÄHNE, B.: Digitale Bildverarbeitung, Springer Verlag, 1991.
4. Manual of Remote Sensing, Ed-in-chief Colwell, R. M, Sheridan Press, 1983.