COLOUR MAP OBJECT SEPARATION

Amândio Cordeiro [(*)], Pedro Pina [(**)]

[(*)] Geographic Engineering Group, LATTEX, FCUL, University of Lisbon.
[(**)] CVRM, IST, Lisbon Technical University (UTL)

agcordeiro@fc.ul.pt, ppina@ist.utl.pt

KEYWORDS: RGB System, Colour Segmentation, Image Analysis, Map Objects Recognition, Mathematical Morphology.

ABSTRACT:

This paper addresses the issue of colour objects separation from map images obtained by current commercial scanners, in the RGB system. The proposed methodology is determined by some tested principles, as follows: **1. No Tonal Change**. Until its isolation, the colour-object maintains its former tonalities. **2. Unique transform.** The *pixels* that do not belong to the objects under consideration are "sent" to the background (value 255, in the 3 components). **3. Colour Spatial Correlation.** The decision of considering a non-object pixel depends only on its "cut value" in a component. A cut value divides the RGB colour cube in two separate volumes or Sections. **4. Spatial Integration**. Different Sections combine to form Regions, as intersections of Sections, for the noise-object and Co-Regions, as unions of Sections, for the isolation of the colour-objects.

1. INTRODUCTION

1.1 Definitions in the RGB System

The very nature of colour systems is a three-dimensional space, according to the CIE (*Commission Internationale de l'Eclairage*) definition. Therefore, we can associate each and every point of this three-dimensional space one tonality, defined by a 3 element vector that describes the Red-Green-Blue (RGB) System. Thus, for each pixel of a colour image we consider an application of $N^2 \rightarrow R^3$:

$$P\ (x, y) \rightarrow Z\ (r, g, b) \mid r, g, b \in [0, 255] \qquad (1)$$

in wich (x, y) specifies a column and a row beginning from the upper left corner of the image or pixels matrix. The r, g and b variables take any integer value in the [0, 255] interval. In addition, we must point out that the system assumes all negative values as zero and as 255 all values above this one. According to these specifications a cube of dimension 255, represents the RGB colourimetric space, as such.

We consider four entities in the RGB cube:

a) **Cut**: Plan defined by a constant coordinate. Example, considering r as a value of the R axis:

$$C_r \mid r = c \rightarrow C_r\ (c)\ ,\ c \in [0, 255] \qquad (2)$$

b) **Section**. All Cuts that are higher or not-higher than a single Cut. There are Upper Sections and Lower Sections. Example:

$$S_g \mid g > c \rightarrow S_{>g}\ (c)\ ,\ S_g \mid g \leq c \rightarrow S_{\leq g}\ (c)\ ,\ c \in [0, 255] \quad (3)$$

Read: Section sup/inf in g, by c.

c) **Region**. Space defined by the intersection of two or more Sections. Example:

$$R_{rg} \mid r > c_1 \wedge g \leq c_2 \rightarrow R_{>r,\ hg}\ (c_1,c_2),\ c_i \in [0, 255] \quad (4)$$

d) **Co-Region**. Space defined by the union of two or more Sections. Example:

$$R^*_{rg} \mid r > c_1 \vee g \leq c_2 \rightarrow R^*_{>r,\ \leq g}\ (c_1,c_2),\ c_i \in [0, 255] \quad (5)$$

This Co-Region designation is introduced as well to suggest the important relation with the previously defined Region. In fact, in the RGB space, both entities are complementary in the order relations, for the same Cuts, i.e.

$$R^*_{r\,g} = [R_{\bar{r}\,\bar{g}}]^c \qquad (6)$$

This is true if, considering the example at c),

$$R_{rg} \mid r > c_1 \wedge g \leq c_2 \rightarrow R_{\bar{r}\,\bar{g}} \mid r \leq c_1 \wedge g > c_2\ ,\ c_i \in [0, 255] \qquad (7)$$

and, looking at the De Morgan laws, we have

$$[R_{\bar{r}\,\bar{g}}]^c \mid r > c_1 \vee g \leq c_2 \text{ equiv. a } R^*_{rg} \mid r > c_1 \vee g \leq c_2\ ,\ c_i \in [0, 255] \qquad (8)$$

which is easily generalized to any number of Sections, as well as Regions implying not only two but three RGB components. For practical purposes the concept of complementary entities in the domain of Co-regions is established: two entities are complementary considering a fixed band if, changing the associated order relation, one goes from one to the other. This is verified, when considering the referred tonal complementary systems, for the colour-objects: Black/Red in the Red band, Black/Green and Red/Brown in the Green band.

1.2 Intersection of exclusions of the sampling space

Colour map objects are characterized by high chromatic complexity, although dominated by close tonalities. This is generally the result of the printing systems, the degradation of the map sheet and the scanning process. Because of this, each **colour plate**[*] is represented by a three-dimensional cloud inside the RGB cube.

[*] Recalling the printing process (Offset).

As immediate application, the identification of what is considered noise, is thus to isolate the corresponding cloud, defining a Region or parallelepiped, inside that cube. Soon after isolation, the next step is to withdraw it from the scene, in order to apply to the "recovered" image the separation process of the colour-objects (Fig. 1).

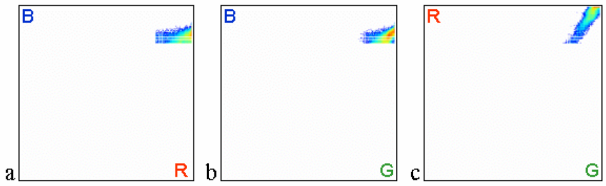

Fig. 1. Scattergrams showing the noise isolated by a Region R (200). The most reddish points reflect different maxima frequencies: a) 7339, b) 2389 and c) 11377.

The general rule in a cartographic document is the existence of various objects sharing somehow the same colour or tonality set. For easier speech understanding over colour separation, we refer very often to "colour-objects", taking it as a mixture of close tonalities. Therefore, one way of defining a map colour-object is its average spectral signature, keeping in mind that, extracting a sample inside the object, it shows currently significant lower values than those we can find at the edges. However, it should also be borne in mind that the limit values are the most important aspect for spatial object definition and tonality ranges.

On the other side, the background noise is normally very close to the RBG White vertex (Fig. 1). In the majority of the cases it is sufficient to model the noise by a surrounding cube containing that vertex. That is why we propose the following criterion for the determination of the Cut value for the correspondent Region:

> **Region R^0 of the background noise**: The Cut value, defining the cube enclosing the background noise, is determined by the greatest border value of the colour-object, that its minimum spectral average value is the greatest (Fig. 3).

In other words, assuming the previous boundary minima as, r^i_f , g^i_f , b^i_f , where the exponent *i* refers to each of the colour-object, we may state:

$$R^{o}_{>r,>g,>b}(c),\ c = \max\,[r^i_f , g^i_f , b^i_f],\ i = 1, 2, 3, \dots \qquad (9)$$

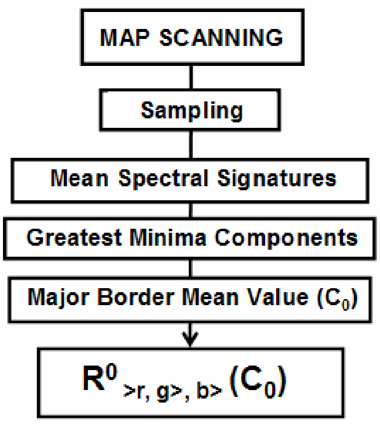

Fig. 2. Steps in the determination of the Noise-Region R^o.

In practice, following the methodology of Fig. 2 in order to determine R^o, it is sufficient to consider the colour-object that in the spectral signature, obtained by sampling, the minimum value among its components be higher than the minimum values of the other objects. Only in the selected object is the Cut value determined, by sampling the higher border values.

Fig. 3 shows an example of spectral signatures for different colour-objects: Black, Red, Blue, Green, and Brown. A broken line, in the same colour, also represents the edge values of the Brown (Contours lines). These values were responsible for the adoption of the Cut value for the background noise: R^o (180). The vertical lines refer to the RGB axes, in that order, from left to right.

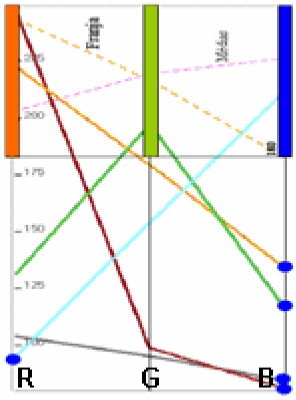

Fig. 3. Spectral signatures with minima components marked with blue dots.

As a region defined by an intersection, only the pixels with a higher value than the Cut value (180) are projected to the background as "white" (255). In the same example (Fig. 3) the lines of the spectral signatures are not entirely cut by the colour vertical bars, covering the complementary cut value of the region R^o. The pixels of the objects in Fig. 4a are retained this way, after noise elimination, as shown in Fig. 4b.

Fig. 4a. Original part of a map, with its yellowish background noise.

The pink broken line in Fig. 3 indicates the average values in each band. They are determined by the great quantity of high values of the background. They have no meaning in the calculations, as they depend on the information of the map.

Fig. 4b. The image from Fig. 4a without background noise.

1.3 Residual noise elimination

Previously we found out that in most cases a unique value defines the noise region, leading to its elimination. Nevertheless, it may happen that a residual background noise may remain, up to the level of 1%. In this case, the Mode filter, settles the problem.

Owing to the fact that even the most elementary Mode filter has the virtue of keeping almost unchanged the colorimetric sampling space, this filter has been selected to wipe out all the remaining noise.

This statistical operator replaces to advantage almost all morphological filters, because of its simplicity and effectiveness, generally attains idempotence* after the first application. Thus, any significant damage can attain the cartographic objects, mainly if we stick with a simple filter dimension of 3x3. In multimodal cases, the lower value is retained, avoiding the adoption of the value 255 in the border of the objects. The operator is applied to each one of the RGB bands.

2. COLOUR OBJECTS SEGMENTATION

2.1 General considerations

Unlikely, to the Region for the noise elimination, we now apply the concept of Co-region for object separation. In this case, the isolation of a colour-object implies the process of sending the other objects to the background colour (255).

When we find a higher interference among the object's tonal sets, between those with greater scattering (which happens more often with the Black layer), it may useful to introduce the spatial component, i.e., to take advantage of the proximity of neighbouring pixels of the same object. In normal cases, it is simply needed to perform the following sampling steps, in a "discriminating" sequence, for the extraction of the three Cut values leading to the definition of the Co-regions.

* Idempotence is reached when no changes are verified after renewed applications of the filter.

2.2 Sampling

The first Cut value to be considered is, naturally, the one established for the Noise-region. This value is at least sufficient for the separation of one object, the one that determined the noise border. It happens very often with the Brown set, because of its high minimum value in the Blue band of the border. As a result, this Cut probably acts as a boundary of the higher values in the Blue band. This is obviously the Blue object (Hydrography).

The determination of the second Cut value necessarily takes into account the existence of the first value leading to the definition of the first object or set of objects, although this case is not very probable. In any case, the number of objects to be isolated is reduced. Normally, this second value leads to the separation of two sets of two objects each.

Finally, the third Cut value is due to lead to the separation of the two objects in each group.

This explanation implies that, basically, we are dealing with maps with five colours, which was very common in recent analogue map series produced by many countries. The variety of colours found in geological series applies mainly to geological formations in well-defined areas. Other thematic mapping should be analyzed *de per si*.

We can classify as **Perfect Tonal Sets** those cartographic chromatic complexes where we can find a biunivoque correspondence between the tonal set of each colour object and a number of distinct Co-regions.

2.3 Union of exclusions of the sampling space

As the Region was defined as an intersection of exclusions of the sampling space, applied to the background noise, the concept of Co-region implies unions of exclusions of the sampling space, in order to isolate groups of colour-objects and, finally, each individual object.

We consider ourselves now prepared to deal with one paradigmatic example, continuing the process that led to the image of Fig. 4b.

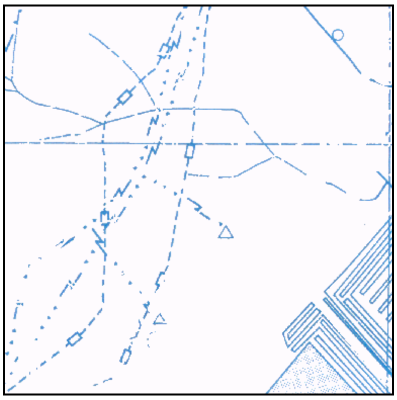

Fig. 5. Blue object isolated from Fig. 4b through a Section (Region) $S_{\leq b}$ (180).

The diagram of Fig. 2 points out the choice of the Cut level for the noise reduction that, in our case (very common) is in the Blue band and refers to the Brown object. We recall that this Region was established as R^o (180). It happens that this level also determines the isolation of the Blue object. This is attained considering a Co-region consisting of a simple Section: $S_{\leq b}$ (180). Therefore, from the noise-free image of Fig. 4b, we come to Fig. 5 showing the Blue object isolated.

To assert the second value we must take account of the one obtained in the first step, considering now the objects left to be isolated. In principle, we may expect that in this second step we will obtain two groups with two colour objects each. Nevertheless, as we can see in Fig. 10a the tonalities of the Blue object are very close to the Brown and Green ones, at the G band. This is why it is advisable to withdraw first the Blue object before we proceed.

If a certain object (Blue), has been extracted by

$$\text{if } I_b \leq c_1 \text{ then } 255 \text{ else } I_x \qquad (11)$$

then the complementary image can be obtained by

$$\text{if } I_b \leq c_1 \text{ then } I_x \text{ else } 255 \qquad (12)$$

remaining valid the principle for any Region or Co-region. The result is Fig. 7.

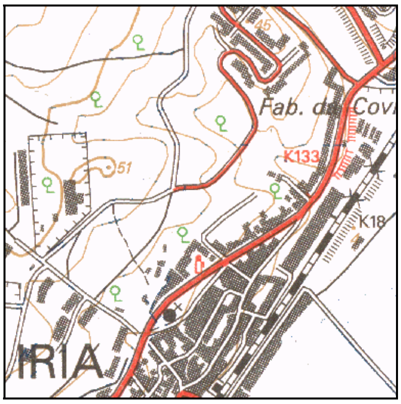

Fig. 7. Image without the Blue object.

This way, we arrived at an image (Fig. 7) which is almost perfect in the sense that it is now possible to separate the remaining objects with only three Cut values.

The following exclusion is dictated by the Cut value (160) that separates the Green and Brown objects on one hand and the Red and Black objects on the other. This follows the complementary principle stated by (11) and (12) and illustrated in Fig. 8 and Fig. 9.

Fig. 8. Inferior section retaining the Black and the Red objects.

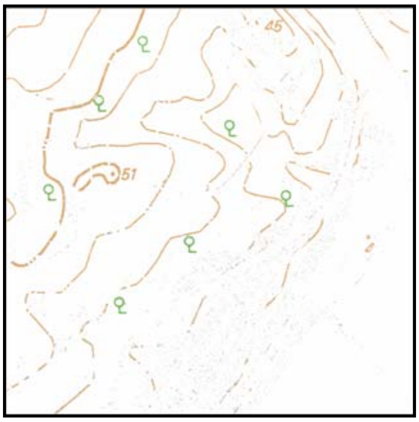

Fig. 9. Superior Section in G band, retaining the Green and Brown objects.

The diagram of Fig. 10 describes how those Cuts can lead to the isolation of the Black plate. The first Cut (180) in the Blue band, indicated in 9b (right side) by a small box with an arrow downwards and with the cardinal 1, suggests the elimination of the Blue object represented in the previous frame a. The Section determined by such Cut, send all pixels having a Blue component higher than 180 to the background (255, 255, 255).

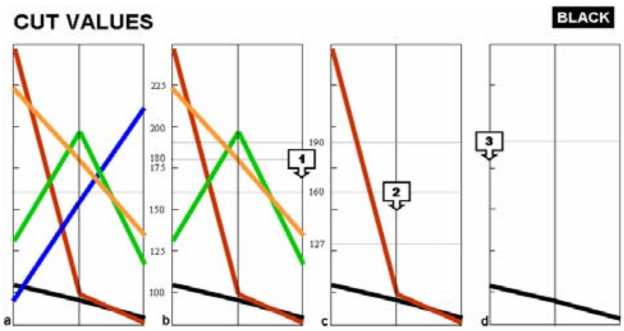

Fig. 10. Successive computation of Cut values and extraction of Black object.

One of the aimed results is the Black object plate in Fig. 11. As an example, the final expression for that plate would be one of the sort:

$$\text{If } (I_r > 190 \text{ or } I_g > 160 \text{ or } I_b > 180) \text{ then } 255 \text{ else } I_x. \qquad (13)$$

This obtains the Region $R^*_{>r,>g,>b}$ (190,160,180).

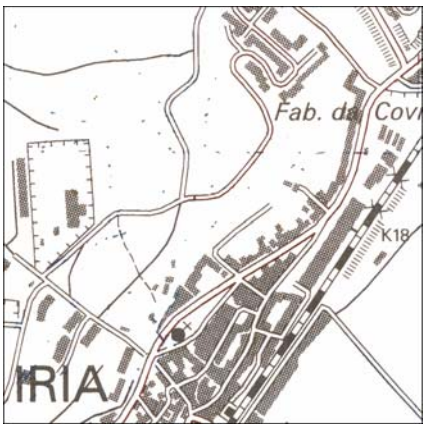

Fig. 11. Final Black object separated from Red object by the Region $R^*_{>r,>g,>b}$ (190,160,180).

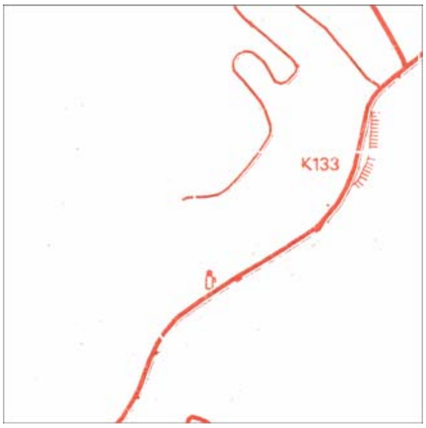

Fig. 12. Final Red object separated from Black object by the Region $R^*_{\le r, >g, >b}$ (190,160,180).

Similarly we can get the Green and Brown plates as showed in Fig. 13.

Fig. 13. Green and Brown by complementary Co-Regions: a) $R^*_{> r, \le g, >b}$ (190,160,180) and b) $R^*_{\le r, \le g, >b}$ (190,160,180).

3. IMPROVEMENT AND NORMALIZATION

3.1 General considerations

It happens that not all the images are considered adequate for subsequent treatment. One of the major concerns is the existence of gaps left by the intersections of colour objects. These gaps arise because the mixtures of inks create tonality sets away from the standard ones. They are generally much darker than the original; of course, this is normally not a problem with the Black plate. We look now for the Blue object in the previous example, as a paradigm.

We can see in Fig.11 that a little portion of the Blue information was "left" in the Black plate. It looks like noise that could be treated as such but, doing this, information would be lost. On the contrary, we can recycle this information and send it back to the Blue plate.

Analyzing the footprints of the Blue in the Black plate, we find out that these "stains" have a significantly higher level than the inner parts of the Black lines, allowing a Cut of 50 to 70 grey levels below the primary cut in B band (Fig. 14).

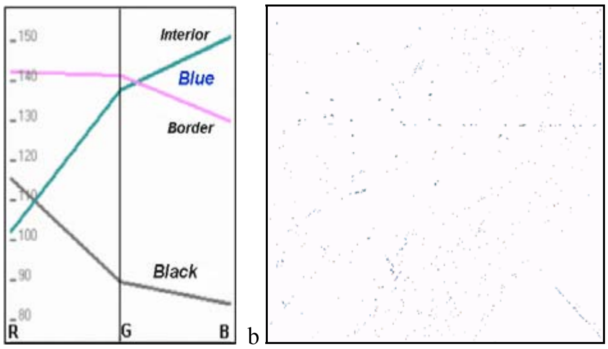

Fig. 14. Different signatures found at the remains of Blue in the Black plate, represented by the blue and rose lines (a) and image of the extracted remains (b).

Adding the Section resulting (Fig. 14) to the first Blue plate (Fig. 5), we get an image with a certain amount of noise which can be easily eliminated by the Mode filter. The final Blue plate is an image with fewer gaps and without noise (Fig. 15).

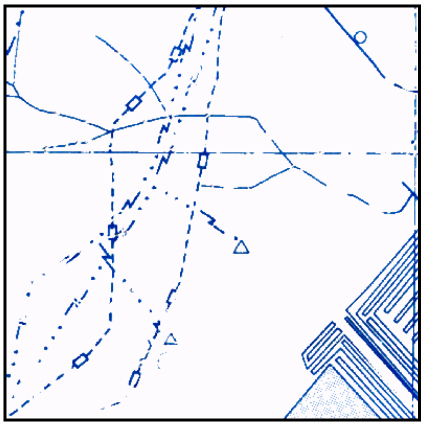

Fig. 15. Blue plate completed by information initially left with Black object.

Information rendered to the Blue plate, does not belong to the Black object. Therefore, we can withdraw that noise and come to a final Black object that can still be improved as final plate. In fact, in the process, some information has been wiped out and we can recover applying the Mathematical Morphology principle of Geodesic Reconstruction.

In the Geodesic Reconstruction, we take as mask the initial image (Fig. 11) and as marker the second image with less information. The initial noise will not be reconstructed and we get a complete Black plate ready for binarization (Fig. 16).

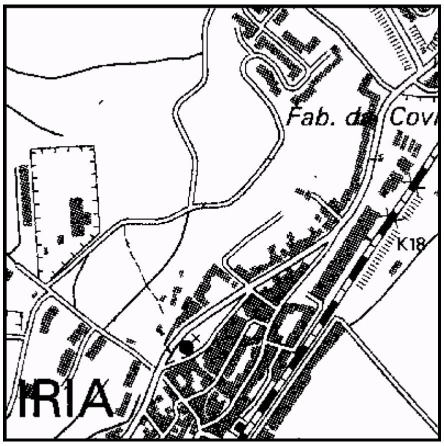

Fig. 16. Black plate completed by information initially left with Blue object.

Stains that seem to be no valuable information appear very rarely and normally are scribing errors, effects of late degradation or abnormal ink overload or superposition.

In addition, we would like to stress the aim of the use of two images, with different Cut values in one and same band, acting one as mask and the other as marker, in a process of Geodesic Reconstruction. This process leads necessarily to the reduction of gaps among objects, to the increase in number of connected objects and to the elimination of the noise introduced in the colour object to be isolated. Moreover, this noise-information can be reintroduced back to its original colour object, concluding a most productive information-recycling loop.

Finally, after isolation of each colour objects, the need may be felt to recover some lost links in an object with a certain degree of incompleteness. It is then advisable to proceed to a dilation of the final object and use it as a mask to extract the whole information. This operation, although obvious, was not referred in the text but was included in the flowchart of Fig. 17, explaining the proposed methodology.

4. CONCLUSIONS

The principles on which this research was built and expounded were already stated in the Abstract. Among them, the **Colour Spatial Correlation** best characterizes our methodology. This principle could be reformulated as follows:

> **Colour Spatial Correlation Principle**: For any pixel, its spatial attribute determines, by itself, a limited and differentiated tonal variability, entirely contained in the sampling space of the image.

We proved also that this principle applies not only to colour objects but also to the noise object, particularly in the field of digitized map objects.

The flowchart suggesting this Methodology is shown in Fig. 17.

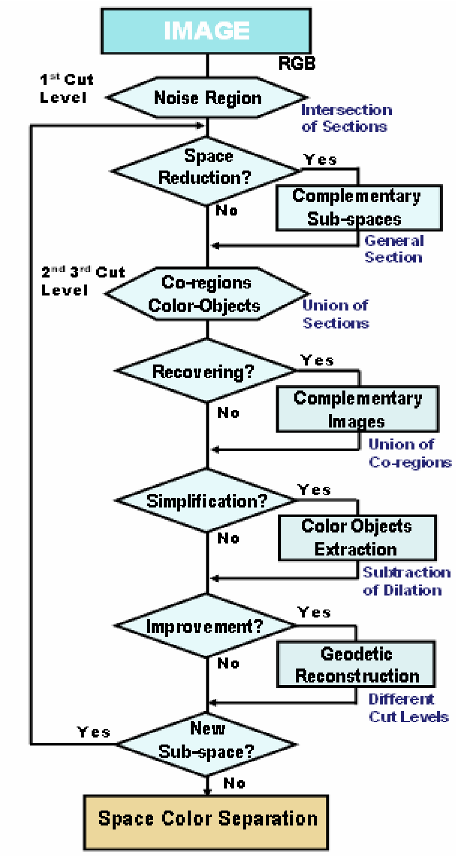

Fig. 17. The **Colour Spatial Correlation** Methodology.

Bibliography

Ablameyko S., 1993, Computer-Aided Cartographical System for Map Digitizing, *IEEE Bol.,* nº 0-8186-4960.

Ablameyko S., Bereishik V., Homenko M., Lagunovsky D., Paramonova N., Patsko O, 2001, A Complete System for Interpretation of Colour Maps. In *Computing and Control Engineering Journal* 12(4): 188-196 AUG 2001.

Ansoult M., Soille P., Loudts J., 1990, Mathematical Morphology: a tool for automated GIS data acquisition from scanned thematic maps, *Photogrammetric Engineering and Remote Sensing*, 56(9):1263--1271.

Arrighi P., Soille P., 1999. From Scanned Topographic Maps to Digital Elevation Models. *Proceeding of Geovision'99, International Symposium on Image Applications in Geology*, University of Liège, Bélgica, p.1.

Cordeiro A.G., Muge F., 1997. *Object Delineation in Remote Sensing*. RECPAD 1997, Coimbra.

Plataniotis K.N., Venetsanopolous, 2000, *Colour Image Processing and Applications*, Springer-Verlag, Berlin.

Soille P., 2003. *Morphological Image Analysis*, 2nd Edition, Springer-Verlag, Berlin.