Vector Objects Preselection and Homogenization

Victor Yu. Chekh, Andrei S. Bobrov

Design Office "Storm"
at National Technical University of Ukraine
"Kiev Polytechnic Institute"

16 Polytechnichna St. 252056 Kiev, Ukraine
Phone: [044] 274 0295, - 442 6077
Fax: [044] 274 1015, - 443 7925
e-mail: chekh@ntu-kpi.kiev.ua, chekh@kbi.kiev.ua
http://www.parallelworlds.com.ua

The novel technology based on the preselection of vector objects is proposed for the acceleration of the digitalization of maps and drawings. Initially the given set of layers must be created in the vector file to be processed. Then the objects of the initial file must be preselected and separated into the corresponding layers using the set of the simple geometrical parameters, which are calculated. Then all objects of each layer can be processed together including attaching of the common fields of attributive databases and their transforming into more natural object types (homogenization). The proposed technology is fast and effective due to the special vector coding and fast calculation of whole parameters.

Introduction

The problem of the automatic recognition of the particular objects on the large images is actual for the wide spectrum of the brunches of technology, particularly in geographic information systems (GIS), remote sensing, nondestructive testing and so forth. Particularly, the well-known terms "digitizing of the raster maps" or "digitizing of drawings" includes not only the exact copying of the center lines or borders but also developing of the database where the tables must be attached with the particular data (geometrical of attributive information) related to each object to be recognized. This take more then half time of all digitizing process.

To accelerate the process of digitalization and also to accelerate the image recognition we propose the method of automatic separation of the vector objects onto the layers accordingly to their particular integral geometric parameters.

Method Description

Let the term "digitalization" hereafter includes not only the process of transforming of the raster picture to the corresponding set of vector objects; moreover, we assume that we already have the initial geometry of the objects to be processed. The problem of the next step of digitalization is to attach the database table for each object or for their particular series.

The multi-layer structure is well-known from the standard GIS and CAD software packages. We propose the tree layer structure to accelerate the objects selection where the database tables can associated only with leaf (output) nodes of the layer tree. The set of the integral geometrical parameters is measured for each object; the objects which satisfies the particular conditions (the parameters values must be in the given frames) will be selected to the corresponding layers, and so on, till the leaf layers. Then the corresponding fields of database tables will be attached automatically.

The set of the parameters depends on the type of the objects to be recognized. The main parameter is the dimension of the vector objects; we recognize 2-D, 1-D (linear) and 0-D objects. The last ones have no geometrical properties just the plane coordinates. The GIS and CAD systems are supported about 20 or so numbers of the 2-D and 1-D object types. Each of them has specific geometrical properties and the specific rules for their calculation.

The described technology is used in the software for automatic digitizing via initial raster images. It makes us possible to take into account only the initial types of geometrical objects, which appears after the base operations - vectoring and skeletonization. We separate initial 2-D regions, polygons, i.e. 2-D objects with polygonal boundaries and 1-D objects - polylines. All other types of objects including 0-D ones will appear as the result of secondary operations with these initial objects. Moreover, for the acceleration of the whole processing the initial objects, which will be processed similarly, must be separated into the same layers.

We use the following parameters for preselection of the 2-D regions:

- general square
- external square
- number of internal boundaries,
- the square and size of circumscribing rectangle
- external perimeter
- and some another ones

and the following ones for the 1-D objects:

- general length
- minimum, maximum and average width
- twisting
- topological charge of the end node
- general orientation
- initial and end orientation
- maximum and average square deflection
- maximum linear deflection
- and some another ones.

We chose namely the parameters above because they are calculated very fast using our internal vector format (linearly to the number of "atom" components such as segments and nodes which can be complex vector objects disassembled onto). The set of the preselection parameters could be supplemented if necessary. Nevertheless the described set is just enough for particular applications even enough complex such as GIS mapping. Whole digitalization including attaching of objects databases turned out 20 times faster using preselection with kind of objects, their style and the geometrical parameters above. Moreover, the number of misrecognized objects comparing with manual mode became less due to minimum operator's errors.

For example, we recognized the lines of different average width (say, roads), hatching of the flat objects as the lines of given width, twisting and orientation, symbols of the given font size as the flat objects with the square and perimeter parameters, and so forth.

It doesn't mean that the described set of parameters is enough for full recognition of all objects could be found on the digital map or drawing. The set of the parameters above doesn't guarantee the unique coding for any possible kind of object if been used as the base set of pattern signs. The most complex objects such as symbols, characters, etc. must be exactly recognized using more complex and therefore more slow methods based on the special signs.

We also use the preselection parameters to estimate is the object looks like rectangle, like circle, like line or arc and so on. It must be noted that the technology of object preselection is successfully used for homogenization, i.e. turning the objects to the most natural view. For example, if the leaf layer must include the circles of given size then its objects which parameters are satisfy the given conditions which mean "the object looks like a circle" can be transformed to the circles from the initial polyline or polygone shape.

Talking about homogenization we must note that it needs not only for decreasing of the vector file size due to more compact format but mostly for supporting of the rest types of vector objects and acceleration of the secondary types of transforms.

Conclusion

The described technology is already software implemented and successfully used for automatic map digitizing. Due to the structure of the used internal vector format the time for preselection is neglectfully less then the time of the automatic transforming of the images from the raster to the vector form. Its features and advantages can be visually demonstrated.

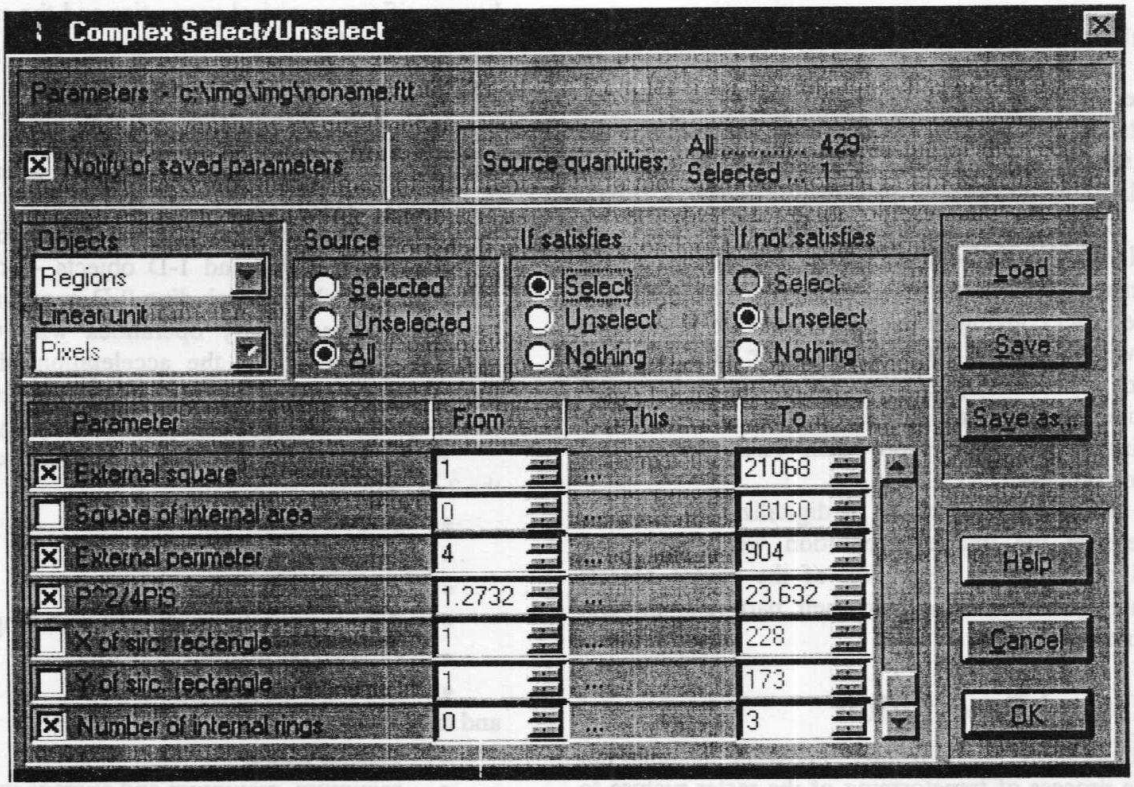

Fig. 1 Dialog panel displays sample parameters for selection of 2-D objects.