

An Introduction to Geographic Information Systems

Automatic vectorization

Stéphane Joost, Marc Soutter, Fernand Kouamé, Amadou Sall



Search MOOC



Video



Automatic vectorization



This lesson focuses on the digitalization automation of vectorial objects from map bottoms or from aerial images for example. We saw in the previous lesson that GIS softwares, in particular QGIS, offer quite performing digitization and editing tools but we realize that the use of these tools can represent a considerable amount of work and that there may be an interest in having methods and tools that enable to automate all or part of this tedious work. We will see some of these methods in this lesson and we will also see that these methods are not a panacea because they also require time to learn how to use it and especially in general to rework the result by hand so that it is often simpler to digitalize vectorial objects directly and simply by hand than going through an automated processing. Identifying objects in particular in an image is an operation which can be greatly facilitated by the application of various preliminary treatments such as filters or a supervised classification. This is the reason why we find various specialized softwares on the market which combine signal processing and object identification.

Notes

Summary



0m 23s

Automatic vectorization

Objectives of the lecture

- To discover the potential and limits of automatic digitization of vector objects

After this lecture you should be able

- To partly automate the vectorization of objects from a scanned document or an image, or on the basis of already existing vector objects

An introduction to Geographic Information Systems

In this lesson we will present the possibilities offered by the tools integrated in the QGIS software. However, we should know that some of these tools, Orfeo in particular, only work well in the case of an advanced mode installation. So this lesson aims to make you discover the potentialities, but also the limits of automating methods of vectorial object digitalization, so that at the end of the lesson you are able to use these methods to automate the data entry from map bottoms, from aerial images or by relying on existing vectorial objects.

Notes

Summary



1m 48s

General principles

Automatic detection of structures

- From an image

Identification of groups of similar pixels

➔ Polygonization

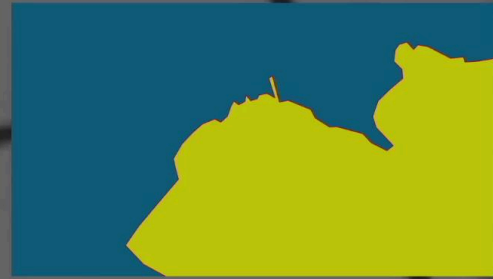
Identification of linear structures

➔ Edge detection

- From a vector layer

Using existing objects as «guides»

➔ Automatic tracing



An Introduction to Geographic Information Systems

We will discuss successively the following topics: first, the general principles of the vectorization automation then 3 specific methods, first the polygonization then edge or segment detection and at the end the automatic tracing. The automatic vectorization is an automatic detection and a structure digitalization which can be carried out from an image for example by identifying groups of similar pixels which is called polygonization. By identifying linear structures and in this case we speak of edge detection or segment detection or else from a vectorial layer using existing elements as a support for digitalization and we then speak of automatic tracing.

Notes

Summary

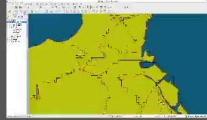


2m 32s

Polygonization

- Clearly delimited homogeneous areas

→ A scanned thematic map



An Introduction to Geographic Information Systems

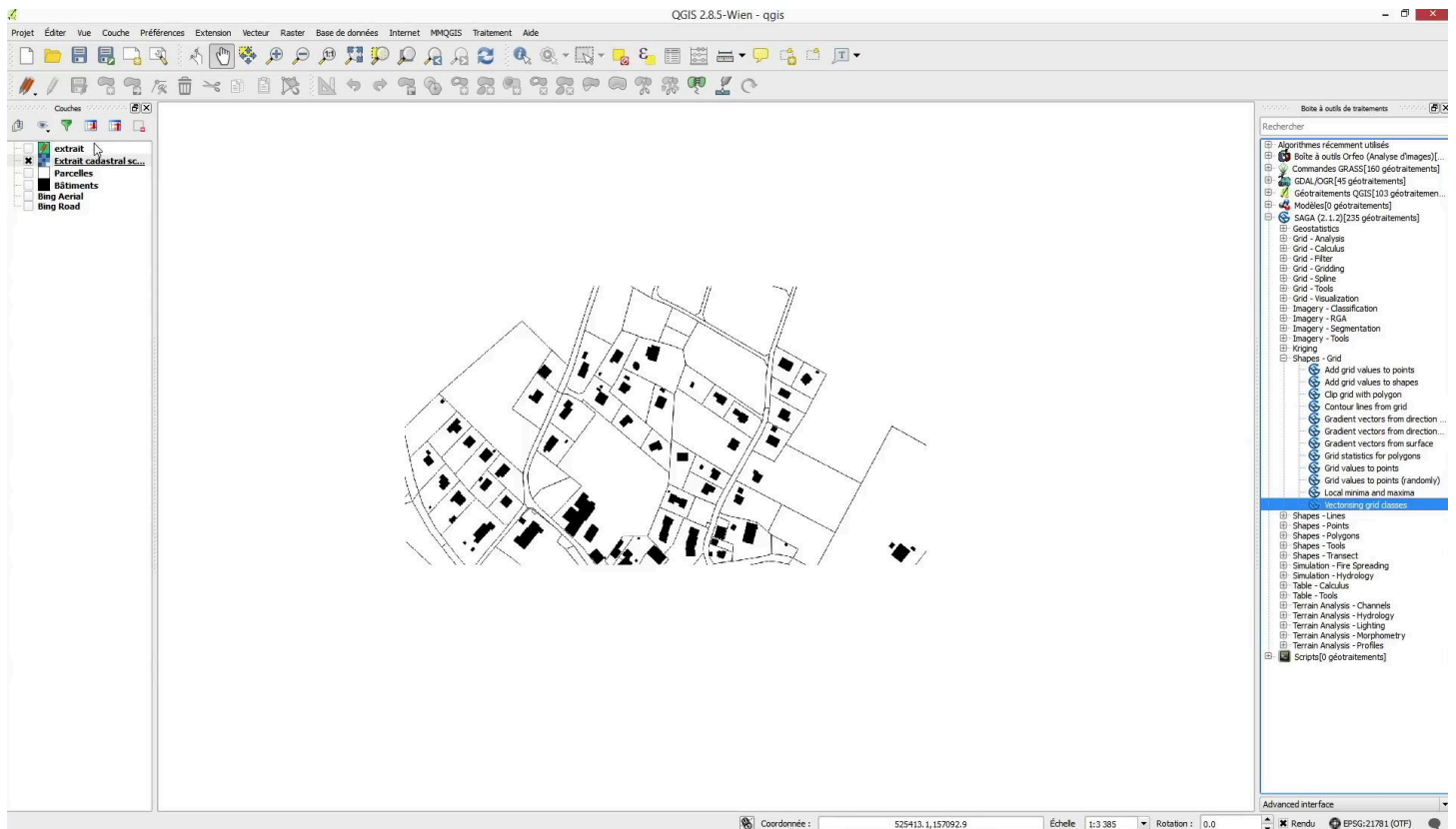
We will discuss successively about several cases of polygonization. First for clearly delimited homogeneous areas, taking the example of a scanned thematic map. We can see the island of Mahé in the Seychelles again for which we have a thematic map representing a variable per scanned district so it presents itself in the form of a regular grid with a rather weak resolution. The vectorization of that map is carried out with the help of this polygonization function and we choose the source so this thematic map in the form of a grid, we define the target, so the Shape file in this case, an ESRI Shapefile which will host the... the vectorial geometries made and we execute the procedure. We see that we obtain a result which reproduces the Seychelles district structure rather well with vectorial objects that we can individually select. If we make this vectorial layer that was created transparent now and that we mask the original raster layer, we can compare the original district limits in red, with the limits obtained by scanning in black and we see that the precision of the adjustment is nevertheless quite relative. The second example is about well-marked linear structures for example a scanned cadastral plan.

Notes

Summary



3m 41s



Here we find one of these plans which, as we can see, corresponds well to the existing cadastre and again the automated digitalization is carried out by this polygonization function in QGIS where we set the layer to polygonize and the destination of the result. Once the operation has been done, we see what we get and it actually is a polygon layer in green in which the main geometries have been recognized but for which the contour of these geometries is characterized by the contour of the starting pixels so a... a shape, staircase shapes and which we may wish to simplify in fact by... by removing the... the excess anchoring points, so here with a coefficient, a tolerance... of simplification of 1 which was chosen a bit by trial and error. We see that this geometry simplification allows to simplify the large shapes which... which are... have straight borders. On the other hand, there remains a large number of small parasitic objects which block the layer. There is another way of proceeding by using in the QGIS toolboxes the SAGA software and especially... in the grid functions the vectorization function.

Notes

Summary



5m 21s

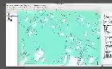
Polygonization

- Clearly delimited homogeneous areas

→ A scanned thematic map

- Well defined linear structures

→ A scanned cadastral plan



An Introduction to Geographic Information Systems

And this vectorization function cannot work with multiband layers as is the case here and it is necessary beforehand to transform this multiband layer into a monoband layer using the QGIS ad hoc function. So here we create this monoband layer of pict type with 2 colors. So, we have a new image raster layer of this cadastre extract but this time with a single channel. And in this case we can apply the SAGA vectorization process and we get a similar result with... the big difference is that the black blocks appear as holes within the polygons so it will always be possible to... to digitalize them afterwards by... by Overlay with... with a layer... another layer. And if we apply again the principle of simplified geometry and we see that we actually get a... a representation that is not so bad but which will require manual adjustments so that the different plots juxtapose correctly etc. The third example is that of complex structures found in an aerial image.

Notes

Summary



7m 11s

Polygonization

- Clearly delimited homogeneous areas

→ A scanned thematic map

- Well defined linear structures

→ A scanned cadastral plan



- Complex structures

→ An aerial image

So here we apply the method of polygonization again proposed by QGIS by choosing the source, the destination as an ESRI Shapefile, the application is this time a little bit slower and we see that we get a layer of violet objects but in fact... whose black border takes all the place and we realize that this polygonization only allow... to identify a small number of pixel clusters which correspond to the paths which have a texture, a very similar color and everywhere else the variability of the image is such that we get approximately one object per pixel.

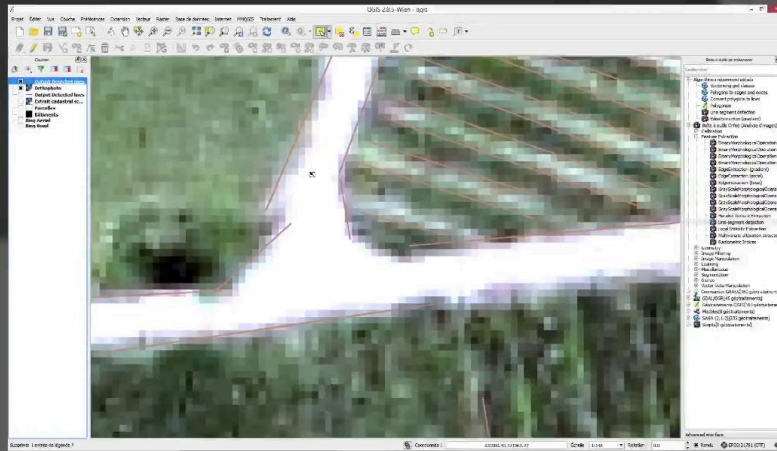
Notes

Summary



8m 49s

Edge detection



Introduction to Geographic Information Systems

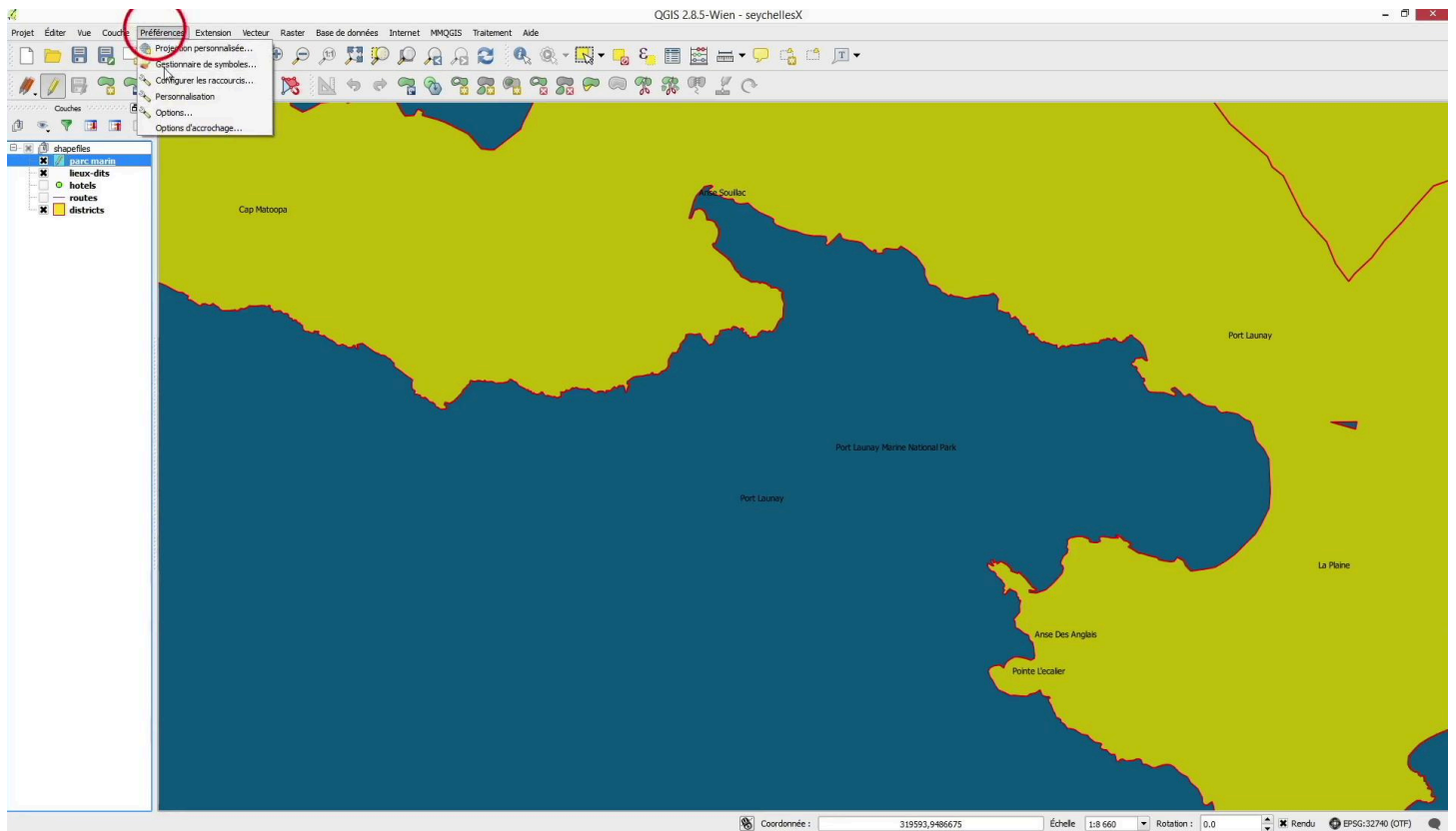
We have seen that the polygonization methods have their limitations in particular when they are applied to aerial photographs. This is the reason why we should seek to use other approaches particularly the detection of edges and segments that seeks to identify linear structures in an image. So we find here our scanned cadastral map for which we will use the Orfeo toolbox methods and in particular this edge detection method which runs and if we mask the original, we will discover the segments that have been identified and created as vectorial objects by this processing method. So we see that in this case, as in the case of the polygonisation, additional manual work will be required to complete automatic digitalization. In the case of the orthophoto which we had earlier, the process of identifying the edges is longer and the result obtained is of the same nature with all the linear structures identified in the image that appears.

Notes

Summary

10m 00s





But if we look a little closer and modify the line thickness of these segments a little bit, to see them clearly, we realize that in fact the linear elements identified correspond approximately to what is in the image but that some alterations will also be necessary. So far we have seen methods of digitalization automation based on image processing, another very useful form of automation is one that relies on existing vectorial objects in the case of automatic tracing. We see here the island of Mahé and more particularly the marine park area of port Launay that we would like to materialize by a... a graphic layer. So we have created this layer of polygons under the name of marine park and we will use this extension called Autotrace to delimit the marine park that follows the coast.

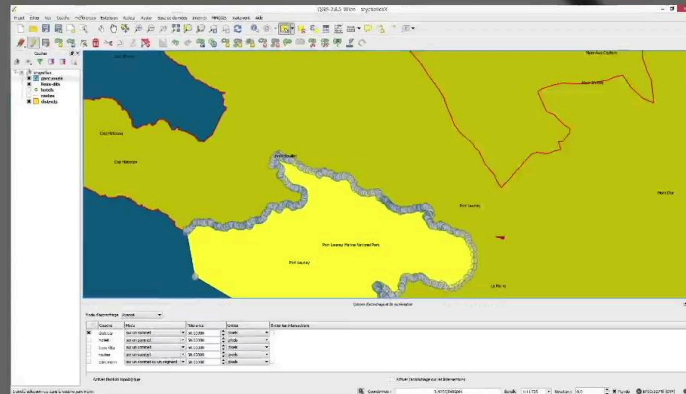
Notes

Summary



11m 21s

Automatic tracing



An Introduction to Geographic Information Systems

In order to do this we need to define in the snap options to validate the district layer as a hanging layer and then using the Autotrace tool select a first point which is going to be the... the beginning of the zone affected by the marine park. We will then search the last point by pressing the shift key to select all the points which are included between those two ends and we see that it allows us to draw an object that will characterize this marine park and which follows very precisely the cutting of the coast as described by the district layer. We see here that all the points indeed have been integrated to this new drawing.

Notes

Summary

12m 42s

