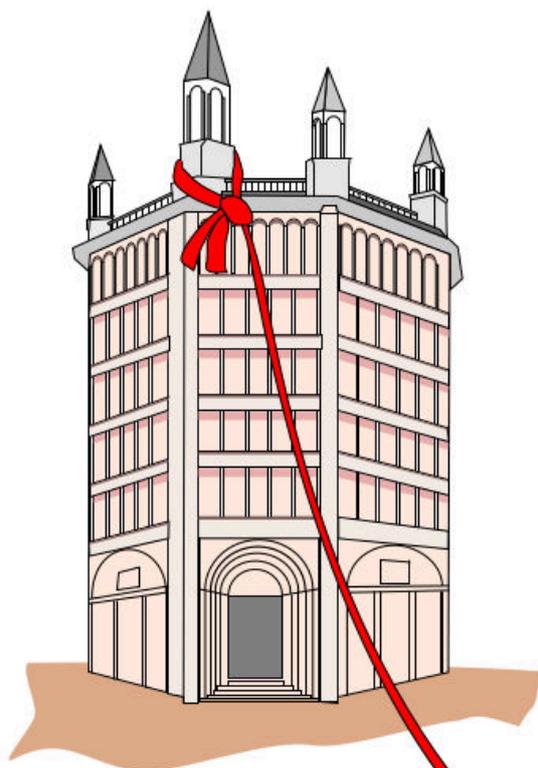


Geomatics Workbooks Vol. 13
Dicembre 2017



XVII GRASS
e GFOSS meeting
Parma 2016

FOSS4G-IT 2017
XVIII GRASS Meeting
X GFOSS Day
OSMit 2017
Genova 2017



Geomatics Workbooks, Volume 13

XVII GRASS e GFOSS Meeting

Parma 2016

FOSS4G-it

XVIII GRASS meeting - X GFOSS Day – OSMit2017

Genova 2017

Geomatics Workbooks, volume 13 - Dicembre 2017

ISSN 1591-092X

Editor: Valentina Marin, Paola Salmona
Geomorfolab – DAD – Università degli Studi di Genova

Editorial Board: Maria Brovelli e Monia Molinari
Laboratorio di Geomatica - Politecnico di Milano

Comitato Scientifico

Alfredo Altobelli
Gerardo Brancucci
Maria Brovelli
Stefano Campus
Corrado Cencetti
Marco Ciolli
Aldo Clerici
Paolo Dabove
Margherita Di Leo
Bianca Federici
Marco Firpo
Ivan Marchesini
Elena Mezzini
Marco Minghini
Monia Molinari
Donatella Muscas
Andrea Scianna
Domenico Sguerso
Clara Tattoni
Alfonso Vitti
Paolo Zatelli

Copertina: Paola Salmona

I Geomatics Workbooks sono disponibili online:

da questo numero all'indirizzo: <http://geomorfolab.arch.unige.it/geomaticsworkbooks>
i numeri pregressi all'indirizzo: <http://geomatica.como.polimi.it/workbooks/>

Contatti

geomorfolab@arch.unige.it
<http://geomorfolab.arch.unige.it/>

Questo volume è stato realizzato utilizzando il software Free Open Source LibreOffice

INDICE

XVII GRASS e GFOSS Meeting - Parma 2016

Webmaps dei caratteri morfologico-sedimentari e antropici del F. Paglia (Bacino del F. Tevere) (Cencetti, De Rosa, Fredduzzi)	<u>1</u>
Alcuni script di shell per il calcolo e la visualizzazione delle principali caratteristiche morfometriche di un corso d'acqua (Clerici, Perego, Chelli)	<u>16</u>
QGIS per la valutazione della qualità morfologica dei Corpi Idrici: l'esempio del T. Vobbia (Mandarino, Botti, Firpo, Martella)	<u>29</u>
Metodologia speditiva per l'individuazione delle aree terrazzate per la conservazione del paesaggio e la riduzione del rischio idrogeologico (Cibrario, Stralla, Solimano, Salmona, Marin, Brancucci)	<u>44</u>
Analisi della fruibilità delle aree costiere: il caso di studio del levante ligure (Marin, Salmona, Volpin, Mor, Bobbio)	<u>58</u>
Automatic segmentation of thin sections with GRASS GIS: first experiments (Angelucci, Chisté, Mologni, Cappellari, Vitti, Zatelli)	<u>77</u>
Gestione e pubblicazione di dati territoriali con GFOSS nella Pubblica Amministrazione: Il caso del Comune di Vicenza (Marzocchi, Fagandini, Beggiato, Berti, Ardini, Cosso, Missaglia) ..	<u>91</u>

FOSS4G-it - Genova 2017

Muoversi in Piemonte: an integrated platform for infomobility services (Arneodo, Castelli, Gagliardi, Sabatelli)	<u>99</u>
GisChatBot: uso delle chat per la costruzione e gestione di strati informativi GIS in maniera collaborativa (Demontis, Pau)	<u>112</u>
Il progetto Buiometria Partecipativa: presentazione generale e stato delle attività (Giacomelli, Nardi)	<u>122</u>
Humanitarian Mapping within a Student Association: PoliMappers (Kilsedar, Oxoli, Frassinelli, Montani, Minghini).....	<u>130</u>
Rendering Large Point Cloud in WebGIS: nuove tecnologie a supporto del rilievo 3D di Bergamo (Gagliano).....	<u>137</u>

XVII GRASS e GFOSS Meeting
Parma 2016

Automatic segmentation of thin sections with GRASS GIS: first experiments

Diego E. Angelucci[^], Paolo Chistè[^], Carlo Mologni[^], Etienne Cappellari*, Alfonso Vitti* and Paolo Zatelli*

[^]Dipartimento di Lettere e Filosofia, Università degli Studi di Trento

*Dipartimento di Ingegneria Civile, Ambientale e Meccanica, Università degli Studi di Trento

Abstract

Micromorphological analyses of soil thin sections from archaeological sites are fundamental tools in geoarchaeology, which includes the study of site formation processes. The thin sections are observed under the optical microscope in polarized light with different lighting conditions (mainly with plane polarized light and with crossed nicols) and digitized with high resolution. The joint analysis of the resulting images allows the detection of significant features for the archaeological interpretation of the site.

This work describes some experiments for the automatic or semi-automatic analysis of this kind of images using modules for image segmentation and filtering available in GRASS GIS.

Single images are segmented using the different approaches, variational and region growing, implemented in GRASS *r.smooth.seg* and *i.segmentl* modules. The resulting images are filtered using threshold values evaluated for each image. Finally, the filtered images with different types of polarization are combined. The resulting map is checked against a reference map created manually by digitizing the contour of relevant areas.

This work presents the first experiments for the optimal choice of the segmentation parameters for the variational approach and the threshold values using automatic or semi-automatic approaches.

Introduction

Archaeological studies are interested not only in objects and concrete remains of human activity, but also in their context, in particular in the soil and sediments containing them.

Among the aims of geoarchaeology is the study of archaeological sites formation processes, analyzing the motivations of the choice of a location for human settlement, the processes governing site formation and the modifications occurring during its occupation and after its abandonment (Renfrew, 1976). Geoarchaeology four main objectives are (Butzer, 1982): the reconstruction of the stratigraphy and chronology of a site, the description of the physical paleo-environment and its modifications, the study of the interactions between humans and the environment and their impact, and the

examination of the formation of an archaeological site due to the interference between anthropic and natural processes. At the micro-scale this is done using laboratory analyses determining physical, chemical, mineralogical, micro-morphological and archeometric features. Some of the features which this technique has found in the sample examined in this paper are (Mologni, 2015):

- tiny shell fragments, with a description of their distribution and chemical and physical transformation after their deposition;
- small bones parts, often mixed with shells fragments;
- traces of combustion on the external parts of some shells.

All the findings above are indicators of human presence on the excavation site.

One fundamental tool used in geoarchaeology is the micro morphological analysis of soil thin sections. Micromorphology (or micropedology) examines undisturbed soil samples using microscopy to identify their components and provide indications on their relationships in space and time (Stoops 2003). Its main advantage is the possibility to distinguish between natural and anthropogenic formation processes while taking into account their mutual relations.

To this end, thin sections of archaeological deposits are obtained from undisturbed samples, after desiccation and impregnation with optically-neutral resins (Stoops 2003).

Thin sections are then observed under the petrographic (aka polarizing) microscope to analyze their fabric (in terms of shape, dimension and frequency of its components), groundmass (texture, homogeneity, color, purity, relationship between coarse and fine material, etc.), coarse mineral components (dimension, selection, roundness, alteration, orientation, distribution, etc.) and so forth (Bullock et al., 1985) (Nesse D. W., 2013).

In the current procedure, homogeneous areas are manually selected by drawing their boundaries while looking simultaneously at different images of the same thin section. Each area is then classified by the operator according to its features as one of the different materials constituting the sample.

To our knowledge, no automatic or semi-automatic approach is available or has even been tested.

The aim of the experiments described in this paper is to implement a procedure for making thin section analysis as automatic as possible.

One possible approach to this problem is the object based image analysis (OBIA) (Blaschke, 2010) used in remote sensing, where image segments are classified according to their geometric and radiometric features. However, there are some differences with respect to the standard OBIA, in particular the availability of images of the same object with different polarizations. For this reason a simpler and bespoke approach has been tested.

Materials and methods

All the tests have been carried out on thin sections obtained from the Middle Palaeolithic site of Figueira Brava (Portugal) (Mologni, 2015). The site was excavated several years ago and gave Neanderthal remains (see Antunes & Cunha 1992). A project for reevaluating its stratigraphy, chronology and archaeological relevance is currently in progress under the direction of João

Zilhão (University of Barcelona, Spain). The specific purpose here is to create a morphological description of shell fragments contained in these samples, assessing their degree of combustion as an indication of anthropic activity. Indications about sample preparation and details about the scanning technique can be found in (Mologni, 2015), here only information relevant to the procedure for the automatic features extraction on the images is provided. Thin sections of sediment with a standard thickness of 30 μ m have been prepared and digitally acquired using a high resolution flat-bed scanner. Different images of the entire thin sections were created by inserting different polarization filters between the thin section and the scanner.

In particular, two polarization modes are used:

1. plane polarized light (PPL from now onwards), using one polarizing filter or two parallel filters; it is possible to observe the color of the minerals, their pleochroism, habit, cleavage, and index of refraction, the shape of the particles, the general structure of the material and its porosity;
2. crossed polarized light (XPL from now onwards), applying two polarizing filters crossed at $\pi/2$: it is possible to observe the mineral isotropy or anisotropy and its possible birefringence.

One thin section (namely sample FB1002), measuring 5.5 x 9.5 cm, has been used as a specimen for the tests, by combining PPL (Fig. 1) and XPL (Fig. 2) scans. Both images are 8503 x 5196 pixels in the RGB bands and must be co-registered to be used together.



Figure 1 plane Polarized Light – PPL scan.



Figure 2: crossed Polarized Light - XPL scan.

For this first experimental study, the evaluation of porosity was chosen, being voids quite easy to identify in thin section.

The selection of the interesting features is currently done manually by an operator observing both images at the same time. The results of this procedure are shown for a particular of the image in Fig. 3 (PPLand) and Fig. 4 (XPL).



Figure 3: manually selected areas on the plane polarized light (PPL) scan (particular).

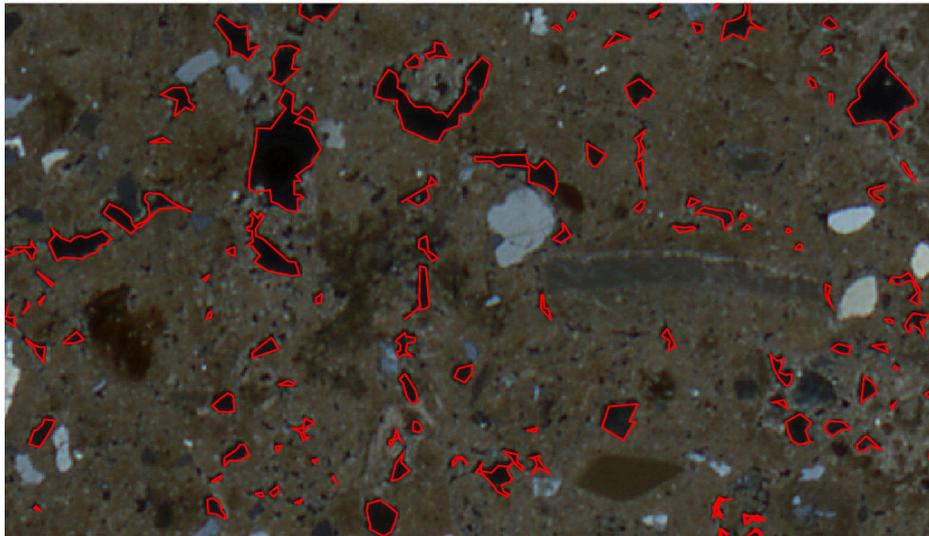


Figure 4: manually selected areas on the crossed polarized light (XPL) scan (particular)

The automatic procedure set up for these experiments consists in three steps:

- 1 segmentation - images are partitioned into homogeneous areas;
- 2 filtering - bright areas are selected on the PPL image and dark areas are selected on the XPL image;
- 3 combination of filtered PPL and XPL - areas selected on both PPL and XPL are marked as valid features.

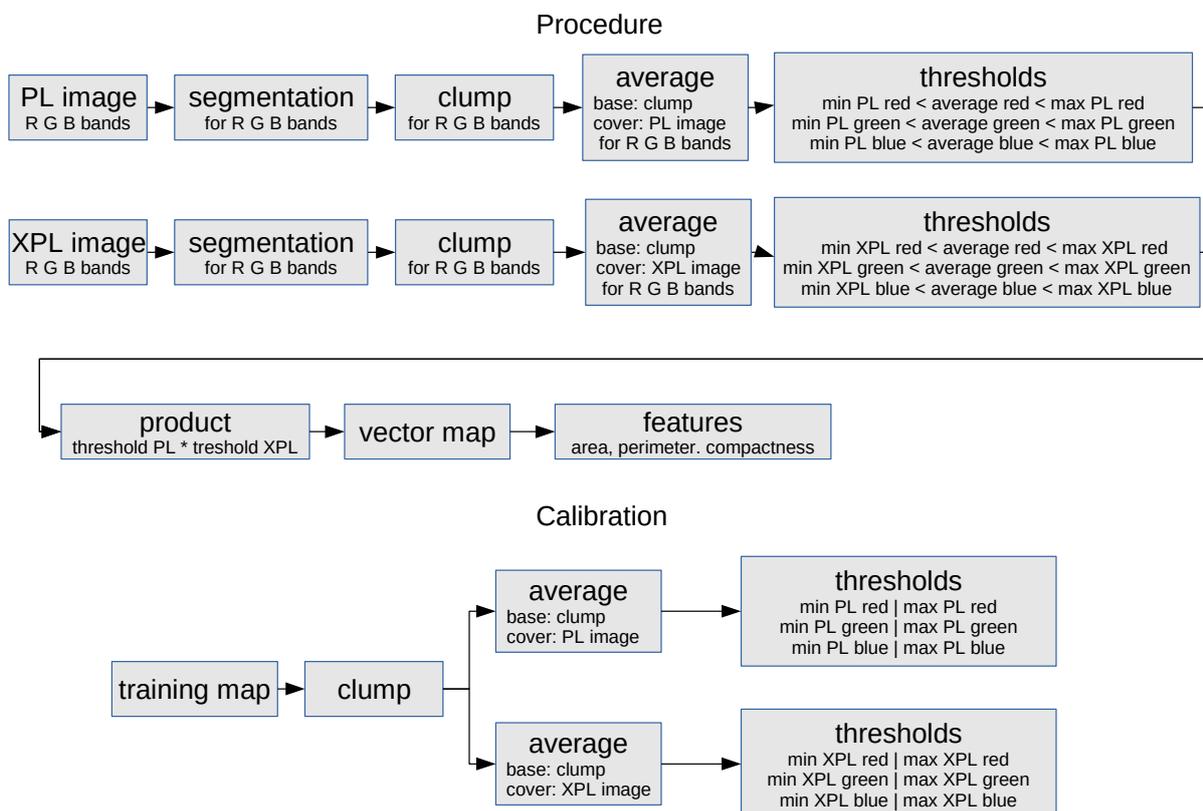


Figure 5: automatic selection and calibration procedures.

Two different approaches have been applied for the first step: a variational model (Vitti, 2012) and a classic region growing and merging algorithm (Pal and Pal, 1993). Both approaches need a calibration phase to determine the optimal combination of parameters for the segmentation of the two images.

The filtering consists in the simple application of thresholds to the 6 raster maps representing the RGB bands of the image, but the 6 threshold values must be carefully chosen to avoid over or under filtering.

The final combination step is carried out by applying simple boolean operators using map algebra. Finally, the result is matched against the manual classification to assess its effectiveness.

The overall procedure, which has been implemented in GRASS GIS (GRASS Development Team, 2012) and completely scripted, is outlined in Figure 5.

Segmentation

The segmentation step has been carried out using GRASS modules *r.smooth.seg* for the variational approach and *i.segment* for the region growing one.

The *r.smooth.seg* GRASS module implements the Mumford and Shah model in dimension 2:

$$MS(u, K) = \int_{\Omega/K} |u - g|^2 dx + \lambda \int_{\Omega/K} |\nabla u|^2 dx + \alpha H^1(K)$$

where $\Omega \subset \mathbb{R}^2$ is a bounded open set, g is the input data, u is the smooth function approximating g , λ and α are positive parameters which are used to control the process, H^1 is the Hausdorff measure (the length, in this case) and K the set of discontinuities. The *r.smooth.seg* GRASS module therefore requires two input parameters that control the segmentation process: α measures the relevance of the request of short boundaries for the segmented regions and λ the importance of uniformity inside the regions, i.e. a smooth solution (Vitti, 2012).

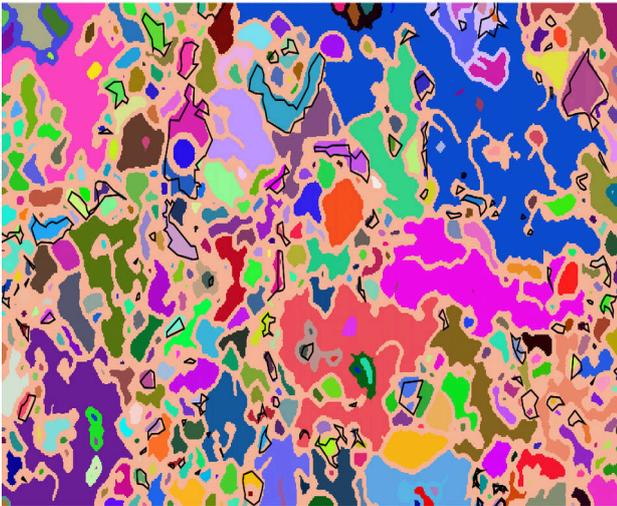
No direct method for the optimal determination of the segmentation parameter is available for 2D signals, therefore tests have been carried out on the two images to identify the best set of α and λ parameters. To reduce the complexity of the procedure, in these tests the same values for the α and λ parameters have been used for all the 6 raster maps representing the RGB bands of the two images. For each set of parameters the whole procedure has been carried out (Table 1) and the set leading to the correct identification of the larger number of pixels has been identified.

α	λ	Valid cells
100	10	43554
1000	10	25
10000	10	0
100	100	17349
1000	100	262238
10000	100	7
100	1000	10
1000	1000	38914
10000	1000	312471

Table 1: segmentation parameters and corresponding numbers of valid cells in the tests.

The best result has been obtained with $\alpha=10\ 000$ and $\lambda=1\ 000$ (Fig. 6a, b and c for PPL image and Fig. 7a, b and c for XPL image).

The *i.segment* module requires only one parameter, controlling the minimum similarity below which segments are merged. The value of this threshold ranges from 0 (each single pixel represents a region) to 1 (the whole image is a single region) and the procedure usually yields better results when used recursively, performing a hierarchical segmentation using the output of the previous step as seed for the next. The algorithm is started with a value of the threshold of 0.01 and continued with increasing threshold values of 0.05, 0.1, 0.2, 0.4 up to 0.5 (Fig. 6d - PPL- and 7d - XPL).



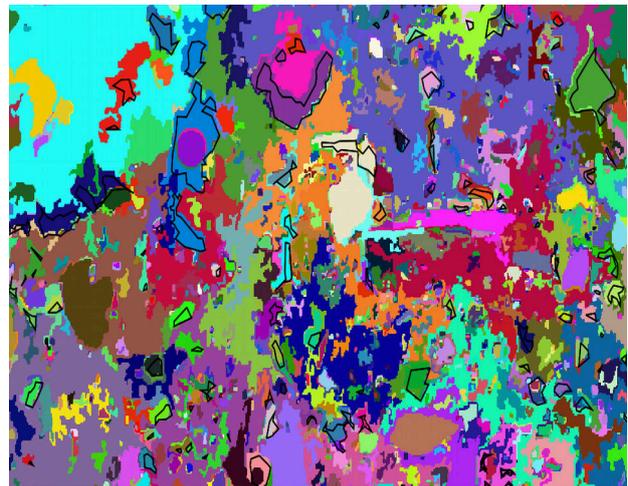
a - clump of the segmented red band



b- clump of the segmented green band



c - clump of the segmented blue band

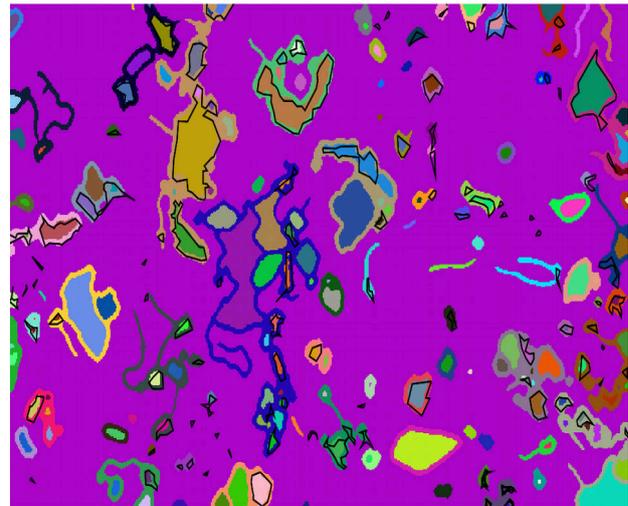


d - clump from i.segment with threshold=0.5

Figure 6: segmented PPL image (particular). Best results from `r.smooth.seg` for red (a), green (b) and blue (c) bands and from `i.segment` with `threshold=0.5`. Black contours identify areas manually selected in the control map.



a - clump of the segmented red band



b- clump of the segmented green band



c - clump of the segmented blue band



d - clump from i.segment with threshold=0.5

Figure 7: segmented XPL image (particular). Best results from `r.smooth.seg` for red (a), green (b) and blue (c) bands and from `i.segment` with `threshold=0.5`. Black contours identify areas manually selected in the control map.

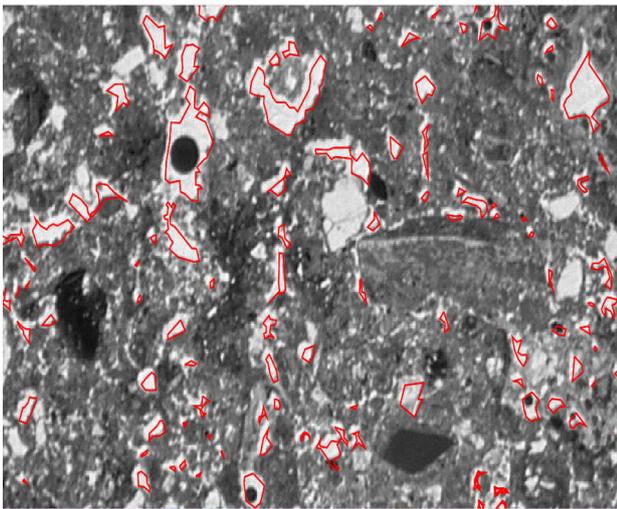
Filtering

Regions identified by the segmentation step are filtered according to their brightness: in the manual procedure the operator selects areas that are bright on the PPL image and dark on the XPL one.

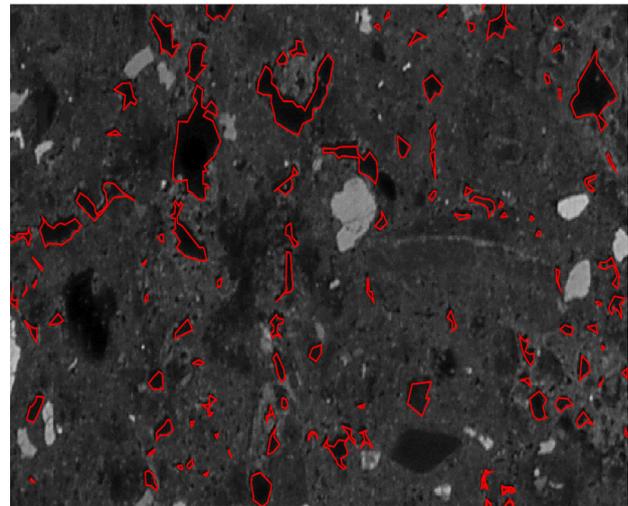
In the automatic procedure this is done by selecting the segmented regions corresponding to average high values on the three bands of the PPL and low values on the bands of the XPL image. It is therefore necessary to set six threshold values.

The first attempt to determine the threshold values is done automatically:

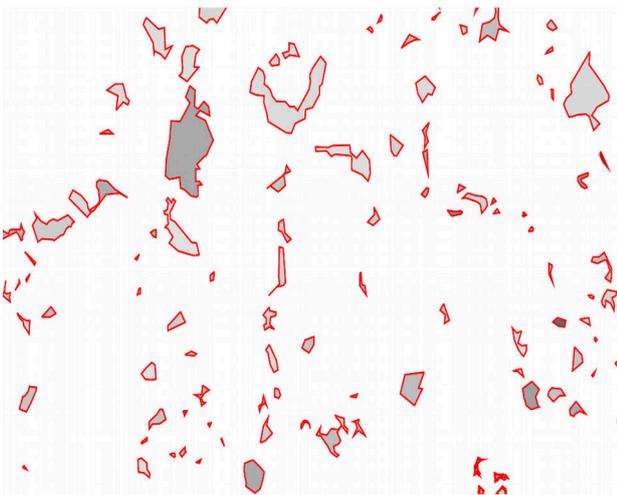
1. 30% of the areas on the control map manually digitized are randomly selected;
2. for each area the mean values on the three RGB bands for the PPL and XPL images are calculated (Fig. 8);
3. the minimum mean values for each band on the PPL and the maximum mean values for each band on the XPL are selected as threshold.



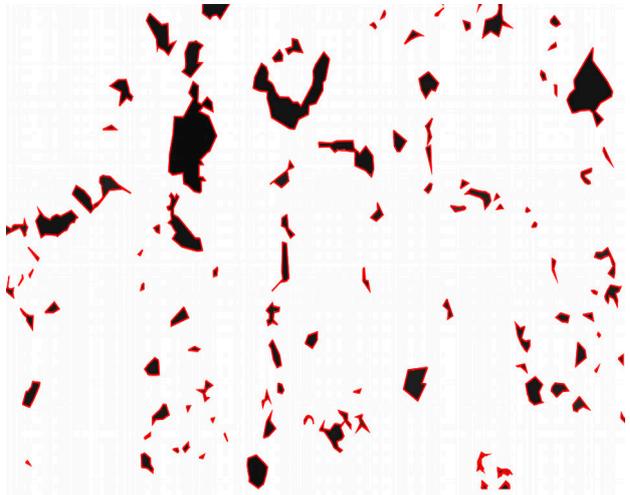
a – PPL image: blue band and control areas (particular).



b- XPL image: blue band and control areas (particular).



c - PPL image: average values for the blue band inside the control areas (particular).

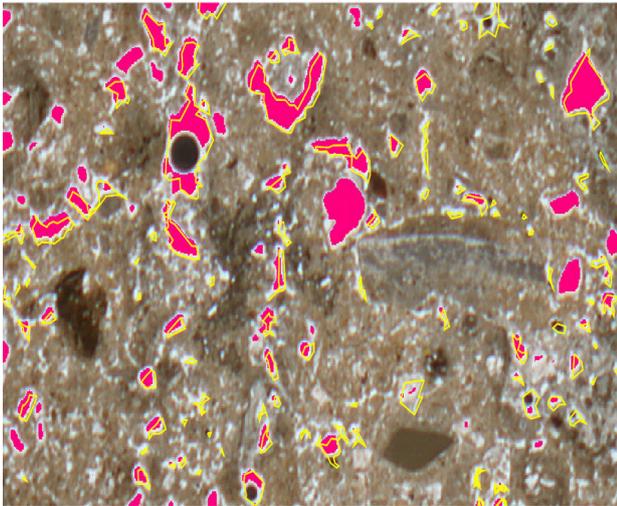


d - XPL image: average values for the blue band inside the control areas (particular).

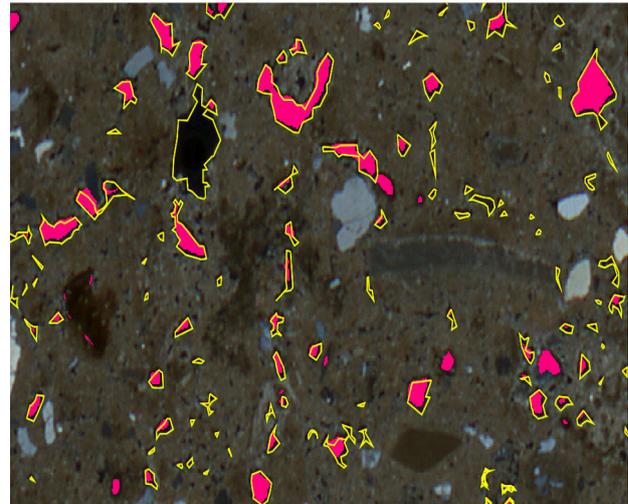
Figure 8: PPL and XPL images (particular): blue bands and control areas (a and b) and mean values for the blue band inside the control areas (c and d).

Threshold values evaluated in this way are typically too loose, due to the inclusion of materials (usually droplets of resin or air bubbles) with different colors in the control areas selected manually. Therefore these values are manually corrected by looking for these inclusions and their brightness values on the images.

Thresholds are then independently applied to the PPL and XPL images (Fig. 9).



a – PPL image: selected areas (in magenta) after filtering (particular).



b- XPL image: selected areas (in magenta) after filtering (particular).

Figure 9: PPL and XPL images (particular): selected areas after the application of the thresholds to the images bands. Yellow contours identify control areas selected manually.

Finally, areas selected on both PPL and XPL are identified by multiplying the two corresponding maps (Fig. 10).

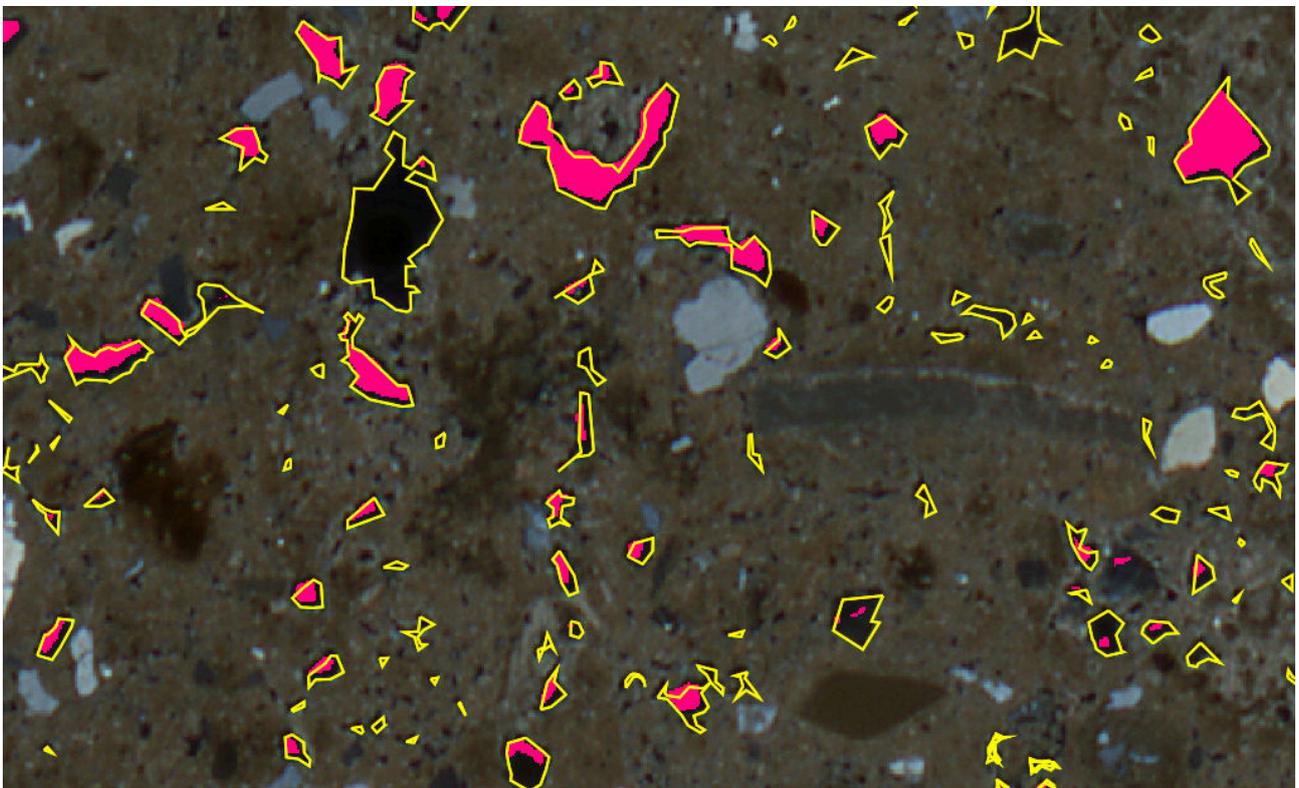


Figure 10: areas selected automatically (magenta) and control areas (yellow contours). XPL image is in the background (particular).

Results

An assessment of the effectiveness of the procedure can be made by evaluating the overlap of control areas, resulting from manual digitalization, and valid areas identified by the automatic procedure.

Tables 2 and 3 report the number of overlapping cells and their percentage.

MAPS: (untitled) (test_a100000b111000@scansioni in scansioni)			
Rasterized vector map from values (mappa_controllo_georect@scansioni			

#	description	Category Information	cell count % cover

1	Control areas		312471 0.71

1	Valid areas.		251268 80.41
*	*no data.		61203 19.59

*	*no data		43869117 99.29

1	Valid areas.		1035022 2.36
*	*no data.		42834095 97.64

TOTAL			44181588 100.00

Table 2: evaluation of the overlap of control areas and areas automatically selected (valid areas).

MAPS: Rasterized vector map from values (mappa_controllo_georect@scansioni in			
(untitled) (test_a100000b111000@scansioni in scansioni)			

#	description	Category Information	cell count % cover

1	Valid areas		1286290 2.91

1	Control areas.		251268 19.53
*	*no data.		1035022 80.47

*	*no data		42895298 97.09

1	Control areas.		61203 0.14
*	*no data.		42834095 99.86

TOTAL			44181588 100.00

Table 3: evaluation of the overlap of areas automatically selected (valid areas) and control areas.

The first report in Table 2 shows that areas in the control map cover 0.71% of the image and even in these first attempts 80.41% of them have been

identified by the automatic procedure. However, the rest 19.59% of the areas have not been detected.

The second report in Table 3 highlights the fact that with the current best parameter set the procedure grossly overestimates the valid areas, covering 2.91% of the image instead of 0.71% in the control map (Table 2), with only 19.53% of them corresponding to areas in the control map.

Conclusions

The automatic procedure properly identifies a good part of correct areas, but it also detects a large number of incorrect areas.

Even on these preliminary tests, some indications have emerged:

1. the variational approach, using GRASS' *r.smooth.seg*, performs better than the *i.segment* region growing algorithm, both in terms of overall results and possibility to fine tune the processes by controlling smoothness and contours length instead of a global similarity measure;
2. the choice of the correct values for the α and λ parameters is fundamental, for the image used in this study the best combination seems to be the one that leads to the maximum regularization, with $\alpha=10\ 000$ and $\lambda=1\ 000$;
3. the image obtained with crossed polarized light (XPL) is usually easier to segment;
4. the green band is the most problematic on the plane polarized light (PPL) and at same time is the one yielding better results on the image with crossed polarized light (XPL);
5. the choice of the 6 thresholds is fundamental and the current procedure for their evaluation is not satisfying;
6. a method to remove inclusions must be devised.

Future development concern the enhancement of already working features of the procedure, such as a better co-registration of the two images, a larger parameter space for the test of the α and λ parameters and the use of different types of polarization.

Possible solutions to the problem of the determination of the thresholds, which will be tested in the future, are:

1. the use of other functions (median, mode, etc.) instead of the average in the control areas; this has partly already been tested;
2. the use of different thresholds for different zones of the images;
3. the removal of inclusions.

Other possible approaches that will be tested in the future are:

1. maximum likelihood classification;
2. the application of machine learning algorithms.

The modules that implement these procedures are already available in GRASS and the segmentation steps already carried out can be used as pre-processing to enhance their results.

Bibliography

- Antunes M.T., Cunha, A.S. (1992). Neanderthalian remains from Figueira Brava cave, Portugal. *Geobios*, 25(5), pp. 681-692.
- Blaschke T (2010) Object based image analysis for remote sensing. *ISPRS Journal of Photogrammetry and Remote Sensing* 65(1):2-16, DOI 10.1016/j.isprsjprs.2009.06.004
- Bullock P.N., Fedoroff N., Jongerius G.J., Stoops G., Tursina, T., Babel, U. (eds.) (1985). *Handbook for Soil Thin Section Description*. Waine Research Publishers, Wolverhampton.
- Butzer K.W. (1982). *Archaeology as Human Ecology. Method and Theory for a Contextual Approach*. Cambridge University Press.
- GRASS Development Team. *Geographic Resources Analysis Support System (GRASS GIS) Software*. Open Source Geospatial Foundation, USA, 2012.
- Mogni C. (2015). Studio micromorfologico preliminare del sito musteriano di Figueira Brava (Portogallo) e proposta di un approccio per l'analisi digitale di sezioni sottili di grande formato. Undergraduate thesis - Corso di Laurea Triennale in Scienze dei Beni Culturali, Indirizzo Archeologico, University of Trento. Anno Accademico 2014/2015
- Nesse D. W., (2013). *Introduction to Optical Mineralogy*, fourth edition. Oxford University Press
- Pal, N.R., Pal, S.K. (1993). A review on image segmentation techniques, *Pattern Recognition*, 26 (9), pp. 1277-1294.
- Renfrew A.C. (1976). *Archaeology and the Earth Sciences*. In: Davidson D.A., Shackley M.L. (eds.) *Geoarchaeology*. Duckworths, London.
- Stoops G. (2003). *Guidelines for analysis and description of soil and regolith thin sections*. Soil science society of America, Inc. Madison, Wisconsin, USA.
- Vitti, A. (2012). The Mumford-Shah variational model for image segmentation: An overview of the theory, implementation and use. *ISPRS Journal of Photogrammetry and Remote Sensing*, 69, 50-64.