A Web-based Collaborative Tool for Land Use Computation on Satellite Images

Maria Grazia Albanesi, Riccardo Amadeo

Dept. of Electrical, Computer and Biomedical Engineering
University of Pavia
Pavia, Italy
e-mail: mariagrazia.albanesi@unipv.it, riccardo.amadeo@gmail.com

Abstract— **In this paper, we present the prototype of a new tool for computing land use on satellite images. Its main feature is the possibility of being used in a collaborative Web-based environment, where multiple expert users cooperate in validating the territory classification of different areas, according to the presence or absence of anthropic activity. For land use estimation, we use a well-known indicator in literature, namely the Anthropentropy Factor (AF). The novelty of the approach is the use of open software libraries, based on Keyhole Markup Language (KML), and Google Application Programming Interfaces (API) for the AF computation; the conceptual approach has been compared to other previously implemented solutions, namely the use of proprietary, ad hoc software tool and the use of Geographic Information System software on European Corine Land Cover data sets. A preliminary result of the work in progress is presented and commented.**

Keywords- Land use; satellite images; anthropentropy factor; Web-based application.

I. INTRODUCTION

The present paper describes a work in progress of a research project related to a great critical issue in environment preservation: land use estimation. Preserving lands, i.e., the wild, natural territory from the negative consequences of an inappropriate, out of control urban expansion, is one of the most important target for biodiversity preservation [1]. The percentage of land that annually is subtracted to natural, wild eco-systems for human activities expansion, has been estimated in the 36 European countries about 112000 ha/year in the period 2000-2006, reaching the percentage of 9% in the most urbanized countries [2].

Land use is not only due to urbanization, i.e., the expansion of rural and urban settlements, but also to the creation of industrial, intensive farming and touristic sites, roads and communication lines. We refer to all these activities as *anthropic activities*, and the consequent land occupation as *anthropic places*. Therefore, in the analysis of the territory we considered an anthropic area if it is occupied by signs of human presence, such as buildings, paved roads and railways, places of intensive agriculture, and industrial settlements, both for production and services.

Despite the great importance of land use estimation, its computation is far to be simple and fast: obtaining data about the territory and validating the area distribution of anthropic places is a long-time consuming process, even if the modern image satellite acquisition and processing have partially simplified the task. In fact, satellite images have to be processed in order to classify each pixel to anthropic places or not, and this is far to be completely automatized by low-level image processing techniques. For this reason, the project here described aims at giving a possible answer to the problem of a fast and easy-to-update process for validating the territory occupation and computing the land use indicator.

The paper is organized as follows: in Section II, a discussion about the choice of land use indicator is given. In Section III, a synthetic definition of AF helps the reader to understand the computational steps realized in the Web-based prototype. In Section IV, a discussion about related works on AF computation is given, in order to distinguish the novelties of the Web-based collaborative solution. In Section V, the details of the prototype, the computer technologies (languages and tools), and data formats are discussed. Conclusion and hints about future work end the paper.

II. THE CHOICE OF LAND USE INDICATOR

In this research, we choose the indicator of land use named *Anthropentropy Factor* (AF in the following) [3]; this neologism is derived from the Greek term *Anthropos* (Ἄνθρωπος) = man, and *entropy* and express, in a quantitative way, the "disorder" introduced in a natural, wild eco-system by the presence of human beings and their related anthropic activities. The peculiarity of this indicator is that it expresses the anthropic impact on land use not only by computing the simple percentage of soil occupied by human activities and urban expansion, but it takes into consideration also the *shape* of the anthropic areas subtracted to nature. In this way, the indicator gives also information about another aspect for land use, i.e., land fragmentation, which is considered by the UN Convention on Biological Diversity [4] as the major threat on species biodiversity preservation, because it limits the wandering and spreading of animals. In fact, after the first step of territory analysis for defining the anthropic areas, information about their shapes and contiguity is taken into consideration by expanding their boundaries in the two dimensions and, consequently, by increasing their size. The extension has two main purposes: first of all, it takes into account the negative effects due to noise and pollution close to the boundaries of the anthropic places. Secondly, if anthropic areas are sufficiently close each other, the enlargement causes an effect of filling small holes, which produces wider areas to be considered in the final computation of the land use indicator. At time of

writing, the AF is the only indicator in literature which take into account the fragmentation of an anthropic place instead only the simple numeric value of its area.

III.   THE ANTHROPENTROPY FACTOR

In this Section, we recall briefly the AF definition, in order to understand how it is computed in the core of the Web based collaborative tool here described. The procedure of the computation of the *AF* consists of the following five steps:

1) Let us consider a generic geographic region bounded by recognized borders; in this project, and in all the previous research activities, we considered as target territory the municipality, because in Italy the municipality is the administrative body in charge of deciding policies for land destination and preservation. Let define as S the area (in squared kilometres) of a target municipality under investigation.

2) Within the target municipality, the satellite image of this territory is analyzed to classify each pixel if belonging to either an anthropic area or to a wild natural area. This step is performed in a semiautomatic way by analysing the satellite images. After an initial automatic pre-processing for boundaries delimitation and road extraction, the classification is performed and validated by human experts.

3) We define a neutral sub-region (Neutral Zone) as the part of target territory containing at inland water (lakes, rivers) extending more than two squared kilometres and/or areas located more than 3,000 m above sea level. Let define NA as the area (in squared kilometres) of the Neutral Zone.

4) Each area occupied by anthropic places is enlarged along its boundaries with a buffer of 50 meters. The reason of this numerical choice is fully discussed in our previous works [3,5] and is here omitted for brevity. The enlargement is conceptually equivalent to the morphological image processing operation of dilation [6] with a circle of radius of 50 meters. We define the union of all the anthropic enlarged areas as Death Zone of the region, i.e., the zone where natural wildness is completely lost (dead) for the human anthropic influence. Let define DA as the area (in squared kilometres) of the Death Zone.

5) We define the Anthropentropy Factor *AF* as the ratio:

$$AF = DA \,/\, (S - NA) \qquad (1)$$

The *AF* is a real number in the range [0-1]. The higher is AF, the more critical is the situation for what concern land use and environmental preservation of wild ecosystems. For sake of completeness, the special case of *NA = S* is not considered, as it would mean that the entire target territory is occupied by water or it is located above 3,000 m above the sea, thus it is not suitable to land use and the computation of FA becomes meaningless.

IV.   EXISTING APPROACHES AND NOVELTY OF WEB-BASED SOLUTION

At time of writing, no related works are present in literature about Web-based approaches for AF computation.

Moreover, contributions are presents for other interesting computations about land use, but they refer to particular aspects of land data analysis and processing, e.g., hydrological change impact assessment [7], or the issue of sharing and integrating different geo-analysis models across an open web environment [8]. For this reason, in order to appreciate the novelties of the work in progress here presented, we chose to compare it to the different approaches used in our previous research activities involving AF indicator [3,5]. The relation between our current work and the previous ones is that the main goal of the research in [3,5] was to define the new indicator and to prove its efficacy in expressing the real situation of land exploitation; on the other hand, in the present paper we focus our attention on some critical issues of AF computation and we propose, as a possible solution, a Web-based collaborative framework on Google Earth satellite maps.

The AF indicator was first proposed in the Italian National project called ACI project (*Antropentropia Comuni Italiani*, i.e., Anthropentropy of Italian Municipalities) [3]. The initial, ambitious goal was to map the entire complete Italian territory in such a way that, for each of the 8092 Italian municipalities, the AF is computed according to (1). However, this goal has been disregarded, as the mapping of land use was possible only for seven of the twenty regions of Italy. In fact, the main problem is data availability, i.e., a suitable description of the territory in order to determine the position and extension of the anthropic places. At this purpose, the ACI project adopted two possible approaches. The first one refers to the data-set obtained by the Corine Land Cover (CLC) project [9]; the definition of CLC data-set started in 1990 and periodically, data are updated. Unfortunately, this approach has two big critical issues: first, Corine data set was not available for all the Italian territory, but only for 7 regions (over a total of 20 regions), namely for the 40,9% of municipalities (3311 over 8092). The second critical issue of this approach is the slowness of data update. As the validation of CLC data is a time-consuming process, the CLC data-sets refer to a description of the land use of several years before. For example, in 2018 the most recent CLC data-set of Italian territory refers to year 2012. The CLC data-set have been processed using standard Geographic Information System (GIS) software to implement the procedure to compute the values of AF according to (1). We called this first solution AF Computation based on *CLC/GIS approach*.

The second approach uses open data (Google Earth maps) and ask to users of a social network to generate the images of Death and Neutral Zones of a territory of a municipality under investigation. The images has been collected at the Computational Sustainability Unit at the Department of Electrical, Computer and Biomedical Engineering of Pavia University, were a software based on Matlab framework has been developed for AF computation. We have called this approach AF computation based on User Generated Content and open data (*UGC/Open data*). The main drawback of this approach is, as every solution based on crowdsourcing, the fact that its success is related to the degree of participation of the user community.

Unfortunately, this was not confirmed in the course of the project, and with this approach only the small percentage of 0,5% of the Italian municipalities was covered.

The novelty of the project here reported is to combine the strength of the two previously described approaches and to overcome their main drawbacks. In fact, the project is based on open, very fast updated data of Google Earth, but their processing is performed in a Web environment by trusted users, i.e., experts able to use a Web based framework to run the software for Death Zone generation and AF computation. Moreover, this software has been completely rewritten in java, and the dependency on proprietary suite of Matlab has been abandoned. Therefore, this solution seems to overcome the main drawback of the first CLC/GIS approach, i.e., the slowness of data update, as Maps in Google Earth are updated annually, as well of the second approach (UGC/Open data), because of the adoption of open source software used by a limited community of experts allow a very fast AF computation, whose efficacy is not related to a social community participation.

The present project uses the Web platform to share data and open software procedures in order to settle a collaborative environment for AF computation. Experts access maps of the territory of all the Italian Municipalities and can share intermediate results or work in collaborative form. For example, more than one expert user can analyse the territory of the same municipality to classify the anthropic areas and apply the dilation operation to generate the boundaries of the Death Zone.

V. THE WEB-BASED COLLABORATIVE PROTOTYPE

In this work in progress, we present here the general framework of the system and a preliminary result of AF computation on an entire municipality.

A. Technologies and tools

Several existing technologies and data sources are combined to realize the Web-based tool for AF computation on Google Earth images. First of all, we use an archive of shape files, downloaded from the official Italian National Institute of Statistics, to code territory boundaries information of all the 8092 Italian municipalities [10].

Then we used the free and open source QGIS software [11] to convert shape files into KML files. KML is an XML-based language for the management of three-dimensional geospatial data and it is used in several popular software, such as Google Earth and Google Maps. Successively, we developed Java code to subdivide the original KML file in multiple files, one for each municipality, in order to load them separately, by the Web tool, in a typical Google map interface. After these preliminary steps of converting and importing files in the proper environment, the software development used Google Maps API's to add drawing layers on a simple map object using the Drawing tools and the library Turfjs [12], which makes available JavaScript functions for advanced geospatial analysis in browsers. The code refers also to the GeoJSON open standard format, designed for representing simple geographical features, along with their non-spatial attributes.

Figure 1.   Graphical controls for AF computation.

Figure 2.   Two anthropic areas (in blue) inside a forest.

By combining graphical primitives with image processing functions, we coded and implemented the Web-based prototype for AF computation.

B. Funcionality and preliminary results

First of all, expert and trusted users are authorized to access the data of the projects by a simple login/password procedure in a browser. The user can choose the municipality for which the AF indicator has to be computed, by starting a new project on its territory, or edit an existing one. The projects can be public and shared with other users.

The graphical tools implemented in the project allow to draw the boundaries of the anthropic areas in the target territory, or edit, cancel or modify them. Anthropic areas are automatically coloured by the graphical tools in blue. Neutral areas are coloured in yellow. By using Google Maps API's we created drawing layers on a simple map object by embedding drawing tools for inserting circles, polygons, polylines, rectangles and by defining figures using WGS84 coordinates, directly attaching them on the map. Moreover, the developed software provides the definition of several polygon attributes, such as fill colour, border colour, and opacity. The information (coordinates, boundaries) related to each anthropic area is added to the map object in a custom data layer. Four graphical controls (see Figure 1) are added to Google Earth maps: *Occupied area* (polygon denoting urbanized, anthropic areas) *Neutral area* (polygon denoting stretches of water, such as lakes, lagoons, whose area is higher than 2 square km or lands with an altitude higher than 3000 m), *Rubber* (polygon that clears the inside area) and Undo (to undo the last graphical operation on the map). Figure 2 shows an example of drawing two polygons, in order to define two anthropic areas of houses inside a forest.

Other implemented functions manage polygon operation, such as intersection, sum, subtraction and expansion. In fact, anthropic areas have to be checked to verify if they do not exit the boundaries of the municipality: in this case the polygon area is reduced to fit the boundaries.

Figure 3.   Drawing overlapping polygons. On the left: an exisitng polygon (in bue) and a new area, delimited by its corners (white dots). On the right: the result of polygon union.

Figure 4.   AF computation on Google Earth map for the municipaliaty of Verbania, Italy.

Moreover, whenever the user add a new polygon, the software verifies if it intersects other polygons already present in the map. If it does, the two polygons are substituted by their union (see Figure 3). This step may last several days and intermediate results may be saved on a server. In a successive session, the analysis may be restarted at the point of the last save operation, by the same user or by other users, in a full collaborative and shared manner. Once the territory has been completely analyzed, the blue areas are ready for the dilation and the definition of the Death Zone. The Java routines enlarge the blue areas according to the dilation operation, in order to generate automatically the Death Zone. Moreover, the software reads the values of S from the internal database and computes the values of DA, NA for the final computation of AF indicator, according to (1). This value can be saved in the project for a further check or saved in the internal database, to associate to the municipality under investigation the value of AF indicator and the year the map refers to. In Figure 4, the first example of the map analysis and computation of the land use indicator on an entire municipality is shown: it refers to the municipality of Verbania (North Italy, Lat. 45°55′16″ N, Long. 8°33′06″ E). The yellow area refer to the lake Lago Maggiore, the red area to the computed Death Zone after dilation. For this municipality, the AF computation is equal to 0.5654, showing a worrying situation: more than half of the territory has been completely anthropized.

IV. CONCLUSION AND FUTURE WORK

As pointed out at the beginning of this paper, this is the description of a work in progress. The first results are very encouraging because the collaborative Web tool has proved to be an agile, simple and efficient framework to easily analyze up-to-date maps of the Italian territory and to compute in a fully automatic way the AF indicator.

The project is being under development to validate the tool on a significant number of municipalities. Once we populated our database of a certain number of classified maps, future work will be related to the use of these data to train a machine learning algorithm to automatically classify unknown maps, without the human supervision.

REFERENCES

[1]   LUCAS: Land Use and Land Cover survey, European Commission on Environment, http://ec.europa.eu/eurostat/statistics-explained/index.php/LUCAS_-_Land_use_and_land_cover_survey [retrieved: March, 2018].

[2]   New guidelines to reduce soil sealing, European Commission Fact Sheet of 2017, http://ec.europa.eu/environment/soil/sealing_guidelines.htm, [retrieved: March, 2018].

[3]   M. G. Albanesi and R. Albanesi, "A New Approach Based on Computer Vision and Collaborative Social Networking for Environmental Preservation: Theory, Tools and Results of Italian ACI Project", Proceedings of The Eighth International Multi-Conference on Computing in the Global Information Technology, Nice (France), 21-26 July 2013, pp. 16-21, Copyright (©) IARIA, 2013

[4]   Fragmentation of ecosystems and habitats by transport infrastructure, http://www.eea.europa.eu/data-and-maps/indicators/fragmentation-of-land-and-forests/eu-ac-fragmentation, [retrieved: March, 2018].

[5]   M. G. Albanesi amd R. Albanesi, "A Decision-making Support System for Land Use Estimation Based on a New Anthropentropy Predictive Model for Environmental Preservation – Theory, Model and Web–based Implementation", International Journal On Advances in Intelligent Systems, v 7 n 1&2, pp. 85-102, ISBN-1942-2679, Copyright (©) IARIA, 2014.

[6]   R. C. Gonzales and R. E. Woods, Digital Image Processing, 3d efition, Pearson Prentice Hall, 2008, Chapter 9, "Morphological Image Processing".

[7]   J.-Y. Choi, B. A. Engel, L. Theller, J. Harbor, "Utilizing Web-based GISand SDSS for Hydrological Land Use Change Impact Assessment", Trans. of the ASAE. Vol. 48(2), pp.815-822 , 2005.

[8]   S.Yue, M. Chen, Y. Wen, G. Lu, "Service-oriented model-encapsulation strategy for sharing and integrating heterogeneous geo-analysis models in an open web environment", ISPRS Journal of Photogrammetry and Remote Sensing, Elsevier, vol. 114, pp. 258-273, April 2016.

[9]   Commission of the European Communities: Corine Land Cover Project, http://www.eea.europa.eu/publications/COR0-landcover [retrieved: March, 2018].

[10] ISTAT On Line Municipality Database: https://www.istat.it/it/archivio/6789 [retrieved: March, 2018].

[11] QGIS software: http://www.qgis.org/, [retrieved: March, 2018].

[12] Turfjs home page, http://turfjs.org/ [retrieved: March, 2018].