SEMI AUTOMATIC REMOTE SENSING IMAGE LAYER GENERATOR BASED ON WEB BASED VISUAL ANALYTICS

Lozano J., Quartulli M., Aginako N., Olaizola I.

Vicomtech-IK4, Mikeletegi Pasealekua 57, 20009 Donostia, Spain
{jlozano; mquartulli; naginako; iolaizola}@vicomtech.org

Abstract

Large scale thematic map layer generation from Remote Sensing (RS) data can benefit the application of Visual Analytics methodologies aimed at characterizing contents by the effective inclusion of the human analyst in the interpretation loop. Web-based interfaces are specially promising in this regard. Applied analytical reasoning by visual representations involves methodological aspects dealing with both the design of interactive visualizations as well as data representation and transformation considerations. In the presented work, a decision tree is implicitly defined and used to generate a classification of the input elements in terms of relevance. This classification is used in the automatic production of the final layer.

Keywords: GIS Technologies, GIS Data Processing, Visual Analytics

INTRODUCTION

Web map servers serve geospatial data usually from a GIS database. As raw geospatial data such as Remote Sensing imagery is often hard to interpret, thematic maps need to be produced for final users in order to make contents easily understandable. Ideally, the map should be generated from raw data and based on specific user goals instead of been generic and produced by laborious manual processes. But the uncertainties that characterize this kind of data warrant the utilization of analysis and modeling approaches that help in the data understanding process. Such tools are provided by Machine Learning solutions that are capable of dealing with highly complex data. Taking into account the vast quantity of geospatial data available, this approach involves a Big Data scale machine learning problem.

Attaining the automation of thematic map creation from raw data requires integrating machine learning capabilities within the map server. Even more, seeking the user implication in the process pushes the design of an interface that serves both as a map presentation tool as well as an interface supporting the learning system.

Our first strategy presented [7] tries to solve this problem by a tile based feature value selection approach. Due to poor characterization of the tiles and the difficulty to address the semantic gap, outcomes demonstrate the difficulty to obtain relevant results. This work builds on the former by testing a new pixel based approach to the same problem.

In the following sections, we briefly review the state of the art in solutions than combine remote-sensing, visual analytics and human interactions, see section 2. Then the methodological approach of our work is presented (Section 3), and in the following section (Section 4) the platform architecture is described. Finally, we conclude with a summarization of the obtained results (Section 5).

STATE OF THE ART

This work involves different topics that are case of study: Remote Sensing image analysis, web mapping services and Visual Analytics. The first one, Remote Sensing image analysis, has experimented an enormous increase thanks to the latest technologies. The analysis of images acquired by RS systems has two principal approaches, object-based and pixel-based, although the future development goes to one kind of hybrid solution.

Within pixel-based solutions, clustering is a classical task in pixel classification problem and k-d tree based nearest neighbor searching technique is commonly used. Ujjwal [1] proposes scalable parallel clustering technique of multi-spectral remote sensing imagery using point symmetry based distance. To compute the point symmetry based distance they use k-d tree based nearest neighbor searching technique. Their results demonstrates better performance comparing to parallel implementation quantitatively and in computing time than K-Means clustering algorithm, on two, SPOT and Indian Remote Sensing satellite images.

Soe W. Myint et al [2] use images from QuickBird image data over Arizona to compare object-based and pixel based approaches. In this study, they made some tests using different combinations of spectral and spatial information. Their study demonstrates that the object-based classifier improves significantly classical per-pixel classifiers results. Costa et al [3] work consists of a supervised per-pixel classification followed by a post-classification processing with image segmentation and semantic map generalization. The results show that segmentation of spatial resolution images and semantic map generalization can be used in an operational context to automatically produce land-cover maps.

Visual Analytics applied to Remote Sensing technologies is another topic that is becoming of great interest for the scientific community. Quan et al [4] introduce a framework and class library to shorten the time and effort needed to develop web applications for visual analytics or geovisual analytics tasks. They provide a collection of geo and information visualization representations. Keel [5] introduces a visual analytics environment for the support of remote-collaborative sense-making activities. The system has computational agents that infer relationships among information items through the analysis of the spatial and temporal organization. Within this kind of works, the user interaction, user friendly environments and user interface design are issues that acquire great relevance.

Finally, inside web mapping services topic, Schröder et al [6] use the input provided by the user in a Bayesian learning framework for supervised classification of the currently open image and also for finding most relevant images (images with large extents of the trained class) across the DB.

METHODOLOGICAL APPROACH

In the following lines we expose the different concepts implemented in our solution for automatic map layer creation from raw data. These concepts support the design and implementation of the system taking advantage of a previously developed platform which is briefly described in order to support taken decisions.

Metric resolution electro-optical urban data

Focused on environmental monitoring at geographic scales, Earth observation main scenes classes were limited to urban areas, forest, agricultural areas, bare soil areas with large extensions and water bodies with decametric resolutions systems. With metric resolution solutions, the development of applications with human scale is allowed, understanding from detailed environmental dynamics to urban areas. This in turn requires new methodologies to deal with the increase of scene classes. In this contribution, we consider data acquired on the city of Rome, Italy, by DigitalGlobe Systems with metric resolution. Given the multimedia derivation of the methodology and in an extreme simplification attempt, we disregard multi spectral information content and only deal with standard tristimulus quick-looks instead.

Previous platform

The previous version of the platform [7] was based on image tiles and on an interactive selection approach to 1D feature value ranges in an N-D feature space. Images were divided on

100x100 pixel tiles and each tile was characterized by the mean of each color component in two color spaces, RGB and HSV. This way we obtained 1D histograms for all the tiles in two different color spaces. The UI allowed the user to select ranges of values separately in each component. This selection was used in real time for filtering the tiles that meet the constrains thereby expressed.

Previous platform drawbacks

The work presented in [7] was based on the selection of ranges of interest for different features. This procedure has such high sensitivity that the selected ranges of feature values often do not agree with the semantic classes of interest of the user. Furthermore, only one range can be selected within each feature space component, being impossible to select separate ranges, limiting selection capabilities.

On the other hand, the tiles are characterized so the created features don't represent the tiles correctly and the searching process generates a significant number of false positives. These are the reasons for the change in approach from explicit to implicit feature hypercube selection done by pixel selection. The use of the platform for users with dissimilar expertise and different media-informatics skills forced us to develop a User Interface with high grade of usability.

From feature hypercube to pixel selection

In principle, a pre-processing step can generate hypercubes of interest in the feature space starting from a selection performed by the user in the image space and therefore implicitly in the underlying image features space. This represents a step forward in solving the UI sensitivity issue.

Solving features representation issue requires going from a representation of the space of interest in terms of feature hypercubes to a description in terms of possibly nested hypercubes representing areas of positive and negative training with respect to the thematic class of interest. This functional improvement, that the performance issues create, in the case of very large datasets can be effectively tackled by recurring to more efficient data structures for representing the structure of the descriptor space and more specifically of how user inputs are represented in this space.

K-d trees

To get an efficient response to the query, in big data environments, the pixel data organization and its indexing are the keys. In this work we have implemented a k-d tree to help the system in indexing and searching operations of the nearest neighbor, that represents the closer item to the query.

A k-d tree is a data structure for space partitioning that organizes the data in a k dimensional Euclidean space. K-d trees are binary trees and every node represents a k-dimensional point. Every node can be considered like a space divider because it only uses a perpendicular hyperplane respect one of the axis of k dimensions. Therefore, every node is associated with one dimension and with its perpendicular hyperplane. Once the k-d tree is built, it is necessary to define a scope determined within the vicinity of the points to find the nearest neighboring points. In a nearest neighbor search of n elements, the computational cost would be $O(n)$, however k-d tree implementations reduce this computation cost to $O(log(n))$ on average.

Positive and Negative Learning

Our first attempt to generate different layers shows us that not all returned pixels are correct. In a semantic sense the quantity of false positives is bigger than expected, decreasing obtained precision value. In other words, the selection of pixels and configured neighborhood radio returns values that adulterate the generated wanted layer.

With the aim to improve the false positive values we introduce a new strategy. With this strategy we try to exclude pixels that are near to selected pixels but do not meet our definition of the semantic class of interest.

To create thematic layer, the user needs to selects different ranges with different priorities in the different feature space regions. This selection can be represented as statistical models, where selection is modeled as a combination of random variables in a feature space with an associated Probability Density Function. This requires the use of proper Probability Density Function Kernel estimation techniques to go from training histograms (e.g. obtained by selecting pixels of interest), to full Probability Density Function estimations. This kernel estimation techniques works by convolution with kernel functions.

With this approach, this first attempt uses different generalization radius parameters, similar to kernel function bandwidth, for each of the positive and negative training points. In this first implementation all radius has the same value. To implement this strategy, we use for a second time the k-d tree data structure. Thus we have a first k-d tree query results, which we called Positive Learning, which we considered the pixels and its neighbors to create a layer, and a second called Negative Learning k-d tree query results, where we store pixels, and its neighbors, which will be deleted in the results returned after the Positive Learning k-d tree was processed.

LAYER GENERATION PLATFORM

Platform architecture

The selected platform is a classical client-server platform. In his case we mount a Flask [8], a microframework for Python based on Werkzeug, Jinja 2 [9], on a server side. In the client side an OpenLayers [10] based web page is served.

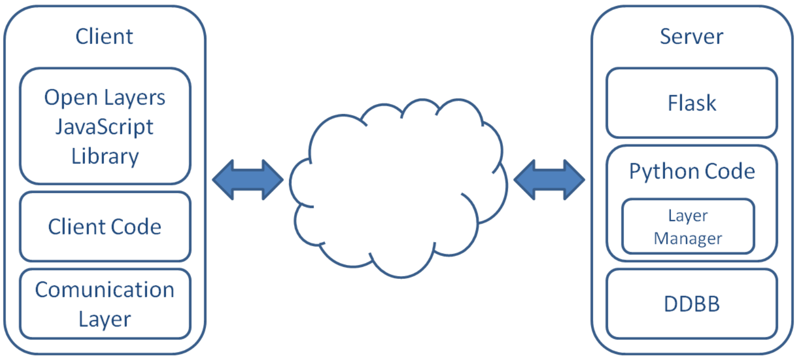

Figure 1: System architecture

OpenLayers is a JavaScript Library to display interactive maps in a web page that makes it easy to display tiles, markers, polylines and polygons, supports most map provider technologies like Google, MapServer, ArcGis, etc, allows to write new plugins and debugging through the source code.

Layer generation process

The server creates one pixel level k-dtree with for the images it has to serve. Once all k-d trees are created the user can interact with the map that the server provides by selecting the pixels which characterize the target searching, and creating a train set of data. In this train set, these pixels could be marked as positive or negative training for the learning process. K-d trees are used on the server to optimize feature space traversal in supervised classification. Classification results are presented as a new thematic layer in the user interface.

Software Engineering trade-offs

We have two important limitations in this case, storage volume and processing power. The implemented platform uses images from a part of the center of Rome in metric resolution. In addition to these base images, the platform needs tiled images for zooming. Added to this we have a layer generation process that creates a new tile in all resolution levels for each created layer. All this images could be an enormous quantity of data that translates into storage and processing costs. Therefore, although pixel-level precision is possible with this approach, a resolution reduction factor is considered for the thematic layers. The implementation of the resolution reduction factor allows us to configure platform precision in easy way and serve our platform in Platform as a Service (PaaS) sites in reduced way.

User Interface

As mentioned above, the main problem in the previous solution interface is the tile-based search where the poor definitions of the features, which characterizing the tile, imply an enormous difficulty to describe the features that characterize the kind of tiles that the users want to get. These changes in the development approach imply a new interface design and new interaction process.

*Figure 2: **Left**: Previous User Interface. **Right**: Redesigned User Interface*

The first big change is the well-organized tile set with a real map appearance instead of randomly organized tiles set. It is possible thanks to the change from tile-based to pixel-based approach that permits to organize the tiles based in their real geolocalization data. The second big change is the pixel selected area, instead of feature of images, where selected pixels for learning process can configure. With these two big changes the usability of the user interfaces increase, making easier the target selection process and therefore the use of it.

About user interface the new implementation provides a higher usability and easier way of interaction, presenting more adequate results for the user needs, and reducing false positive cases dramatically.

CONCLUSIONS

The obtained results show the improvement of the platform and the classification process. To validate the generated layer, we create a ground truth map adapted to our necessities. Since false positive cases were reduced, the obtained layer is semantically better defined. All this validates the pixel based and k-d tree combination approach, considering the obtained results satisfactory in terms of accuracy and response time.

The usability of the user interface has been extended, allowing being semantically more precise when choosing targets and therefore underlying the capacity of the platform for thematic layer creation.

Figure 3: Redesigned user interface with semantic defined N-D response

REFERENCES

[1] Ujjwal Maulik and Anasua Sarkar, Efficient parallel algorithm for pixel classification in remote sensing imagery, Springer US - GeoInformatica. 2012, 391-407.

[2]Soe W. Myint and Patricia Gober and Anthony Brazel and Susanne Grossman-Clarke and Qihao Weng, “Per-pixel vs. object-based classification of urban land cover extraction using high spatial resolution imagery”, Remote Sensing of Environment, Volume 115, Issue 5, 15 May 2011, Pages 1145-1161, ISSN 0034-4257.

[3] Hugo Costa , Hugo Carreño, Fernando BaÇao and Mario Caetano, “Combining Per-pixel and Object-based Classifications for Mapping Land Cover over Large Areas”, Int. J. Remote Sens, Volume 35, 20 January 2014, Pages 738-753, ISSN 0143-1161.

[4] Quan Van Ho and Patrik Lundblad and Aström and Mikael Jern, “A Web-enabled Visualization Toolkit for Geovisual Analytics”, Information Visualization, Volume 11, January 2012, Pages 22-42, ISSN 1473-8716.

[5] Paul E Keel, “Collaborative Visual Analytics: Inferring from the Spatial Organization and Collaborative Use of Information”, IEEE VAST, Volume 11, January 2012, Pages 22-42, ISSN 1473-8716.

[6] Schroder, M.; Rehrauer, H.; Seidel, Klaus; Datcu, M., "Interactive learning and probabilistic retrieval in remote sensing image archives", Geoscience and Remote Sensing, IEEE Transactions on , Sep 2000, vol.38, no.5, pp.2288,2298, doi: 10.1109/36.868886.

[7] Javier Lozano and Marco Quartulli and Iñigo Tamayo and Maider Laka and Igor Garcia Olaizola, Visual Analytics for Built-up Area Understanding from Metric Resolution Earth Observation Data, ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Science, 2013, pp 151-154,

[8] http://flask.pocoo.org/docs/

[9] http://jinja.pocoo.org/docs/

[10] http://openlayers.org/

AUTHORS BIOGRAPHY

Javier Lozano Silva studied Technical Engineering in Telecommunication Systems at the School of Engineering of Mondragon, Mondragon Unibertsitatea (2000-2004), he completed his studies and graduated from the faculty of Telecommunication Engineering in between 2004-2007, also at the School of Engineering of Mondragon, Mondragon Unibertsitatea. Afterwards, he developed the End of Degree Project in Orona EIC S.COOP about “The enhancement of the functionalities of Teleservicio 2005“. Since 2008 February he is working at Vicomtech-Ik4 research center (www.vicomtech.org), where he designs and develops projects of the Digital Television and Multimedia Services area. In 2009 he received the DEA (Diploma of Advanced Studies) by the System Engineering and Automation Department from the UPV/EHU University. He is currently working on his doctoral thesis.

Dr. Marco Quartulli received the laurea degree in physics from the University of Bari, Italy, in 1997 and a PhD in EE and CS from the University of Siegen, Germany, in 2005. He worked from 1997 to 2010 on remote sensing ground segment engineering, image analysis, archives and mining for Advanced Computer Systems, Italy. From 2000 to 2003, he was with the Image Analysis Group at the Remote Sensing Technology Institute of the German Aerospace Center (DLR) in Germany, working on metric resolution synthetic aperture radar image understanding in urban environments and content-based image retrieval. Since 2010, he has joined the Digital Television department of Vicomtech in Spain.

Naiara Aginako Bengoa got a degree in Telecommunications engineering from the University of the Basque Country in 2005. From 2003 till 2005 she collaborated the Signal Processing and Communication Group of the Electronics and Telecommunications department (http://det.bi.ehu.es/). She develops and manages research projects in Vicomtech-IK4 since 2005, in the Digital Media department. Since three years and a half she is also teaching at the Politechnique School of Donostia. She is currently pursuing her PhD in image analysis and content based retrieval of images and video. Her work as a researcher counts on a number of publications and on two patents.

Dr. Igor G. Olaizola received a six years degree in Electronic and Control Engineering from the University of Navarra, Spain (2001) and a PhD in Infomatics from the Faculty of Informatics of San Sebastián, in the University of Basque Country (2014). He developed his master thesis at Fraunhofer Institut für Integrierte Schaltungen (IIS), Erlangen Germany 2001 where worked for a year as research assistant on several projects related to MPEG standard audio decoding. He is member of Vicomtech technological centre since 2002. Nowadays he is the head of the Digital Interactive TV and Multimedia Services department. In 2006 he also participated as a technology consultant in Vilau (www.vilau.net) company for one year. His PhD focused on automatic multimodal indexing and management of multimedia content.