The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume XL-1/W5, 2015
International Conference on Sensors & Models in Remote Sensing & Photogrammetry, 23–25 Nov 2015, Kish Island, Iran

SEMANTIC INDEXING OF TERRASAR-X AND IN SITU DATA FOR URBAN ANALYTICS

D. Espinoza Molina [a, *], K. Alonso[a], M. Datcu[a]

[a] German Aerospace Center (DLR), Earth Observation Center, Remote Sensing Institute,
Oberpfaffenhofen, 82234 Weßling, Germany – Daniela.EspinozaMolina@dlr.de –Mihai.datcu@dlr.de

KEY WORDS: Earth-Observation images, data mining, semantic definition, land use land cover, analytics

ABSTRACT:

This paper presents the semantic indexing of TerraSAR-X images and in situ data. Image processing together with machine learning methods, relevance feedback techniques, and human expertise are used to annotate the image content into a land use land cover catalogue. All the generated information is stored into a geo-database supporting the link between different types of information and the computation of queries and analytics. We used 11 TerraSAR-X scenes over Germany and LUCAS as in situ data. The semantic index is composed of about 73 land use land cover categories found in TerraSAR-X test dataset and 84 categories found in LUCAS dataset.

1. INTRODUCTION

The continuous image acquisition and advances in storage technology have led to tremendous growth in very large and detailed image databases. These databases, if analysed, can reveal useful information to the human users (Hsu et al., 2002). However, search techniques are required in order to take advantage of the huge image archives. Image data mining systems have been introduced to deal with finding and retrieving scenes of interest. In this context, several implementations using different approaches such as image retrieval based on image content (Datcu et al., 2003),(Shyu et al., 2007) have been proposed. In the next generation of search engines, semantic concepts were integrated to the image content in order to improve the retrieval and to partially solve the semantic gap caused by the different understanding be- tween humans and machines (Rasiwasia et al., 2007). Nowadays the tendency is to use several types of information for querying and exploiting the image archives as for example combining metadata, content and semantics was proposed in (Espinoza- Molina and Datcu, 2013). All those advanced search engines are able to find hidden information in the image archives and retrieve big amount of data. However there is a need of dissemination tools for understanding and analysis of the results. Therefore, the new challenge is the way of visualization and presentation of the results. Methods like data visualization, visual data exploration, and visual analytics play an important role in the data mining process and presentation of query results (Keim et al., 2009).

Earth-Observation (EO) images are broadly used to create several types of applications as for example Land Use and Land Cover (LULC) classifications, urban mapping, disaster assessment, monitoring environmental changes and trends in urban development, urban analytics, etc. Previously, EO images were mainly used in macroscale urban mapping. Currently, the avail- ability of high resolution Optical and Synthetic Aperture Radar (SAR) data has assisted in recognizing more details within an urban scene like road detections, building extraction, man-made object recognition, etc. For instance, TerraSAR-X system (TX- GS-DD-3302, Issue: 1.6) offers high resolution SAR data, in which buildings, roads, vegetation area and man-made structures are clearly distinguishable and can be indexed in a LULC catalogue. Moreover, with this increase in resolution, pixel sizes have become smaller than the objects on ground enabling new object oriented automatic recognition and indexing techniques. In addition, EO images may be complemented with other kind of information as for example geographical information in vector format coming from geographical information systems, thematic databases composed of alpha numeric information, data collected in situ, etc. Regarding in situ data, the European Com- mission has consolidated efforts to carry out a survey on the state and the dynamics of changes in land use and cover in the European Union called LUCAS: Land Use/Cover Area frame statistical Survey (Commission, 2000). LUCAS is a good example of showing the integration of several types of information since it comprises geographical information (latitude/longitude coordinates) of the points, the thematic data explaining content of the terrain, and documented the land cover/use of the point by taking photographs of the explored point.

In this paper, we propose to generate a semantic index using TerraSAR-X images and in situ data towards land cover and land use analytics applications. One goal of the paper is to create an advanced index of the TerraSAR-X image content by using ma- chine learning methods and relevance feedback. A second goal is to integrate in situ data like LUCAS dataset. The semantic indexes of both datasets will allow exploring and exploiting the image content using semantics and geographical information in order to respond questions such as the distribution of land cover and land use categories by cities. The structure of this paper is the following: Section II describes the semantic definition and indexing of the image content and in situ data. Section III focuses on the experimental results. Finally, Section IV concludes the paper.

2. SEMANTIC DEFINITION AND INDEXING OF THE IMAGE CONTENT

Along the years many approaches to semantically describe the image content have been presented. The study of (Liu et al., 2007) summarized and remarked the importance of high-level semantics for content-based image retrieval. Here, the authors identified five major categories for reducing the semantic gap: (1) using object ontology to define high-level concepts; (2) using machine learning methods to associate low-level features with query concepts; (3) using relevance feedback to learn users intention; (4) generating semantic template to support high-level image retrieval; (5) fusing the evidences from HTML

*Corresponding author. This is useful to know for communication with the appropriate person in cases with more than one author

text and the visual content of images for WWW image retrieval. In the framework of defining semantics by ontologies, (Steggink and Snoek, 2011) proposed the image annotation using a game which includes semantic structure by means of the WordNet ontology. In the context of machine learning, the work of (Lienou et al., 2010) presented the annotation of large image databases based on the supervised classification of the patches and the integration of spatial information between the patches. Here, the semantic concepts were defined by the user. Our approach is focused on machine learning methods and user interaction (relevance feed- back) to generate semantic indexes of the image content. Further applications can be created based on such as indexes as for example urban analytics.

In this work the procedure followed is: (1) generation of the LULC catalogue (indexing) of the TSX image content using machine learning methods and relevance feedback; (2) processing and indexing of LUCAS data; and (3) computation of queries and analytics using of both datasets.

Step 1: Semantic indexing of the TerraSAR-X image content

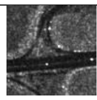
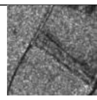
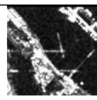
The formulation of high-level semantic features may require the use of formal tools such as supervised or unsupervised machine learning techniques. In the case of supervised learning, the goal is to predict the value of an outcome result (for example, semantic category label) based on a set of input data. In the case of unsupervised learning, the goal is to describe how the input data are organized or clustered (Liu et al., 2007). Support Vector Ma- chines (SVMs) are a group of supervised learning methods that can be applied to classification or regression. A SVM is often used to learn high-level concepts from low-level image features. It has been used for object recognition, text classification, and can be applied to image classification. The following steps were per- formed to define LULC semantic categories from TerraSAR-X images.

1. Analysis and extraction of the image content: this step involves tiling the images to generate a pyramid with multi-size of patches, extracting the metadata from the sources, converting the patch content into primitive features, and storing all the processed information into the geo-database. The primitive feature extraction is based on two methods Gabor filters (Manjunath and Ma, 1996) and Weber local descriptors (Chen et al., 2010),(Cui et al., 2013a). The results of this process are the image content descriptors in the form of vectors (e.g., feature vectors), the high resolution quick-looks, and the metadata entries. A geo-database scheme is designed in order to support all the information.

2. Semantic definition via machine learning methods: When all the generated information is available in the geo-database a new process starts: the semantic definition by using the support vector machine and human supervision. Here, the SVM uses a large pool of unlabelled data (test data) and only a small set of labelled data (training data) to predict an image semantic class. Starting from a limited number of labeled data, active learning selects the most informative samples to speed up the convergence of accuracy and to reduce the manual effort of labeling (Espinoza- Molina and Datcu, 2013). Active learning has two core components: the sample selection strategy and the model learning, which are repeated until convergence. The sample selection is performed with the help of an expert, who defines the training data to be used. The expert selects a set of patches and gives positive and negative feedback examples. A positive example means that the patch contains the desired content. Later, the list of positive and negative samples is passed as training data to the support vector machine. The SVM creates a model based on the training data, using this model it will be able to predict whether another patch belongs to the desired category or not. At the beginning of the procedure, only a few labelled instances are available (training data) then a coarse classifier is learnt and applied to the test data. After that, the iteration of the two components is repeated until the classification results are satisfactory. The expert decides when to stop the interaction and store the new defined semantic class, incorporating the concept of relevance feedback (Zhang et al., 2001).

In order to define a new semantic LULC class, we rely on the LULC taxonomy presented in (Dumitru et al., 2014), which described the possible LULC classes that can be retrieved from TerraSAR-X products. Here the semantic categories are grouped in a two-level hierarchical taxonomy, where the main categories describe general land use and land cover classes (i.e. urban area, water bodies, forest, bare ground, agriculture), while the secondary categories represent specific characteristics of the main categories as for example high density urban area, industrial areas, forest, tress, lakes, see, ocean, etc. The complete taxonomy is fully described in (Dumitru et al., 2014). Table 1 shows examples of LULC categories based on TerraSAR-X images.

Table 1: Example of land use land cover categories of TerraSAR-X images

			
Industrial area	Roads	Agricultural Land	Bridges

Step 2: Processing and indexing of in situ data

In the second part, we selected LUCAS as in situ data, which means that the data is gathered through direct observations by the surveyors on the ground all over the European Union during 2006 to 2013. A surveyor recorded the geographical information (latitude/longitude coordinates) and the thematic information such as the content of the terrain (grass, crops, etc.). In addition, a photo- graph sequence was taken for each location successively looking at North, East, South, and West (clockwise rotation). LUCAS land cover classification is a 3 level hierarchical scheme. The main level contains 8 land cover categories: artificial land, crop- land, woodland, shrub land, grassland, bare land, water and wetland. In total, there are 84 categories for land cover. The thematic information in LUCAS comes in csv files which can be directly uploaded into the geo-database. Moreover, the csv files describe information about the points where the survey was performed and the LULC classes. There are between 4 and 5 photos for each point and they are in jpg format. The processing starts uploading the points and LULC categories from csv files into the database. Later using a script the link between the photos with their respective points is created in the database; and finally the relationship between the photos and the LULC semantic categories (annotation of the photos) is established and stored in the geo-database. Table 2 shows examples of LULC categories created using LU- CAS data.

Step 3: Query and Analytics

In the third part, since all the information is stored in a geo-database, this will allow to link different sources of information together so important relations between the data can be seen. Linking the patches with their semantic annotations and geographical location, for instance, allows finding the LULC categories of a specific location. Moreover, the different information sources (TSX and LUCAS) are easily integrated

Table 2: Example of land use land cover categories of LUCAS data

			
Grassland without tree/shrub cover	Non built-up linear features	Buildings with 1 to 3 floors	Maize

in the database and their annotations can be jointly used. Having the TSX patches and the defined LULC categories together with LUCAS photos and their semantic annotations, an analysis based on spatial queries is per- formed to understand the relation between objects with different semantic present in TerraSAR-X images and LUCAS data. The queries are based on Standard Query Language and their results are exported to csv files. These files can be used for computing some analytics which are presented in the form of pie or bar charts.

3. EXPERIMENTAL RESULTS

3.1 Description of the dataset

The test data set is composed of TerraSAR-X images and LUCAS data over Germany.

3.1.1 TerraSAR-X dataset The TSX image dataset comprises 11 Multi-look Ground range Detected (MGD) TerraSAR-X L1b products radiometrically enhanced (RE) over Germany. The dataset contains around 1000 metadata entries. The images are high resolution spotlight mode with pixel spacing equals to 1.25 meters, and resolution of about 2.9 meters.

3.1.2 Land Use and Land Cover in situ data Our experiments were conducted using LUCAS data over Germany in 2009, which is composed of about 22.000 observed points, 95.000 photos with size of 1600×1200 pixels, and 84 land use/land cover categories.

3.2 Procedure

The process starts reading the annotation xml file from the TSX L1-b product and extracts relevant information like geographical coordinates. Later, each image is cut into patches with 160×160 pixel size resulting in about 10.000 very high resolution patches. Next, the high resolution quick-look of each patch is generated. In the following, the primitive features are extracted using Gabor filter and Weber local descriptor method, thus each patch is characterized by two feature vectors. To finalize the content analysis all the generated information is stored into the geo-database.

Once the information is available in the database, we performed the semantic annotation of the image content. The tool presented in (Cui et al., 2013b) was used to support the annotation process. This tool allows the users to search patches of interest in a large repository via the Graphical User Interface (GUI), a list of patches with their respective quick look is shown to the user. The tool allows ranking the suggested images which are expected to be grouped in a class of relevance. When a relevant class is found, the user concludes whether the retrieved patches belong or not to the desired semantic category so the annotation of this set of patches is generated. The user introduces a semantic description to the retrieved class and the tool groups the patches accordingly. As result of the annotation, the 10.000 patches are linked with 73 semantic categories forming a land use land cover catalogue.

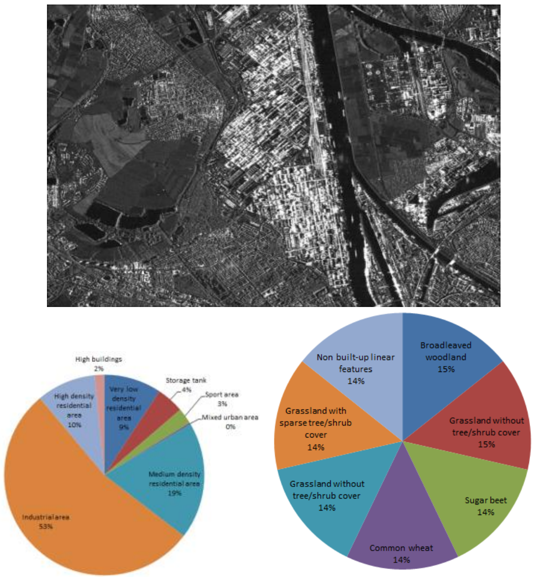


Figure 1: Distribution of the land use and land cover categories of the TerraSAR-X scene found using (Left) the SVM and (right) LUCAS data.

3.3 Examples of Queries and Analytics

Analytics helps the end-user to better understand the content of data and the relation between different variables. The following figures present some examples of analytics based on land cover and land use distributions using both datasets. Figure 1 shows a TerraSAR-X image and the distribution of LULC categories found by using the SVM and LUCAS annotations. The pie chart at the left part of Figure 1 shows that semantic categories like high building, industrial area, sport area, etc. were found using human expertise and machine learning method while categories such as common wheat, grass land were defined using LUCAS data. In order to find the categories of LUCAS in the region, a geographic query using the four image corner coordinates as parameters was performed. It retrieved the annotations and computed the percentages of coverage; the results are summarized in the pie chart at the right part of Figure 1.

The second example presented in Figure 2 summarizes the distribution of the LULC classes in both datasets. For the sake of simplicity only the main categories are presented. Upper part indicates the main categories found in TerraSAR-X data and their coverage while the lower part shows the distribution of categories discovered in LUCAS. We can observe that TerraSAR-X and LU- CAS datasets have high diversity of categories which are not uniformly distributed. In the case of LUCAS, the highest category is *Cropland* with 52 percentage followed by *Grassland* while in the case of TerraSAR-X, the major distribution corresponds to *Urban area* with 47 percentage.

The last example presented in Figure 3 shows the LULC distribution in the different German regions according to LUCAS data. Here, it can be seen the Bayern has the highest number of *Crop- land* and *Grassland* followed by Baden-Württemberg. The cate- gory *Wetland* is the lowest annotated in the dataset.

4. CONCLUSION

In this paper we presented our approach for semantic definition of the TerraSAR-X image content using machine learning methods and relevance feedback, and the processing and

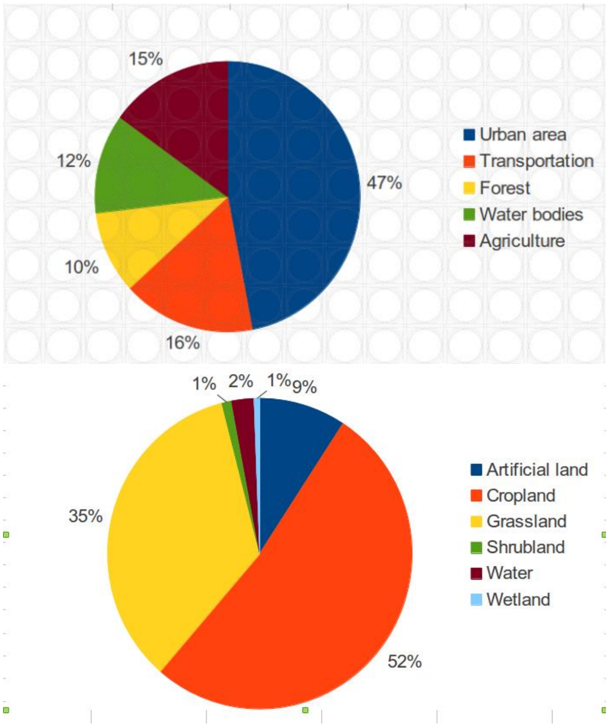


Figure 2: Land Use and Land Cover distribution using TerraSAR-X images and LUCAS over Germany

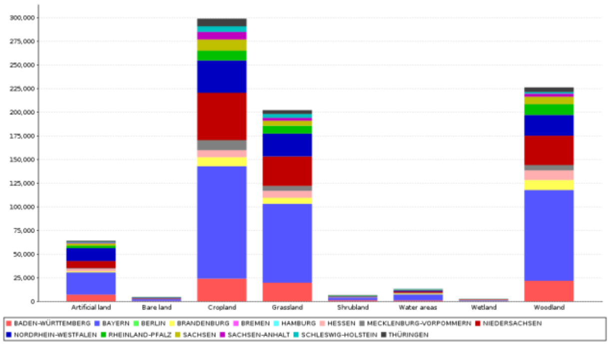


Figure 3: Land Use and Land Cover distribution in different German regions using LUCAS data

integration of in situ data in order to have a semantic index of both datasets stored in a geo-database that can later be used for other types of applications i.e. statistics, analytics, etc. As conclusion, we can remark that the use of auxiliary data coming from observations in situ helps to improve the LULC semantic categories found in TerraSAR-X images, since the data in situ contain several reliable entries about the land use land cover can be considered as ground truth. Moreover, combining both data types, interesting applications like urban statistics and analytics can be achieved. As further work remains the classification of the image content based on LUCAS annotations and the analysis of relation between both dataset for automatically annotation of the image content.

BIBLIOGRAPHY

REFERENCES

Chen, J., Shan, S., He, C., Zhao, G., Pietikainen, M., Chen, X. and Gao, W., 2010. WLD: A Robust Local Image Descriptor. IEEE Transactions on Pattern Analysis and Machine Intelligence 32(9), pp. 1705–1720.

Commission, E., 2000. Land Cover Use statistics: LUCAS. http://epp.eurostat.ec.europa.eu/portal/page/portal/lucas/introduction.

Cui, S., Dumitru, C. and Datcu, M., 2013a. Ratio-Detector-Based Feature Extraction for Very High Resolution SAR Image Patch Indexing. IEEE Geoscience and Remote Sensing Letters 10(5), pp. 1175–1179.

Cui, S., Dumitru, O. and Datcu, M., 2013b. Semantic annotation in earth observation based on active learning. International Journal of Image and Data Fusion (ahead-of-print), pp. 1–23.

Datcu, M., Daschiel, H., Pelizzari, A., Quartulli, M., Galoppo, A., Colapicchioni, A., Pastori, M., Seidel, K., Marchetti, P. and D'Elia, S., 2003. Information mining in remote sensing image archives: system concepts. IEEE Transactions on Geoscience and Remote Sensing 41(12), pp. 2923 – 2936.

Dumitru, C. O., Shiyong Cui, G. S. and Datcu, M., 2014. A Taxonomy for High Resolution SAR Images. In: Image Information Mining: Geospatial Intelligence from Earth Observation Conference (ESA-EUSC-JRC 2014), pp. 89–92.

Espinoza-Molina, D. and Datcu, M., 2013. Earth-Observation Image Retrieval based on content, semantics and metadata. IEEE Transactions on Geoscience and Remote Sensing.

Hsu, W., Lee, M. and Zhang, J., 2002. Image mining: Trends and developments. Journal of Intelligent Information Systems 19(1), pp. 7–23.

Keim, D. A., Mansmann, F., Stoffel, A. and Ziegler, H., 2009. Visual analytics. Encyclopedia of Database Systems pp. 1–6.

Lienou, M., Maitre, H. and Datcu, M., 2010. Semantic Annotation of Satellite Images using Latent Dirichlet Allocation. IEEE Geoscience and Remote Sensing Letters.

Liu, Y., Zhang, D., Lu, G. and Ma, W.-Y., 2007. A survey of content-based image retrieval with high-level semantics. Pattern Recognition 40(1), pp. 262 – 282.

Manjunath, B. S. and Ma, W. Y., 1996. Texture features for browsing and retrieval of image data. IEEE Transactions on Pattern Analysis and Machine Intelligence 18(8), pp. 837–842.

Rasiwasia, N., Moreno, P. and Vasconcelos, N., 2007. Bridging the Gap: Query by Semantic Example. IEEE Transactions on Multimedia 9(5), pp. 923 –938.

Shyu, C.-R., Klaric, M., Scott, G. J., Barb, A. S., Davis, C. H. and Palaniappan, K., 2007. GeoIRIS: Geospatial Information Retrieval and Indexing System-Content Mining, Semantics Modeling, and Complex Queries. IEEE Transactions on Geoscience and Remote Sensing 45(4), pp. 839–852.

Steggink, J. and Snoek, C. G. M., 2011. Adding semantics to image-region annotations with the name-it-game. Multimedia Systems pp. 367–378.

TX-GS-DD-3302, Issue: 1.6. Terrasar-x basic products specification document. Technical report, DLR.

Zhang, L., Lin, F. and Zhang, B., 2001. Support vector machine learning for image retrieval. In: Proceedings in International Conference on Image Processing, 2001, Vol. 2, pp. 721–724 vol.2.