GABOR AND WEBER FEATURE EXTRACTION PERFORMANCE BASED ON URBAN ATLAS GROUND TRUTH

Mihaela STAN[1], Anca POPESCU[2], Dan Alexandru STOICHESCU[3]

This paper compares the performance of two feature extraction methods applied on data acquired by the ESA satellite Sentinel-1. The feature extraction methods used in our experiments were previously tested on high and very high resolution SAR data and were reported to be able to discriminate between a relevant numbers of land cover classes. Based on the available resolution (10x10m) of Sentinel-1 Interferometric Wide images compared to the resolution of TerraSAR-X (1m) the number of detected classes is expected to be much lower. The results are quantitatively assessed by employing optical data (Urban Atlas) as reference and for visual support.

Keywords: feature extraction, image classification, Sentinel-1, Gabor filter, Weber Local Descriptor, SVM

1. Introduction

SENTINEL-1 (S-1) is an imaging radar mission providing continuous all-weather, day-and-night imagery at C-band. The main operational mode features a wide swath (250 km) with high geometric (typically 20 m Level-1 product resolution) and radiometric resolutions, suitable for most applications.

In this paper, we present a benchmark for S-1 images classification by using statistical local descriptors such as Gabor filter [1] and Weber Local Descriptor [2], combined with a support vector machine (SVM) classifier.

Both feature extraction methods that we used in our experiments, were previously tested on high and very high resolution SAR data (TerraSAR-X). The aim of these tests was to obtain a classification of image texture, by extracting the maximum amount of information from the satellites images. The results were grouped in semantic classes such as buildings, water, vegetation, forest etc. [4]

[1] PhD student, Faculty of Electronics, Telecommunications and Information Technology, University POLITEHNICA of Bucharest, Romania, e-mail: stan.mihaela.andreea@gmail.com
[2] PhD As. Prof., Faculty of Electronics, Telecommunications and Information Technology, University POLITEHNICA of Bucharest, Romania, e-mail: ancaandreeapopescu@gmail.com
[3] Prof., Faculty of Electronics, Telecommunications and Information Technology, University POLITEHNICA of Bucharest, Romania, e-mail: stoich@elia.pub.ro

2. Sentinel-1 images

The first SENTINEL-1 (S-1) satellite was launched in April 2014. The S-1 Synthetic Aperture Radar (SAR) instrument may acquire data in four exclusive modes illustrated on Fig. 1: Strip map (SM), Interferometric Wide swath (IW), Extra Wide swath (EW), and Wave (WV). Additionally the data provided can be in four differently processed formats as well as three different resolutions where the resolution strongly depends on the selected mode. We have performed our experiments using images in the IW acquisition mode, that have a spatial resolution of 20x22m (range x azimuth) and a pixel spacing of 10x10 pixels.

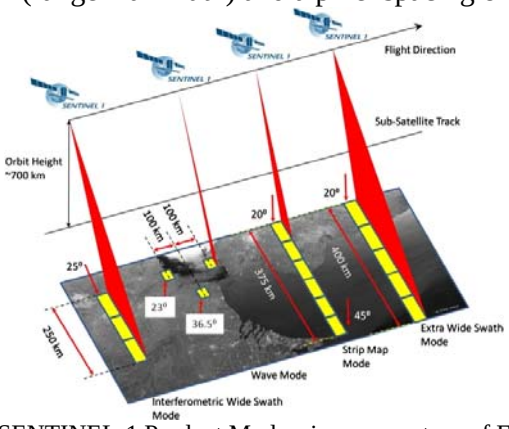

Fig. 1. SENTINEL-1 Product Modes; image courtesy of ESA [6]

3. Feature extraction

In the literature there are multiple feature extraction methods that are used for classification of SAR images, mostly based on the spatial distribution of pixel intensities, such as co-occurence [9], Markov Random Fields [8], or computed at different scales such as wavelet-based methods [10]. The next sections present the methods employed in our experiments, namely the Gabor filtering and Weber Local Descriptors.

3.1 Gabor filtering

A Gabor filter is a linear filter used for edge detection and it is characterized by a preferred orientation and a preferred spatial frequency. Frequency and orientation representations of Gabor filters are similar to those of the human visual system, and they have been found to be particularly appropriate for texture representation. Thus, the 2D Gabor filter can be defined by the equation (1) where the parameters $\lambda, \theta, \varphi, \sigma, \gamma$ represent wavelength, orientation angle (in radians), phase offset, standard deviation and filter scale.

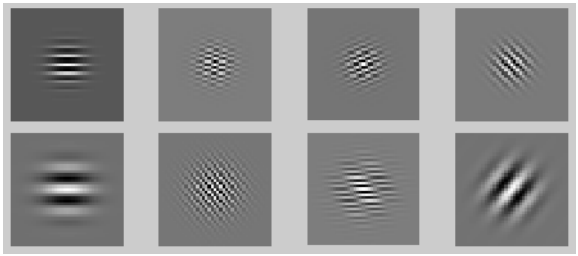

Fig. 2. Examples of Gabor filters in the spatial domain with orientation 2 and scale 4, theta=90

The Gabor filter can be defined as:

$$g_{\lambda,\theta,\varphi,\sigma,\gamma}(x,y) = e^{-\left(x'^2+\gamma^2 y^2\right)/2\sigma^2}\cos\left(2\pi\frac{x'}{\lambda}+\varphi\right) \tag{1}$$

where:

$$\begin{aligned} x' &= x\cdot\cos\theta + y\cdot\sin\theta \\ y' &= -x\cdot\sin\theta + y\cdot\cos\theta \end{aligned} \tag{2}$$

$$\frac{\sigma}{\lambda} = \frac{1}{\pi}\sqrt{\frac{\ln 2}{2}\cdot\frac{2^b+1}{2^b-1}} \tag{3}$$

$$b = \log_2\frac{\frac{\sigma}{\lambda}\pi+\sqrt{\frac{\ln 2}{2}}}{\frac{\sigma}{\lambda}\pi-\sqrt{\frac{\ln 2}{2}}} \tag{4}$$

3.2 Weber Local Descriptors (WLD)

The Weber Local Descriptor is inspired by the SIFT (scale invariant feature transform) which is invariant to scaling and rotations, and is based on the Weber law which is described by equation (5) and is known as Weber ratio. The law defines constant threshold for which stimulus variation is noticeable by a human:

$$\frac{\Delta \mathrm{I}}{\mathrm{I}} = k \tag{5}$$

Where ΔI represents the increment threshold, I represents the initial stimulus intensity, and k signifies that the proportion on the left side of the equation remains constant despite variations in the I term. WLD is a descriptor with two components: differential excitation (ξ) and orientation (θ).

For optical data the equations are given below. The input image is filtered for the computation of the intensity differences between a pixel and its neighbors. The filters $f_{00}, f_{01}, f_{10}, and\ f_{11}$ are defined in [11]. The outputs of the filters are denoted by v, while x_c refers to the current pixel:

$$v_s^{00} = \sum_{i=0}^{p-1}(\Delta x_i) = \sum_{i=0}^{p-1}(x_i - x_c) \tag{6}$$

$$G_{ratio}(x_c) = \frac{v_s^{00}}{v_s^{01}} \tag{7}$$

The differential excitation is defined as:

$$\zeta(x_c) = \arctan\left(\frac{v_s^{00}}{v_s^{01}}\right) = \arctan\left(\sum_{i=0}^{p-1}\frac{(x_i - x_c)}{x_c}\right) \tag{8}$$

and the orientation is defined as:

$$\theta(x_c) = \arctan\left(\frac{v_s^{11}}{v_s^{10}}\right) \tag{9}$$

Where

$$v_s^{10} = x_5 - x_1 \text{ and } v_s^{11} = x_7 - x_3 \tag{10}$$

In our analyses we use S-1 images that are SAR acquisitions. In this case, an adaptive WLD is employed, that takes into account the ratio of the mean of two no-overlapping neighborhoods on the opposite sides of the point being analyzed. To detect all possible edges, the ratio detector must be applied in all possible directions [3]. By taking into account those changes, the differential excitation is computed using equation (11) and the orientation by using equation (12).

$$\zeta(x_c) = \arctan\left(\sum_{i=0}^{d}\sum_{j=1}^{2}\frac{\left(\mu_j^i - x_c\right)}{x_c}\right) \tag{11}$$

μ_j^i is the local mean and d is the number of directions taken into account.

$$\theta(x_c) = \arctan 2\left(\frac{\mu_1^1 - \mu_2^1}{\mu_1^3 - \mu_2^3}\right) \tag{12}$$

5. Experiments and Results

5.1 Dataset Description

For the performance evaluation of the image descriptors on Sentinel-1 data we selected a number of relevant scenes acquired in the IW mode. The selection of the image mode was based on the spatial resolution (20m x 22m) and pixel spacing (10m x 10m), based on the rationale that the descriptors had been previously tested on lower resolution data (ASAR) and higher resolution data

(TerraSAR-X). Moreover, the IW mode is the main acquisition mode over land for the Sentinel-1 mission, and therefore the majority of datasets for land cover applications are available in IW mode.

Since our interest is to determine the number of relevant classes which can be extracted from such data, the scenes were selected in such a manner as to contain as many classes as possible, and with high intra-class heterogeneity. The following European cities were selected as our test sites: Bucharest, Munich and Paris. The regions of interest contain the three capitals and their outskirts, in order to ensure the variability of both urban and natural classes.

Fig. 3. Bucharest Test Site

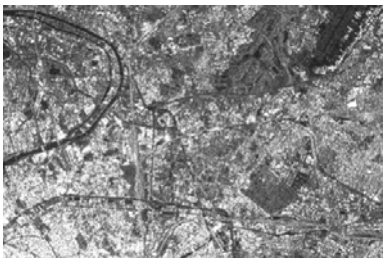

Fig. 4. Paris Test Site

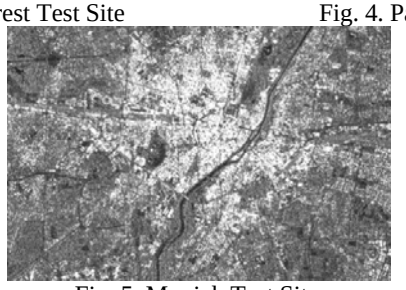

Fig. 5. Munich Test Site

5.2 Dataset Processing and Manipulation

To obtain the classification of S-1 images we performed the following-steps: image pre-processing (patch extraction using a regular grid with no overlapping), patch-wise feature computation and data classification (use of classification algorithms in order to group results into classes).

Based on the information content in the test images we were able to distinguish between classes including vegetation, water and urban fabric. In order to ensure error propagation and to obtain comparable and unitary results on the whole dataset, the first processing step implied forming a synthetic image comprising all of the selected areas. Then, we applied the feature extraction and classification on this new super-image and observed the results.

With respect to the patch extraction procedure, previous works reported that the feature extraction on high and very high resolution SAR data (TerraSAR-X) is

performed on patches between 100x100 and 200x200 pixels, as the retrievable information was very detailed. In our case, the image content has a lower degree of heterogeneity, so we choose a smaller window size used for our analysis: 20x20 and 30x30 pixels. The chosen window size ensures that the area on the ground is on the same order of magnitude as the one in the reported literature, thus the number of effective classes (objects) in one patch is comparable.

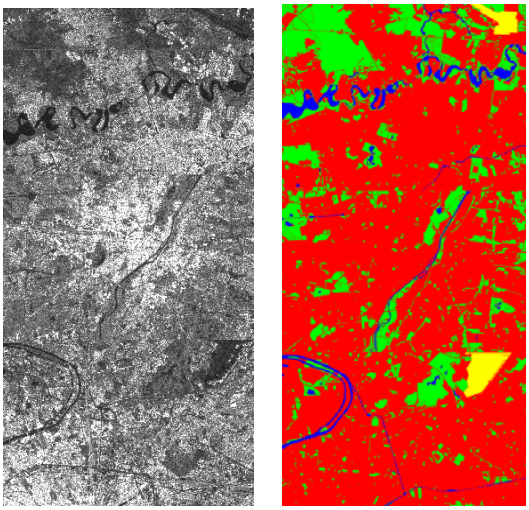

Fig. 6. Bucharest, Munich, Paris – Sentinel-1 and Urban Atlas

5.3 Results and discussions

After we created the patches and extracted the features using the Gabor and Weber methods, we applied a semi-automatic SVM (Support Vector Machine) classifier developed by DLR in order to label the image patches. The tool is semi-automatic in the sense that it learns a specific class from the positive and negative examples given by the human operator. The learning process relies on an iterative interaction between the expert and the tool. In the end, the detected positive samples are labeled (one label per session), and excluded from the following learning stage.

Finally, we computed the precision and recall (equation 13) by using as ground truth Urban Atlas data. We created the ground truth data by semantically annotating each class using as visual support optical data available in Urban Atlas and assigning each patch to the appropriate class. Fig.3 displays the concatenated Urban Atlas image with all the three cities. The results that we obtained are presented in tables 1 and 2.

Two of the problems encountered in satellite images are big data storage requirements and long time required for processing. One way to optimize this is to

reduce the resolution of the images. Therefore we performed a 50% sub-sampling of the original data (both horizontally and vertically). This effectively reduced the size of the images by four. Then we applied the same methods for extracting and class determination on the sub-sampled images. The results that we obtained are presented in tables 3 and 4.

$$precision = \frac{tp}{tp + fp} \quad recall = \frac{tp}{tp + fn}$$

Table 1

Precision and recall results for Gabor

	Class	Patch size 20 pixels		Patch size 30 pixels	
		Precision [%]	Recall [%]	Precision [%]	Recall [%]
Gabor	Buildings	89.09	97.48	95.25	95.37
	Vegetation	80.20	57.56	81.26	79.73
	Water	96.36	30.72	79.46	83.05

Table 2

Precision and recall results for WLD

	Class	Patch size 20 pixels		Patch size 30 pixels	
		Precision [%]	Recall [%]	Precision [%]	Recall [%]
WLD	Buildings	87.71	94.24	84.96	90.51
	Vegetation	59.21	46.83	55.93	43.97
	Water	68.05	28.40	39.16	31.81

Table3

Precision and recall results for Gabor (sub-sampled)

	Class	Patch size 20 pixels		Patch size 30 pixels	
		Precision [%]	Recall [%]	Precision [%]	Recall [%]
Gabor (sub-sampled)	Buildings	89.42	96.68	88.18	98.39
	Vegetation	73.54	54.26	91.47	47.36
	Water	96.31	45.50	83.75	76.13

Table 4

Precision and recall results for WLD (sub-sampled)

	Class	Patch size 20 pixels		Patch size 30 pixels	
		Precision [%]	Recall [%]	Precision [%]	Recall [%]
WLD (sub-sampled)	Buildings	91.15	87.92	88.47	91.12
	Vegetation	51.09	61.86	60	54.49
	Water	69.90	64.63	82.80	73.86

6. Conclusions

We are the first ones that applied the Gabor and Weber filters on S-1 images acquired in IW mode and compared which method offers better results for determine the number of relevant classes which can be extracted by using this type of images. Thus, from the obtained results we can conclude that in most of the cases the Gabor method gives better results than Weber, i.e. better precision

and better recall. It can also be observed that the detection of water surfaces was generally worse than the detection of buildings and vegetation. This is due to the fact that water surfaces are much smaller than the others and have not enough detail to distinguish them. While sub-sampling reduced the size of the data, we observed that it didn't significantly change the performance of the algorithms; hence time and storage can be optimized by sub-sampling the data before processing.

Acknowledgments

I would like to thank Professor Mihai Datcu for his expert advice throughout this article. The work has been funded by the Sectoral Operational Program Human Resources Development 2007-2013 of the Ministry of European Funds through the Financial Agreements POSDRU/159/1.5/S/134398 and *POSDRU/159/1.5/S/132395*.

R E F E R E N C E S

[1]. *P. Kruizinga, N. Petkov, S.E. Grigorescu*, "Comparison of texture features based on Gabor filters", Proceedings of the 10th International Conference on Image Analysis and Processing, Venice, Italy, 1999.

[2]. *C. Jie, S. Shiguang, H. Chu,Z. Guoying, P.Matti,C. Xilin, G. Wen* , "WLD: A robust Local Image Descriptor", IEEE Transactions on Pattern Analysis and machine intelligence, 2009.

[3]. *S. Cui, C. O. Dumitru, M. Datcu*, "Very High Resolution SAR Image Indexing Based on Ratio Operator", IEEE, GRS Letter, February 9, 2012.

[4]. *A.A. Popescu, I. Gavat, M. Datcu*, "Contextual Descriptors for Scene Classes in Very High Resolution SAR Images", IEEE Geoscience and Remote Sensing Letters, Vol. 9, No.1, January 2012.

[5]. *F.-A.Georgescu, C. Vaduva, M. Datcu, D. Raducanu*, "A framework for benchmarking of feature extraction methods in earth observation image analysis", Proceedings of ESA-EUSC-JRC 2014 - 9th Conference on Image Information Mining Conference: The Sentinels Era, Bucharest, 2014, pp. 81-84.

[6]. https://sentinel.esa.int/web/sentinel/user-guides/sentinel-1-sar/acquisition-modes

[7]. http://www.eea.europa.eu/data-and-maps/figures/urban-atlas

[8]. *H. Deng and A. Clausi*,"Unsupervised Image segmentation using a simple MRF model with a new implementation scheme," Pattern Recognition, vol. 37, pp.2323-2335, 2004.

[9]. *R.M.Haralick*,"Statistical and structural approaches to texture," IEEE Proc., vol.67,no.5, pp.786-804, May 1979.

[10]. *C. M. Pun and M. C. Lee*, "Log-polar wavelet energy signatures for rotation and scale invariant texture classification," IEEE Trans. Pattern Anal. Mach. Intell., vol.21, no. 4, pp. 291-310, Apr 1999.

[11]. *Jie Chen, Shiguang Shan, Chu He, Guoying Zhao, Matti Pietikainen, Xilin Chen, and Wen Gao*, "WLD: A Robust Local Image Descriptor", IEEE Transactions On Pattern Analysis And Machine Intelligence, Vol. 32, No. 9, September 2010, pp. 1705-1720.