

# A Novel Statistical-Based Scale-Independent Approach to Unsupervised Water Segmentation of SAR Images

Francesco Asaro

Department of Electronics, Information and Bioengineering  
Politecnico di Milano Milan, Italy  
francesco.asaro@polimi.it

**Abstract**—In this work, a novel approach to water segmentation of Synthetic Aperture Radar images is presented. The proposed methodology, suitable for any VV-VH dual-pol data source, it is based only on the statistical and morphological content of the two polarimetric channels. No external information or supervision is required to successfully complete the segmentation task, allowing a complete automation of the process. The methodology has been applied to Sentinel-1 Interferometric Wide-Swath data and validated using very high-resolution optical images acquired by an Unmanned Aerial System over two Italian rivers, during different seasons. The performances of the algorithm have been estimated carrying out a pixel-based validation, deriving from the binary confusion matrix two classical figures of merit such as Global accuracy and F1-score. In the validation procedure, it has been assessed an overall Global Accuracy of 0.92 and an overall F1-score of 0.81, suggesting that the presented methodology applied to S-1 data is well-suited for the monitoring of rivers characterized by a wet channel width greater than 60 m.

**Index Terms**—SAR, water mapping, segmentation, Sentinel-1

## I. INTRODUCTION

Water mapping is one of the main research topics in Synthetic Aperture Radar (SAR) imaging, with applications in a wide number of different areas of the environmental monitoring, spanning from water resource management to inland water monitoring [1]. Beside the classical advantages (i.e. active illumination, atmosphere penetration), SAR imaging has among all the other imaging systems the best capabilities in detecting water surfaces. When illuminated, water surfaces reflect the signal according to a single bounce mechanism, resulting in low amplitude areas in the SAR image. Thus, when water is present in the scene, it is possible to observe in the image histogram a bimodal distribution. The two modalities are generated from the two different scattering mechanism of water and land surfaces. In principle, it is possible to segment water surfaces in the scene identifying the associated distribution from the histogram and applying a threshold [2]. This approach corresponds to the underlying principle of a very high number of methodologies, which try to overcome through more sophisticated adaptation (e.g. split-based, tiling, statistical modelling) the severe mixture of the distributions, that occurs due to the speckle power and land-water ratio in the scene [3], [4]. Here, I propose an approach for unsupervised water segmentation that is not affected by

image scale and allows the unbiased classification of water and land. The remainder of this work is organized as follow: Section II describes the proposed methodology, in Section III are presented the validation results and Section IV hosts a brief discussion.

## II. THE PROPOSED METHOD

Are assumed two multi-temporal stacks  $S_{VV}$  and  $S_{VH}$  of co-registered SAR amplitude images, characterized by a spatial resolution  $\delta a$  in azimuth and  $\delta r$  in range, acquired under the same observation geometry and equally spaced in time (e.g. every six days) of an interval  $\Delta t$ , corresponding to the revisit time of the constellation. The pre-processing required to generate the stack is not in the scope of this work and will not be analyzed.

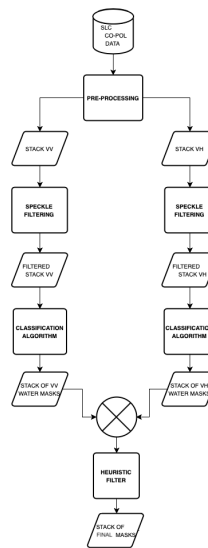


Fig. 1. Flowchart of the described steps.

The two stacks are filtered in order to reduce the speckle power (Section II-A) and classified (Section II-B) separately, producing for each epoch two different water masks for VV and VH polarization. The two masks are then combined with a heuristic filter (Section II-C) in a single final mask, refinement of the first two classifications as shown in Figure 1.

### A. Speckle Filtering

In single look SAR images, the speckle power is such to hinder the correct segmentation of distributed targets, thus whatever the classification approach it is necessary to first increase the Equivalent Number of Look (ENL), filtering the image. Averaging  $n$  independent, identically distributed, from here *homogeneous*, pixels it is possible to reduce the speckle power of a factor  $\sqrt{n}$ . The choice of the speckle filter does not modify the functioning principle of the classification algorithm, but it affects the final performance of the whole methodology. Thus, it is presented a simple resolution-saving, nonlinear, spatiotemporal filter optimized for the water segmentation task. It is considered a tridimensional convolution kernel  $K$  of size  $h, w$  and  $d$ , which operates over the stack  $S$ . The spatial dimensions of the kernel are set accordingly with the following:

$$h = \frac{\delta L}{\delta a} \quad (1)$$

$$w = \frac{\delta L}{\delta r} \quad (2)$$

where  $\delta L$  is the minimum target size for the segmentation task. Furthermore, in order to exploit the multi-temporality of the stack, it is assumed that water bodies area extent function has a step-like dynamics, rather than impulse-like, in the interval  $\Delta T$  [5], when interested by external forcing (e.g. rain, snow melting). Under this hypothesis, setting  $d = 3$ , at least  $2 * h * w$  samples are homogeneous. In order to have a more robust estimate of the filtered amplitude not affected, in the worst case, by the  $h * w$  heterogeneous sample the median is chosen as operator over the sampled population.

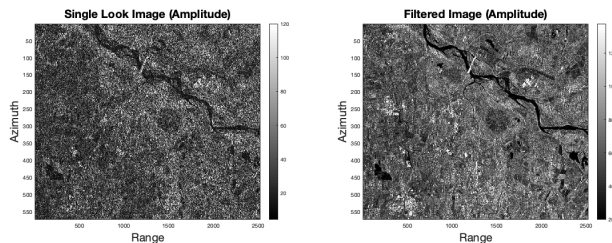


Fig. 2. Single look image (left) and its filtered version (right).

In Figure 2 it is shown a comparison between a single look Sentinel-1 (S-1) Interferometric Wide-Swath (IW) 20 – by – 5m resolution amplitude image (left) and its filtered version (right), obtained convolving the multitemporal stack with a 1 – by – 5 – by – 3 median window.

### B. Classification

Usually, a binary segmentation task is modelled as a problem where the image is composed of two classes  $\Omega_1$  and  $\Omega_2$ , described by two probability density functions (pdf)  $pdf_1$  and  $pdf_2$ . As introduced in the beginning, in this framework such an approach suffers the severe mixture of the two pdfs, and it does not allow to recover  $\Omega_1$  and  $\Omega_2$ . Rather, here it is assumed

that the image is composed by a number  $N$  of subclasses  $\omega_i$ , which belong to  $\Omega_1$  or  $\Omega_2$ . The subclasses are built following these steps:

- The image dynamic range (DR) upper bound is saturated to cut out extreme values generated from dihedral scattering;
- The saturated DR is divided in a number  $N$  of radiometric intervals of width  $\delta x$ , such that pixels in the intervals change at the feature scale  $\delta L$ . The radiometric intervals composed by less of  $N_{min}$  pixels are rejected;
- Iteratively, the  $\omega_i$  subclass is built collecting the radiometric values of all the pixels in the 3 – by – 3 neighborhood of the pixels belonging to the radiometric interval  $i$ ;
- Iteratively, the  $pdf_i$  describing the  $\omega_i$  sub-class is built through a simple gaussian model, estimating the sample mean and variance:

$$pdf_i = G(\mu_i, \sigma_i) \quad (3)$$

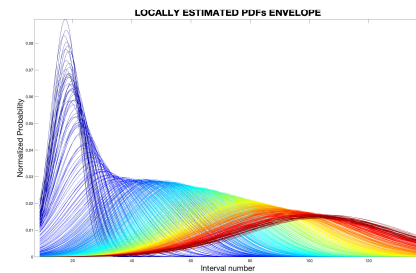


Fig. 3. Ensemble of the  $N$  locally estimated pdfs. Colors refer to the amplitude value of the  $i^{th}$  interval upper bound.

The  $N$  pdfs from Equation 3 are evaluated over the amplitude image and the maximum probability image is built as follow:

$$I_{max_{l,j}} = \max_{i \in [1,N]} G(I_{i,j}, \mu_i, \sigma_i) \quad (4)$$

where  $l$  and  $j$  refer to the pixel coordinates of the image.

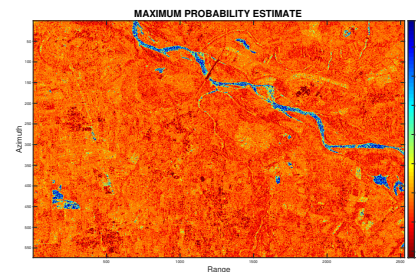


Fig. 4. Image of maximum probability as from Equation 4.

The image of maximum probability can be binarized, since it does not suffer of the issues encountered in the amplitude domain.

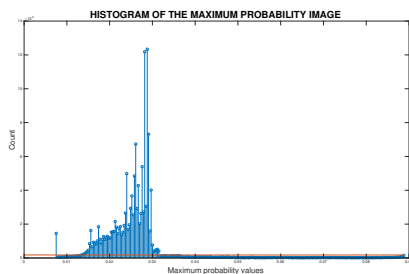


Fig. 5. Histogram relative to the image of maximum probability (Figure 4).

As shown in Figure 5 there is only one peak associated to the background of the image, water pixels are mapped unbiasedly to very high values of probability (right of the peak). This behavior is controlled by the residual speckle in the image, which, being multiplicative, is lower for the water and higher for other distributed features, meaning in a lower variance for water sub-classes (as can be seen also in Figure 3). In this methodology, the homogeneity of water features also contributes to maximize the separation from the background in the probability domain, further lowering the variance of the subclasses.

### C. Heuristic Filter

The classification algorithm is run on both the  $S_{VV}$  and  $S_{VH}$  stacks, producing a stack of logical water masks. For each epoch, the VV mask and the VH mask are combined and filtered to obtain a unique refined classification. In order to separate the pixels which have been classified positively in both polarizations from the pixels which have been classified positively just in once, the two masks are summed together. Pixels from the first group are accepted in the final classification only if they are at least adjacent to another pixel of their group. This criterion, based on the continuity of water surfaces, allows removing outliers introduced by the fluctuation of the residual speckle or by misclassification, such as patches of saturated soils. Pixels from the second group are accepted in the final classification only if they are at least adjacent to one pixel of the refined first group. This second criterion allows to filter out outliers introduced by depolarization effects introduced by the geometry of the targets.

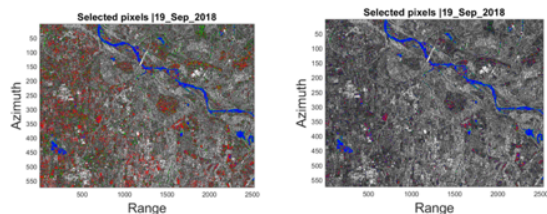


Fig. 6. Sum of VV and VH classification masks (left) and filtered classification mask (right). VV-VH positive blue, VV positive red and VH positive green.

In Figure 6 it is shown the result of the heuristic filtering

operation. In conclusion, in the post-processing step the stack of water masks is geocoded in cartographic or geographic coordinate system.

## III. VALIDATION

The algorithm has been validated applying the methodology described in Section II to S-1 IW-TOPS data and comparing the output classifications to ground truths derived from photointerpretation of very-high resolution images acquired by an Unmanned Aerial System (UAS).

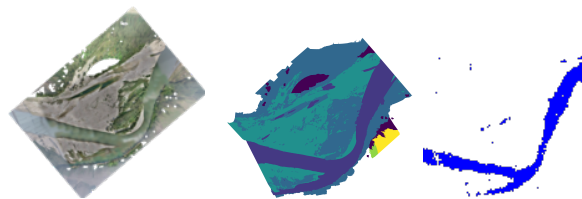


Fig. 7. UAS acquisition (left), ground-truth from photointerpretation (center) and water mask (right).

### A. Validation Area

Po and downstream Sesia rivers (Piedmont, Italy) have been chosen as test areas. Different scenes have been acquired during 2017 and 2018, according to different seasonality. Acquisitions are listed in the following table.

ID	UAS	S-1
SC_2017_07	19/07/17	19/07/17
PO_2017_09	16/09/17	14/09/17
SC_2018_0	16/04/18	18/04/18
PO_2018_04	17/04/18	16/04/18
SC_2018_09	22/09/18	21/09/18
PO_2018_09	20/09/18	19/09/18

### B. Validation Strategy

The validation has been performed on a pixel basis undersampling with a mode interpolation the 40cm resolution ground truth on the same grid of the 20m resolution S-1 classification.

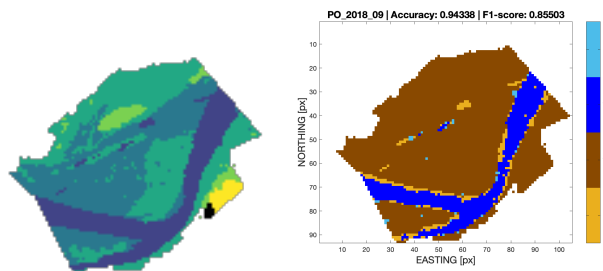


Fig. 8. Undersampled ground-truth (left), pixel based classification (right): turquoise false positives, blue true positive, light brown false negative, dark brown true negative.

### C. Validation Results

As shown in Figure 8, for every classification ground-truth pair the confusion matrix is computed as well as the Global Accuracy and the F-1 score, defined as follow:

$$ACC = \frac{TP + TN}{TP + TN + FP + FN} \quad (5)$$

$$F1 = \frac{2 * TP}{2 * TP + FP + FN} \quad (6)$$

Computing the confusion matrix over the whole classification dataset, it has been assessed an overall Global Accuracy of 0.92 and an overall F-1 score 0.81. The results for the single observation are listed in the following table.

ID	ACC	F1
SC_2017_07	0.888	0.667
PO_2017_09	0.941	0.81
SC_2018_04	0.920	0.773
PO_2018_04	0.914	0.853
SC_2018_09	0.887	0.702
PO_2018_09	0.943	0.855

### IV. DISCUSSION

From the scatterplot Global accuracy versus F-1 score shown in Figure 9 it is evident that the performances depend from the wet channel width of the monitored river (dot color).

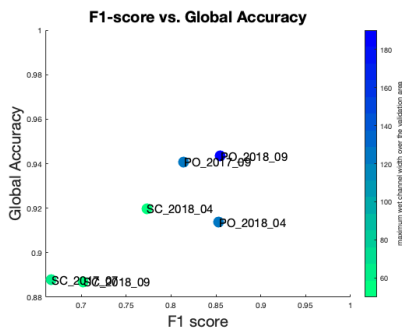


Fig. 9. Scatterplot of Global Accuracy and F-1 score. The color of the dots refers to the wet channel width of the river during the acquisition.

The performances which have been assessed in the validation phase are satisfying for the purpose of continuously monitoring rivers characterized by a channel width greater than 60m. This result is more significant in the light of the fact that the procedure can be fully automatized and it may allow further studies on the possibility of retrieving the discharge from remote-sensed data [6].

### ACKNOWLEDGEMENTS

The author would like to thank Simone Bizzi, Barbara Belletti, Martina Bussettini, Marco Casaioli, Barbara Lastoria, Giulia Marchetti and Stefano Mariani who have worked in order to acquire the UAS images and produce the ground-truth data. I also would like to thank Professor C. Prati for the fruitful discussions.

### REFERENCES

- [1] S. Bizzi, L. Demarchi, R. C. Grabowski, C. J. Weissteiner, and W. Van de Bund, The use of remote sensing to characterise hydromorphological properties of European rivers, *Aquat Sci*, vol. 78, no. 1, pp. 5770, Jan. 2016.
- [2] J.-B. Henry, P. Chastanet, K. Fellah, and Y.-L. Desnos, Envisat multipolarized ASAR data for flood mapping, *International Journal of Remote Sensing*, vol. 27, no. 10, pp. 19211929, May 2006.
- [3] P. Matgen, R. Hostache, G. Schumann, L. Pfister, L. Hoffmann, and H. H. G. Savenije, Towards an automated SAR-based flood monitoring system: Lessons learned from two case studies, *Physics and Chemistry of the Earth, Parts A/B/C*, vol. 36, no. 7, pp. 241252, Jan. 2011.
- [4] M. Chini, R. Hostache, L. Giustarini, and P. Matgen, A Hierarchical Split-Based Approach for Parametric Thresholding of SAR Images: Flood Inundation as a Test Case, *IEEE Transactions on Geoscience and Remote Sensing*, vol. 55, no. 12, pp. 69756988, Dec. 2017.
- [5] P. E. Black, *Watershed Functions1*, JAWRA Journal of the American Water Resources Association, vol. 33, no. 1, pp. 111, 1997.
- [6] P. Matgen, R. Hostache, G. Schumann, L. Pfister, L. Hoffmann, and H. H. G. Savenije, Towards an automated SAR-based flood monitoring system: Lessons learned from two case studies, *Physics and Chemistry of the Earth, Parts A/B/C*, vol. 36, no. 7, pp. 241252, Jan. 2011.