

URBAN AREAS ENHANCEMENT IN MULTITEMPORAL SAR RGB IMAGES THROUGH A FEEDBACK SYSTEM

Donato Amitrano, Gerardo Di Martino, Antonio Iodice, Daniele Riccio, Giuseppe Ruello

Department of Electrical Engineering and Information Technology,
University of Napoli Federico II, Via Claudio 21, 80125 Napoli, Italy

ABSTRACT

In this paper we introduce a feedback system for the enhancement of urban areas in Level-1 α multitemporal RGB composite. In particular, our method focus on the interferometric coherence, whose estimator performances depends on the dimension of the computation window. The proposed method allows for the mitigation of this problem, reducing speckle and increasing the resolution and the accuracy of the output maps through the generation of an adaptive window, bringing benefits also in the process of buildings extraction.

I. INTRODUCTION

Urban areas monitoring through satellite imagery is a powerful tool for large scale analysis and planning activities concerning with growth tracking, surveillance of critical infrastructures, three-dimensional modeling, risk analysis and damage detection in case of natural disasters. This is witnessed by the huge literature dedicated to this topics in the last years [1]–[3].

In this paper, we focus on the extraction of urban areas from SAR images. We think that this operation should be achievable with simple techniques (being often carried out by nonexpert users) with limited computational burden. In fact, this activity assumes relevance when applied to large scales trying to preserve the original resolution, according to the limitations dictated by the presence of speckle. In that regard, the availability of multitemporal datasets allows for the application of despeckling procedures which preserve the spatial resolution [4]. Therefore, it is now possible to map urban areas with resolution in the order of few meters.

We present a feedback system for the MAP3 framework introduced in [5]. This processing chain rules a new family of multitemporal sub-symbolic products named as Level-1 α , having a semantic and informative content intermediate between the classic Level-1 and Level-2 products.

The principal characteristic of this new class of products is that the association between the color and the scene object is fixed given the climatic condition. In such representation, urban areas constitute a particular stable feature and appear in white since the high contribution of all the bands of intensity and coherence involved in the color composite. However, the use of the interferometric coherence poses

some questions about the effectiveness of its estimator varying the considered scene feature. A raw estimation of this characteristic of the image can affect its semantic interpretation and successive automatic processing steps aimed, for example, to the extraction of urban areas, especially when dealing with high-resolution imagery.

The work is organized as follows. In Section II we discuss the influence of the size of the mean window in the estimation of the interferometric coherence. The implementation of the feedback system is presented in Section III. In Section IV, the benefits of the presented method are explored through a buildings extraction application. Conclusions are drawn at the end of the work.

II. ON THE ROLE OF THE MEAN WINDOW IN THE COHERENCE CALCULATION

In the classic SAR literature, the interferometric coherence extraction is principally related with the interferometric processing for DEMs retrieval or with differential interferometry techniques for deformation mapping. However, the coherent characteristic of SAR sensors can find different application fields such as image enhancement [5], classification [6] and urban areas mapping [7].

The interferometric coherence is computed through the well known relation:

$$\gamma = \frac{E[I_1 \cdot I_2^*]}{\sqrt{E[|I_1|^2] \cdot E[|I_2|^2]}}, \quad (1)$$

where $E[\cdot]$ and $*$ indicate the mathematical expectation and the complex conjugation operations, and I_1 and I_2 are the master and slave images, respectively.

Equation 1 requires the selection of the dimension of the mean window, which determines the performances of the estimator with respect to the scene objects. Rough surfaces (e.g., the sea surface) exhibits a stochastic backscattering, which is expected to produce incoherent signals. An insufficient dimension of the window would produce a noisy coherence, with ambiguous results. Man-made structures, instead, exhibit a high coherence. An estimator with an excessive mean window dimension could decreases its performances. In fact, urban areas are strongly inhomogeneous

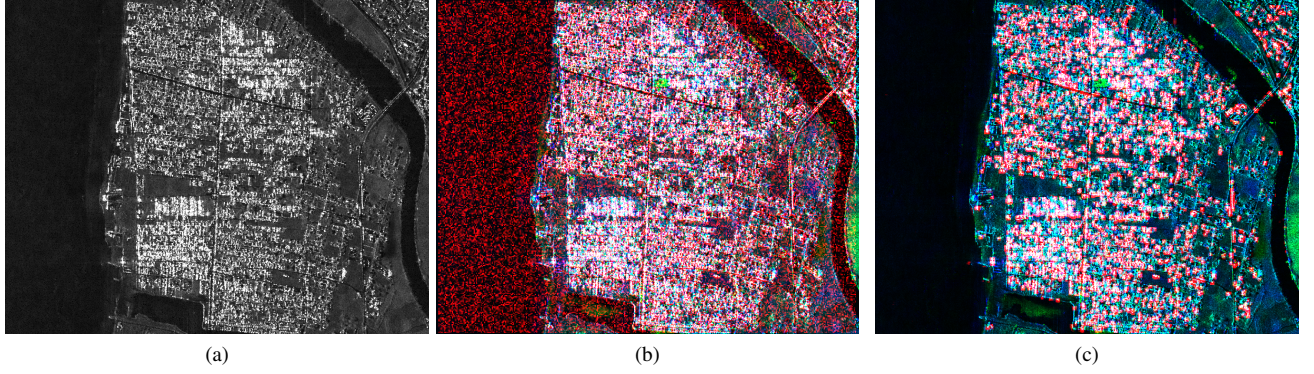


Fig. 1: Castel Volturno town (Italy): (a) De Grandi filtered intensity map and Level-1 α representations computed setting the coherence window dimension to (b) three pixels and (d) eleven pixels.

and, especially if the urban texture is not dense, the computation window could include features like vegetation or roads together with buildings, thus causing a fall in the resulting coherence.

These concepts should be more clear considering Fig. 1, in which a portion of the town of Castel Volturno (Italy) is considered). In particular, in Fig. 1a, the De Grandi-filtered intensity product is shown. In Fig. 1b a Level-1 α product of the scene is depicted. The reference image (blue band) has been acquired on December 2009, while the test image on August 2010; thus, about eight months have been passed between them. We expect that only stable features, such as buildings, keep high values of the interferometric coherence. However, the result of the estimation using a three pixels side window does not coincide with our expectation, as shown in Fig. 1b. In this representation, urban areas are well-identified, according to the rationale of the Level-1 α products. Anyway, some anomalies arise. As an example, wide portions of the sea surface exhibit high coherence values. This phenomenon, more than be physically not possible, introduces at visual level an unacceptable granularity. The use of an eleven pixel square window (see Fig. 1c) guarantees an adequate average for stochastic signals, paid with a worse estimation of deterministic signals coming from urban areas. In both cases the values for the upper and lower bounds in the quantization process have been set to 0.35 and 0.6, respectively.

Anyway, as shown in the detail of Fig. 2, the urban area is better identified using the three pixel window (see Fig. 2a). In the same way, an unacceptable speckle arises with respect to the case of the eleven pixel window (see Fig. 2b). Thus, a strategy for adapting the window dimension to the scene features is needed.

III. FEEDBACK IMPLEMENTATION

In order to adapt the characteristic of the coherence window to the scene's objects, we propose a feedback system, whose rationale is depicted in Fig. 3. The Level-1 α products

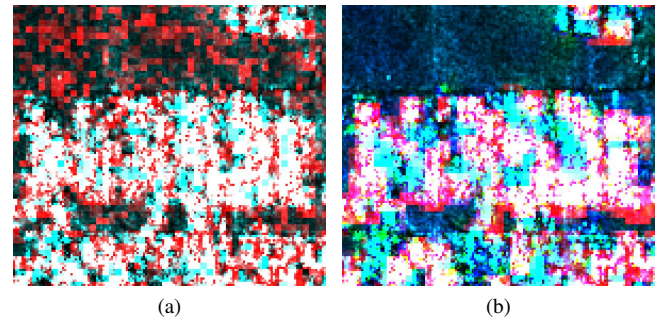


Fig. 2: Castel Volturno town (Italy), detail of a group of buildings: Level-1 α representations computed with coherence window dimension of (a) 3 pixels and (b) 11 pixels.

outputted by MAP3 [5] are used to feed an expert system driven by an *a priori* knowledge structured as follows:

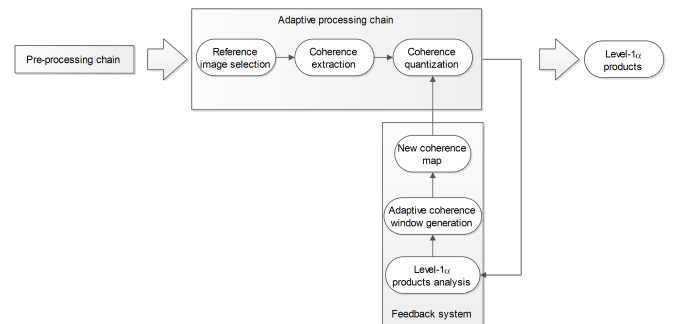


Fig. 3: MAP3 processing chain with feedback system.

- Level-1 α products analysis: the characteristics of intensity and coherence of the MAP3 output are evaluated;
- Adaptive coherence window generation: an adaptive coherence window is implemented basing on the above analysis; in such way the coherence estimate by equation (1) is optimized for the considered target since the coherence window dimension decreases in presence of highly reflective targets;

- New coherence map: the final coherence map is assembled and given as input in the quantization process for the generation of the refined Level-1 α products.

In Fig. 4a and In Fig. 4b we show a subset of the considered Level-1 α product before and after the feedback application, respectively. Qualitatively, it is remarkable how, more coherent targets are retrieved without introducing speckle. Quantitatively, the application of the feedback causes an increasing in the mean coherence for the considered subset from 0.43 (considering the 11 pixel window) to 0.5. Thus, the gain is of about 15%.

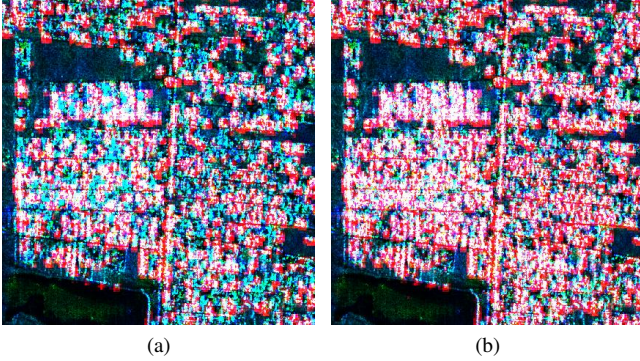


Fig. 4: A subset of the considered Level-1 α product (a) before and (b) after the feedback application.

IV. BUILDINGS EXTRACTION

Level-1 α imagery offer the possibility to quickly extract some image features through the exploitation of opportune bands combination, which is a technique widely adopted dealing with optical images.

Here, we propose a simple formulation suitable for buildings extraction:

$$BI = \frac{R \cdot G \cdot B}{255^3}, \quad BI \in [0, 1], \quad (2)$$

where the acronym BI stands for Building Index. The physical explanation is readily provided. In fact, buildings are typically highly reflective and strongly coherent. Thus, this feature has a response to the BI closer to unit. Features such as trees or layover, which can cause misclassifications in intensity-based extractions due to their high backscattering (especially when performed on high-resolution images), are typically not coherent and then pushed towards values of the BI close/equal to zero.

This formulation is particularly suited for nonexpert SAR users, which are typically reluctant in the usage of complicate techniques, and looks towards large scale applications since its negligible computational burden.

In Fig. 5 we show the result of the proposed framework. In Fig. 5a we show a buildings mask obtained via thresholding of the map resulting by the application of (2) to the original

Level-1 α product (see Fig. 4a), composed using the 11 pixels coherence window. In Fig. 5a the mask has been obtained from the refined Level-1 α product (see Fig. 4b). In both cases, we adopted a threshold value of 0.5. Finally, in Fig. 5c we overlaid the two aforementioned masks. In this representation, green pixels represent “building points” detected both before and after the refinement; red pixels indicates detections made only after the refinement; yellow pixels represent detections made only before the refinement.

Qualitatively, the application of the feedback produces a richer buildings mask, as shown in Fig. 5c.

Quantitatively, in the considered subset we found only 56 detections belonging to the no-feedback case (blue pixels in Fig. 5c). The number of new “building points” detected thanks to the feedback application is about 11800 (red pixels in Fig. 5c) while those detected in both cases is about 27700 (green pixels in Fig. 5c). Thus, the obtained gain with the feedback application is in the order of 40% with respect to the no-feedback case.

IV-A. Semantic queries

Level-1 α products are defined in a space C of 255^3 colors, and practically thus belong to the world of continuous variables. In this representation, each pixel is defined by an RGB triplet which is as such a-semantic.

Anyway, Level-1 α imagery allows for moving toward object-based analysis. The first step is the introduction of a basic semantic which could be, as suggested in [8], a color label, i.e. a categorical variable. Here, we used a Kohonen-like network [9] for re-quantize the refined Level-1 α product in N categories (i.e. colors), each of which associated with a label rather than an RGB code.

In Fig. 6a we show the result of the re-projection from the space C into 49 categories. At visual level, the reader should appreciate as very little differences arise with respect to the continuous product depicted in Fig. 4b. Operatively, the semantic query for buildings extraction equivalent to (2) is to require that “building points” exhibit white color. In such way, we are looking to points whose BI is equal to 1. In order to set the equivalent of a threshold, other colors can be considered, such as light pink.

In Fig. 6b we show the result of the application of this semantic query to the refined Level-1 α product. In this representation, green points have been detected both with the semantic query and by thresholding the map obtained by the application of (2) to the continuous product, red points have been detected only in the continuous case and yellow points only through the semantic query. We registered an agreement between the BI-derived mask and that derived via the semantic query in the order of 91%.

V. CONCLUSIONS

In this paper we presented a feedback system for the enhancement of urban areas in Level-1 α imagery. The proposed method implements an expert system based on the

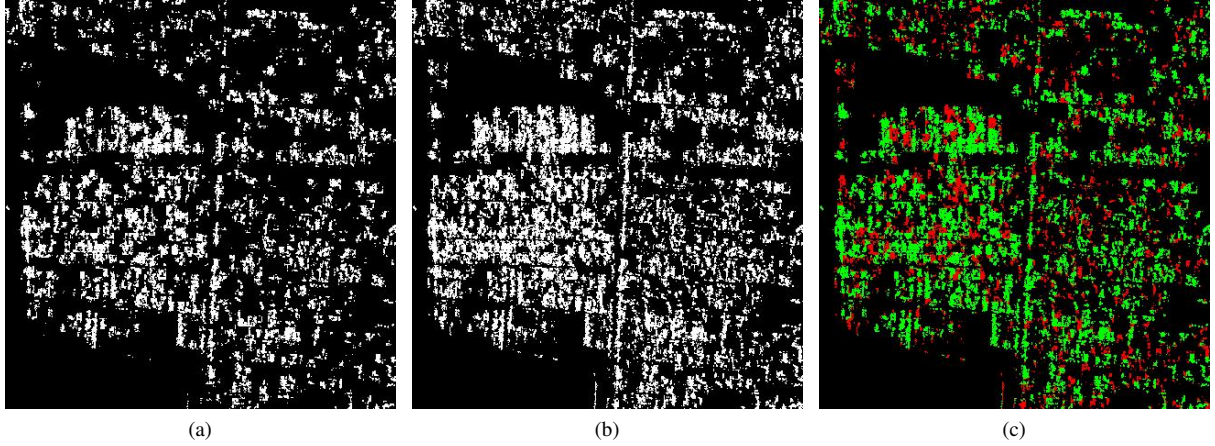


Fig. 5: Buildings masks obtained (a) before and (b) the feedback application. (c) Overlay between the two masks.

characteristic of coherence and reflectivity of the products outputted by the MAP3 framework allowing for mitigate the effects related with the choice of the coherence window dimension. The obtained results testify the goodness of the proposed method both from semantic and operative standpoints.

From the semantic standpoint, the feedback allows for an enhancement of the urban areas through the retrieval of more coherent target thanks to the adaptive coherence window.

From the operative standpoint, the proposed approach allows for a more precise mapping of the buildings in the study area through the implementation of a simple and automatic procedure based on the characteristics of Level-1 α imagery.

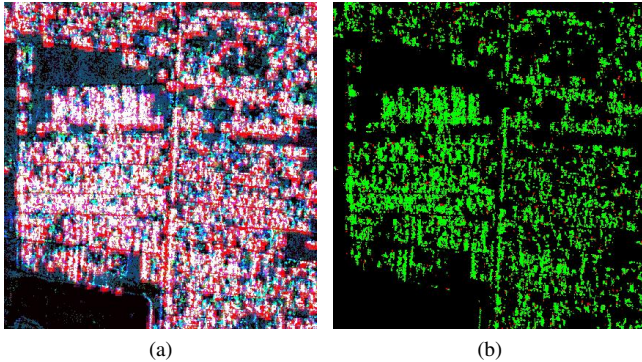


Fig. 6: (a) Refined Level-1 α re-quantized in 49 categories and (b) result of the application of the semantic query.

VI. ACKNOWLEDGMENTS

The authors thank the Italian Aerospace Research Center (CIRA) for providing the dataset used in this work.

VII. REFERENCES

- [1] A. R. Brenner and L. Roessing, "Radar Imaging of Urban Areas by Means of Very High-Resolution SAR and Interferometric SAR," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 10, pp. 2971–2982, 2008.
- [2] G. Di Martino, A. Iodice, D. Riccio, and G. Ruello, "Power Spectra of Very High Resolution SAR Amplitude Images of Urban Areas," *IEEE J. Sel. Topics Appl. Earth Observ.*, vol. 7, no. 7, pp. 2723–2731, 2014.
- [3] P. Gamba and G. Lisini, "Fast and Efficient Urban Extent Extraction Using ASAR Wide Swath Mode Data," *IEEE J. Sel. Topics Appl. Earth Observ.*, vol. 6, no. 5, pp. 2184–2195, 2013.
- [4] G. F. De Grandi, M. Leysen, J.-S. Lee, and D. Schuler, "Radar reflectivity estimation using multiple SAR scenes of the same target: technique and applications," in *IEEE Int. Geosci. Remote Sens. Symp.*, 1997, pp. 1047–1050.
- [5] D. Amitrano, G. Di Martino, A. Iodice, D. Riccio, and G. Ruello, "A New Framework for SAR Multitemporal Data RGB Representation: Rationale and Products," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 1, pp. 117–133, 2015.
- [6] R. Gaetano, D. Amitrano, G. Masi, G. Poggi, A. Verdoliva, G. Ruello, and G. Scarpa, "Exploration of Multitemporal COSMO-SkyMed Data Via Tree-Structured MRF Segmentation," *IEEE J. Sel. Topics. Appl. Earth Observ.*, vol. 7, no. 7, pp. 2763–2775, 2014.
- [7] D. Polli, F. Dell'Acqua, and P. Gamba, "Low-resolution urban area outlining in satellite SAR images," in *Joint Urban Remote Sensing Event*, 2009, pp. 1–5.
- [8] A. Baraldi, "Satellite Image Automatic Mapper (SIAM) - A Turnkey Software Executable for Automatic Near Real-Time Multi-Sensor Multi-Resolution Spectral Rule-Based Preliminary Classification of Spaceborne Multi-Spectral Images," *Recent Patents on Space Technology*, vol. 1, no. 2, pp. 81–106, 2011.
- [9] T. Kohonen, *Self-Organizing Maps*. Berlin, Heidelberg: Springer-Verlag, 2001.

[1] A. R. Brenner and L. Roessing, "Radar Imaging of Urban Areas by Means of Very High-Resolution SAR