

URBAN AREAS EXTRACTION FROM MULTITEMPORAL SAR RGB IMAGES USING INTERFEROMETRIC COHERENCE AND TEXTURAL INFORMATION

Donato Amitrano⁽¹⁾, Francesca Cecinati⁽²⁾, Gerardo Di Martino⁽¹⁾, Antonio Iodice⁽¹⁾,
Deniele Riccio⁽¹⁾, Giuseppe Ruello⁽¹⁾

⁽¹⁾ University of Napoli Federico II, Department of Electrical Engineering and Information Technology,
Email: {donato.amitrano; gerardo.dimartino; iodice; dariccio; ruello}@unina.it

⁽²⁾ University of Bristol, Department of Civil Engineering, Email: francesca.cecinati@bristol.ac.uk

ABSTRACT

In this paper, we introduce an innovative methodology for urban areas extraction based on a new class of multitemporal colour composite products. As first, we propose a method for enhancing the coherence characteristic of the image adapting the dimension of the main window to the scene target. Thus, we show how to link this property with a colour attribute, which is combined with textural information to form the final fuzzy urban areas map.

1. INTRODUCTION

Satellite imagery is a powerful tool for urban areas monitoring and for large-scale analysis concerning with planning, growth tracking, surveillance of critical structures, three-dimensional modelling, risk analysis and damage detection in case of natural disasters. In fact, in the last years, the remote sensing community dedicated a huge literature to this topic [1], [2].

In this paper, we focus on the extraction of urban areas from SAR images. 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 algorithms which preserve the spatial resolution [3]. Therefore, it is now possible to map urban areas with resolution in the order of few meters.

Here, we present an innovative methodology for mapping urban areas. As first, we present a feedback system for the MAP3 framework introduced in [4]. Then, we show how to combine a colour attribute with a textural information for feature extraction.

This methodology is based on Level-1 α imagery, which constitute a new family of colour composite multitemporal products having a semantic and informative content intermediate between the classic Level-1 and Level-2 products.

The principal characteristic of these products is that the association between the colour 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 colour 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. This problem is discussed in Section 2, in which a method for enhancing the performances of the coherence estimator is proposed. In Section 3, we show how introduce a basic semantic through Self-Organizing Maps and how to combine the image colour attribute with texture information for extracting urban areas. Conclusions are drawn at the end of the work.

2. COHERENCE FINE TUNING

2.1. The role of the mean window in the calculation of the interferometric coherence

In SAR literature, the interferometric coherence is principally linked with DEMs retrieval and deformation mapping. However, this characteristic of SAR sensors can find other applications such as image enhancement [4], classification [5] and urban areas extraction [6].

The interferometric coherence is computed by relation:

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

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

Equation (1) requires the selection of the dimension of the mean window. This parameter determines the performances of the estimator varying the scene objects. As an example, rough surfaces, such as the sea, exhibit a stochastic backscattering, which is expected to produce non-coherent signals. In this case, a small mean window could produce a noisy estimation, which is hard to be interpreted physically. On the contrary, man-made

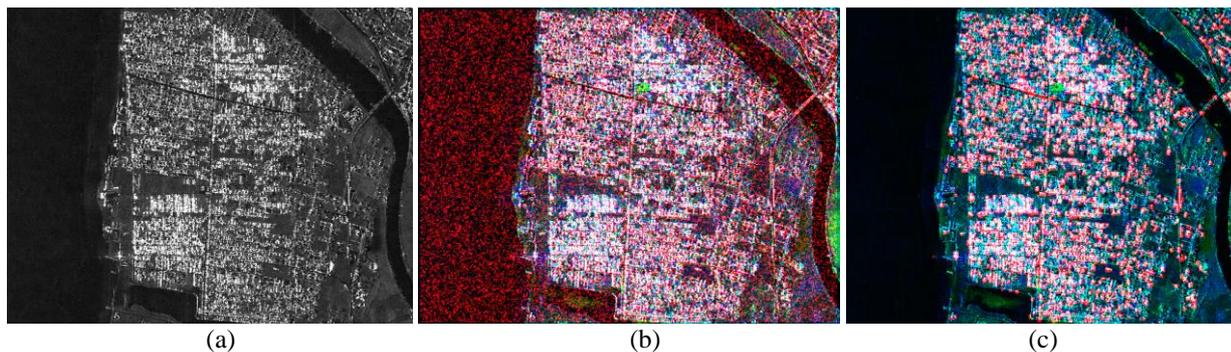


Fig. 1. Castel Volturno (Italy): (a) De Grandi filtered intensity map and Level-1 α representations computed setting the coherence window to (b) three pixels and (c) eleven pixels.

structures typically show high coherence values, but in this case a small mean window is more advisable; in fact, urban areas are strongly inhomogeneous and, especially when the texture is not dense, a large window could include feature like vegetation or roads together with buildings, thus causing a fall in the resulting coherence.

These concepts should be clarified looking at Fig. 1, in which we considered a portion of the town of Castel Volturno (Italy). 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 (green band) 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 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, but 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 pixels square window (see Fig. 1c) guarantees an adequate average for stochastic signals. On the other hand, a worse estimation of deterministic signals coming from urban areas is obtained. Thus, a strategy for adapting the window dimension to the scene features is needed.

2.2. Feedback implementation

In order to optimize the performances of the coherence estimator defined in (1) as a function of scene's objects, we propose a feedback system, whose block diagram is depicted in Fig. 2. The Level-1 α products outputted by MAP3 [4] are used to feed an expert system driven by an a priori knowledge and structured as follows:

- Level-1 α products analysis: the characteristics of intensity and coherence of the MAP3 output are evaluated;

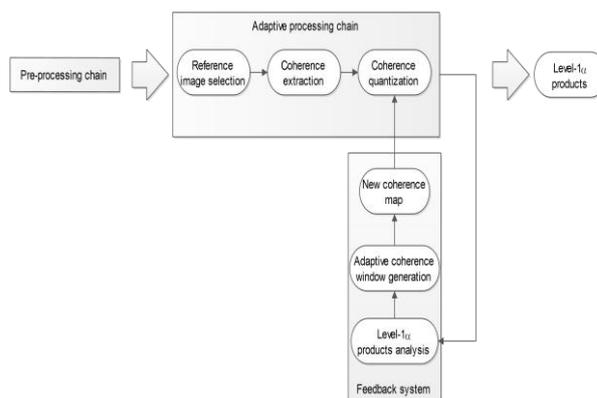


Fig. 2. MAP3 chain with coherence feedback.

- 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.

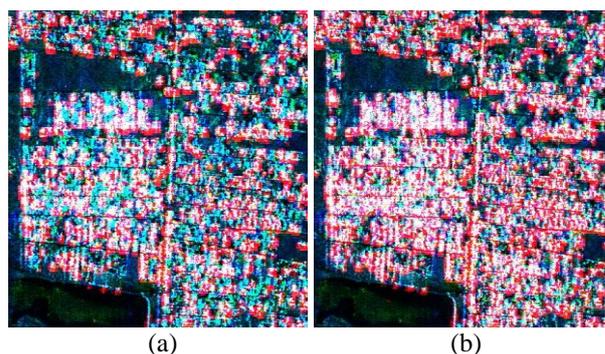


Fig. 3. A subset of the Castel Volturno city (a) before and (b) after the application of the coherence feedback.

In Fig. 3a and Fig. 3b, we show a subset of the considered Level-1 α product before and after the feedback application. The reader should easily appreciate how the representation of the urban area is qualitatively enhanced by the proposed feedback.

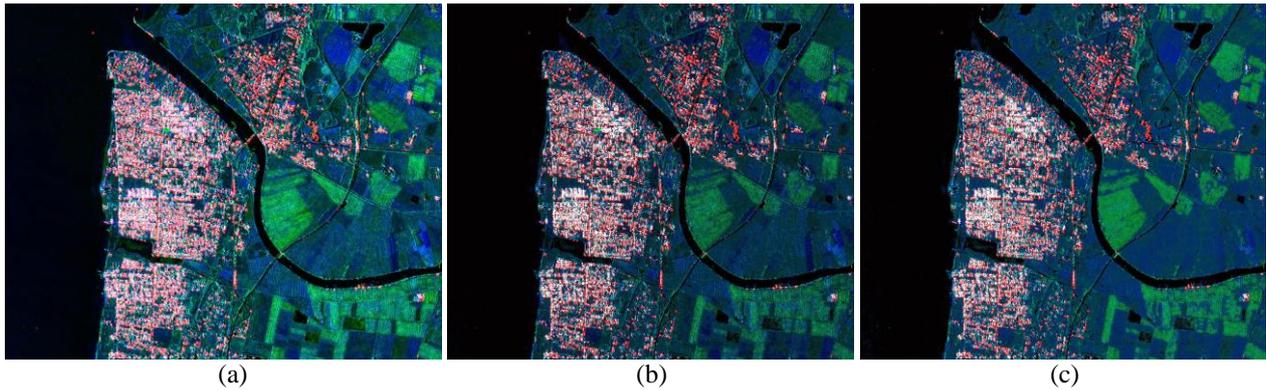


Fig. 4. Castel Volturno: (a) continuous Level-1 α product and its (b) 49-colours and (c) 22-colours categorical versions.

3. URBAN AREAS EXTRACTION

3.1. Data re-quantization

Level-1 α products are defined in a space C of 255^3 colours, which can be practically assumed as continuous. In this representation, each pixel is defined by a RGB triplet which has to be evaluated in order to extract features or image properties such as texture.

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 [7], a colour label. Here, we used a Kohonen-like network [8] for re-quantize the refined Level-1 α product in N categories (i.e. colours), each of them associated with a label rather than a RGB code.

In Fig. 4, we show the result of the above described re-quantization. In particular, in Fig. 4a, the continuous Level-1 α product is depicted, while in Fig. 4b we report its 49-colours version. The reader should easily appreciate as, qualitatively, the information conveyed by the categorical representation about urban areas and terrains is preserved.

This 49-colours representation can be further simplified by merging very similar colours, obtaining the 22-colours representation. Even in this case, no loss of information are appreciable, nevertheless the conversion from 255^3 to 22 colours.

The re-quantization procedure allows for obtaining an extremely simplified representation of the input datum in which the assigned categories can be used for feature extraction through semantic queries.

3.2. Semantic queries for urban areas extraction

According to the rationale of the Level-1 α product, urban areas are identified by pixels of white or red colours [4]. This is due to the high contribution of the interferometric coherence in the colour composite. Therefore, it is possible to extract this feature by querying the colour image attribute [6].

In Fig. 5, we show the output of the request for pixels having a colour attribute ‘‘White’’ or ‘‘Red’’.

3.3. Texture extraction

Texture is an image property which can be exploited in combination with colour attributes for extracting urban areas. In fact, this feature is usually characterized by high density of classes due to its inhomogeneity.

Here, we used a very simple rule for separating high texture areas from low texture ones. It was introduced by Nagao and Matsuyama in [9]. According to this rule, a pixel (i.e. the central pixel of a typically square moving window) is classified as high-texture area if in a window of fixed dimension N fall more than $2N$ contour points. This rule is applied on a contour map obtained through the application of the second Laplacian to the categorical image [10].

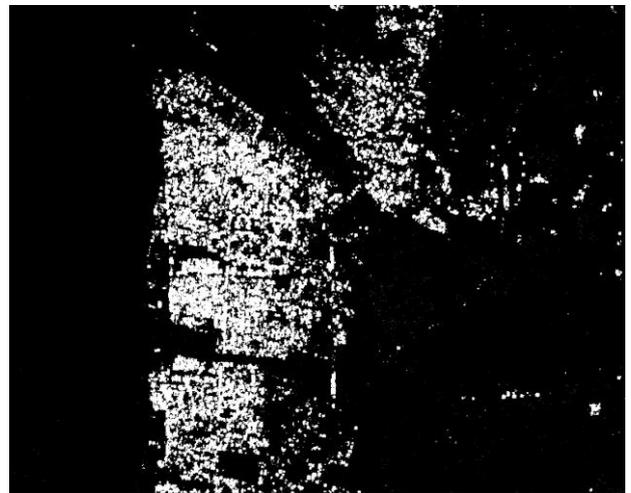


Fig 5. Result of the semantic query for ‘‘White’’ or ‘‘Red’’ pixels.

In Fig. 6, we report the texture map obtained through the above described procedure. It is remarkable that crops, as well as urban areas, at few meters resolution could exhibit high texture.

3.4. Convergence of evidences

The colour and texture evidences are now combined for obtaining the final urban areas map [11]. To this end,

we propose a fuzzy rule, in which the more weight is given to the colour attribute. In fact, as stated above, high texture can be found not only on urban areas but also on cultivated fields. Thus, we consider the texture as a weak evidence. On the contrary, the colour attribute, which is directly related to the interferometric coherence, is considered as a strong evidence. In particular, the white colour is assumed as the strongest evidence. In Fig. 7 the output urban areas map is shown. Brightest pixels are those in which there is the highest probability to find urban agglomerates.

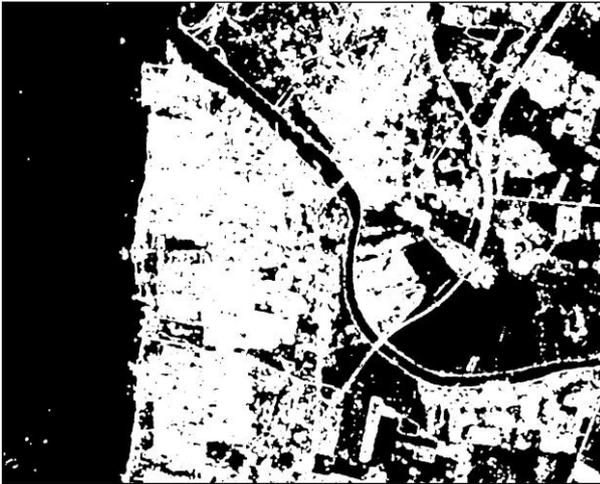


Fig 6. Texture map obtained by the application of the Nagao and Matsuyama rule.

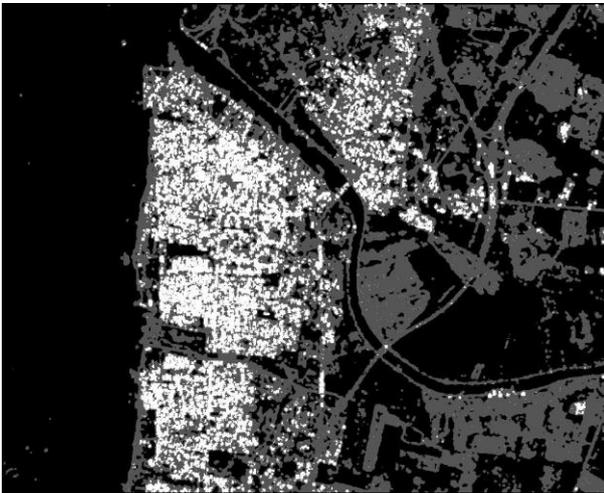


Fig 7. Final fuzzy urban areas map.

4. CONCLUSIONS

In this paper, we presented an innovative methodology for urban areas extraction, which exploits the combination of two image properties such as interferometric coherence and texture.

As first, we proposed a method for improving the performances of the coherence estimator adapting the window size to the scene target. This allowed for obtaining an enhancement of urban areas within Level-

1α imagery bringing considerable benefit in its interpretation and in the quality of the conveyed information.

Refined Level- 1α images can then be used in an object-based reasoning. We showed how to introduce a basic semantic through a Kohonen network. The colour attribute and the image texture constitute evidences that can be combined effectively for the extraction of urban areas.

5. ACKNOWLEDGMENTS

The authors thank the Italian Aerospace Research Center (CIRA) for providing the Castel Volturno multitemporal dataset.

6. REFERENCES

1. Brenner, A. R. & Roessing, L. (2008). Radar imaging of urban areas by means of very high-resolution SAR and interferometric SAR. *IEEE Trans. Geosci. Remote Sens.* **46**(10), 2971–2982.
2. Gamba, P. & Lisini, G. (2013). Fast and efficient urban extent extraction using ASAR wide swath mode data. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **6**(5), 2184–2195
3. De Grandi G. F., Leysen, M., Lee, J.-S. & Schuler, D. (1997). Radar reflectivity estimation using multiple SAR scenes of the same target: technique and applications. In Proc. IEEE International Geoscience and Remote Sensing Symposium.
4. Amitrano, D., Di Martino G., Iodice, A., Riccio D. & Ruello, G. (2015). A New Framework for SAR Multitemporal Data RGB Representation: Rationale and Products. *IEEE Trans. Geosci. Remote Sens.* **53**(1), 117–133.
5. Gaetano, R., Amitrano, D., Masi, G., Poggi, G., Ruello, G., Verdoliva, L. & Scarpa, G. (2014). Exploration of Multitemporal COSMO-SkyMed Data via Interactive Tree-Structured MRF Segmentation. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **7**(7), 2763–2775.
6. Amitrano, D., Di Martino, G., Iodice, A., Riccio, D. & Ruello, G. (2015). Urban Areas Enhancement in Multitemporal SAR RGB Images Through a Feedback System. In Proc. IEEE Joint Urban Remote Sensing Event.
7. A. Baraldi. (2015). 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 Sp. Technol.* **1**(2), 81–106.
8. Kohonen, T. *Self-Organizing Maps*. (2001) Berlin, Heidelberg: Springer-Verlag.
 9. Nagao N. & Matsuyama T. (1980). *A Structural Analysis of Complex Aerial Photographs*. New York: Plenum Press.
 10. Marr D. & Hildreth E. (1980). Theory of edge detection. *Proc. R. Soc. B.* **207**(1167). 187–217.
 11. Matsuyama T. & Hwang V. S.-S. (1990). *SIGMA - A Knowledge-Based Aerial Image Understanding System*. New York: Plenum Press.