
A Human-Centered Framework for the Understanding of Synthetic Aperture Radar Images

DONATO AMITRANO



UNIVERSITÀ DEGLI STUDI DI NAPOLI FEDERICO II

Contents

1	Introduction	5
1.1	Definitions	5
1.2	The power of images	9
1.3	General remote sensing process and proposed approach . . .	11
1.4	Book contents	14
2	Change-detection-oriented products: Level-1α	17
2.1	Introduction	17
2.2	Multitemporal SAR in remote sensing literature	20
2.3	MAP3 workflow	23
	2.3.1 Pre-processing chain	24
	2.3.2 Adaptive processing chain	31
2.4	Multitemporal Level-1 α products	33
2.5	Application of MAP3 in temperate environment	38
2.6	Conclusions and discussions	44
3	Classification-oriented products: Level-1β	53
3.1	Introduction	53
3.2	Background	56
3.3	General workflow and pre-processing chain	59
3.4	Synthesis definition	59
	3.4.1 Multitemporal analysis	60
	3.4.2 Multitemporal synthesis	61

3.5	Products and their physical interpretation	63
3.5.1	Scenario 1 - Monitoring seasonal crops in temperate environment with Mediterranean climate	63
3.5.2	Scenario 2: Level-1 β products in regions with temperate continental climate	68
3.5.3	Scenario 3: Level-1 β products in semi-arid environment	72
3.6	Conclusions and discussions	75
4	Semantic products	81
4.1	Introduction	81
4.2	Self-organizing maps	81
4.3	Semantic self-organizing maps	84
4.4	Conclusions	88
5	Applications	91
5.1	Introduction	91
5.2	Water resources management in semi-arid regions	92
5.2.1	Methodology and case study characterization	93
5.2.2	Water bodies extraction	94
5.2.3	Basins bathymetry	97
5.2.4	Summary	101
5.3	Classification	103
5.3.1	Using simple classification tools with Level-1 α images	103
5.3.2	Crops monitoring using neural networks	117
5.3.3	Land cover mapping	125
5.4	Features extraction	130
5.4.1	A seasonal water index based on Level-1 α products .	130
5.4.2	Buildings extraction	141
5.4.3	Object-based image-analysis	160
5.5	Conclusions and discussions	167
6	Conclusions	177

Chapter 1

Introduction

1.1 Definitions

Digital reconstruction of the world for data analysis, monitoring, planning and response to natural and man-induced hazards is the most important aspect of sensing. Today, we can exploit sensors acquiring data throughout the whole electromagnetic spectrum. However, when acquisitions are made beyond the visible, the human-machine interface becomes fundamental for a correct interpretation and understanding of data, especially in multi-disciplinary contexts.

Remote sensing is one of them, because it involves a large variety of professionals with different expertise and background. This technique has been defined many times in the past literature. As an example, in Lintz and Simonett (1976), one can read:

“Remote sensing is the acquisition of physical data of an object without touch or contact.”

This is a very general and basic definition, in which the emerging topic is the gathering of information at a distance. A more articulated one has been provided by White (1977):

“Remote sensing, though not precisely defined, includes all methods of obtaining pictures or other forms of electromagnetic records of the Earth’s surface from a distance, and the treatment and processing of the picture data.”

Campbell (2002) wrote:

“Remote sensing is the practice of deriving information about the Earth’s land and water surfaces using images acquired from an overhead perspective, using electromagnetic radiation in one or more regions of the electromagnetic spectrum, reflected or emitted from the Earth’s surface.”

The last two definitions highlight the topic’s most important themes: again, the gathering of information at a distance; then, the use of pictures.

In everyday life, using pictures is quite common. Today, the spread of digital technologies caused our lives to be immersed by images coming from television, newspapers, internet and so on. Indeed, our need and inclination to express feelings or to describe the surrounding environment by images of any nature is innate. Think about, as an example, to cave paintings, which represent the first proofs of human activities in many parts of the world. However, without coming back to the prehistory, we can recall a saying, attributable to an advertisement of the early XX century, to enforce the importance of images in our life. It plays as follows:

“Use a picture. It’s worth a thousand words.”

Like other sayings, it is inspired by everyday life, in which it has a clear meaning. However, we can find the same claim and a technical explanation in one of the most referenced textbooks in remote sensing. In fact, J. B. Campbell begun his *Introduction to Remote Sensing* with the following paragraph:

“A picture is worth a thousand word. Is this true, and if so, why? Pictures concisely convey information about positions, sizes, and interrelationships between objects. [...] humans possess a high level

of proficiency in deriving information from such images, (therefore) we experience little difficulty in interpreting even those scenes that are visually complex. We are so competent in such task that it is only when we attempt to replicate these capabilities using computer programs [...] that we realize how powerful our abilities are to derive this kind of intricate information.” (Campbell, 2002)

Indeed, these considerations are true for natural images, i.e. those acquired with consumer devices. Remote sensing images constitute a special class of pictures, often acquired exploiting radiation not visible to human eye, for which many issues has to be taken into account, such as (Campbell, 2002):

- Image presentation,
- Unfamiliar scales and resolutions,
- Overhead views from airborne or spaceborne platforms,
- Use of several regions of the electromagnetic spectrum.

Remote sensing images have qualities that make our ability to extract information from them not innate. Therefore, an effort is necessary to understand data and/or to make data understandable for their end-users.

In Figure 1.1 and Figure 1.2, the reader can find an example concerning these considerations. If a natural image is considered (see Figure 1.1), we have no difficult to interpret the complex patterns of color, light and dark in order to recognize the scene objects. Moreover, at a higher level, we are able to understand their arrangement and interrelationships in order to reconstruct the picture meaning not emerging by chromatic analysis.

On the other hand, if we look at a satellite image (see Figure 1.2), an effort must be made to recognize the patterns corresponding to sea surface, lakes, cities, vegetation and so on. Of course, in this case the interpretation is rather easy, because the image has been acquired in the visible spectrum. Therefore, interpretative issues are mainly related with unusual scale, resolution, and top of the roof visualization. Moreover, it is also remarkable that this representation is nowadays part of our everyday experience thanks to softwares like Google Earth. This introduce an



Figure 1.1. Humans have an innate proficiency to extract information from complex patterns of light and dark in order to recognize objects. At a higher level, we can interpret the arrangement and the interrelationships between scene's objects in order to reconstruct the picture meaning not emerging by the analysis of colors.

element which is not negligible in image interpretation, that is the habit to a certain semantic context.

However, as argued by Schroder et al. (1998), the information contained in satellite images is enormous and usually not limited in the visible spectrum. Indeed, probably the most interesting information is the one we can not see with nude eyes concerning, as an example, land surface temperature, water pollution, biomass, soil moisture, etc.

The purpose of remote sensing scientists is therefore to develop methods and techniques suitable to extract this information through the elaboration of image data. In this book, we will mainly deal with images acquired at microwave wavelength by synthetic aperture radars (SARs). Data acquired by these sensors are critical for interpretation because images are usually presented in such way they lose their principal charac-

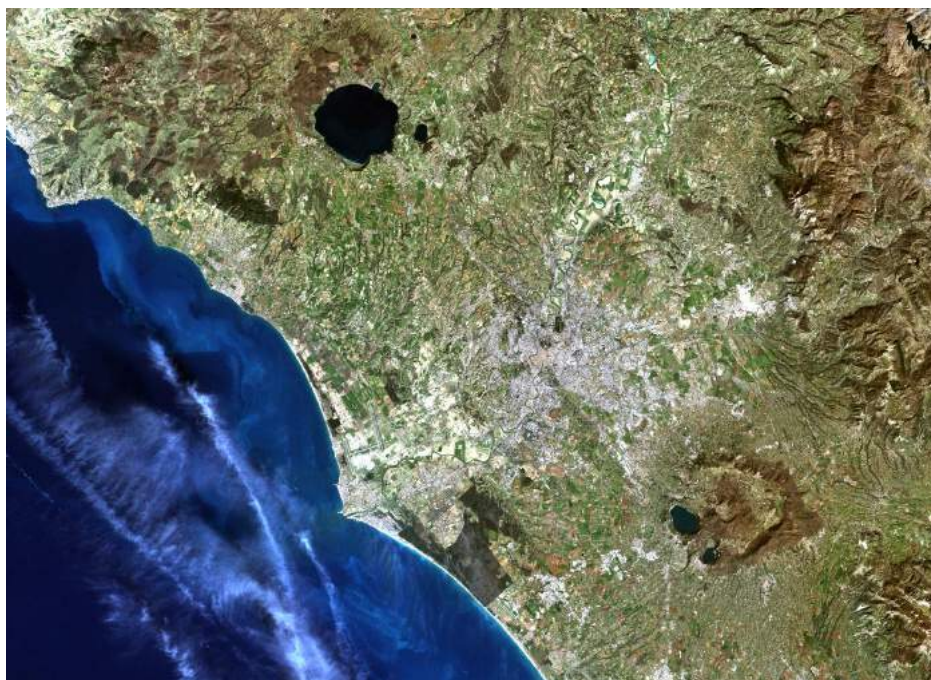


Figure 1.2. A satellite image acquired in the visible spectrum over the southern Italy. It can be interpreted roughly with limited expertise. However, the information content is much higher than that we can catch with nude eye using intuition and everyday experience. Moreover, unusual image presentation limits our understanding capabilities. Therefore, it is necessary to acquire knowledge, also using computer algorithms, to extract the meaning and the relationships between recorded patterns.

teristics, that is the one we mainly exploit in everyday life: the capability to tell a story.

1.2 The power of images

In the previous Section, we introduced some definitions of remote sensing in which the importance of images was claimed. This concept can be enforced by reading Fischer et al. (1976). They wrote:

“Remote sensing has been variously defined but basically it is the art or science of telling something about an object without touching it.”

A similar claim was also provided by Lillesand and Kiefer (1994). With respect to the definitions provided in Section 1.1, here it arises the concept of remote sensing as the art of telling a story about the imaged objects. Indeed, this is the primary objective of using images. As an example, if one looks at the picture reported in Figure 1.3a, the message it wants to convey is quite clear: it is about love, happiness, appeal of the city. These are the things one wants to feel planning a travel. In the same way, if Figure 1.3b is considered, the conveyed feelings are about family, happiness, safety, future. These are the things one wants to feel going into a bank to start a loan for buying a house. It is also remarkable that these messages are transferred to the observer just by images. In fact, if these pictures would be part of an advertisement, the observer would have no need to read its text to catch them.

One of the issues preventing the diffusion of SAR data is that the story we want to tell (and sell) to our customers (i.e. the end-users) is quite unpleasant. In order to prove this claim, consider Figure 1.4. In this picture, a SAR image in single look complex (SLC) format, that is the format provided by space agencies/data provider, is depicted. It is rather clear that the extraction of information is difficult, even for users with high expertise with radar imaging. Therefore, a non-trivial effort is necessary to make data understandable (and attractive) for their customers.

Indeed, the recent SAR literature paid few attention to the problem of data representation, privileging the extraction of information through automatic algorithms. In other words, the trend is the exclusion of users from the processing chain. This brought to a fruitful development of algorithms and techniques for data analysis. Nevertheless, this did not correspond with a widespread usage of SAR images in applications and/or in the end-user community. However, despite of the limited diffusion of SAR, the investment of the international community on these sensors is huge. As an example, the Italian SAR constellation COSMO-SkyMed represented the highest technological investment of Italian government with one billion euros. Sentinel-1 costed to the European community

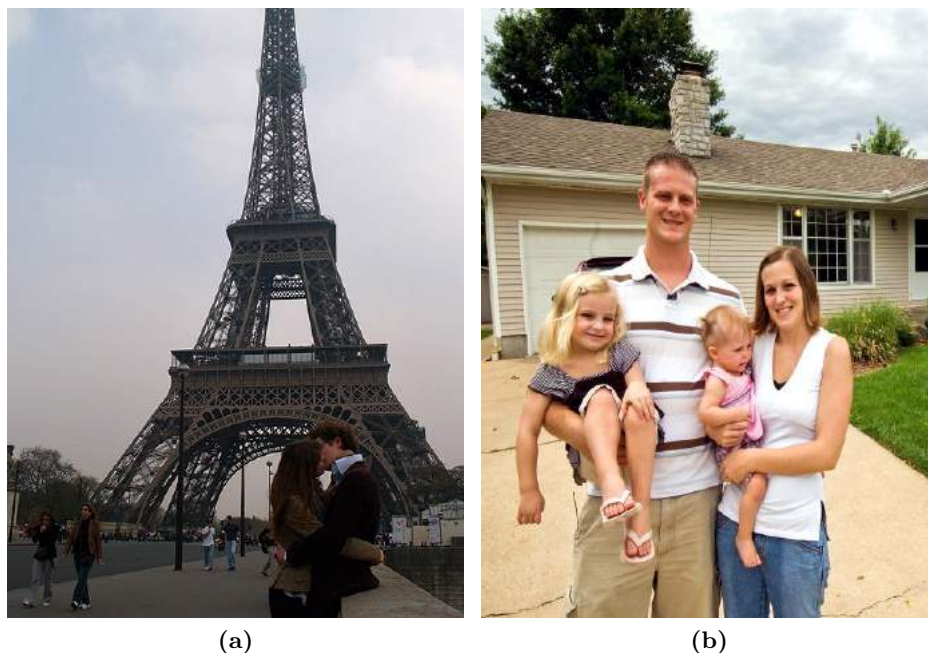


Figure 1.3. Images can concisely convey a message to the observer. (a) Love, happiness and joy are the things one wants to feel planning a travel. (b) Family, happiness, safety and future are the things one wants to feel starting a house loan.

about 300 million euros. Therefore, an effort is necessary in the SAR community to favor the diffusion of data to the general public.

1.3 General remote sensing process and proposed approach

In Figure 1.5, a general remote sensing processing chain is depicted (Campbell, 2002). It starts from *physical objects*, i.e. what scientists want to examine. Knowledge about them resides within specific disciplines, such as geology, forestry, geography, hydrology and so on.

Sensor data represent the world as filtered by a sensor, which records the electromagnetic radiation emitted or reflected from the landscape. In-

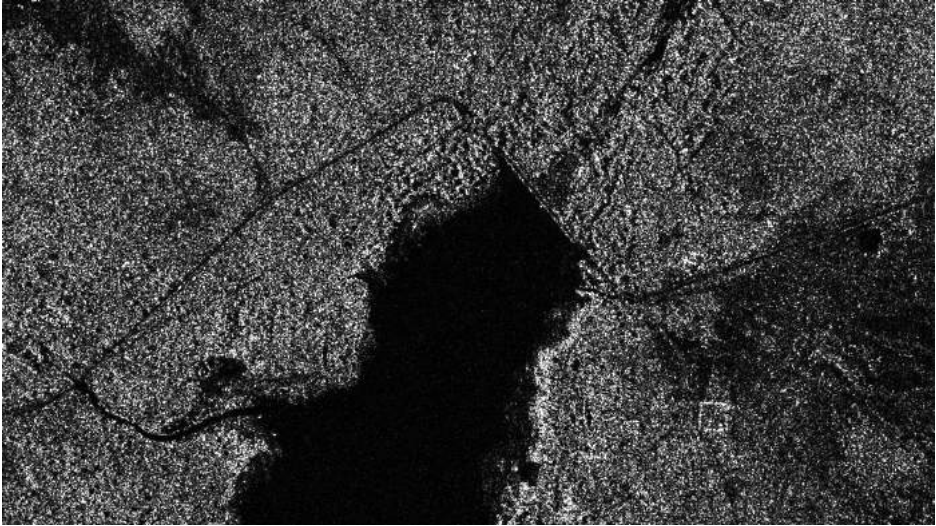


Figure 1.4. A SAR image in single look complex format, i.e. as provided by space agencies. Information extraction is quite difficult from such image, even for users with high expertise with radar imaging.

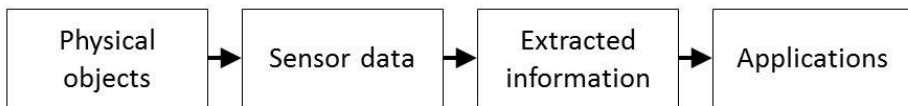


Figure 1.5. General remote sensing processing chain: from the physical object to the application.

interpretation of these data brings to the *extracted information*, and then to the *applications*, in which data coming from different sources are fused, usually in a geographic information system (GIS) environment, to get the information the analyst is looking for. Most of the modern remote sensing literature, especially dealing with SAR data, implements this passage using frameworks/algorithms in which users play (in some cases) only in the set-up phase, being often excluded from decisions. This is also due to the difficulties in data interpretation, which discourage users to try to interact with data. As a consequence, the information extraction process is

usually left to automatic algorithms, in which the user is a mere executor. This is equivalent to treat data as a simple numeric matrix, ignoring that numbers are organized to form structures, objects, patterns which deserve to be understood by the observer. We think that this is wrong for two reasons.

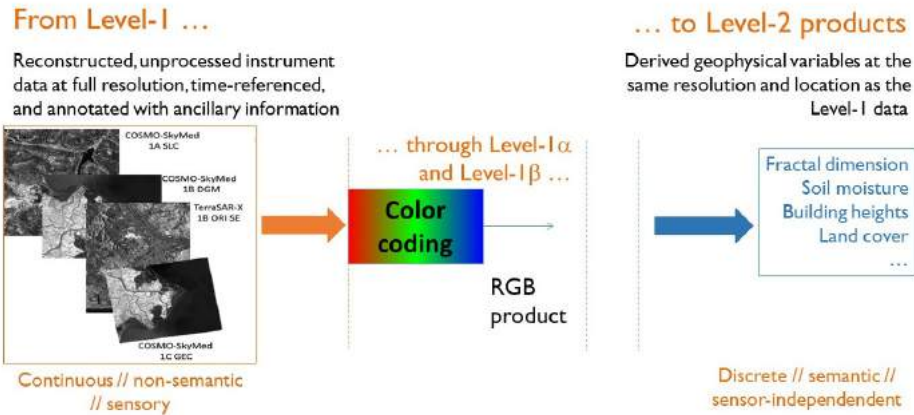


Figure 1.6. Proposed processing chain: an intermediate product is used as vehicle to obtain the physical information the analyst is looking for. This product has two main characteristics: interpretability and possibility to be processed with simple algorithms.

The first is purely philosophical: an image has the reason to exist only if there is someone that observes and tries to catch information from it. The second is more practical. In fact, in order to obtain the remotely sensed image, as an example the SAR SLC image depicted in Figure 1.4, a great effort, also economic, is needed. Starting from the project of the sensor and continuing with its launch, data acquisition and focusing, a lot of work has been done to obtain such image. Therefore it is worth to try to catch information from it even at visual level. In fact, a higher understanding of the image increases analyst's awareness of problem he/she is facing, and could be helpful in the design of the best information extraction process.

Some authors expressed the necessity to design more user-oriented and user-centered frameworks (Madhok and Landgrebe, 2002; Datcu and Seidel, 2005; Amitrano et al., 2015, 2016). The diagram depicted in Figure 1.6

provided an example of such framework, specified for SAR data. In literature, Level-1 products are those usually provided by space agencies/data provider.

In this case, an intermediate product has been introduced as a vehicle from the sensor data to the physical parameter (the extracted information in the diagram of Figure 1.5) the analyst is looking for. This product has two fundamental characteristics: i) it facilitates users' information mining thanks to a suitable RGB displaying, and ii) it is processable with simple algorithms. The purpose is to simplify the entire processing chain, loading the expertise required to process SAR data in the building of the RGB product, and delivering to users (who often have low/moderate expertise with radar imaging) a product they can easily understand and process with simple algorithmic tools.

1.4 Book contents

In this Book, we present an innovative methodology for a model-based and user-oriented representation of SAR data. In particular, in Chapter 2, a change-detection-oriented product, we named as Level-1 α product, is introduced. In Chapter 3, a classification-oriented product (we named as Level-1 β product) obtained by the fusion of N images belonging to the same time series is discussed. In Chapter 4 we will show how introduce a basic semantic in the RGB products using neural networks. Applications are discussed in Chapter 5. Conclusions are drawn at the end of the Book.

The reader interested in SAR fundamentals can refer to the past distinguished literature, such as, among the others, Franceschetti and Lanari (1999); Oliver and Quegan (1998); Franceschetti et al. (1992); Moreira et al. (2013); Bamler and Hartl (1998).

References

- Amitrano, D., Di Martino, G., Iodice, A., Riccio, D., and Ruello, G. (2015). A New Framework for SAR Multitemporal Data RGB Representation: Rationale and Products. *IEEE Transactions on Geoscience and Remote Sensing*, 53(1):117–133.
- Amitrano, D., Di Martino, G., Iodice, A., Riccio, D., and Ruello, G. (2016). An end-user-oriented framework for the classification of multitemporal SAR images. *International Journal of Remote Sensing*, 37(1):248–261.
- Bamler, R. and Hartl, P. (1998). Synthetic Aperture Radar Interferometry. *Inverse Problems*, 14:R1–R54.
- Campbell, J. B. (2002). *Introduction to Remote Sensing*. The Guilford Press, New York, third edition.
- Datcu, M. and Seidel, K. (2005). Human-Centered Concepts for Exploration and Understanding of Earth Observation Images. *IEEE Transactions on Geoscience and Remote Sensing*, 43(3):52–59.
- Fischer, W. A., Hemphill, W. R., and Kover, A. (1976). Progress in remote sensing (1972-1796). *Photogrammetria*, 32:33–72.
- Franceschetti, G. and Lanari, R. (1999). *Synthetic Aperture Radar Processing*. CRC Press, Boca Raton, FL.
- Franceschetti, G., Migliaccio, M., Riccio, D., and Schirinzi, G. (1992). SARAS: A Synthetic Aperture Radar (SAR) Raw Signal Simulator. *IEEE Transactions on Geoscience and Remote Sensing*, 30(1):110–123.
- Lillesand, T. M. and Kiefer, R. W. (1994). *Remote Sensing and Image Interpretation*. John Wiley & Sons, Inc., New York.
- Lintz, J. and Simonett, D. S. (1976). *Remote Sensing of Environment*. Addison-Wesley, Reading, MA.
- Madhok, V. and Landgrebe, D. A. (2002). A Process Model for Remote Sensing Data Analysis. *IEEE Transactions on Geoscience and Remote Sensing*, 40(3):680–686.
- Moreira, A., Prats-Iraola, P., Younis, M., Krieger, G., Hajnsek, I., and Papathanassiou, K. (2013). A tutorial on synthetic aperture radar. *IEEE Geoscience and Remote Sensing Magazine*, 1(1):6–43.
- Oliver, C. J. and Quegan, S. (1998). *Understanding Synthetic Aperture Radar Images*. Artech House, Nordwood, MA.

- Schroder, M., Rehrauer, H., Seidel, K., and Datcu, M. (1998). Spatial Information Retrieval from Remote-Sensing Images - Part II : Gibbs-Markov Random Fields. *IEEE Transactions on Geoscience and Remote Sensing*, 36(5):1446–1455.
- White, L. P. (1977). *Aerial Photography and Remote Sensing for Soil Survey*. Clarendon Press, Oxford, UK.

Change-detection-oriented products: Level-1 α

2.1 Introduction

Remote sensing technologies have a huge potential in environmental monitoring and constitute a powerful instrument for large-scale investigations and planning. In particular, SAR sensors play a decisive role in this context due to their all-weather and all-time capability. Moreover, the available short revisit time makes them very attractive to work with time series of images for the study of dynamic phenomena. The analysis of such data can be set in the frame of the image-fusion theory, introducing relevant challenges concerning the extraction, representation, and interpretation of information.

Dealing with multitemporal data, the crucial point is the definition of a reliable processing which should allow:

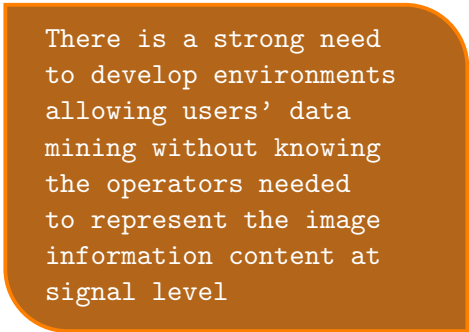
- The achievement of a set of spatially, temporally and radiometrically comparable images;
- The best level of interaction between the user and the machine during the decisional phase.

Indeed, the recent remote sensing literature mainly focused on the

first item of this list, aiming at the development of algorithms having the maximum degree of automation. Therefore, limited attention was paid to the human-machine interface. In that regard, it is quite clear that one of the issues preventing an easy interpretation of SAR data is the grayscale displaying. Therefore, a way to realize a good interface between the operator and the data he/she is analyzing is the RGB visualization. However, giving more importance to user's experience means to define frameworks in which the centrality is reserved to the analyst rather than to algorithms.

The necessity to define such general and analyst-driven frameworks has been highlighted by various authors (Almendros Jimenez et al., 2013; Madhok and Landgrebe, 2002; Datcu and Seidel, 2005; Amitrano et al., 2015b, 2016, 2015a; Gaetano et al., 2014), who expressed the need to focus the processing on users' skills, making the mathematical rigor, though of fundamental importance, secondary with respect to analysis and interpretation processes Madhok and Landgrebe (2002).

Many of the modern algorithms for SAR data processing “through the optimization of mathematical criteria are often sub-optimal in the sense that the output image is cluttered/fuzzy/noisy, visually unpleasing” (Madhok and Landgrebe, 2002) and difficult to be interpreted without a high technical expertise. Therefore, one of the aim of this Book is to emphasize



There is a strong need to develop environments allowing users' data mining without knowing the operators needed to represent the image information content at signal level

the human-computer interaction as a value, provided that the respective strengths (see Table 2.1) are enhanced (Madhok and Landgrebe, 2002; Shneiderman and Plaisant, 2005). This means that any remote sensing data processing technique, although looking for the maximum degree of automation, should never overlook users' capability of judgment and native intuition, as argued by Madhok and Landgrebe (2002) and Datcu and Seidel (2005).

There is a strong need to develop environments allowing users' data mining without “knowing the operators needed to represent the image in-

Table 2.1. Human's and machine's peculiarities in data analysis.

The human	The machine
Can draw upon experience and adapt his decision to specific situations	Can execute computationally complex algorithms and repetitive programmed actions
Can reason deductively	Can process several items simultaneously
Can generalize from observations to build analytical models or decision rules	Can implement the generalizations
Can interpret subjectively data	Generates output conforming to pre-implemented models
It is a source of knowledge	Has a short response time and high computational speed

formation content at signal level” (Datcu and Seidel, 2005). In fact, the capacity to understand data is still rather limited, because the incremental value of developing sophisticated data-analysis algorithms is counterbalanced by the difficulty in the dissemination of the knowledge required to use them effectively. Hence, it is crucial that remotely sensed data become accessible to people who may have the analytical ability but not necessarily the mathematical background to understand the physics that generated the image (Madhok and Landgrebe, 2002).

An important task is to develop new methodologies to transform the computer from a machine used for running complicate algorithms into an interface for the communication with other people and machines. In other words, one of the urgent priorities is to model the information either at a semantic level or at an electromagnetic/physical level, privileging aspects that are often post-posed to the development of the specific technique.

The idea is to introduce a general framework for building a new class of user-oriented products, characterized by high interpretability and manageability (even by non-expert users), to be added to the so-called Level-1 and Level-2 products today available from SAR data providers. This Chapter is focused on the definition of such products, we named as *Level-1 α products*. They represent an intermediate level between the already available Level-1 (see Figure 2.1) and Level-2 products. These products, unlike the

generalist Level-1, are application-oriented due to the adaptive nature of the processing that generates them.

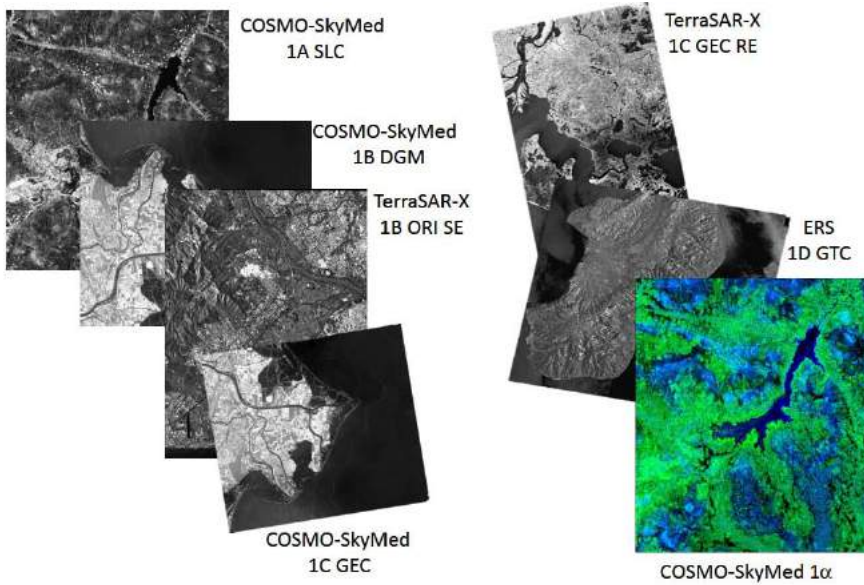


Figure 2.1. Available Level-1 products, from 1A to 1 α : SLC - Single Look Complex; DGM - Detected Ground projected Multilooked; ORI - Orthorectified Radar Image; GEC - Geocoded Ellipsoid Corrected; GTC - Geocoded Terrain Corrected; SE - Spatially Enhanced; RE - Radiometrically Enhanced.

2.2 Multitemporal SAR in remote sensing literature

The design of a new framework for multitemporal SAR data processing could not overlook the past literature. A systematic bibliographic study helped to properly collocate all the examined works, and to identify the properties and characteristics to be taken into account for the building of the new product.

Multitemporal SAR data have been widely exploited in the development of techniques oriented to change detection, classification, segmen-

tation and multisensor data fusion. Various authors used these data to validate empirical scattering models (see as an example Pulliainen et al., 1999; Kurvonen et al., 1999; Pulliainen et al., 1996).

Nagler and Rott (2000) studied snow response to estimate its water content and to extract classification maps. Soil moisture retrieval has been faced in Moran et al. (2000) and Le Hégarat-Masclé et al. (2002). Empirical scattering models for crops monitoring have been developed in Kurosu et al. (1995) and Shao et al. (2001).

In the last decade, multitemporal SAR data have been often used for classification, segmentation, change detection, and land-cover-oriented applications, even with the use of data fusion techniques. Lombardo et al. (2003) proposed a technique for the fusion of multitemporal SAR data with a single multispectral image for unsupervised classification purposes. A similar approach was proposed by Errico et al. (2015) for environmental hazards monitoring. Camps-Valls et al. (2008) also developed a multitemporal and multisource technique for classification and change detection. Davidson and Ouchi (2003) proposed to use multitemporal SAR data for extracting segmentation maps. Classification-oriented applications have been developed by Bruzzone et al. (2004) and Skriver et al. (2011).

A lot of literature can be found dealing with change detection issues, either in rural or in urban areas (Bovolo and Bruzzone, 2005; Du et al., 2012; Gamba et al., 2006; Bazi et al., 2005). Other applications involved hydrological and agricultural monitoring (Amitrano et al., 2014b,a, 2013b) and studies on natural disasters, especially regarding floods (Pulvirenti et al., 2011; Dellepiane and Angiati, 2012; Martinez and Le Toan, 2007) and fires (Siegert and Hoffmann, 1998).

From the examined literature, it arose that multitemporal SAR data processing should take into account the six key aspects listed below (see also Amitrano et al., 2015b):

- **Reproducibility:** the processing chain should be transparent in all its steps and reproducible on different datasets with slightly variations of the required parameters;
- **Automation:** the processing chain should be manageable by users with different expertise and background; it should have a limited number of free parameters to be set;

- **Adaptability:** the processing chain should be suitable for a large number of applications;
- **Reversibility:** the processing chain should preserve the electromagnetic and physics characteristics of the analyzed scene, giving rise to invertible models;
- **Robustness:** the processing chain should be implementable using data acquired by different sensors, in different modalities and in different areas and/or climatic conditions;
- **Visualization and interpretation:** the output maps should be comfortable with human view and displayed in such a way to ease their interpretation (Pohl and Van Genderen, 1998);

None of the examined works have all these properties.

Adaptability is the major limitation for semi-empirical models such as those proposed in Pullinen et al. (1999, 1996); Kurvonen et al. (1999); Kurosu et al. (1995) and Shao et al. (2001). Inversion issues are faced only by Nagler and Rott (2000); Moran et al. (2000) and Le Hégarat-Masclé et al. (2002). Automation

is a fundamental requirement for modern remote sensing data analysis. As a matter of fact, all the examined works share a high degree of automation. Reproducibility, instead, is the weakness of many modern remote sensing algorithms, which are characterized by great mathematical insight and strong parametric nature.

Indeed, one of the most annoying issues experienced by users handling SAR data is the grayscale visualization. In fact, humans have the habit to deal with color images, for which a fast comprehension can be realized (Jacobson et al., 2007). Therefore, some authors proposed the use of RGB compositions in order to favour data interpretation and users' experience. Siegert and Hoffmann (1998) used RGB images to classify the damage induced by fires. Color images have been also exploited for the generation of land cover maps (Martinez and Le Toan, 2007; U. Wegmüller and

Reproducibility, automation, adaptability, reversibility, robustness and visualization are the keywords for designing a multitemporal SAR framework

C. Werner, 1997; Amitrano et al., 2013b), water resources management (Amitrano et al., 2014b) and for floods analysis (Dellepiane and Angiati, 2012). However, a suitable processing chain for obtaining RGB images stable in the association color-object has not been defined in the analyzed literature.

2.3 MAP3 workflow

The processing chain we are going to define has been designed by taking into account of the six keywords listed in the previous Section and arose by the performed literature review. Its objective is to define a coded processing in order to obtain multitemporal RGB products stable in the color-object association and easy to be interpreted and manageable even by operators with a limited technical expertise with SAR imagery.

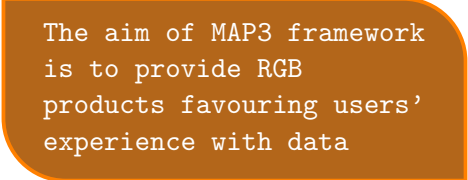
The proposed methodology can be conveniently organized in three major blocks of activities, as shown in the block diagram of Figure 2.2.

The first block of activities concerns the pre-processing chain, involving data coregistration, de-speckling and calibration.

The second block deals with the adaptive processing in which the combination of images better emphasizing the characteristics of the study area is selected.

Finally, the last block is relevant to representation and analysis.

Because of the multitemporal and adaptive nature of the proposed framework, we named it MAP3 (Multitemporal Adaptive Processing), where the number three recalls the number of blocks. As previously pointed out, the design of the blocks was guided by the six properties listed in Section 2.2. Moreover, the objective is to build a simple framework, with the minimum number of required operations to obtain the products, and characterized by the availability (in literature and in commercial and/or open source software suites) of the algorithms necessary to implement the single operations. Innovation was necessary in the cross-calibration step, for which an innovative methodology will be introduced



The aim of MAP3 framework is to provide RGB products favouring users' experience with data

in Section 2.3.1.

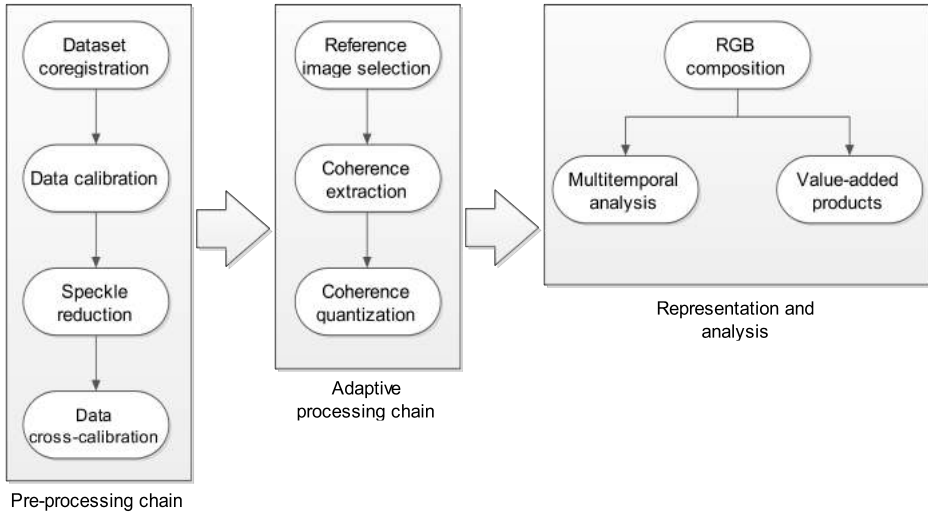


Figure 2.2. Block diagram of the MAP3 framework.

2.3.1 Pre-processing chain

Coregistration and speckle reduction

Data coregistration and speckle reduction are the first two steps of the processing chain, as shown in Figure 2.2. Data must be spatially registered with sub-pixel precision (typically 1/8 or better), as required by the evaluation of the interferometric coherence. Standard three-steps algorithms (based on orbital shifts, cross-correlation shifts and coherence shifts) usually allows for the achievement of such requirement.

Despeckling is mandatory, especially if the objective is to provide products aimed at improve the user experience with SAR data. In multitemporal analysis, it assumes a key importance, since the detrimental effects of speckle could significantly alter the discrimination of the features along the temporal axis Di Martino et al. (2014).

Working with multitemporal data offers the opportunity to exploit filters that preserve the original spatial resolution. Several studies have been proposed on this issue (De Grandi et al., 1997; Ciuc et al., 2001;

S. Quegan and T. Le Toan and J. J. Yu and F. Ribbes and N. Floury, 2000; Yu and Acton, 2002). In this work, the despeckling is performed by employing the optimal weighting multitemporal De Grandi filtering (De Grandi et al., 1997) because of its availability on commercial software and excellent declared performance in terms of Equivalent Number of Looks (ENL) on simulated SAR data.

An example of the application of this filter is shown in Figure 2.3, where we compare the SAR image of the Laaba basin (Burkina Faso) before (see Figure 2.3a) and after (see Figure 2.3b) the application of despeckling. At a visual inspection, a significant reduction of the speckle has been obtained. Quantitatively, the image after despeckling has about 12 ENL. No loss in spatial resolution or deterioration of the boundaries is appreciable. In addition, the contrast between the water surface and the surrounding environment is increased.

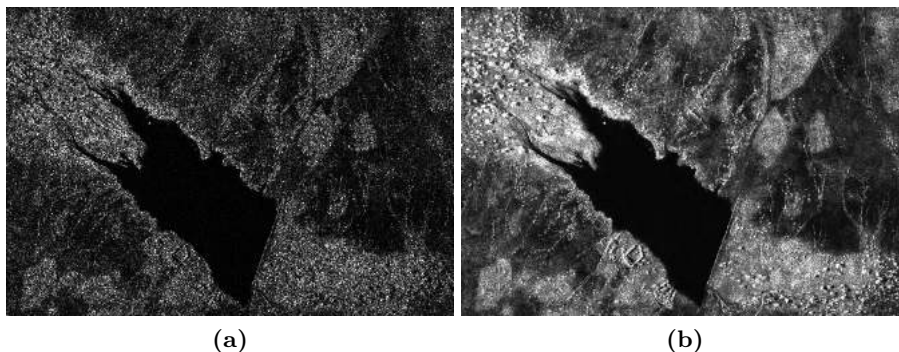


Figure 2.3. Burkina Faso, Laaba basin. SAR intensity image acquired on 09 October 2011 (a) before and (b) after the application of multitemporal De Grandi despeckling.

Data calibration

One of the major challenges in multitemporal SAR data processing is the achievement of a set of radiometrically comparable images. As an example, COSMO-SkyMed SLC hold a radiometric accuracy smaller than 1 dB (Torre et al., 2011), provided the application of a calibration coefficient that can be calculated by ancillary data (see e-geos, 2012). Note that

these coefficients refer to the SLC data, hence they should be applied to the power image before the filtering step.

Finer calibration can be obtained using passive or active instruments (S. Falzini and V. Speziale and E. De Viti, 2007; Riccio et al., 2012), if available, or using image-based techniques, such as the one proposed by D’Aria et al. (2010), based on permanent scatterers (PSs).

Indeed, the process of calibration of data is very delicate and could be affected by several sources of errors. For a complete discussion on this topic the reader should refer to Freeman (1992).

Data cross-calibration: The VALE method

In order to combine the time series images in a color composite, it is necessary that data are expressed in a common scale with a number of levels (usually 256) suitable with human visual perception.

Dellepiane and Angiati (2012) solved this problem applying a histogram clipping corresponding with the percentile that best preserves the image entropy. Successively, data are transformed from floating point value to unsigned byte datatype and equalized.

This method requires that “only marginal changes (statistically speaking) have occurred from one date to the next” Dellepiane and Angiati (2012). This is not the case in many applications, especially if agricultural cycles are under observation and/or the scene is subject to strong seasonal variations (Amitrano et al., 2014c). In addition, equalization is a nonlinear and image-dependent process, which could alter the intensity ratios between the images.

These considerations suggest that a reliable criterion of comparability between a series of images must be the conservation of the amplitude ratios. In other words, the processing must guarantee that the same reflectivity values are represented in the same histogram bin for all the images of the time series.

Assuming that data are perfectly calibrated, it is possible to establish a metric for the entire time series selecting a common threshold for clipping all the histograms. This allows for the preservation of the amplitude ratios between the images. To this end, we propose the following processing chain, which defines the variable amplitude levels equalization (VALE) method (see Amitrano et al., 2015b):

1. Let N_I be the number of images belonging to the time series and $M_i, i \in [1, 2, \dots, N_I]$ the maximum amplitude value of the i -th image; the vector $\mathbf{M} = [M_1, M_2, \dots, M_{N_I}]$ is defined.
2. Let be p the position of the minimum in \mathbf{M} , i.e. $M_p = \min\{\mathbf{M}\}$. The p -th element of the series is then the image with the minimum dynamics and is assumed as reference for establishing the amplitude threshold. This ensures that the threshold value exists in all the images of the series;
3. All the histograms of the series are clipped at an amplitude value $\hat{A} = qM_p$, with $q < 1$ evaluated on the p -th image. This ensures the best compromise between the entropy preservation and the minimization of the percentage of saturated pixels. The optimum percentile is established based on considerations on SAR image probability density function (pdf).
4. All the histograms will be clipped at the same amplitude value qM_p , corresponding to the same binsize value $\hat{b} = qM_p/255$. In this way, the metrics of the time series is established.

The selection of the best q value is now in order. As known, the pdf of SLC SAR images is strongly asymmetric and characterized by a very long right tail. Hence, the skewness parameter exhibits positive values much greater than zero. This means that most of the information is grouped in a limited percentile of the cumulative histogram. Image entropy can guide in the search of this percentile. In fact, as stated by Dellepiane and Angiati (2012), when q tends to unit, the entropy has an abrupt fall, i.e. a large number of amplitude values are compressed in a few histogram bins.

Our experiments (see Amitrano et al., 2015b), revealed that the best clipping value is $q = 0.98$. In other words, the greatest part of the information of a SAR image is contained in the 98% of the cumulative histogram. In fact, although the maximum entropy could be reached at a lower percentile, the choice of the 98th percentile allows for the saturation of a smaller percentage of pixels with negligible loss in entropy.

VALE method ensures the comparability of images by preserving amplitude ratios

In the following, we provide a brief discussion about the importance of a correct data cross-calibration for the interpretation of the dynamics of a time series.

The two SAR images depicted in Figure 2.4 have been acquired over a small basin in Burkina Faso. In this area, the landscape is extremely variable due to semi-arid climate, which is characterized by a long dry season and a short and intense rainy season, as detailed in following. Here, we present the same scene in two very different conditions. In fact, the image reported in Figure 2.4a has been acquired on August 2010, during the rainy season. The image reported in Figure 2.4b has been acquired on March 2011, during the dry season.

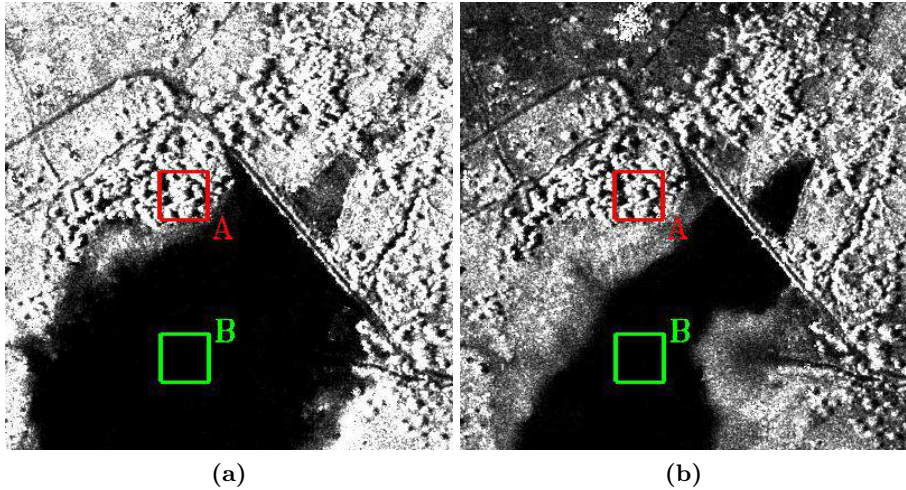


Figure 2.4. Burkina faso, Particular of the Tougou basin for the acquisitions of (a) 31/08/2010 and (b) 27/03/2011. The red rectangle (subset A) is relative to a homogeneously and persistently vegetated area; the green rectangle (subset B) is taken in a surface-water area.

On these pictures, two colored rectangles have been drawn. The red rectangle (subset A) concerns an area in which the vegetation is homogeneous and persistent all over the year. The green rectangle (subset B), concerns persistent surface water. Both the subsets are relative to a homogeneous area in which the electromagnetic response is expected to be almost stable, despite the strong variability of the surrounding scene.

We want to evaluate the effect of different cross-calibration methods on the amplitude ratios of these images. To this end, we processed the available dataset using the following chains:

- Chain A: percentile-based clipping and rescaling between $[0, 255]$, as proposed by Dellepiane and Angiati (2012). The histograms have been clipped at the same percentile of the cumulative histogram, thus resulting in a non-uniform binsize in the rescaling procedure.
- Chain B: the above described VALE method. It calls for: calibration using metadata-based coefficients, amplitude threshold-based clipping and rescaling between $[0, 255]$. Using VALE, the histograms are clipped at the same intensity value. This means that the corresponding percentile in the cumulative histogram could be different, but the binsize in the rescaling procedure is the same for all the images.

In Figure 2.5 we show the scatterplot in the plan mean vs. standard deviation for the above introduced images processed with chains A and B. We have indicated with T1 and T2 the acquisitions of August 2010 and March 2011, respectively.

It arises that the two cross-calibration methods bring to conflicting results. In fact, the class vegetation for the T1 image (August 2010, wet season) is placed to the left of the class vegetation for the T2 image (March 2011, dry season) in the case of chain A processing (proposed by Dellepiane and Angiati (2012)) and to the right in the case of chain B processing (VALE). The physically consistent result is obtained using VALE. In fact, we expect a higher electromagnetic response during the wet season due to volumetric backscattering given by the growth of the leaves. Surface water areas, unlike vegetated ones, are only slightly affected by the choice of the clipping method. Therefore, as an example, for flooding applications, as the one proposed by Dellepiane and Angiati (2012), both the chains are equally suitable.

However, VALE allows for a better separation between the classes because the centers of mass of the populations concerning water and vegetation are more distant in both the images if processed with VALE (i.e. chain B).

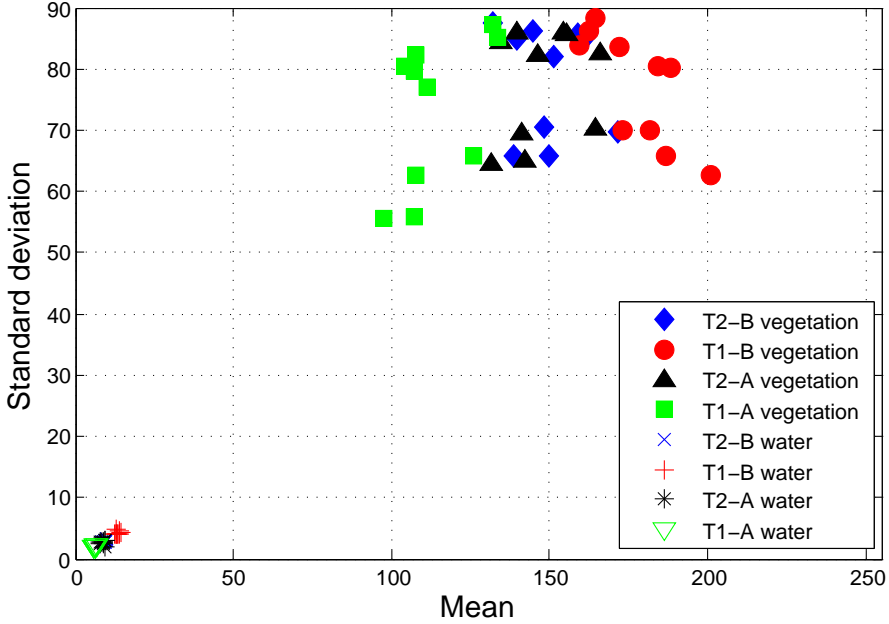


Figure 2.5. Scatterplot in the plan mean vs. standard deviation for ten sample subsets taken on the acquisition of August 2010 (T1) and March 2011 (T2). They are representative of homogeneous vegetated areas and water surface area and have been processed with percentile clipping, as proposed by Dellepiane and Angiati (2012) (chain A) and VALE (chain B).

In Table 2.2 we report the entropy analysis for the two test images. The entropy has been calculated with the following relation (Shannon, 1948):

$$H = \sum_{n=1}^N -P_n \log_2 P_n \quad (2.1)$$

where P_n is the normalized probability of the n -th quantization level of the histogram and N the total number of bins.

The analysis reported in Table 2.2 highlights that VALE ensures high values of entropy for the processed images. In fact, images processed with the VALE have an entropy value close to the theoretical maximum (i.e. 8),

Table 2.2. Images entropy as a function of the clipping percentile. In chain A (as proposed by Dellepiane and Angiati (2012)), all the histograms of the time series have been clipped at the same percentile. Using chain B (VALE) the percentile value refers only to the reference image of the time series. Image T2 refers to March 2011, while image T1 to August 2010.

Percentile	A-Dellepiane and Angiati (2012)		B-VALE	
	T1	T2	T1	T2
90 th	7.544	7.454	7.425	6.103
91 th	7.580	7.458	7.460	6.286
92 th	7.603	7.489	7.485	6.454
93 th	7.615	7.508	7.502	6.606
94 th	7.620	7.515	7.511	6.743
95 th	7.617	7.511	7.507	7.076
96 th	7.582	7.499	7.498	7.163
97 th	7.544	7.459	7.452	7.364
98 th	7.343	7.270	7.320	7.571
99 th	2.840	2.398	2.842	3.634
100 th	$9.8E^{-4}$	$1E^{-3}$	$1.3E^{-3}$	$1.5E^{-3}$

which would be assumed by an uniform distribution. This confirms VALE robustness, since it preserves the amplitude ratios between the images keeping their informative content.

2.3.2 Adaptive processing chain

Level-1 α images are application-oriented products. In other words, the best band combination and disposition in the RGB channels must be tailored on the scene and on the objective of the analyst.

Semi-arid environments constitute a particularly illustrative example of the characteristics of these products.

A semi-arid environment is characterized by extreme climatic conditions, constituted by a long dry season, in which vegetation reaches its minimum and basins are empty, and an intense wet season, in which the abundant rainfalls cause a strong growth of vegetation and crops (Amirano et al., 2014c). In Western Africa, the dry season typically goes from

October to May, while the the wet season from June to September.

In Figure 2.6 the average of monthly rainfall for three control stations located in Burkina Faso is reported. About 75% of the rain falls during the months of July, August, and September, while from November to April rainfall is lower than 20 mm per month Amitrano et al. (2014c). The response of the scene to the rain forcing depends on the land cover:

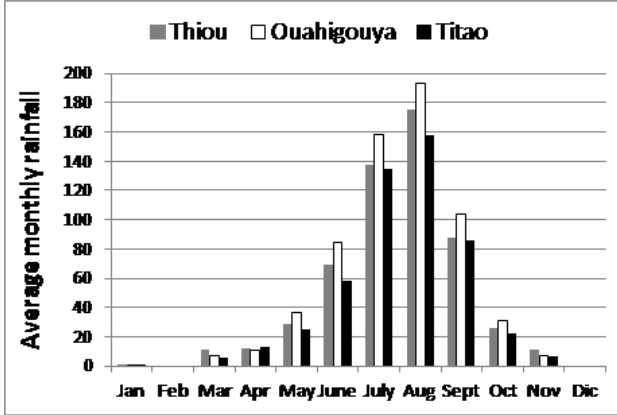


Figure 2.6. Average monthly rainfall computed over the period 1974-2011 for three control stations located in Burkina Faso. About 75% of the rain falls during the months of July, August, and September, while from November to April the rainfall is lower than 20 mm per month.

- Cultivated areas experience a growth of vegetation correlated with the rain trend. The same behavior holds for uncultivated areas that are not eroded; at the end of the dry season, most of the terrains are almost completely bare.
- Eroded areas have lost the capability of absorbing water, so that their dielectric characteristics are stable during all the year (Amitrano et al., 2013a);
- Reservoirs fill up rapidly during the wet season and empty almost completely during the dry season. This happens especially for small basins located in rural areas.

The electromagnetic reflectivity of the scene depends on these phenomena. In fact, an increased backscattering is expected on vegetated areas due to volumetric effects interaction Fung (1979). Vice versa, a low response is expected on bare soils and ponds.

Interferometric coherence extraction and quantization

As known, the interferometric coherence is given by the relation:

$$\gamma = \left| \frac{E [I_1 I_2^*]}{\sqrt{E [|I_1|^2] E [|I_2|^2]}} \right|, \quad (2.2)$$

where the operator $E[\cdot]$ denotes mathematical expectation and the apex $*$ the conjugate operation.

The interferometric coherence is an indicator of the phase stability of a target. Hence, dealing with multitemporal series, in which the images could have very large temporal baselines, we expect high values for γ only in correspondence with fixed man-made structures. In other words, when a long term interferometric coherence is computed, the result is an almost binary mask composed by “coherent” and “not coherent” points. Therefore, we impose that points whose interferometric coherence is below an assigned threshold are classified as “not coherent” and placed in the first histogram bin in the rescaling operation. For prevalently rural areas, a reasonable threshold value turns out to be in the order of 0.45.

2.4 Multitemporal Level-1 α products

The pre-processing chain described in Section 2.3.1 guarantee the comparability of intensity channels. According to the third block of Figure 2.2, it is now possible to build the multitemporal RGB Level-1 α products.

The key aspect for highlighting the desired features is the identification of a reference scenario, in which the physical parameters to be retrieved have extreme values. In semi-arid environment, this situation occurs at the end of the dry season, usually in April, when the environment is almost completely dry. Therefore, it is possible to detect changes in the landscape by comparison with respect to the reference acquisition.

Band disposition is fundamental for enhancing the feature of interests of the scene. In semi-arid environments, basically, they are vegetation, water and small human settlements. In this context, the best band disposition is the following (Amitrano et al., 2015b):

- Red band: Coherence map;
- Green band: Test image;
- Blue band: Dry season image.

In Figure 2.7 and Figure 2.8 we show two examples of the images obtained by running MAP3 on a dataset acquired by the sensor COSMO-SkyMed over a semi-arid environment, like Burkina Faso. In both images, the blue band (reference image) has been reserved to the acquisition of 28 April 2011, made at the peak of dry season.

Level-1 α products are bi-temporal images in which changes in the landscape can be identified by comparison with respect to a reference acquisition

The interpretation of the color meaning is now in order.

In Figure 2.7 the green band has been acquired on 31 August 2010, during the wet season. As previously explained, in this season the abundant rainfalls favour a strong growth of vegetation and crops. This causes an enhancement of backscattering due to volumetric effects. Therefore, vegetated areas are displayed in green.

Areas covered by seasonal surface water, i.e. present only at the time of the acquisition of the test image, appear in blue. In fact, during the wet season the ponds are filled up by intense rainfalls, the water acts as a reflector, and the corresponding backscattering (green band) is significantly weaker with respect to that of the reference dry situation (blue band). Permanent water appears as a black area within the basin, because of the low electromagnetic response in both dry and wet season. The adjective “permanent” specifies that this feature is present on the scene all over the year.

The product depicted in Figure 2.8 was obtained through a composition of two images belonging to the dry season. The test image (green band) was acquired on 27 March 2011. In this case, cyan is the dominant color because the electromagnetic response of the two scenes is almost unchanged, except for areas in which the counter-season agriculture is exercised, in the immediate neighborhood of the available water source. Therefore, those areas are characterized by green pixels. As explained

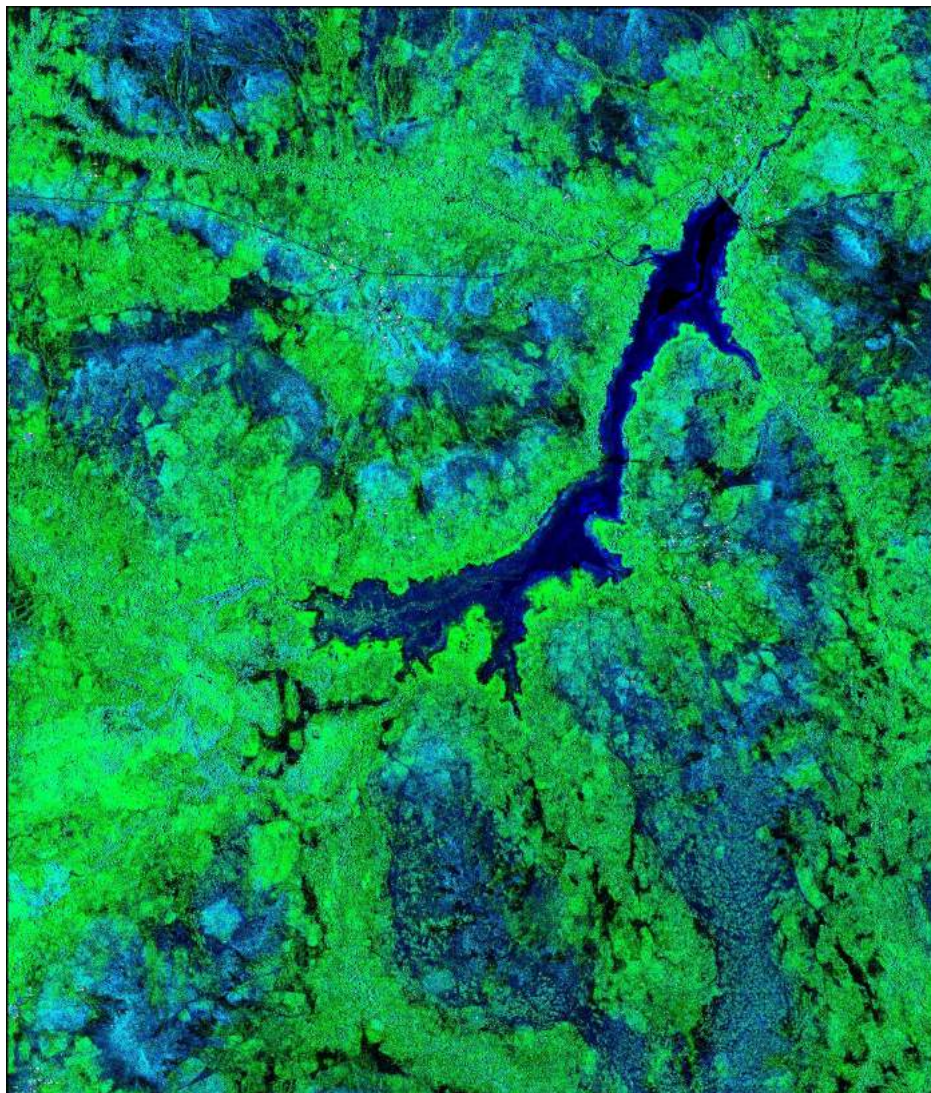


Figure 2.7. Tougou basin (Burkina Faso), COSMO-SkyMed Level-1 α product. Blue band: 28/04/2011 (dry season); Green band: 31/08/2010 (wet season); Red band: interferometric coherence. The adopted bands combination and disposition allows for the enhancement of the most important characteristics of the scene, i.e. growing vegetation and water. These features are rendered in natural colors, i.e. green and blue.

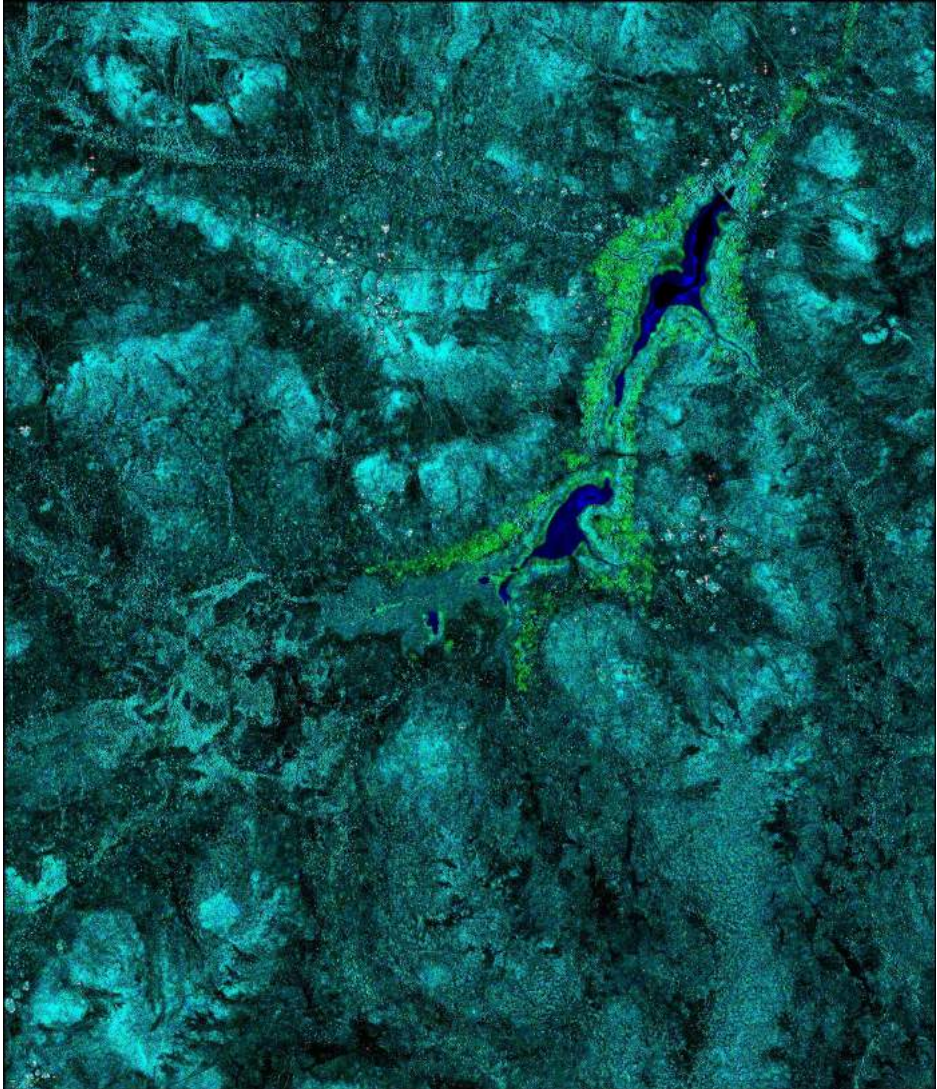


Figure 2.8. Tougou basin (Burkina Faso), COSMO-SkyMed Level-1 α product. Blue band: 28/04/2011 (dry season); Green band: 27/03/2011, (dry season); Red band: interferometric coherence. Using as reference and test images two acquisition made during the dry season, the dominant color is the cyan, meaning a balance in the surfaces backscattering. This colorimetric response is associable to bare soils.

above, seasonal water is displayed in blue color due to the dominance of terrain scattering.

Level-1 α representation also provide immediate information about the location of villages, otherwise hardly identifiable, as shown in Figure 2.9. In fact, human settlements appear as bright points because they are characterized by high backscattering all over the year and high values of the interferometric coherence, due to their phase stability. This contribution allows for separate man-made structures from highly reflective natural targets. Note that man-made targets detection can be performed using any available Level-1 α product due to the stability of such features.

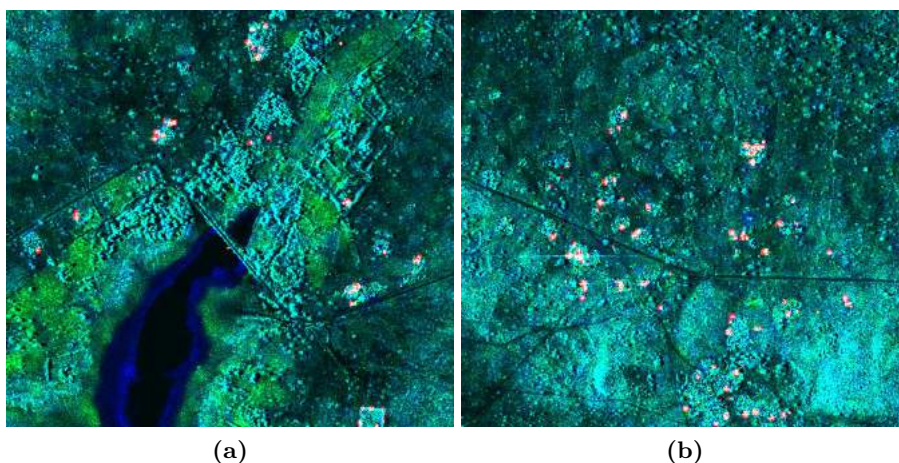


Figure 2.9. Detail concerning small human settlements. In Level-1 α imagery man-made structures are rendered in white due to the high contribution of all the bands involved in the composition.

In Figure 2.10 we provide a synthetic instrument for the interpretation of the images. Vegetation is displayed in green due to volumetric scattering; seasonal water is rendered in blue due to the dominant backscattering of terrain. The interferometric coherence introduces the third dimension, with a high contribution of the red band for man-made structures which, as a consequence, will appear as bright targets on the maps.

In Figure 2.11 the Level-1 α products for the entire analyzed time series are shown. From this representation, covering about one year and a half, the potentiality of the proposed products in the understanding of the scene

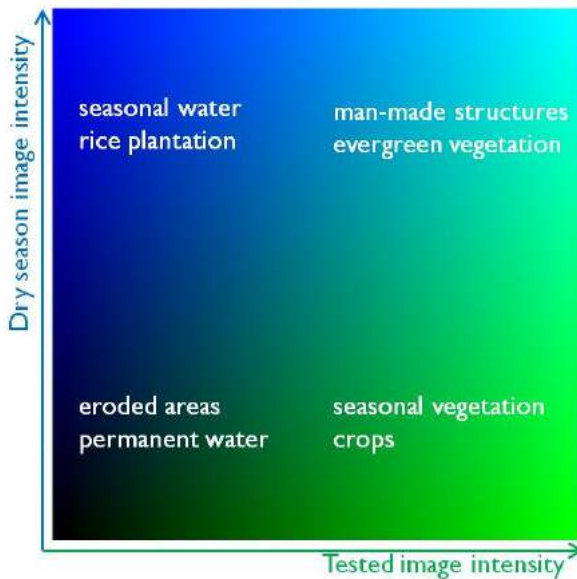


Figure 2.10. Interpretative tool for Level-1 α products in semi-arid environment (green-blue or test image-reference image plan). Vegetation is displayed in green due to volumetric scattering. Seasonal water is rendered in blue due to the dominant backscattering of terrain. Coherence introduces a contribution of the red band on man-made targets, making them distinguishable from highly reflective natural targets.

dynamics is emphasized. The cycles of vegetation and inland water can be quantitatively correlated to the rainfalls. In addition, the increasing of vegetated surfaces during the wet season and the water amount in the intakes can be easily monitored exploiting change detection with respect to the reference image.

In synthesis, the key aspect in Level-1 α product is the emphasis they provide concerning dynamic and cyclic phenomena. The great advantage introduced by this representation is the immediacy with whom the information is conveyed to the observer, especially for those features for which the association color-object is close to his/her expectation.

2.5 Application of MAP3 in temperate environment

Semi-arid environments constituted a very illustrative framework for explaining the potentiality of MAP3 processing. In this Section, we will prove the robustness of this processing chain for variations of the scene and of the climatic zone by processing a time series of images acquired by

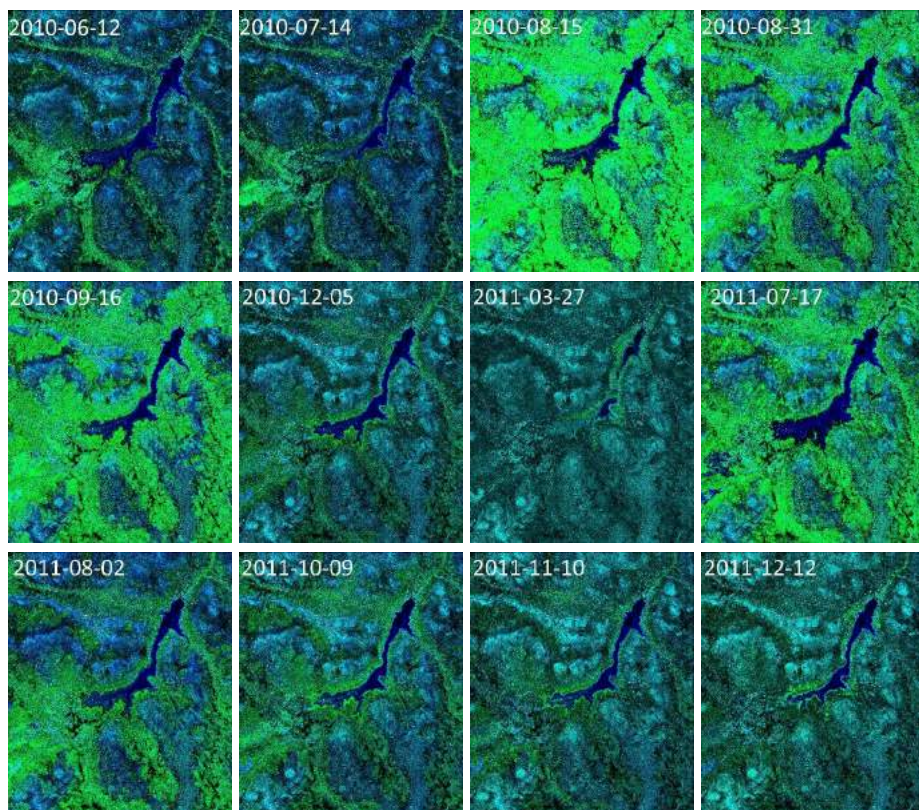


Figure 2.11. Tougou basin (Burkina Faso), multitemporal Level-1 α COSMO-SkyMed series. Reference image (blue band): 2011/04/28 (dry season). Test images from left to right and up to down: 2010/06/12, 2010/07/14, 2010/08/15, 2010/08/31, 2010/09/16, 2010/12/05, 2011/03/27, 2011/07/17, 2011/08/02, 2011/10/09, 2011/11/10, 2011/12/12.

the sensor COSMO-SkyMed on a temperate environment, like Southern Italy.

The results presented below have been obtained with the same parameters (i.e. histogram clipping parameter, interferometric coherence threshold, bands combination and their disposition) adopted for the semi-arid Burkina Faso scene. Of course, the selection of the reference acquisition is guided by different criteria, depending on the phenomena the analyst

wants to study.

The solution presented in Figure 2.12 has been built taking as reference image the acquisition of 14 December 2009, and it has been loaded on the blue band. The reference image has been acquired on 27 August 2010, and it has been loaded on the green band. In this

representation, which could be suitable, as an example, for monitoring the growth of summer crops, the colorimetric response of the features previously analyzed for the Burkina Faso scene, is stable. In particular:

The colorimetric response of Level-1 α images is stable for a given feature varying the scene and the climatic zone

- Permanent water surfaces (i.e. the sea) appear as dark areas because of the low response of all the involved channels;
- A balance in the response of the intensity channels is associable to unchanged land cover;
- Urban areas are represented as bright areas due to the high contributions of both intensity and interferometric coherence channels;
- Cultivated areas exhibit a color that depends on the state of the terrain in the two dates and on the type of cultivation. Because of the high degree of anthropization, in this case the agronomic interpretation is not straightforward and needs to be supported by specific expertise. However, vegetation growing during summer is associable to the green color due to volumetric scattering. Similarly, winter cultivations will exhibit a dominance of the blue band response.

It is worthwhile to note that in highly urbanized areas, buildings are not the only coherent targets. Indeed, buildings are among the possible coherent targets. In fact, guardrails and street lamps also appear as bright points, as shown in Figure 2.13a. Moreover, azimuth ambiguities present a coherent behavior since they are the replicas of stable targets (Di Martino et al., 2014). This is evident looking at the the sea surface in front of the city in Figure 2.12, whose detail is provided in Figure 2.13b. Another concern relevant to the interferometric coherence band is the “red crown” surrounding bright targets. This is due to the usage of a coherence window



Figure 2.12. Castel Volturno (Italy), Level-1 α product. Blue band: 14/12/2009; Green band: 27/08/2010; Red band: interferometric coherence. The colorimetric response of the product is consistent with that analyzed for semi-arid Burkina Faso. In fact, the same phenomenon/object/land cover can be associated to the same color in both the representations. As an example, green means growing vegetation, black represents permanent surface water, white stands for man-made targets and cyan indicates unchanged land cover.

of fixed dimensions (Amitrano et al., 2015c) and causes this quantity to be calculated mixing together stable and unstable features. This results in a non-null coherence value obtained on features like shadows or soils, which, being characterized by low backscattering, are rendered in rather pure red color.

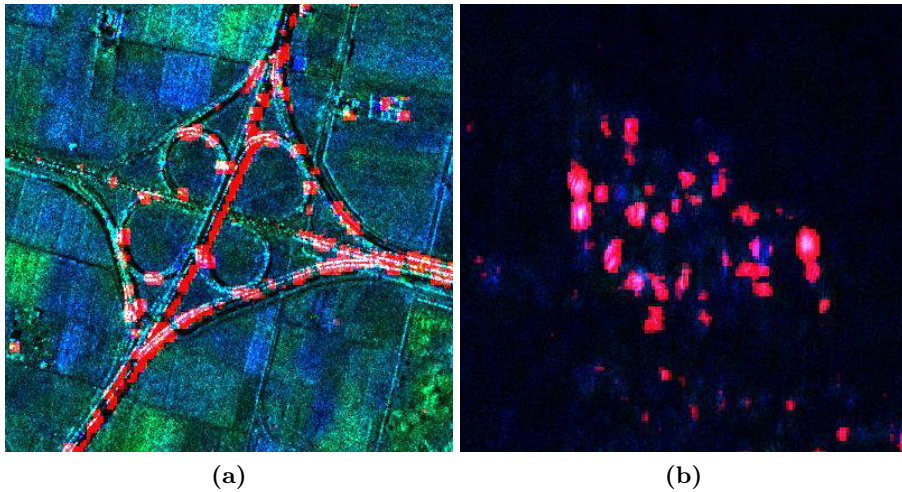


Figure 2.13. In Level-1 α imagery (a) guardrails and street lamps appear as bright targets, as well as any other stable man-made feature one can find in highly urbanized areas. Moreover, (b) azimuth ambiguities present a coherent behavior since they are the replicas of stable targets. Another concern relevant to the coherence band is the “red crown” surrounding bright targets. This is due to the usage of a coherence window of fixed dimensions and results in a non-null coherence value obtained on features like shadows or soils, which, being characterized by low backscattering, are rendered in rather pure red color.

It is possible to argue that the representation provided in Figure 2.12, although having a well-defined semantic content, is not so favourable for users, due to its high blue content. Indeed, this is true, since the blue color is present in natural images only if pictures of the sky or of the sea are shot.

In temperate environment, it is possible to obtain a representation closer to the natural color palette by exchanging the role of the reference image and coherence bands

However, in order to obtain a representation closer to the natural color palette, a very simple operation can be performed. It consists in exchanging the role played in the RGB composition by the reference image band and by the interferometric coherence band, loading the first on the red band and the latter on the blue band. The obtained result is shown in

Figure 2.14. The change in the bands disposition allows for retrieving a significant red component on terrains, resulting in a rendering closer to the natural color palette.



Figure 2.14. Castel Volturno (Italy), Level-1 α product. Blue band: interferometric coherence; Green band: 27/08/2010; Red band: 14/12/2009. In a temperate environment, the inversion of the role played in the RGB composition by the reference image band and by the interferometric coherence band allows for retrieving a significant red component on terrains, resulting in a rendering closer to the natural color palette.

However, this operation should be done only dealing with scenes acquired in temperate environments. In fact, doing this in semi-arid environment would cause seasonal water to appear in red, and this would be

very visually unpleasant and confusing.

2.6 Conclusions and discussions

Multitemporal SAR analysis is a powerful and attractive technique for environmental monitoring and planning. It introduces challenging issues concerning to data comparability, representation and information extraction.

In particular, representation issues have been widely explored in the past literature, but never coded in all its aspects. In this Chapter, the problem was formalized. The MAP3 framework, mixing state-of-art techniques with novel cross-calibration method (VALE) and innovative RGB compositions, brings to the extraction of a new family of multitemporal products, the Level-1 α products. These products are characterized by a context-adaptive nature, unlike the already available and generalist Level-1 products.

MAP3 framework, mixing state-of-art techniques with innovative cross-calibration method and RGB compositions, brings to the extraction of context-adaptive Level-1 α products, which represent an advance with respect to generalist Level-1 products

The MAP3 framework has been designed pointing to the minimization of the complexity and to the exploitation of the available algorithms. Innovation was necessary in the cross-calibration phase, where we introduced the VALE method.

The proposed framework faced the challenges introduced in Section 2.2 providing the following solutions:

- **Reproducibility:** a well-defined and transparent framework for the building of Level-1 α products has been presented;
- **Automation:** the processing chain is highly automated. Only few parameters has to be set, i.e. the coherence quantization threshold and the clipping percentile of the reference image histogram. The first parameter is selected basing on the characteristics of the scene.

In particular, for prevalently rural areas, a threshold of 0.45 is a reasonable *zero - th* quantization level. The second parameter should be determined through an entropy analysis of the reference image. However, the 98-*th* percentile of the reference image is a parameter which can suit with several applications;

- **Adaptability:** the processing chain can be adapted to different climatic zones and scenarios;
- **Reversibility:** the proposed framework ensures the preservation of the electromagnetic information thanks to an accurate calibration of the channels involved in the processing. If the inversion of the channels is needed, it should be performed before any quantization or after classification;
- **Robustness:** the processing chain restituted products in which the color-association is stable for a given feature; no variation of the required parameters was applied;
- **Visualization and interpretation:** we exploited the RGB composition of suitable cross-calibrated channels in order to provide a comfortable and interpretable experience of data. In such way, Level-1 α products realize a significant improvement in SAR images understanding and features extraction applications.

The adaptive processing chain makes possible to model the proposed framework on the specific available dataset. We showed experiments concerning images acquired in semi-arid and temperate environments. In semi-arid environment, we choose as reference band an acquisition made at the peak of the dry season and loaded it on the blue band. This choice allowed us to enhance changes in vegetation cover and water availability. In temperate environment, we showed how the change in the disposition of the reference image can be used to obtain color composites rather close to the natural color palette. The balance of the channels involved in the multitemporal processing is ensured by the novel cross-calibration method that we called VALE.

Level-1 α products shorten the distance between Level-1 products today available from data providers and Level-2 products, carrying an explicit physical information. In fact, these products can be easily managed by a large variety of users and are classification-ready, due to the clear

separation between the scene classes. These aspects, together with the interpretative immediacy given by the RGB composition, makes the proposed framework an attractive multidisciplinary tool for remote sensing based/aided applications.

The MAP3 framework is robust and flexible. In fact, on the one hand it can be replicated in different climatic zones and for different sensors; in this case, the adaptive processing can be different but the results in terms of the enhancement of the semantic content of the images hold whatever the climate condition. On the other hand the techniques embodied in each block (for example filtering or calibration) can be substituted by others proposed in literature or implemented *ex novo*. Moreover, other blocks can be inserted in the processing chain in order to adapt MAP3 to the specific application.

References

- Almendros Jimenez, J. M., Domene, L., and Piedra-Fernandez, J. A. (2013). A Framework for Ocean Satellite Image Classification Based on Ontologies. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 6(2):1048–1063.
- Amitrano, D., Cecinati, F., Di Martino, G., Iodice, A., Riccio, D., and Ruello, G. (2015a). Sentinel-1 Multitemporal SAR Products. In *IEEE International Geoscience and Remote Sensing Symposium*, pages 3973–3976.
- Amitrano, D., Ciervo, F., Di Martino, G., Papa, M. N., Iodice, A., Koussoubé, Y., Mitidieri, F., Riccio, D., and Ruello, G. (2014a). Modeling Watershed Response in Semiarid Regions with High Resolution Synthetic Aperture Radars. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 7(7):2732–2745.
- Amitrano, D., Di Martino, G., Iodice, A., Mitidieri, F., Papa, M. N., Riccio, D., and Ruello, G. (2014b). Sentinel-1 for Monitoring Reservoirs: A Performance Analysis. *Remote Sensing*, 6:10676–10693.
- Amitrano, D., Di Martino, G., Iodice, A., Riccio, D., and Ruello, G. (2015b). A New Framework for SAR Multitemporal Data RGB Representation: Rationale and Products. *IEEE Transactions on Geoscience and Remote Sensing*, 53(1):117–133.
- Amitrano, D., Di Martino, G., Iodice, A., Riccio, D., and Ruello, G. (2016). An end-user-oriented framework for the classification of multitemporal SAR images. *International Journal of Remote Sensing*, 37(1):248–261.
- Amitrano, D., Di Martino, G., Iodice, A., Riccio, D., Ruello, G., Ciervo, F., Papa, M. N., and Koussoubé, Y. (2013a). High resolution SAR for the monitoring of reservoirs sedimentation and soil erosion in semi arid regions. In *IEEE International Geoscience and Remote Sensing Symposium*, pages 911–914.
- Amitrano, D., Di Martino, G., Iodice, A., Riccio, D., Ruello, G., Ciervo, F., Papa, M. N., and Koussoubé, Y. (2013b). Synthetic Aperture Radar for Humanitarian Purposes: Products and Opportunities. In *IEEE Global Humanitarian Technology Conference*, pages 546–551.
- Amitrano, D., Di Martino, G., Iodice, A., Riccio, D., Ruello, G., Papa, M. N., Ciervo, F., and Koussoubé, Y. (2014c). Effectiveness of high-

- resolution SAR for water resource management in low-income semi-arid countries. *International Journal of Remote Sensing*, 35(1):70–88.
- Amitrano, D., G. Di Martino, Iodice, A., Riccio, D., and Ruello, G. (2015c). Urban Areas Enhancement in Multitemporal SAR RGB Images Through a Feedback System. In *IEEE Joint Urban Remote Sensing Event*.
- Bazi, Y., Bruzzone, L., and Melgani, F. (2005). An Unsupervised Approach Based on the Generalized Gaussian Model to Automatic Change Detection in Multitemporal SAR. *IEEE Transactions on Geoscience and Remote Sensing*, 43(4):874–887.
- Bovolo, F. and Bruzzone, L. (2005). A Detail-Preserving Scale-Driven Approach to Change Detection in Multitemporal SAR Images. *IEEE Transactions on Geoscience and Remote Sensing*, 43(12):2963–2972.
- Bruzzone, L., Marconcini, M., Wegmuller, U., and Wiesmann, A. (2004). An Advanced System for the Automatic Classification of Multitemporal SAR Images. *IEEE Transactions on Geoscience and Remote Sensing*, 42(6):1321–1334.
- Camps-Valls, G., Gomez-Chova, L., Munoz-Mari, J., Rojo-Alvarez, J., and Martinez-Ramon, M. (2008). Kernel-Based Framework for Multitemporal and Multisource Remote Sensing Data Classification and Change Detection. *IEEE Transactions on Geoscience and Remote Sensing*, 46(6):1822–1835.
- Ciuc, M., Bolon, P., Trouvé, E., Buzuloiu, V., and Rudant, J. P. (2001). Adaptive neighborhood speckle removal in multitemporal SAR images. *Applied Optics*, 40(32):5954–5966.
- D’Aria, D., Ferretti, A., Monti Guarnieri, A., and Tebaldini, S. (2010). SAR Calibration Aided by Permanent Scatterers. *IEEE Transactions on Geoscience and Remote Sensing*, 48(4):2076–2086.
- Datcu, M. and Seidel, K. (2005). Human-Centered Concepts for Exploration and Understanding of Earth Observation Images. *IEEE Transactions on Geoscience and Remote Sensing*, 43(3):52–59.
- Davidson, G. and Ouchi, K. (2003). Segmentation of SAR images using multitemporal information. *IEE Proceedings Radar, Sonar and Navigation*, 150(5):367–374.
- De Grandi, G. F., Leysen, M., J.-S. Lee, and Schuler, D. (1997). Radar reflectivity estimation using multiple SAR scenes of the same target:

- technique and applications. In *IEEE International Geoscience and Remote Sensing Symposium*, pages 1047–1050.
- Dellepiane, S. G. and Angiati, E. (2012). A New Method for Cross-Normalization and Multitemporal Visualization of SAR Images for the Detection of Flooded Areas. *IEEE Transactions on Geoscience and Remote Sensing*, 50(7):2765–2779.
- Di Martino, G., Iodice, A., Riccio, D., and Ruello, G. (2014). Filtering of Azimuth Ambiguity in Stripmap Synthetic Aperture Radar Images. *IEEE J. Sel. Topics Appl. Earth Observ.*, 7(9):3967–3978.
- Di Martino, G., Poderico, M., Poggi, G., Riccio, D., and Verdoliva, L. (2014). Benchmarking Framework for SAR Despeckling. *IEEE Transactions on Geoscience and Remote Sensing*, 52(3):1596–1615.
- Du, P., Liu, S., Gamba, P., Tan, K., and Xia, J. (2012). Fusion of Difference Images for Change Detection Over Urban Areas. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 5(4):1076–1086.
- e-geos (2012). COSMO-SkyMed Image Calibration. Online, last access: 12 January 2016. Retrieved from http://www.e-geos.it/products/pdf/COSMO-SkyMed-Image_Calibration.pdf.
- Errico, A., Angelino, C. V., Cicala, L., Persechino, G., Ferrara, C., Lega, M., Vallario, A., Parente, C., Masi, G., Gaetano, R., Scarpa, G., Amitrano, D., Ruello, G., Verdoliva, L., and Poggi, G. (2015). Detection of environmental hazards through the feature-based fusion of optical and SAR data: a case study in southern Italy. *International Journal of Remote Sensing*, 36(13):3345–3367.
- Freeman, A. (1992). SAR Calibration: an Overview. *IEEE Transactions on Geoscience and Remote Sensing*, 30(6):1107–1121.
- Fung, A. K. (1979). Scattering from a Vegetation Layer. *IEEE Transactions on Geoscience Electronics*, 17(1):1–6.
- Gaetano, R., Amitrano, D., Masi, G., Poggi, G., Verdoliva, A., Ruello, G., and Scarpa, G. (2014). Exploration of Multitemporal COSMO-SkyMed Data Via Tree-Structured MRF Segmentation. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 7(7):2763–2775.
- Gamba, P., Dell’Acqua, F., and Lisini, G. (2006). Change Detection of Multitemporal SAR Data in Urban Areas Combining Feature-Based and

- Pixel-Based Techniques. *IEEE Transactions on Geoscience and Remote Sensing*, 44(10):2820–2827.
- Jacobson, N. P., Gupta, M. R., and Cole, J. B. (2007). Linear Fusion of Image Sets for Display. *IEEE Transactions on Geoscience and Remote Sensing*, 45(10):3277–3288.
- Kurosu, T., Fujita, M., and Chiba, K. (1995). Monitoring of Rice Crop Growth from Space Using the ERS-1 C-band SAR. *IEEE Transactions on Geoscience and Remote Sensing*, 33(4):1092–1096.
- Kurvonen, L., Pulliainen, J. T., and Hallikainen, M. T. (1999). Retrieval of Biomass in Boreal Forests from Multitemporal ERS-1 and JERS-1 SAR Images. *IEEE Transactions on Geoscience and Remote Sensing*, 37(1):198–205.
- Le Hégarat-Masclé, S., Zribi, M., Alem, F., Weisse, A., and Loumagne, C. (2002). Soil Moisture Estimation from ERS/SAR Data: Toward an Operational Methodology. *IEEE Transactions on Geoscience and Remote Sensing*, 40(12):2647–2658.
- Lombardo, P., Oliver, C. J., T. Macrí Pellizzeri, and Meloni, M. (2003). A New Maximum-Likelihood Joint Segmentation Technique for Multitemporal SAR and Multiband Optical Images. *IEEE Transactions on Geoscience and Remote Sensing*, 41(11):2500–2518.
- Madhok, V. and Landgrebe, D. A. (2002). A Process Model for Remote Sensing Data Analysis. *IEEE Transactions on Geoscience and Remote Sensing*, 40(3):680–686.
- Martinez, J. and Le Toan, T. (2007). Mapping of flood dynamics and spatial distribution of vegetation in the Amazon floodplain using multitemporal SAR data. *Remote Sensing of Environment*, 108(3):209–223.
- Moran, M. S., Hymer, D. C., Qi, J., and Sano, E. E. (2000). Soil moisture evaluation using multi-temporal synthetic aperture radar (SAR) in semiarid rangeland. *Agricultural and Forest Meteorology*, 105:69–80.
- Nagler, T. and Rott, H. (2000). Retrieval of Wet Snow by Mean of Multitemporal SAR Data. *IEEE Transactions on Geoscience and Remote Sensing*, 38(2):754–765.
- Pohl, C. and Van Genderen, J. L. (1998). Multisensor image fusion in remote sensing: Concepts, methods and applications. *International Journal of Remote Sensing*, 19(5):823–854.
- Pulliainen, J. T., Kurvonen, L., and Hallikainen, M. T. (1999). Multitem-

- poral Behavior of L- and C- Band SAR Observations of Boreal Forests. *IEEE Transactions on Geoscience and Remote Sensing*, 37(2):927–937.
- Pulliainen, J. T., Mikkela, P. J., Hallikainen, M. T., and Ikonen, J.-P. (1996). Seasonal Dynamics of C-Band Backscatter of Boreal Forests with Applications to Biomass and Soil Moisture Estimation. *IEEE Transactions on Geoscience and Remote Sensing*, 34(3):758–770.
- Pulvirenti, L., Chini, M., Pierdicca, N., Guerriero, L., and Ferrazzoli, P. (2011). Flood monitoring using multi-temporal COSMO-SkyMed data: Image segmentation and signature interpretation. *Remote Sensing of Environment*, 115:990–1002.
- Riccio, D., Di Martino, G., Iodice, A., Koussoubé, Y., Pinelli, A. D., and Ruello, G. (2012). Calibration of COSMO-SkyMed SAR data for hydrology applications. In *IEEE International Geoscience and Remote Sensing Symposium*, pages 844–847.
- S. Falzini and V. Speziale and E. De Viti (2007). COSMO-SkyMed active calibrator: A sophisticated tool for SAR image calibration. In *IEEE International Geoscience and Remote Sensing Symposium*, pages 1577–1580.
- S. Quegan and T. Le Toan and J. J. Yu and F. Ribbes and N. Floury (2000). Multitemporal ERS SAR analysis applied to forest mapping. *IEEE Transactions on Geoscience and Remote Sensing*, 38(2):741–753.
- Shannon, C. E. (1948). A mathematical theory of communication. *Bell System Technical Journal*, 27:379–423.
- Shao, Y., Fan, X., Liu, H., Xiao, J., Ross, S., Brisco, B., Brown, R., and Staples, G. (2001). Rice monitoring and production estimation using multitemporal RADARSAT. *Remote Sensing of Environment*, 76(3):310–325.
- Shneiderman, B. and Plaisant, C. (2005). *Designing the user interface: strategies for effective human-computer interaction*. Paerson Education, Inc., Boston.
- Siegert, F. and Hoffmann, A. A. (1998). The 1998 Forest Fires in East Kalimantan (Indonesia): A Quantitative Evaluation Using High Resolution, Multitemporal ERS-2 SAR Images and NOAA-AVHRR Hotspot Data. *Remote Sensing of Environment*, 72(1):64–77.
- Skriver, H., Mattia, F., Satalino, G., Balenzano, A., Pauwels, V., Verhoest, N., and Davidson, M. (2011). Crop Classification Using Short-Revisit

- Multitemporal SAR Data. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 4(2):423–431.
- Torre, A., Calabrese, D., and Porfilio, M. (2011). COSMO-SkyMed: Image quality achievements. In *Proceedings of the 5th International Conference on Recent Advances in Space Technologies*, pages 861–864.
- U. Wegmüller and C. Werner (1997). Retrieval of Vegetation Parameters with SAR Interferometry. *IEEE Transactions on Geoscience and Remote Sensing*, 35(1):18–24.
- Yu, Y. and Acton, S. (2002). Speckle Reducing Anisotropic Diffusion. *IEEE Transactions on Image Processing*, 11(11):1260–1270.

Chapter 3

Classification-oriented products: Level-1 β

3.1 Introduction

The use of SAR data has been so far limited because the information content held in the images rarely emerges without a technical expertise. As argued in Chapter 2, the development of new products devoted to provide user-friendly representations of the physical information is a necessary condition for the full exploitation of SAR sensors.

High-level processing is a key task for enhancing the interpretability and the underlying information content of remote sensing images. This problem assumes different connotations depending on the sensor acquiring the data. As an example, in hyperspectral imagery, principal components analysis is probably the most popular linear projection method to find the more suitable representation for human users. However, it presents several drawbacks (Jacobson and Gupta, 2005). Therefore, refined methods have been proposed in literature to overcome its limitations (Jacobson and Gupta, 2005; Jacobson et al., 2007; Wilson et al., 1997) and to make possible the displaying of the information contained in N channels (where N can be in the order of few hundred) on a standard tristimulus RGB device through a consistent dimensionality reduction.

Dealing with SAR images, data interpretability problems are prin-

cipally related with: i) the geometrical distortion induced by the side-looking acquisition mode, ii) the scattering mechanism resulting from the interaction between the incident electromagnetic field and the physical surfaces and iii) the gray-scale displaying, at least in non-polarimetric modality. In fact, humans can easily interpret color images, which lead to fast searching and comprehension of data (Jacobson et al., 2007).

Multiple acquisitions of the same scene represent a reasonable mean for increasing the dimensionality of SAR data through the combination of information collected along the time axis. The fusion problem is brilliantly focused by Pohl and Van Genderen (1998), where the authors suggested a series of questions an image fusion technique should answer. Indeed, these questions refer to the problem of multisensor data fusion. Therefore, we slightly reworked the aforementioned framework as follows:

1. What is the objective/application of the user?
2. What are the necessary pre-processing steps involved?
3. Which combination of the data is the most successful?
4. Which is the “best” technique to fuse these data for that particular application?

Our objective is the definition of a new class of RGB SAR products exploiting multitemporal. This answers to the first question, establishing the objective of the whole processing to data representation.

Dealing with multitemporal SAR data, several valuable solutions can be found in the very last literature concerning, as an example, forest monitoring (Dong et al., 2015), flooding events (Dumitru et al., 2015) and land cover mapping (Antropov et al., 2014). Other, more general, suggestions and answers can be found in the past distinguished literature (Richards and Jia, 2006; Franceschetti and Lanari, 1999; Campbell, 2002). Here, we want to emphasize the usefulness of an appropriate and comfortable representation. In fact, it makes data more attractive and visual data mining easier, allowing for a better design of the information extraction process. In fact, a better knowledge/understanding of the scene implies a higher awareness of the process necessary for features extraction.

The above listed questions can be summarized in the following one (see Schroder et al., 1998):

5. How the enormous amount of information contained in remotely-sensed images can be extracted?

We will not provide a complete answer to this question. In fact, we mainly deal with a part of the whole problem, that is data representation. However, as stated in Sonka et al. (1999),

“A good knowledge representation design is the most important part of solving the understanding problem.”

Sharing this position, in Chapter 2, we introduced an intermediate product level, i.e. the Level-1 α , between the classic Level-1 and Level-2 products. These products are change-detection oriented because they are bi-temporal, although obtained through a fully multitemporal processing. However, Level-1 α processing is not the only solution to combine multi-temporal information in RGB images. In this Chapter, we linger on how to provide an unique representation of a time series, enhancing the dynamics of the scene in order to provide a classification-oriented product. We will refer to this new class of products as Level-1 β products (Amitrano et al., 2015a).

As for Level-1 α processing, the objective is twofold: i) to favour interpretation and visual data mining; ii) to provide products suitable to be processed with standard algorithms, which are the most popular in the end-users community. These characteristics lower the expertise required to handle data and we think this is crucial, especially if the operator has a limited experience with SAR. In such way, the retrieval of the physical information (which is usually seen as Level-2 product) is simplified since it takes place starting from a level which is higher than 1. As an example, if the objective is crop classification, the discrimination between summery and wintery cultivations is a good starting point. This is not an information of Level-2 but is more than a Level-1 and could be helpful for moving toward superior levels.

A good knowledge
representation design
is the most important
part of solving the
understanding problem

The design of a suitable knowledge representation, especially dealing

with data acquired beyond the visible spectrum, requires a certain confidence with both the *sensed world*, i.e. the world as seen by the sensor, and of the *perceived world*, i.e. the world as understood by users. In other words, in the design of a processing chain answering to the four questions above introduced, two kinds of knowledge coexist: an objective knowledge, given by the electromagnetic models, and a subjective knowledge, related to the users.

3.2 Background

Mendel (1995) argued that two distinct forms of knowledge exist: *objective knowledge* and *subjective knowledge*. SAR images interpretation is mainly a matter of subjective knowledge. An expert SAR user has more possibilities to successfully interpret data thanks to his/her habit to that semantic context. A basic representation of SAR data usually makes difficult even at visual level the extraction of information for many multidisciplinary users, which exactly for this reason often turn toward other sources of data which they can more easily manage, or leave the process of extraction of information to automatic algorithms.

This brings to the concept of *emergent semantics*. According Santini et al. (2001), the informative content, or the *meaning*, of an image is assumed to be not an intrinsic property but an emergent characteristic through the interaction with users. Therefore:

- The meaning of the image becomes *contextual*, i.e. depends on the condition in which a particular query is made and on the user performing it;
- The knowledge about the image is assessed by the user experience which, in this context, is built also by exploration/inspection of data.

From these considerations, it arises that the center of an image understanding process should be the user, rather than an algorithm. This appears more evident if the data analysis process is approached from a semiotic viewpoint.

In Figure 3.1 we show the Peirce's semiotic triangle (Atkin, 2013). It schematizes how a concept is formed in our mind. Pierce claims that this

mechanism consists of three inter-related elements: a sign, an object and an interpretant. The sign is everything can be perceived. In our case it is the image or the electromagnetic response of scenes' objects. The object is what the sign concerns with and it exists independently on the sign. It consists in the physical object on the scene. Finally, the interpretant is the understanding that the observer reaches of some sign/object relation. The interpretant is in turn a sign which refers to a further object and interpretant, giving rise to the concept of *infinite semiosis* (Nöth, 1995), in which signs are step by step enriched in their informative content by the previous interpretant.

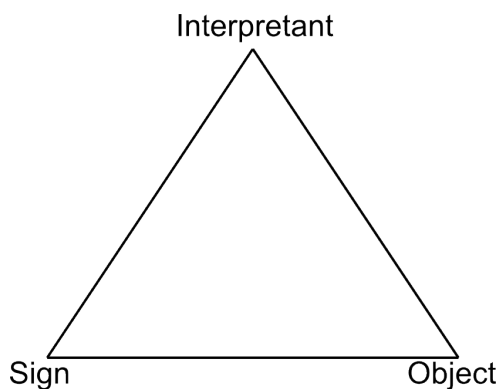


Figure 3.1. Peirce's semiotic triangle, depicting how a concept is formed in our mind. It consists of three inter-related elements: a sign, an object and an interpretant. The sign is everything can be perceived. In our case it is the image or the electromagnetic response of scenes' objects. The object is what the sign concerns with and it exists independently on the sign. It consists in the physical object on the scene. Finally, the interpretant is the understanding that the observer reaches of some sign/object relation.

This interpretation of the cognitive process highlights how the machine can be only a support (although in many cases indispensable) of the human activity. In fact, it can not participate to the cycle depicted in Figure 3.1, apart in helping a better formation of the interpretant through its capacity of executing complex tasks in a short time (see also Madhok and Landgrebe, 2002, and Table 2.1).

In Chapter 2, MAP3 gave an example of the potentialities of such an user-driven framework. The mean we used to concretize the human-machine interaction was the fusion of the information contained in a N SAR images in a series of $N-1$ bi-temporal RGB products. Here the fusion has the objective to provide an unique RGB frame (we called Level-1 β

product) representative of the scene dynamics. Therefore, we will speak about a *synthesis* rather than a fusion. In fact, this word can assume the twofold meaning of *combination of separate things into a complete whole* and of *overviewing* a series of information which can be useful as a first step for the understanding of dynamic phenomena, segmentation and mining.

In a context in which the automatic extraction of information from large databases is still rather limited, Level-1 α and Level-1 β products, by lowering the expertise level required to correctly interpret data, aim to help the human photo-interpreter, which is in many cases the figure designated for assigning the scene semantic. In fact, a suitable representation is a powerful mean for the right association between the signs and the objects through the semantic emergent from the interaction with data.

Level-1 α and Level-1 β products are “rather general purpose” data representations because the quantity (and the quality) of information one can mine from data is always dependent on the type of representation chosen for it. In fact, as argued by Marr (1982)

“A representation is a formal system for making explicit certain entities of types of information.”

In other words, data can be represented in different ways, and the choice made greatly affects the information extraction process, since it makes explicit some information at the expense of other that is pushed back and could be quite hard to recover (Marr, 1982).

Summarizing, in this Section we explained the philosophical background Level-1 α and Level-1 β products refer to. At the end of the Book, the reader should be successful in figuring out an operative model that, placing the users at the center of the processing chain, allows for an effective management of several applications, replying to the fifth question introduced in Section 3.1.

Being the automatic extraction of information still limited, RGB products help the human photo-interpreter, which is in many cases designated to assign the scene semantic

3.3 General workflow and pre-processing chain

The general workflow to generate the proposed products is depicted in the block diagram of Figure 3.2. The pre-processing block answers the second question posed in Section 3.1 and involves all the operations necessary for obtaining a set of data suitable for the multitemporal fusion, as detailed in Section 2.3.1 and by Amitrano et al. (2015b). The only difference with respect to the chain illustrated by that references is in data normalization. Even in this case, the VALE method was adopted to reduce the images dynamics to its more relevant part through a saturation of the spikes due to (typically man-made) high reflective targets. However, here we keep the floating point values of the data, postponing the quantization step at the end of the synthesis, as detailed in the following sections.



Figure 3.2. General workflow of the multitemporal Level-1 β synthesis. The pre-processin chain is the same illustrated in Section 2.3.1. The only difference concerns the VALE normalization. In fact, in this case we keep the floating point values of the data, postponing the quantization step at the end of the synthesis.

3.4 Synthesis definition

Reference Jacobson et al. (2007) proposed some goals the fusion of a set of images should achieve for preserving information and enabling interpretability. Here, we want to highlight the four properties of *summarization*, *consistent rendering*, *computational ease* and *natural palette* recalling some concepts expressed in the aforementioned work.

As for the summarization, an effective synthesis should accurately summarize the original dataset, transferring to the user information he/she can not reach otherwise. This is possible if the rendering of data is consistent, i.e. the user is able to easily produce the association sign-object or color-feature.

To this end, the natural color palette is of course the best intermediary. However, while it is achievable with good approximation through an appropriate weighting of N bands of multispectral/hyperspectral data, for SAR images it is still far away to be produced. Hence, an advance towards this representation, with

Level-1 β products have been designed with the aim to meet the four properties of summarization, consistent rendering, computational ease and natural palette

short computational times in order to also enable real-time or near real-time applications, is desirable. Level-1 α imagery gave a first solution to this challenge, as shown in Section 2.4 and Section 2.5. Level-1 β products we are going to define have the same objective, trying to give an effective answer to this requests by solving appropriately the challenges relevant to the selection of the more suitable bands and to their fusion. This makes possible to meet the four propriety discussed above and also represents the subject of the third and fourth question introduced in Section 3.1.

3.4.1 Multitemporal analysis

In this Section, an answer to the third question raised in Section 3.1 will be provided: what is the most successful combination of data in order to output a RGB representation meeting the four properties of summarization, consistent rendering, computational ease and natural palette?

The idea is to compare (and combine) the mean intensity values of the time series with some variability indicators. The usage of the variance is a simple choice in order to evaluate the deviation with respect to the mean behaviour.

Another interesting information about the scene dynamics is represented by the maximum excursion of backscattered energy during the entire time series. It allows for the identification of outliers. Therefore, we use as third element of our synthesis the saturation map, as defined by Bruzzone et al. (2004), which exploits the information derived by the absolute maximums σ_{\max} and minimums σ_{\min} bands as follows:

$$\sigma_s = \frac{\sigma_{\max} - \sigma_{\min}}{\sigma_{\max}}, \quad \sigma_s \in [0, 1], \quad (3.1)$$

The drawback of this quantity is a pdf typically strongly skewed to the right. This results, in the final composition, in a sort of watermark of the color corresponding with that band covering the entire image. In order to avoid this visually unpleasant effect and reach a more balanced composition, we consider a saturation index defined by:

$$\hat{\sigma}_s = \frac{\sigma_{\max} - \sigma_{\min}}{\sigma_{\max} + \sigma_{\min}}, \quad \hat{\sigma}_s \in [0, 1]. \quad (3.2)$$

This formulation of the saturation index allows for reducing the skewness of the distribution with respect to Equation (3.1) making possible to obtain a higher balance of the RGB channels.

3.4.2 Multitemporal synthesis

As detailed in the previous Section, the quantities that will be used for the multitemporal synthesis are the mean intensity, the variance and the saturation index expressed in (3.2). The answer to the fourth question posed in Section 3.1, i.e. how to fuse the selected data, is now in order.

One of the basic problems in computer vision is to allow the observer to segment the image into meaningful regions (Khellaf et al., 1991) preventing the emergence of bright saturated regions, which has been judged to be distracting and confusing Jacobson et al. (2007). Hence, it is fundamental that the visualization favour the pre-attentive processing (Healey et al., 1995), i.e. the unconscious accumulation of information from the environment. As a consequence, we design our fusion mechanism in such way it gives as output highly contrasted images (which makes easier edges individuation), with limited occurrences of saturated regions, and a good classes separability.

Level-1 β synthesis involves the time series mean, variance, saturation index and the mean interferometric coherence

The fusion is implemented by maximizing the entropy of the channels involved in the composition. In particular, this is realized by clipping recursively the image pdf (at both left and right edges) for different percentages of the cumulative histogram until the maximum entropy value is reached. This procedure allows for obtaining more stretched histograms and a higher contrast. In fact, the flatter the histogram the higher the entropy.

The synthesis has a fourth participant, i.e. the interferometric coherence. This quantity is useful for separating high-reflective natural targets from man-made surfaces. In order to insert this information in our RGB composite we proceeded as follows:

- a. A master image is fixed and assumed as reference for the entire series;
- b. The interferometric coherence between the master image and all the slaves is computed;
- c. The mean coherence value γ is extracted;
- d. The mean coherence map is linearly quantized between a user-defined maximum and minimum. Reasonable values for these parameters are $\in [0.3, 0.4]$ for the minimum and $\in [0.5, 0.6]$ for the maximum. This means that, assuming $\gamma_{\min} = 0.3$ and $\gamma_{\max} = 0.5$ all the pixels whose coherence value is below 0.3 are set to null coherence; pixels with a coherence value between $[0.3, 0.5]$ are linearly distributed between $[0, 255]$; pixels whose coherence value is above 0.5 are set to 255. In the following we refer to this map as time series coherence map.
- e. The obtained time series coherence map is used as a mask within the saturation index map. In such a way, where the time series coherence values are zero (natural surfaces), the composite map exhibits the saturation index value. Otherwise the composite map returns the time series coherence value.

In the following sections, when the saturation index map is referred, the reader should take in mind the above described modification performed on its original definition.

As for the combination of the selected bands, we propose a solution trying to satisfy the requirements of consistent rendering and natural palette

introduced in Section 3.4. Therefore, the products shown in the following sections have been built with the following sequence:

- Red band: time series variance;
- Green band: time series mean;
- Blue band: saturation index map powered with the time series coherence map as explained above.

3.5 Products and their physical interpretation

In this Section, we propose three applicative scenarios for explaining the characteristics of Level-1 β products. In particular, we will deal with three scenes acquired by different sensors in different climatic areas, in order to highlight the robustness and the semantic stability of the defined synthesis.

3.5.1 Scenario 1 - Monitoring seasonal crops in temperate environment with Mediterranean climate

In this scenario, our objective is the detection of summery cultivations in Southern Italy, an area characterized by a temperate Mediterranean climate. To this end, we used images acquired during the summer. Indeed, since the sowing time is slightly moved up and the harvest could be delayed of some weeks, we considered 9 images between 5 April 2010 and 28 October 2010. The obtained Level-1 β product is shown in Figure 3.3.

Level-1 β products are application-oriented. The combination of images belonging to summer is suitable for monitoring crops growing in this season

Before examining the monitoring activity, it is worthwhile to linger on this representation which allows for highlighting some characteristics of Level-1 β products and the identification of some phenomena, as listed below:

- Sea appears in almost pure blue since the low contribution of the mean and variance bands and the high values of the saturation index



Figure 3.3. Castel Volturno (Italy), COSMO-SkyMed Level-1 β product obtained by synthesis of 9 images acquired between April 2010 and October 2010 for summer crops monitoring. Red band: time series variance; Green band: mean backscattering; Blue band: saturation index powered by the mean interferometric coherence.

due to different sea states that produce different backscatter in time. Indeed, this behavior (i.e. the blue color) can be observed on other weak scatterers, which could exhibit an unstable response such as roads, shadows, or surfaces in backslope;

- A more stagnant surface water, typical of rivers and irrigation tanks turns the response of the composition toward the black (see Fig-

ure 3.4a);

- Some irrigation tanks are subject to a more intensive usage during summer and dry up completely. Hence, their response acquires a strong red (i.e. variance) component which, red combined with an increased contribution of the mean band, results in a violet color, as shown in Figure 3.4b. In the following we will refer to these structures as “temporary tanks”;
- Man-made structures appears in cyan due to the combined contribution of the interferometric coherence and mean. The former contribution, in particular, allows for the distinction of such targets from the pine-grow which is characterized by a strong response of the mean, as shown in Figure 3.4c and Figure 3.4d.

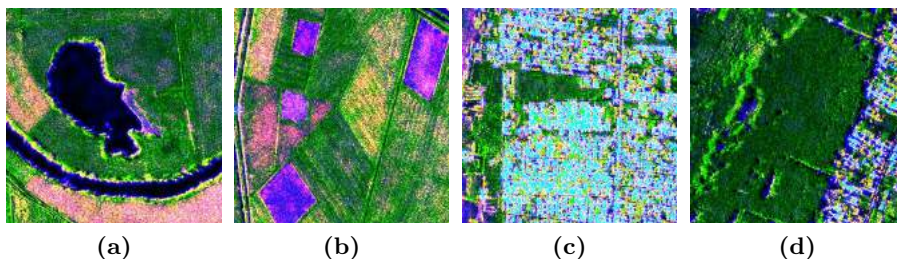


Figure 3.4. Castel Volturno, particulars of the Level-1 β product: (a) stagnant water in rivers and in “permanent” irrigation tanks appear in black due to the low response of all the involved channels. (b) “Temporary” irrigation tanks are rendered in violet due to significant contributions of variance and saturation index. (c) Man-made structures are displayed in cyan due to the high contribution of mean backscattering and mean interferometric coherence. (d) The pine grove has a green color due to its backscattering stability.

As a general comment, a high contribution of the blue band can be found both on very stable targets (such as buildings, due to the contribution of the time series coherence), and on very dynamic objects (such as cropfields or sea surface, due to the contribution of the saturation index). Indeed these two phenomenon, in principle ambiguous if only the blue band is considered, are well separable. In fact, on stable (man-made) targets, a negligible contribution of variance is observed. Conversely, a high variance is expected on agricultural land or dynamic natural surfaces.

ner of the scene, as well as the stripe of sparse vegetation and brushwoods about on the seaside are in the same way identifiable. Actually, the classes brushwoods and sparse vegetation are indistinguishable on the Level-1 β product, but they form a whole near the coast that is clearly separate by the adjacent urban area and conifer woods.

It is worthwhile to remark that the product we are analyzing has been produced considering a series of images belonging to summer. Hence, seasonal cultivations are expected to have high values of the mean intensity and variance bands due to the enhancement of backscattering derived by the volumetric contribution of the plants growth (Fung, 1979). Indeed, the saturation index is also expected to be high due to harvesting. Therefore, summery cultivations exhibit high values in all the bands of the composition, but their balance is dictated by cultivation type, terrain roughness, foliage density, plants height and fruits dimensions.

This means that there is no unique colour association that identify all agricultural fields. However, given that the variance contribution is significant, a higher contribution of the mean intensity (green band) results into a color which goes into yellow. A dominance of the saturation index (blue band) turns the response of the composition into a pink tonality. As an example, if the field appears in yellow, a high foliage density or taller plants have to be expected, as in the case of orchards (see the upper part of thematic map of Figure 3.5a).

Coming back to the overlay depicted in Figure 3.5b, the reader should note as the color of the contour for the classes “Springy/summery grain cereals”, “Springy/summery vegetable crops”, “Springy/summery industrial crops” and “Protected crops - vegetables and fruits” is the same (yellow), i.e. these categories have been grouped in a macro-class. This helps in highlighting as the Level-1 β product is congruent on average with the CORINE thematic map, since the fields whose response turn into pink or yellow are enclosed in the contour of that macro-class.

The classes “Autumnal/wintery grain cereals” and “Grasslands”, individuated in Figure 3.5b by red and amethyst contours can be also fused in an unique class. In fact, grasslands are objects almost stable and therefore their response is dominated by the mean intensity band (green). During summer, wintery-cultivated fields are usually destined to fallow and hence they are most likely covered by vegetation. As a consequence, their re-

sponse turns into green, too. In our case, grasslands and wintery cultivations are grouped in the center of the scene (see the thematic map of Figure 3.5a) and this is confirmed by the Level-1 β product which, in that zone, exhibits a large dominance of the mean intensity band response.

The above listed outcomes hold for sensor variations.

In Figure 3.6a and Figure 3.6b we show two Level-1 β products relevant to the city of Cirò Marina (Italy). In both cases, the SLC images were acquired by the sensor TerraSAR-X. In particular, the product depicted in Figure 3.6a has been obtained by fusing a time series belonging to April-October 2008. The product depicted in Figure 3.6b is the result of the fusion of a time series belonging to the same period of the year 2009. In order to make the two time series comparable, a metric has to be established. Therefore, we assumed as reference the 2008 product, fixing the maximum and minimum values of the quantities involved in the building of the Level-1 β product of the successive year. In other words, when data of the 2009 image are transformed from floating point data format to unsigned byte, the zero-*th* and 255-*th* amplitude levels coincide for each band with those of the reference image.

The reader can immediately appreciate how the semantic of these compositions is coherent with that illustrated above for the Castel Volturno dataset. The principal cultivation of the study area is the grapevine, which is a summer cultivation. Thanks to the above considerations, vineyards are easily identifiable in the coastal stripe of Figure 3.6a in the fields with a yellow response. Moreover, due to the calibration, variations in the colorimetric response in the two products can be exploited for studying the behavior of the cultivations during the two farming seasons, as detailed in Chapter 5.

3.5.2 Scenario 2: Level-1 β products in regions with temperate continental climate

Sentinel-1 started to acquire data on April 2014. The success of the mission requires the development of repeatable and reliable processing able to produce attractive products for the end-user community. The potential of the mission has been widely explored in the remote sensing literature (Malenovský et al., 2012). Here, we will show that Level-1 β imagery is fully compatible with the characteristics of the new sensor.

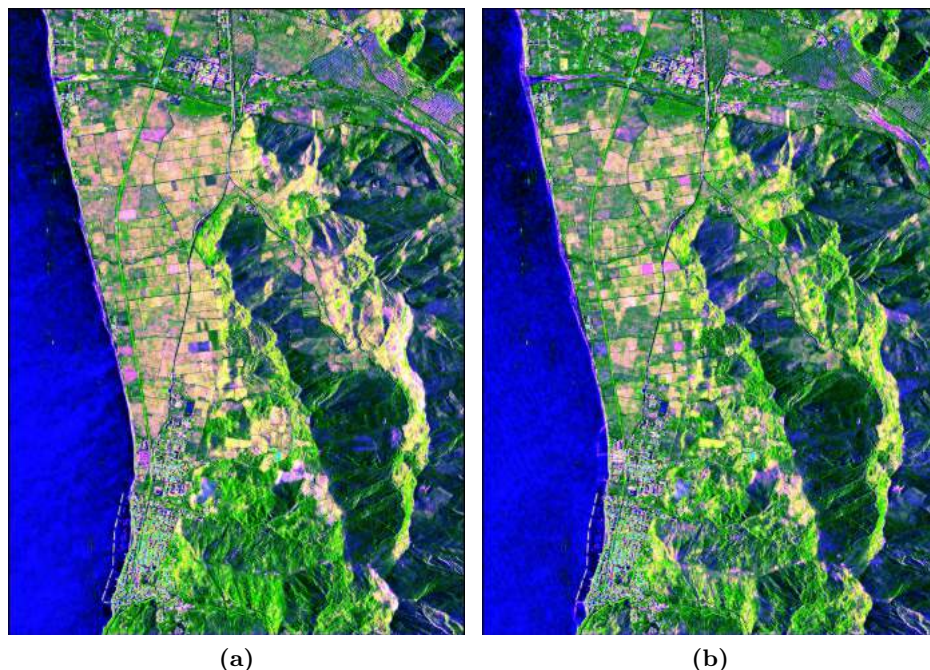


Figure 3.6. Ciró Marina (Italy): (a) April-October 2008 and (b) April-October 2009 TerraSAR-X Level-1 β products.

To prove this claim, we processed a multitemporal series of six images acquired between 3 October 2014 and 2 December 2014 in interferometric wide swath (IWS) mode over the Saxony region, Germany. Preliminary processing, such as TOPS (De Zan and Monti Guarnieri, 2006) deburst and calibration, has been carried out thanks to ESA Sentinel-1 Toolbox software.

In Figure 3.7 we show the full Level-1 β product of the study area. The spatial resolution is 15 meters. The yellow dot on the map identifies the city of Dresden. We reported in the same image the 100 meters resolution CORINE land cover corresponding with the imaged area (in the top left corner) and a wider view catching part of the central Europe (in the bottom right corner).

Macroscopically, the reader should note as the colors restituted by the composition are constantly associated to the same image feature. Cities

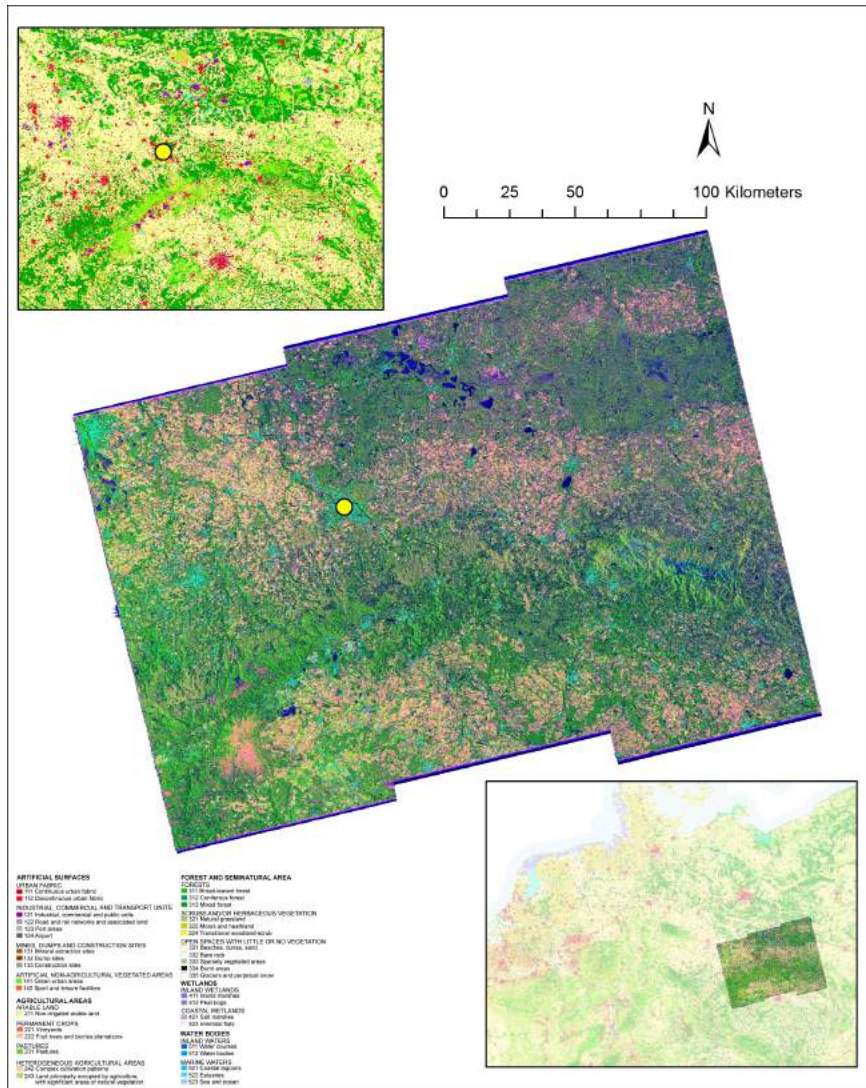


Figure 3.7. Sentinel-1 full Saxony Level-1 β product with 15 meters spatial resolution. The yellow dot identifies the city of Dresden. On the top left corner the 100 meters resolution CORINE land cover is reported. A wider view of part of the center Europe is provided in the bottom right corner. The semantic carried by the Level-1 β product is congruent with the classes reported in the reference map and it is constant for sensor variations.

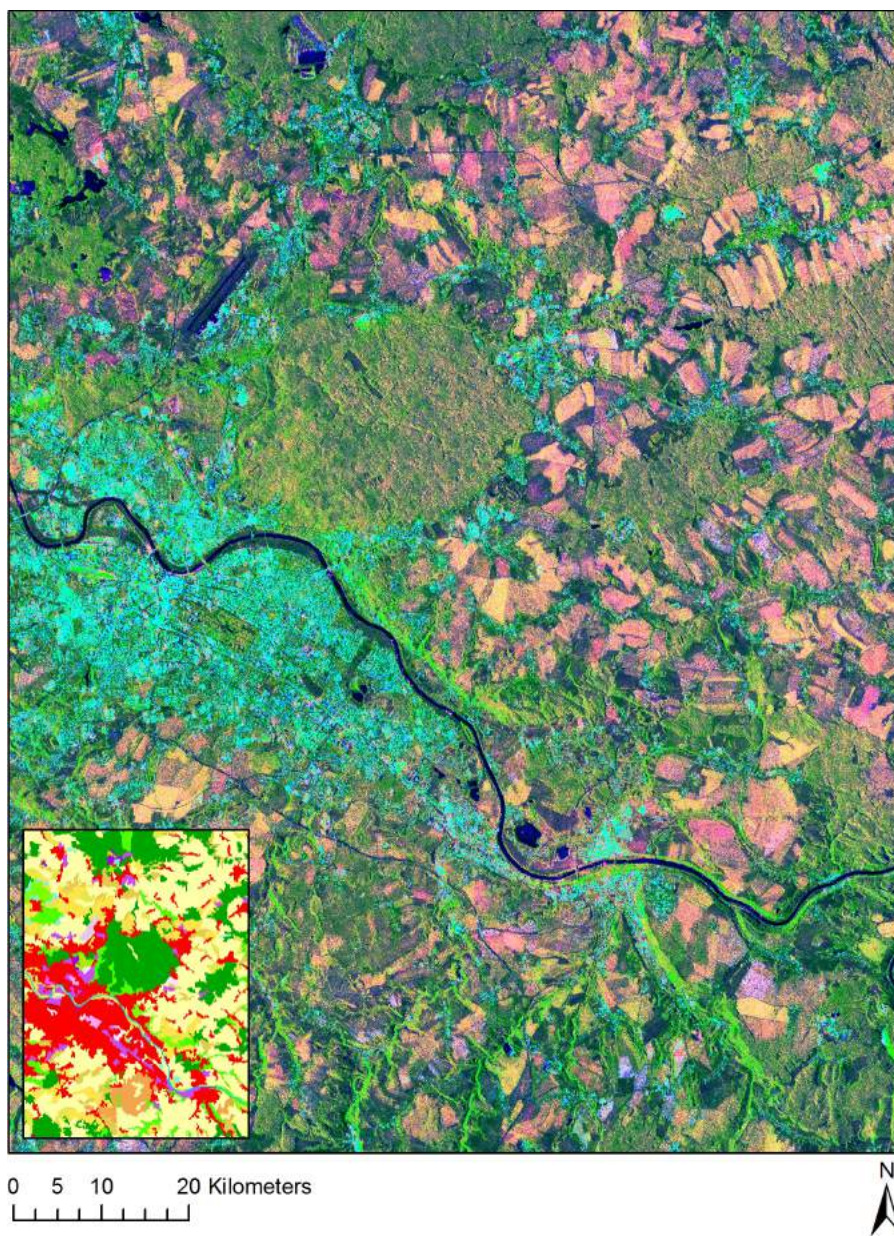


Figure 3.8. Sentinel-1 Level-1 β product of the Dresden city area with its relevant 100 meters resolution CORINE land cover.

are rendered in cyan. In the middle of the image, at South of Dresden, a wide stripe of woods area is identifiable since the response of the composition has a strong green component. Looking at the CORINE land cover in the top left corner of Figure 3.7, the dominant classes in this area are “Broad-leaved forest” and “Conifer Forest”, rendered in green and dark green, respectively.

At East and West of the city of Dresden (see the light yellow class on the CORINE land cover), croplands are distinguishable, as well as at South, above the woodland stripe. At North-West of Dresden, instead, another wide woody area can be identified with some water bodies of different size.

Summarizing, the proposed Sentinel-1 Level-1 β product respects macroscopically the semantic indicated by the CORINE land cover.

The same reasoning can be repeated on a finer scale looking at the close up of the Dresden city depicted in Figure 3.8. The corresponding portion of the 100 meters resolution CORINE land cover is displayed in the bottom left corner of the image. This representation should allow the reader to better appreciate the matching between the semantic extractable from the Level-1 β product and that established by the CORINE land cover.

As an example, it is easy to link the area at North-West of the Elbe river in which the green response is dominant with the conifer forest indicated in dark green within the CORINE land cover. In the same way, the Dresden city park, rendered in light pink on the reference land cover, is clearly distinguishable in Level-1 β imagery thanks to its green response. In synthesis, basing on the considerations provided in Section 3.5.1 discussing about the Castel Volturno scene, the reader should be able to reconstruct the semantic of the Level-1 β response of the city of Dresden. This testifies the robustness of our framework in the association color-object (or sign-object in the Peirce view), and its suitability to Sentinel-1 imagery.

3.5.3 Scenario 3: Level-1 β products in semi-arid environment

In this Section, we discuss the reliability of the proposed framework varying completely the climatic zone. In this case we deal with a semi-arid environment (see Section 2.3.2) where agricultural activities are strongly

related to the cycle of rainfalls (Amitrano et al., 2014b), determining the effectiveness of seasonal cultivations and the quantity of water which can be harvested for facing the dry season (Amitrano et al., 2014a).

In Figure 3.9 we show a Level-1 β product relevant with the Laaba basin obtained by processing eight COSMO-SkyMed spotlight images between 3 July 2014 and 4 August 2014, i.e. in the middle of the rainy season (Amitrano et al., 2014b). In this scenario, it is of interest to monitor water surface and cultivation dynamics, due to their importance for local community wellness, as well as the presence of small human settlements, since updated maps are usually not available.

Level-1 β imagery provides useful information regarding all these activities. As an example, the Laaba basin is clearly distinguishable at the center of the image. As previously hinted, it is highly influenced by rainfall cycles. In fact, it dries up completely during the dry, season due to evapotranspiration and human consumption, and it is filled by rain water during the wet season (Amitrano et al., 2014a). This behavior is almost constant in different years. However, basin contours are strongly variable even during the wet season, since the available water is used for irrigation and human activity. This emerges analyzing the Level-1 β product in the basin area. In fact, in its bottom part, wide areas characterized by a pink response are present (see also Figure 3.10a), due to the combination of variance and saturation index bands. Physically, because of the climatic conditions of the imaged area, it is possible to associate this response to regions characterized by an unstable water coverage.

The same phenomenon characterizes bright regions within the basin area and on its contour. However, in this case, the high contribution of the mean indicates that that portion of terrain is not covered by water in the most part of the considered acquisitions. Thus if a mask of the maximum basin extension is available, Level-1 β imagery allows for the extrac-

tion of information about the water coverage extent in the considered

The semantic carried by Level-1 β images is congruent with the classes indicated in the available reference maps and hold for sensor and climatic condition variations



Figure 3.9. Laaba basin (Burkina Faso), spotlight COSMO-SkyMed Level-1 β product obtained by fusion of eight spotlight COSMO-SkyMed images acquired between 3 July 2014 and 4 August 2014.

time interval.

In Figure 3.10b and Figure 3.10c the reader should note how both the change in acquisition modality and climatic area does not affect the response of features such as cultivated fields (Figure 3.10b) and “urban areas” (Figure 3.10c). We used quotation marks since in this case we are talking about very small settlements composed by few constructions which behave as stable targets.

In Figure 3.10d we show a detail relevant to a feature characteristic of a semi-arid environment, that is an eroded area (Amitrano et al., 2013), i.e. a region which lost its capability to retain and absorb water and which is, for this reason, completely barren and characterized by a low reflectivity. Thus the response of Level-1 β imagery for this feature turns out toward the black or a very dark blue because of a little contribution of the saturation index.

It is worthwhile to remark that dealing with semi-arid environments, in which scenes are mostly natural and, during the rainy season, characterized by a strong dynamic, the temporal sampling of the images involved in the Level-1 β composition is very important, as well as the considered time span. In fact, a thin set of images or a long time interval could give rise to anomalous response due to the effects of variance and/or of the saturation index.

3.6 Conclusions and discussions

In this Chapter, we introduced a new class of multitemporal SAR RGB products. Level-1 β products are complementary to the bi-temporal Level-1 α products discussed in Chapter 2, since they fuse N acquisition in an unique frame representative of the dynamics of the scene.

Level-1 β products have been designed in such a way to help the human photo-interpreter, thanks to a rendering as much as possible consistent with human view. As for Level-1 α products, the purpose is to lower the expertise required to correctly interpret data. In fact, in a scenario in which the automatic extraction of information from large databases is rather limited, users are in many cases appointed to extract information from the analyzed scene, i.e. to transfer a semantic to the imaged objects. Level-1 β products encounter this necessity, which is more and

More attention to users' experience is needed for a wider diffusion of SAR data in the end-user community. This is fundamental for the full exploitation of the new missions, whose data are freely available for the scientific community

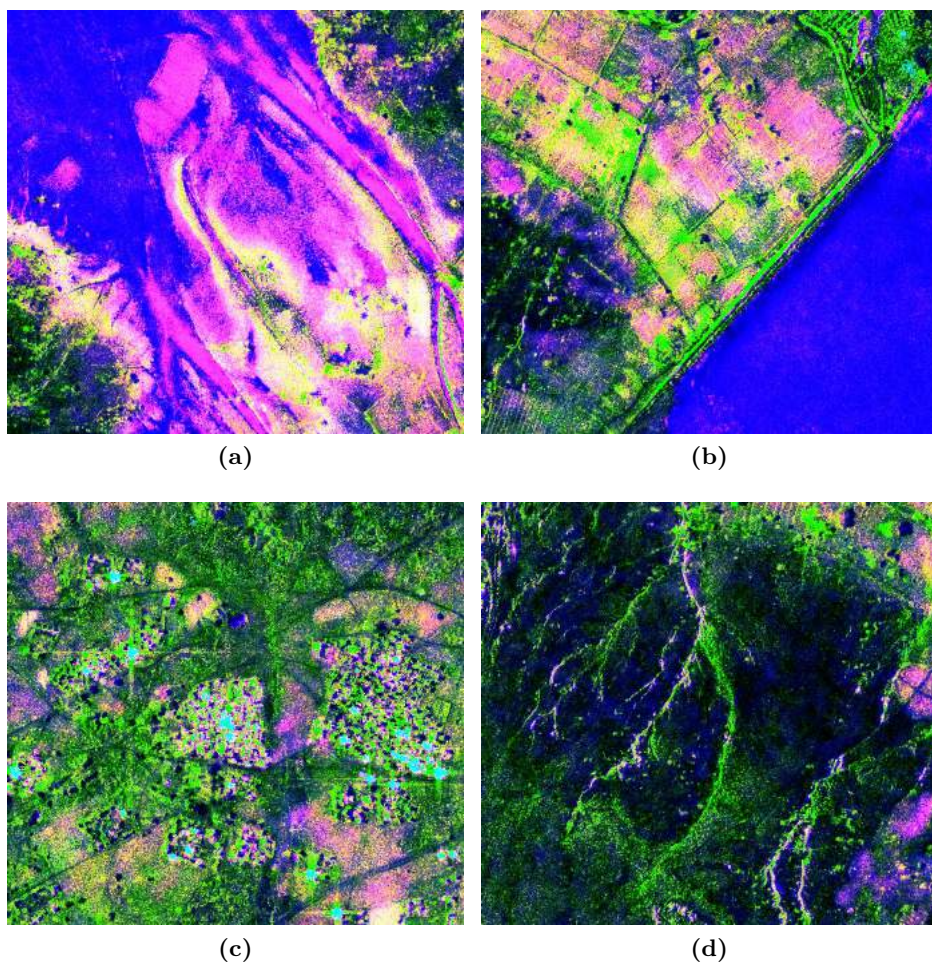


Figure 3.10. Laaba basin (Burkina Faso), details of the Level-1 β product shown in Figure 3.9: (a) an area characterized by instable water coverage, (b) crops in the nearby of the dam, (c) a small human settlement and (d) an eroded area.

more felt in the end-users community, making the human-machine interaction more easy and comfortable.

We tested the reliability of our products by implementing the Level-1 β processing chain using images acquired by three different sensors on four scenes (Castel Volturno, Ciró, Laaba and Dresden). All the available

acquisition modalities (stripmap, spotlight and scan) were tested, as well as different climatic zones: temperate with Mediterranean/continental climate and semiarid climate. The obtained results confirmed the semantic stability of the proposed products, i.e. the consistency of the association between the displayed colors and the objects on the scene.

In Section 3.1, we recalled five questions proposed in the past literature concerning the reliability of a data fusion technique. We proposed an answer for all of them with regard to the problem of data representation. However, question 5 (concerning how to extract the enormous amount of information from remotely sensed data) is only partially answered. It could not be otherwise, considering that Level-1 β processing, as well as Level-1 α , deals with a part of the whole problem, that is data representation.

However, the message we want to convey is that more attention to users' experience is needed for a wider diffusion of SAR data in the end-user community, in which the level of technical expertise could be not appropriate to manage classic SAR images and/or algorithms for information extraction. This is fundamental for the full exploitation of the new missions, whose data are freely available for the scientific community.

References

- Amitrano, D., Cecinati, F., Di Martino, G., Iodice, A., Riccio, D., and Ruello, G. (2015a). Sentinel-1 Multitemporal SAR Products. In *IEEE International Geoscience and Remote Sensing Symposium*, pages 3973–3976.
- Amitrano, D., Ciervo, F., Di Martino, G., Papa, M. N., Iodice, A., Koussoube, Y., Mitidieri, F., Riccio, D., and Ruello, G. (2014a). Modeling Watershed Response in Semiarid Regions with High Resolution Synthetic Aperture Radars. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 7(7):2732–2745.
- Amitrano, D., Di Martino, G., Iodice, A., Riccio, D., and Ruello, G. (2015b). A New Framework for SAR Multitemporal Data RGB Representation: Rationale and Products. *IEEE Transactions on Geoscience and Remote Sensing*, 53(1):117–133.
- Amitrano, D., Di Martino, G., Iodice, A., Riccio, D., Ruello, G., Ciervo, F., Papa, M. N., and Koussoube, Y. (2013). High resolution SAR for the monitoring of reservoirs sedimentation and soil erosion in semi arid regions. In *IEEE International Geoscience and Remote Sensing Symposium*, pages 911–914.
- Amitrano, D., Di Martino, G., Iodice, A., Riccio, D., Ruello, G., Papa, M. N., Ciervo, F., and Koussoube, Y. (2014b). Effectiveness of high-resolution SAR for water resource management in low-income semi-arid countries. *International Journal of Remote Sensing*, 35(1):70–88.
- Antropov, O., Rauste, Y., Astola, H., Praks, J., Hame, T., and Halikainen, M. (2014). Land Cover and Soil Type Mapping From Spaceborne PolSAR Data at L-Band With Probabilistic Neural Network. *IEEE Transactions on Geoscience and Remote Sensing*, 52(9):5256–5270.
- Atkin, A. (2013). Peirce’s theory of signs. In Zalta, E. N., editor, *The Stanford Encyclopedia of Philosophy*. Summer 2013 edition.
- Bossard, M., Feranec, J., and Otahel, J. (2000). CORINE land cover technical guide: Addendum 2000. Technical report, European Environment Agency Copenhagen.
- Bruzzone, L., Marconcini, M., Wegmuller, U., and Wiesmann, A. (2004). An Advanced System for the Automatic Classification of Multitemporal

- SAR Images. *IEEE Transactions on Geoscience and Remote Sensing*, 42(6):1321–1334.
- Campbell, J. B. (2002). *Introduction to Remote Sensing*. The Guilford Press, New York, third edition edition.
- De Zan, F. and Monti Guarnieri, A. (2006). TOPSAR: Terrain Observation by Progressive Scans. *IEEE Transactions on Geoscience and Remote Sensing*, 44(9):2352–2360.
- Dong, X., Quegan, S., Yumiko, U., , Hu, C., and Zeng, T. (2015). Feasibility Study of C- and L-band SAR Time Series Data in Tracking Indonesian Plantation and Natural Forest Cover Changes. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 8(7):3692–3699.
- Dumitru, C., Cui, S., Faur, D., and Datcu, M. (2015). Data Analytics for Rapid Mapping: Case Study of a Flooding Event in Germany and the Tsunami in Japan Using Very High Resolution SAR Images. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 8(1):114–129.
- Franceschetti, G. and Lanari, R. (1999). *Synthetic Aperture Radar Processing*. CRC Press, Boca Raton, FL.
- Fung, A. K. (1979). Scattering from a Vegetation Layer. *IEEE Transactions on Geoscience Electronics*, 17(1):1–6.
- Healey, C., Booth, K. S., and Enns, J. (1995). Visualizing real-time multivariate data using preattentive processing. *ACM Transactions on Modeling and Computer Simulation*, 5(3):190–221.
- Jacobson, N. P. and Gupta, M. R. (2005). Design Goals and Solutions for Display of Hyperspectral Images. *IEEE Transactions on Geoscience and Remote Sensing*, 43(11):2684–2692.
- Jacobson, N. P., Gupta, M. R., and Cole, J. B. (2007). Linear Fusion of Image Sets for Display. *IEEE Transactions on Geoscience and Remote Sensing*, 45(10):3277–3288.
- Khellaf, A., Beghdadi, A., and Dupoisot, H. (1991). Entropic Contrast Enhancement. *IEEE Transactions on Medical Imaging*, 10(4):589–592.
- Madhok, V. and Landgrebe, D. A. (2002). A Process Model for Remote Sensing Data Analysis. *IEEE Transactions on Geoscience and Remote Sensing*, 40(3):680–686.
- Malenovský, Z., Rott, H., Cihlar, J., Schaepman, M. E., García-Santos,

- G., Fernandes, R., and Berger, M. (2012). Sentinels for science: Potential of Sentinel-1, -2, and -3 missions for scientific observations of ocean, cryosphere, and land. *Remote Sensing of Environment*, 120:91–101.
- Marr, D. (1982). *Vision*. W. H. Freeman, San Francisco.
- Mendel, J. M. (1995). Fuzzy Logic Systems for Engineering: A Tutorial. *Proceeding of the IEEE*, 83(3):345–377.
- Nöth, W. (1995). *Handbook of semiotics*. Indiana University Press, Bloomington, IN 47404-3797 USA.
- Pohl, C. and Van Genderen, J. L. (1998). Multisensor image fusion in remote sensing: Concepts, methods and applications. *International Journal of Remote Sensing*, 19(5):823–854.
- Richards, J. A. and Jia, X. (2006). *Remote Sensing Digital Image Analysis - An Introduction*. Springer-Verlag, Berlin, Germany, fourth edition.
- Santini, S., Gupta, A., and Jain, R. (2001). Emergent semantics through interaction in image databases. *IEEE Transactions on Knowledge and Data Engineering*, 13(3):337–351.
- Schroder, M., Rehrauer, H., Seidel, K., and Datcu, M. (1998). Spatial Information Retrieval from Remote-Sensing Images - Part II : Gibbs-Markov Random Fields. *IEEE Transactions on Geoscience and Remote Sensing*, 36(5):1446–1455.
- Sonka, M., Hlavac, V., and Boyle, R. (1999). *Image Processing, Analysis and Machine Vision*. Brooks/Cole Publishing Company, Pacific Grove, CA, 93950, USA.
- Wilson, T. A., Rogers, S. K., and Kabrisky, M. (1997). Perceptual-Based Image Fusion for Hyperspectral Data. *IEEE Transactions on Geoscience and Remote Sensing*, 35(4):1007–1017.

Chapter 4

Semantic products

4.1 Introduction

In this Chapter, we will show how to introduce a basic semantic in the Level-1 α and Level-1 β products discussed in the previous chapters. In particular, Kohonen's self-organizing maps (Kohonen, 2001) will be used to map the input RGB product in a feature space of limited dimension. A label will be attached to each of the obtained clusters in order to make the output product semantic. The label will consist in a verbal attribute recalling a physical attribute of the cluster, in this case its color.

The Chapter is organized as follows. In Section 4.2, self-organizing maps are introduced. Semantic self-organizing maps are introduced in Section 4.3. Conclusions are drawn at the end of the Chapter.

4.2 Self-organizing maps

Self-organizing maps (SOMs), are a machine learning technique of the Artificial Neural Networks (ANN) family. They are used for the classification of the most diverse data types. Their application, since their introduction by Kohonen, has been experimented in different sectors, like gene expression analysis (Törönen et al., 1999), financial diagnosis (Serrano-Cinca, 1996), synoptic climatology (Hewitson and Crane, 2002), microbial community dynamics Dollhopf et al. (2013), bankrupt prediction (Kiviluoto,

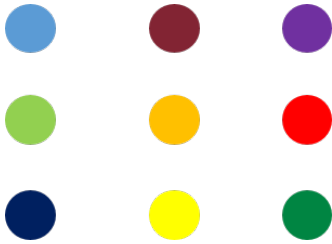


Figure 4.1. Randomly initialized SOM. In this example, nodes constituted by a RGB triplet are used. However, this neural net can be used to classify the most diverse and heterogeneous datatypes. The number of nodes is user-defined. They are usually connected with a rectangular structure, as the one depicted in figure.

1998), and political science (Pearson and Cooper, 2012). SOMs have been widely applied to remote sensing since the '90s (Atkinson and Tatnall, 1997) as well, and new applications are still studied and proposed today. This widespread use of the SOMs is due to the extreme flexibility of the tool, which can be adapted/modified/integrated to work with different data types/structures (Kamal et al., 2010) and learning techniques (Patra and Bruzzone, 2014). The robustness to large amounts of data makes them a suitable instrument for unsupervised or semi-supervised classification in a big data scenario, which is, and will increasingly more be in the future, a key issue in remote sensing.

Dealing with SAR data, SOMs have been employed, as an example, for polarimetric data classification (Ito and Omatu, 1998), change detection (Neagoe et al., 2014), and flood mapping Skakun (2010). Here, the SOM will be used to map the input RGB product in a feature space with a limited number of elements. Each element will be labeled with a verbal attribute recalling a physical property, i.e. its color.

A SOM is composed by a user-defined number of randomly initialized nodes of the same type of the data to be classified (see Figure 4.1). The nodes are connected by a (usually) rectangular structure and are trained using a pre-defined number of sample vectors randomly selected into the input data. Each time a training vector is presented to the network (see Figure 4.2), the most similar node is detected and identified as the best matching unit (BMU). The BMU and its neighbor, defined by a radius, are update to become more similar to the presented training sets. This operation is repeated for several iterations (called epochs). At the end of each epoch the neighbor of the BMU as well as the learning rate are decreased. In such way, after many epochs, the SOM becomes stable and the obtained nodes (i.e. colors) can be used to classify data.

In the classic Kohonen's scheme, SOM's neurons are typically randomly initialized. Consequently, for a given map and network parameters, the SOM output will be different for each run. However, as explained before, SOMs, can be easily modified to be adapted to specific data. In our case, we established a data-driven seed for generating the initial set of neurons. In such way, the input map fixes the output SOM for a given set of the required network parameters.

The training sets necessary for running the SOM are chosen based on the input RGB product. In particular, a matrix of $M \times 3$ RGB triplets is randomly generated using the previously introduced seed. In order to consider a larger variety of combinations of the primary

The flexibility of SOMs allows to modify easily their basic schema to adapt the tool to the specific input data

colors, M is greater than the pre-established number T of training vectors. These random triplets are made consistent with the re-quantization problem by computing pixel-wise the Euclidean distance between the i -th training set and the input RGB product. Therefore, among the M available tones, the T that are closer to a color existing in the input RGB products are chosen as training sets.

One of the knowledge required for understanding remote sensing images concerns the mapping of scene features into the image (Matsuyama and Hwang, 1990). Exploiting this knowledge, we implemented a modification to the classic Kohonen's scheme based on the input products. As an example, consider a Level-1 α product as in Figure 2.12, i.e. with the coherence band loaded on the red channel. In this case, we impose the presence of white, black and red colors among the training sets to be used in the competitive phase. In fact, these colors are associate to precise classes (such as urban areas, water surfaces and low-backscattering areas) which are likely to be present in every acquisition, even if, maybe, with small occurrence with respect to other classes. In order to ensure the presence of such colors in almost pure tonality in the final SOM, when the relevant training sets are presented to the network, it behaves as in a Learning Vector Quantization (LVQ) schema (Kohonen, 2001), in which only the winning neuron is updated with a high learning rate.

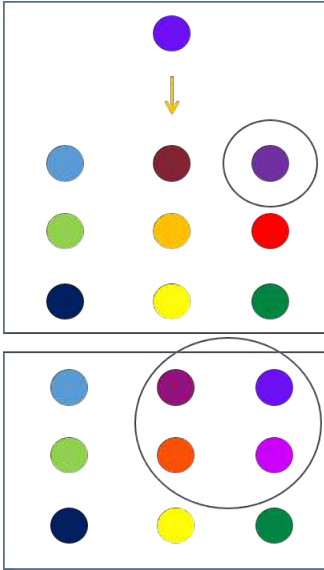


Figure 4.2. SOM's nodes are trained using a pre-defined number of sample vectors randomly selected into the input data. Each time a training vector is presented to the network the most similar node is detected and identified as the best matching unit (BMU). The BMU and its neighbor, defined by a radius, are update to become more similar to the presented training sets. This operation is repeated for several iterations (called epochs). At the end of each epoch the neighbor of the BMU as well as the learning rate are decreased. In such way, after many epochs, the SOM becomes stable and the obtained nodes (i.e. colors) can be used to classify data. This is the classic Kohonen's scheme. It will be slightly modified to be adapted to the characteristics of the input RGB products.

4.3 Semantic self-organizing maps

As previously hinted, the objective of SOM clustering is not only to map the input RGB product into a feature space of limited dimension, but also to attach to each cluster a verbal label related to its color. In such way, the SOM is made semantic (SSOM).

To this end, a HTML color database is considered. At the end of the training phase, the Euclidean distance between the output SOM and the database elements is computed. Finally, for each SOM element, the name of the less distant color of the database is assigned.

SOM are used to map the input product from the RGB space into a feature space with a limited number of labeled elements

In Figure 4.4 we show the results of SSOM clustering. In particular, in Figure 4.3a, a Level-1 α product of the city of Castel Volturno is shown (see also 2). In Figure 4.3b, its 25-elements clustering is depicted. The reader should appreciate that, despite a significant reduction of the space

in which the products are defined, they look very similar. This allows for an immediate semantic transferring between the two products.

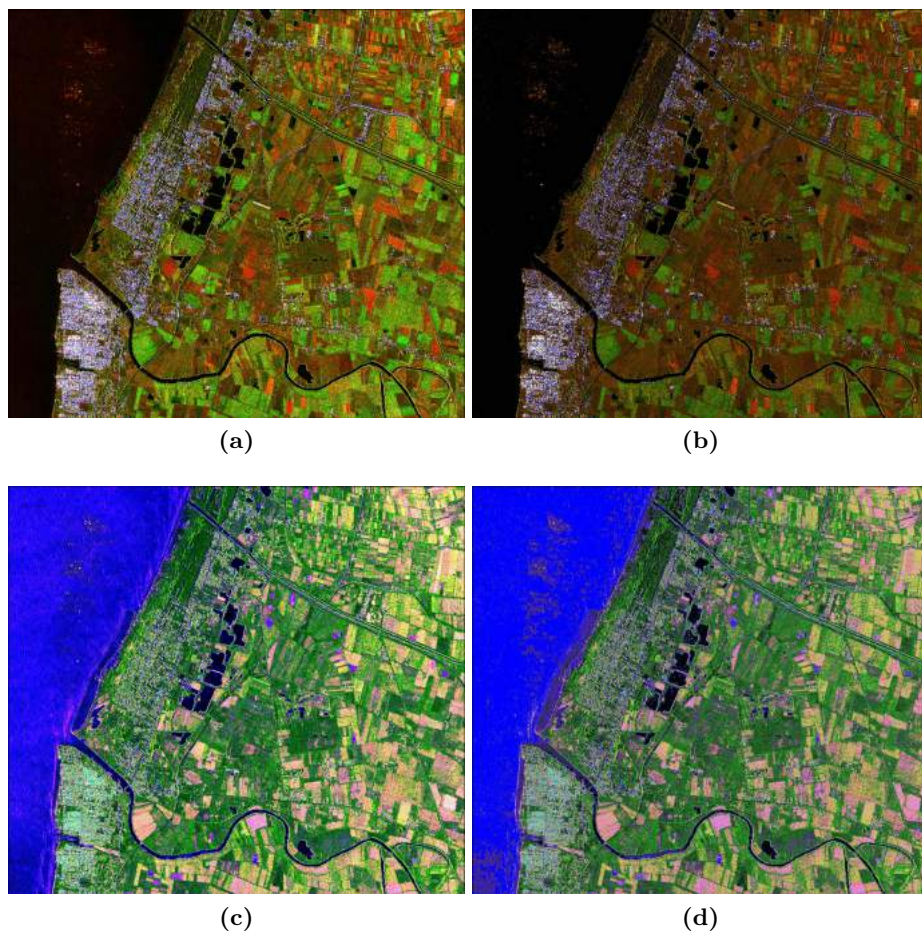


Figure 4.3. Castel Volturno (Italy): Level-1 α product (a) before and (b) after 25-elements SSOM clustering. Level-1 β product (c) before and (d) after 25-elements SSOM clustering. The usage of the SOM allows for the preservation of the chromatic response of the input RGB product. This allows for an immediate semantic transferring from the RGB product to the classified one.

The same reasoning can be made looking at the Level-1 β product of the Castel Volturno city analyzed in Chapter 3 and depicted in Figure 4.3c.

Cluster label
White
Orange
Chocolate
Orange Red
Blue
Gold
Dark Goldenrod
Dark Goldenrod
Saddle Brown
Brown
Lawn Green
Olive Drab
Olive Drab
Saddle Brown
Saddle Brown
Lime Green
Olive Drab
Olive Drab
Saddle Brown
Dark Forest Green
Lime Green
Forest Green
Forest Green
Dark Forest Green
Black

Table 4.1. Cluster labels assigned by the SSOM for the Level-1 α product depicted in Figure 4.3b.

Cluster label
Blue
Royal Blue
Cyan
Medium Purple
Plum
Dark Slate Blue
Steel Blue
Cadet Blue
Dark Orchid
Pale Violet Red
Dark Slate Gray
Dim Gray
Dim Gray
Sienna
Chocolate
Black
Dark Olive Green
Olive Drab
Olive Drab
Peru
Forest Green
Forest Green
Lime Green
Yellow Green
Dark Khaki

Table 4.2. Cluster labels assigned by the SSOM for the Level-1 β product depicted in Figure 4.3d.

Its 25-elements clustering is depicted in Figure 4.3d. Even in this case, the reader can appreciate few differences in the two products. In fact, the chromatic content of the input RGB product is largely preserved in the clustered one thanks to the usage of the SOM.

In Table 4.1 and Table 4.2, the cluster labels assigned by the SSOM

during the clustering of the Level-1 α and of the Level-1 β products depicted in Figure 4.3b and Figure 4.3d, respectively. It is possible that the algorithm assigns to more colors the same label due to their similarity. In this case, that colors can be immediately considered for the fusion, since they are likely to represent the same phenomenon.

These labels, identifying a certain scattering mechanism, can be used for simple classification/pre-classification activities. To this end, consider the Level-1 β subset of the Castel Volturno scene reported in Figure 4.4a. According to Level-1 β products rationale, it belongs to an area mainly dedicated to growing crops. In fact, this feature, as explained in Chapter 3 is rendered in yellowish/pinkish color. In order to identify it, a simple semantic query can be made.

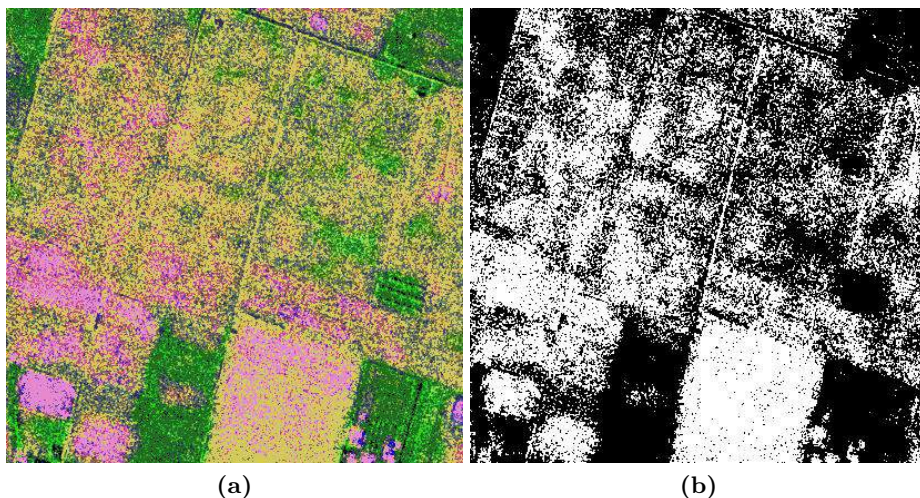


Figure 4.4. Semantic queries allows for easily identifying image features through the color attribute. (a) Level-1 β product and (b) its growing crops mask obtained isolating the clusters having attributes “Plum”, “Pale Violet Red”, and “Dark Khaki”.

In fact, it is possible to isolate the clusters concerning growing crops simply by querying the attacher color label. In Figure 4.4b, the mask obtained by selecting only cluster having “Plum”, “Pale Violet Red”, and “Dark Khaki” colors is reported. These colors have a strong red component (i.e. a strong variance component), which is dominant compared with

the green component (i.e. the mean backscattering component). Therefore, they are likely to represent a changing landscape due to the growth of plants. In this procedure, we selected only clusters with occurrence higher than 5% of the scene, discarding those with dominant green component.

It is remarkable that the usage of semantic queries allows for avoiding to run statistical classifiers that, having typically a strong parametric nature, could require a non-trivial technical expertise to be used. Therefore, it represents an end-user-oriented solution having, moreover, a negligible computational burden. A richer characterization of the objects using shape and spatial relationships can be used to refine the obtained results and/or to run more complicated applications, as detailed in Chapter 5. In alternative, semantic queries could act as a pre-classifier, to be used to retrieve more specific classes (as an example the type of cultivation) using scattering-based statistical methods.

The color attribute introduced through the SOM can be exploited in classification activities by querying the image at semantic level

4.4 Conclusions

In this Chapter, we showed how to introduce a basic semantic in SAR data using SOM clustering. In particular, we slightly modified the classic Kohonen's schema to make the neural net consistent with vector quantization of SAR-derived data. The objective was to map the input color-composite from the RGB space into a feature space with a limited number of (labeled) elements. The label is constituted by a verbal attribute recalling a physical property of the class, in particular its color.

We showed how to exploit the cluster label for querying the image at semantic level in an end-user-oriented environment for classification. In alternative, semantic processing could be used as pre-classifier, to be exploited to retrieve more specific classes (as an example the type of cultivation) using scattering-based statistical methods.

References

- Atkinson, P. M. and Tatnall, R. L. (1997). Introduction Neural networks in remote sensing. *International Journal of Remote Sensing*, 18(4):699–709.
- Dollhopf, S., Hashsham, S., and Tiedje, J. (2013). Interpreting 16S rDNA T-RFLP Data: Application of Self-Organizing Maps and Principal Component Analysis to Describe Community Dynamics and Convergence. *Microbial Ecology*, 42(4):823–836.
- Hewitson, B. C. and Crane, R. G. (2002). Self-organizing maps: applications to synoptic climatology. *Climate Research*, 22:13–26.
- Ito, Y. and Omatu, S. (1998). Polarimetric SAR data classification using competitive neural networks. *International Journal of Remote Sensing*, 19(14):2265–2684.
- Kamal, M. M., Pasmore, P. J., and Shepherd, I. D. H. (2010). Integration of geographic information system and RADARSAT synthetic aperture radar data using a self-organizing map network as compensation for real-time ground data in automatic image classification. *Journal of Applied Remote Sensing*, 4(1):1–13.
- Kiviluoto, K. (1998). Predicting bankruptcies with the self-organizing map. *Neurocomputing*, 21(1-3):191–201.
- Kohonen, T. (2001). *Self-Organizing Maps*. Springer-Verlag, Berlin, Heidelberg.
- Matsuyama, T. and Hwang, V. S.-H. (1990). *SIGMA - A Knowledge-Based Aerial Image Understanding System*. Plenum Press, New York.
- Neaogoe, V.-E., Stoica, R.-M., Ciurea, A.-I., Bruzzone, L., and Bovolo, F. (2014). Concurrent Self-Organizing Maps for Supervised/Unsupervised Change Detection in Remote Sensing Images. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 7(8):3525–3533.
- Patra, S. and Bruzzone, L. (2014). A Novel SOM-SVM-Based Active Learning Technique for Remote Sensing Image Classification. *IEEE Transactions on Geoscience and Remote Sensing*, 52(11):606–616.
- Pearson, P. T. and Cooper, C. I. (2012). Using Self Organizing Maps to Analyze Demographics and Swing State Voting in the 2008 U.S. Presidential Election. In Mana, N., Schwenker, F., and Trentin, E., editors, *Artificial Neural Networks in Pattern Recognition*, volume 7477

- of *Lecture Notes in Computer Science*, pages 201–212. Springer Berlin Heidelberg.
- Serrano-Cinca, C. (1996). Self organizing neural networks for financial diagnosis. *Decision Support Systems*, 17(3):227–238.
- Skakun, S. (2010). A Neural Network Approach to Flood Mapping Using Satellite Imagery. *Computing and Informatics*, 29:1013–1024.
- Törönen, P., Kolehmainen, M., Wong, G., and Castrén, E. (1999). Analysis of gene expression data using self-organizing maps. *FEBS Letters*, 45(2):142–146.

Chapter 5

Applications

5.1 Introduction

In this Chapter, some applications involving multitemporal SAR, with particular emphasis on the usage of Level-1 α and Level-1 β products, will be discussed. In particular, water resources management in semi-arid regions through the synergic use of SAR data and hydrological modeling is presented in Section 5.2. In this application, SAR images are used to extract the bathymetry of small reservoirs and to retrieve a semi-empirical relation between basins' surface area and retained volume.

In Section 5.3, classification using Level-1 α and Level-1 β products is discussed. In particular, in Section 5.3.1, a simple supervised classifier is used to classify a Level-1 α product. In Section 5.3.2, a two years crops monitoring application using Level-1 β products in combination with neural networks is faced. In Section 5.3.3, land cover mapping through neural networks starting from a Level-1 β products is performed.

Features extraction is discussed in Section 5.4. In particular, in Section 5.4.1, an index based on Level-1 α products for detecting temporary water surfaces is introduced. In Section 5.4.2, buildings extraction through a new radiometric index is discussed, together with a technique for the enhancement of built-up areas in Level-1 α images. In Section 5.4.3, a precision agriculture application is faced exploiting semantic RGB products and object-based image analysis.

5.2 Water resources management in semi-arid regions

In semi-arid regions, small reservoirs are widely employed for facing water scarcity and climatic variability (Boelee et al., 2009; Cecchi et al., 2009). In Burkina Faso, it is estimated that almost 1700 small reservoirs are actually used for irrigation, livestock, and other uses. However, despite of their crucial importance, they are not appropriately monitored and, in many cases, not catalogued or the relative data not kept up-to date (Cecchi et al., 2009). Moreover, small reservoirs are often built or modified for the initiative of local communities and even basic data as their location and capacity are not available. For these reasons it is extremely hard to study their impact on the territory and to optimize their management.

In order to improve the management of water resources it is necessary to provide hydrological models suitable for the specific environments under study. In literature, there is a lack of environmental and hydrological data concerning semi-arid West Africa. Therefore, it is necessary to improve data availability through the development of cost-effective monitoring techniques and to adapt the hydrological modeling to the limited available data (Amitrano et al., 2014a).

The use of remote sensing in this context could really be a breakthrough, allowing for a significant reduction of costs and time needed for achieving crucial information for effective and integrated water management (Zhang and Jia, 2013). So far, the use of multispectral (Liebe et al., 2005;

Mutiti et al., 2010) and low-resolution radars (Liebe et al., 2009; Anzor et al., 2009) for reservoir monitoring evidenced great potentialities, but also poor practical results. In fact, the main limit for a continuous monitoring with passive sensors (i.e. optical sensors) turned out to be the dependence of the results on the cloud cover, which is particularly critic in wet season. Conversely, SAR systems are independent of illumination and atmospheric conditions. However, their usage for small reservoirs monitor-

Multitemporal SAR is a powerful and cost-effective tool for water resources monitoring in low-income semi-arid countries

ing has been limited by the resolution of the available data, at least until the launch of the new generation of high-resolution sensors, which allowed for a significant reduction of the dimension of the observable surfaces, up to few thousand square meters (Amitrano et al., 2014a,c).

The increased resolution (both spatial and temporal) and radiometric quality of the available data allows for overcoming most of the existing limits, opening the way to new fields of applications of SAR that could provide humanitarian, technological and economic benefits.

5.2.1 Methodology and case study characterization

The study area is located in northern Burkina Faso, and it is entirely enclosed in the Volta basin. The region is characterized by a semi-arid climate, with precipitation concentrated in the June-September rainy season, as detailed in Section 2.3.2. This climate calls for appropriate strategies for water storage in the rainy season and water use in dry season.

One of the oldest and most popular technique for water harvesting is the construction of dams for the creation of small reservoirs. The northern Burkina Faso is characterized by an increasing demographic pressure and by soils that are prone to erosion, causing sedimentation phenomena that reduce the available water volume. The number of reservoirs frequently changes, due to uncontrolled construction and ruptures, therefore an updated inventory of the existing reservoirs with traditional techniques is hard to obtain. Reservoirs monitoring would be of great importance in order to optimize their management, also because their effects on the overall water distribution of the Volta basin is still unclear and it is object of controversies with neighboring countries.

The block diagram presented in Figure 5.1 describes the overall methodology. More details on each block are provided in the following sections. It is worth to note that the modularity of the approach allows for choosing the most appropriate technique in accordance with the desired results and the working environment. Here, most of the choices have been dictated by the need to obtain accurate results with the simplest techniques. In such way, a framework easily reproducible even by non-expert SAR users is obtained.

Along with the obtained products, in the last column of the block diagram, a set of applications have been cited to show the potentiality of

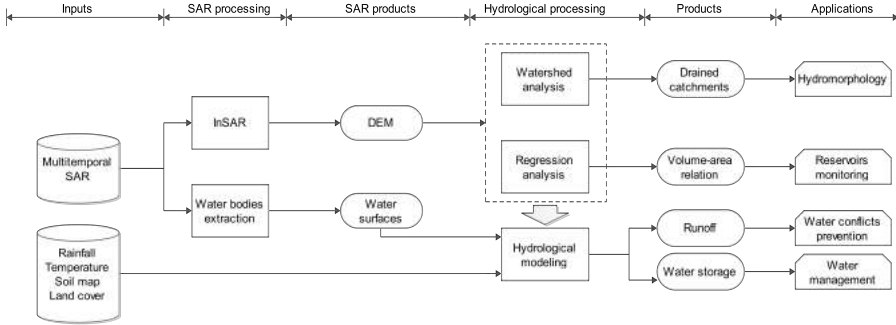


Figure 5.1. Block diagram of the adopted processing chain. Basin contours obtained through SAR processing and bathymetric data retrieved using a high resolution DEM are used to feed an hydrogeological modeling for monitoring the available water volume.

the approach (Amitrano et al., 2014a).

5.2.2 Water bodies extraction

In this Section, the processing steps for basins' shoreline extraction is discussed. For basic multitemporal SAR processing, the reader is addressed to Section 2.3.1.

The adopted technique for basin contours extraction is very simple. It consists in a thresholding of the intensity image in order to separate the class “water” from the class “no water”. The selection of the threshold value is guided by the histogram evaluation of a subset taken around the basin under analysis. The constrain to respect is the selection of a subset in which the presence of surface water is relevant; in fact, in this case the histogram of the intensity map exhibits a bimodal distribution.

In Figure 5.2 we considered a subset around the Gouinre basin and reported the relevant histogram. Thanks to multitemporal despeckling, low reflectivity objects are well-separated from the rest of the scene and this makes the image pdf bimodal. This distribution emerges after an appropriate histogram clipping. In fact, the histogram depicted is relevant to the 98th percentile of the pdf. This parameter has been selected with an entropy-guided criterion as detailed by Amitrano et al. (2015c).

Histogram clipping allows for an easy identification of the optimal

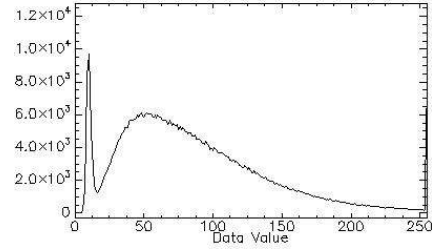


Figure 5.2. Gouinre basin (Burkina Faso). (a) SAR intensity map after multitemporal De Grandi despeckling with (b) its 98th percentile histogram. The bi-modal pdf allows for an easy individuation of the optimal threshold for separating water surface from land features

threshold, which is given by the local minimum between the two distribution modes. This threshold is basin- and image-adapted, since water surfaces are not necessarily static objects and their reflectivity could be influenced by local phenomena.

The threshold-guided segmentation allows for a quick preliminary estimate of the water surface, as shown in the example depicted in Figure 5.3. In particular, in (Figure 5.3a), the SAR image of the Laaba basin after multitemporal De Grandi despeckling and histogram clipping is depicted; the binary mask obtained via thresholding of the intensity map is reported in (Figure 5.3b).

The quality of the obtained mask, which could be affected by granularity (principally due to the presence of eroded areas) and/or holes within the basin area, can be significantly improved by using morphological operators. Here, we applied a mode filter (Amitrano et al., 2016) followed by a closing operator (Ronse and Serra, 2010). The result of these operations are reported Figure 5.3c and Figure 5.3d, respectively.

Finally, the shoreline is extracted with the Roberts' operator (Shrivakshan and Chandrasekar, 2012), as shown in Figure 5.3e. The result of the whole processing chain is shown in Figure 5.3f, in which the obtained contour has been superimposed to the amplitude map. Qualitatively, a

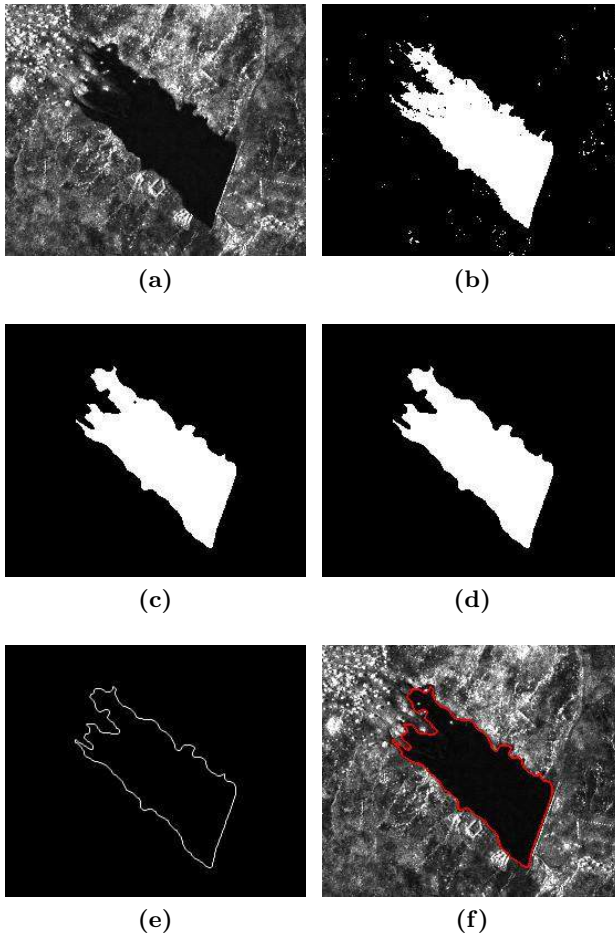


Figure 5.3. Shoreline extraction workflow: (a) intensity De Grandi-filtered map after the 98th percentile histogram clipping, (b) threshold-guided segmentation. The quality of the obtained mask, which could be affected by granularity and/or holes within the basin area (principally due to the simplicity of the adopted method) can be improved through mathematical morphology. Here, we applied a (c) mode filter, followed by (d) a closing morphological operator. Finally, (e) the Roberts edge detector is used for extracting the basin contour which are (f) superimposed to the intensity map, as a result of the whole processing.

good matching between the extracted shoreline and the intensity map can be appreciated.

The extraction of the shorelines has been performed with such a semi-automatic procedure for all the available acquisition for two basins in the nearby of the city of Ouahigouya and the basins of Laaba, Tougou, Aoérana and Derhogo for a total of six sites (see Figure 5.4). The results of this procedure are shown in Figure 5.5 to Figure 5.6 for Laaba and Derhogo Tougou basins, respectively.

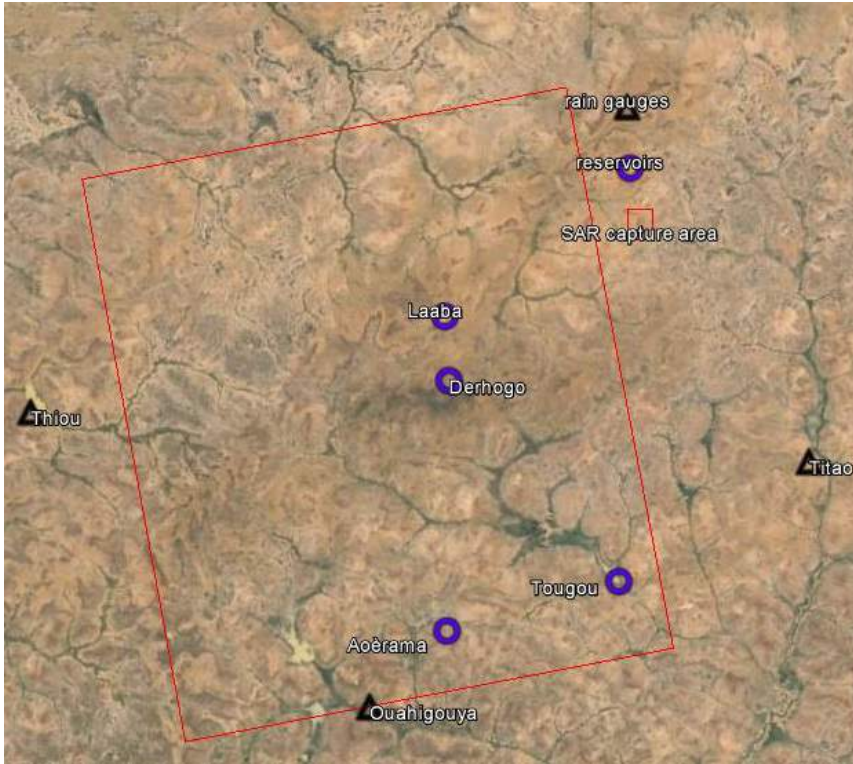


Figure 5.4. Map of the monitored sites: we selected two basins in the nearby of the city of Ouahigouya and the basins of Laaba, Tougou, Aoèrama and Derhogo for a total of six sites.

5.2.3 Basins bathymetry

Once the basin contour is known and re-projected in a cartographic system, the water volume contained into the basin can be computed considering each pixel of the water mask as a water column whose height h_{wc} is given by:

$$h_{wc} = h_c - h, \quad (5.1)$$

where h_c is the elevation of the equipotential surface identified by the basin contour, and h is the DEM height corresponding to the considered pixel. The DEM has been obtained via standard interferometric processing

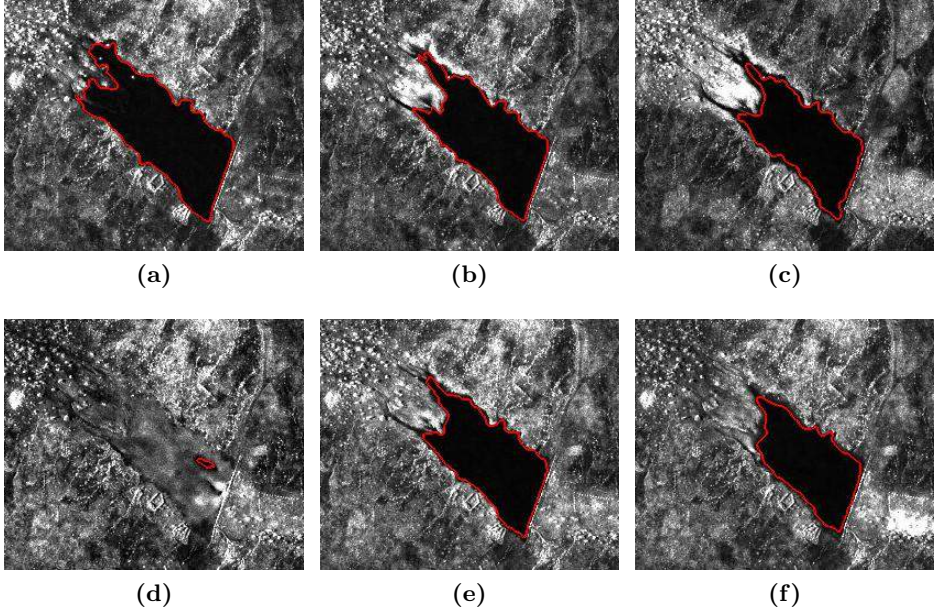


Figure 5.5. Laaba basin (Burkina Faso), shorelines for the acquisitions of (a) 2010/06/12, (b) 2010/07/14, (c) 2010/08/31, (d) 2011/03/27, (e) 2011/11/10, (f) 2011/12/12 (f). The water availability follows the cycle of rainfalls.

of a couple of images acquired at the peak of the dry season of the year 2011 and has a resolution of 9 meters (Amitrano et al., 2014c).

The water volume contained into the basin is given by the summation of all the elementary contributions brought by the water columns:

$$V = \sum_{i=1}^N S_i \times h_{wc_i}, \quad (5.2)$$

where S_i is the basin surface of the i -th resolution element belonging to the water mask and N is the total number of pixels belonging to the water mask.

In Figure 5.7 we show the results of the analysis for the six considered basins. As widely discussed in the following Section, these diagrams are strictly related to the seasonal variation of rainfall with an abrupt increment of volume at the beginning of the rainy season and a continue

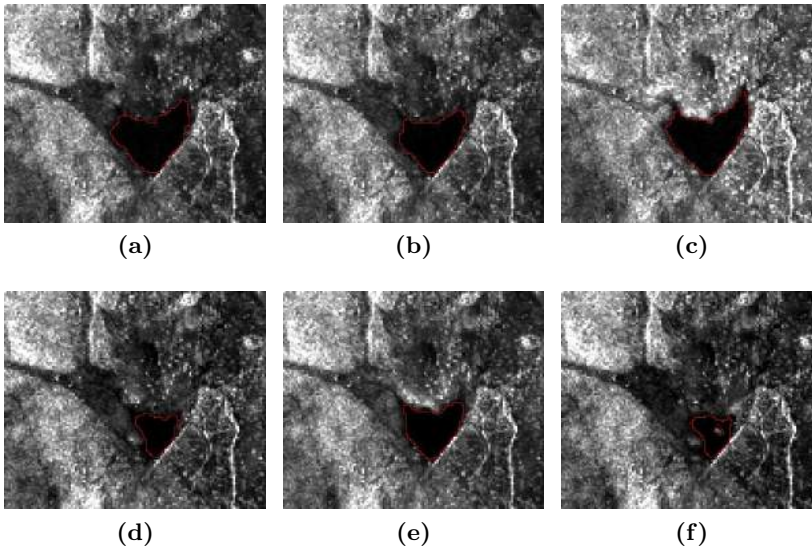


Figure 5.6. Derhogo basin (Burkina Faso), shorelines for the acquisitions of (a) 2010/06/12, (b) 2010/07/14, (c) 2010/08/31, (d) 2011/03/27, (e) 2011/11/10, (f) 2011/12/12 (f). The water availability follows the cycle of rainfalls.

decrease in the dry season.

Note that the Tougo basin does not dry up completely during the dry season (Amitrano et al., 2014a). Therefore, its bathymetry can not be estimated in the part of the basin covered by water during the acquisitions used for DEM retrieval. As a consequence, while the basin surface is correctly determined, the estimated volume represent the variation with respect to the minimum and not the entire available water volume.

In Table 5.1, dam positions and basins' maximum surface and volume availability are provided.

A relation between reservoirs' storage volumes and surface areas can be derived exploiting the obtained database. These relations are very useful since the reservoirs' surface areas can be always estimated with a good precision by satellite or aerial imagery, while the volume measurement requires more demanding and expensive bathymetric surveys.

Area-volume relations have been developed in literature, both theoretically and empirically. Based on an extensive bathymetric survey in

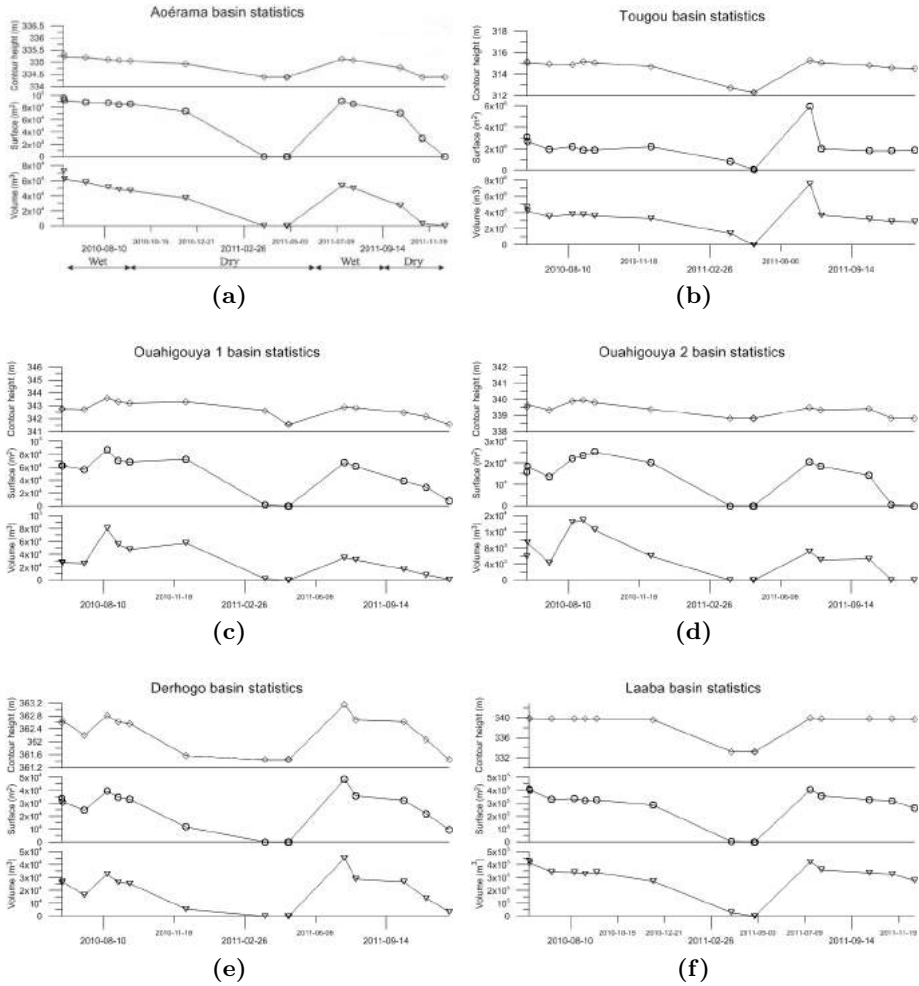


Figure 5.7. Contour height, surface and water volume for the six analyzed basins calculated for each available SAR acquisition. The monitored period goes from June 2010 to December 2011.

Upper East Region of Ghana, Liebe et al. (2005) obtained the following relationship:

$$\text{Volume} = 0.00857 \text{Area}^{1.4367} \left[m^3 \right]. \quad (5.3)$$

Table 5.1. Reservoirs database.

Reservoir	Position	Surface max (m^2)	Volume max (m^3)
Laaba	$13^{\circ}52'23.12''N, 2^{\circ}.20'44.12''W$	407430	427215
Tougou	$13^{\circ}40'48.68''N, 2^{\circ}12'52.72''W$	5918022	7511779
Ouahigouya 1	$13^{\circ}36'21.44''N, 2^{\circ}24'42.31''W$	86994	80286
Ouahigouya 2	$13^{\circ}37'78.96''N, 2^{\circ}24'57.97''W$	25029	14871
Aorama	$13^{\circ}40'10.61''N, 2^{\circ}21'4.66''W$	96147	72723
Derhogo	$13^{\circ}49'36.53''N, 2^{\circ}20'34.15''W$	48519	44954

The regression analysis of the reservoirs volumes and areas obtained using SAR-derived data allowed for the derivation of the following relation: (Amitrano et al., 2014a):

$$\text{Volume} = 0.10120\text{Area}^{1.1670} [m^3]. \quad (5.4)$$

The Tougou basin was not used in this derivation. In fact, as previously explained, it does not dry up completely, thus preventing the DEM extraction for the areas covered by surface water.

As shown in Figure 5.8, there is only a slight difference between (5.3) and (5.4). This confirms that, thanks to the morphological and morphometrical regularity of the regions, the area-based volume estimation is possible with good approximation.

5.2.4 Summary

In semiarid regions small reservoirs form a set of well-distributed and easily accessible water sources that are used for agriculture, domestic use, and livestock. Small reservoirs are widely used to reduce the people's vulnerability to drought and improve their livelihoods.

The retrieved area-volume relation is a quick and effective tool for water resources monitoring

In order to optimize the management of these reservoirs, the access to information is a crucial problem that calls for cheap, reliable and con-

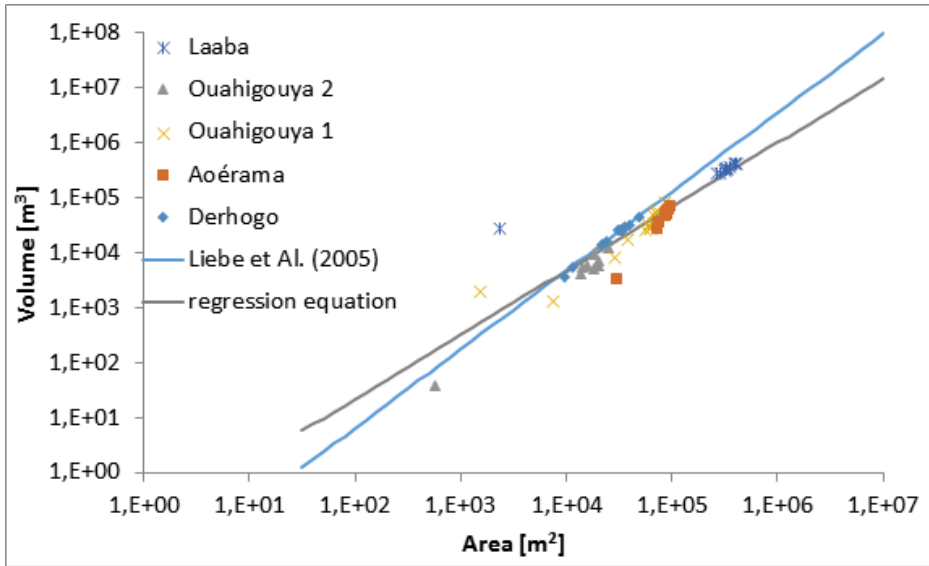


Figure 5.8. Reservoirs' storage volume as a function of their surface area in a log-log plane.

tinuous monitoring. The use of SAR sensors is, in this context, a very powerful tool thanks to their all weather and all time characteristics and high spatial and temporal resolution. They provided a detailed, continuous and wide observation of the territory that deserves to be exploited in regions where in situ measurements are impervious and expensive.

Multitemporal SAR processing of COSMO-SkyMed stripmap images allowed for retrieving the extension of reservoirs through a simple segmentation technique. The surface area information, combined with a high-resolution DEM obtained via standard interferometric processing, allowed for the estimation of the available water volumes for each of the monitored sites. These data was used for regression analysis of reservoir volumes and areas in order to retrieve an equation relating these two parameters. In fact, if the surface area can be easily obtained through SAR imagery, the estimation of the available volume requires expensive and difficult surveys. Therefore, the retrieved relation represents a quick and effective tool for obtaining information about the water resources in the study area.

5.3 Classification

5.3.1 Using simple classification tools with Level-1 α images

Introduction

Since the beginning of the SAR history, the scientific community aimed to extract information from data with the purpose to classify objects on the imaged scenes.

One of the first definitions for this activity was given by Harger (1973):

“Classification consists in the discrimination and identification of randomly reflecting areas, characterized by a certain reflectivity spectral density.”

This definition highlights two aspects:

- *Discrimination*, i.e. the separation of objects which share similar characteristics of backscattering at the operating wavelength. This operation is carried out at a physical/signal level;
- *Identification*, i.e. the process of the assignment of a label (or a class) to a group of scatterers identified as similar. This operation is carried out at a semantic level and is powered by the knowledge of the SAR backscattering mechanism.

A more articulate definition of classification is provided by Campbell and Wynne (2011):

“Classification is the assignment of objects [...] to classes based on their appearance on the imagery. Often a distinction is made between three levels of confidence and precision. Detection is the determination of the presence or absence of a feature. Recognition implies [...] that the object can be assigned an identity in a general class or category. Finally, Identification means that the identity of an object or feature can be specified with enough confidence and detail to place it in a very specific class.”

This definition highlights the *appearance* of the objects, and thus the physical models which allows for the interpretation of the world representation, i.e. the image, as filtered by the sensor.

In recent years, the increased volume of available data and calculating power raised the user requests concerning classification procedures. Thus, while in '70s and '80s the rules were essentially dictated by the SAR community and by research needs, today the pressure on this activity comes from a wide variety of professionals and scientists belonging to different disciplines who posed new user-requirements to the SAR community.

In this context, several outstanding solutions were found. Some authors exploited multitemporal datasets for extracting time-dependent variables used to discriminate the behavior of scene features (Bruzzone et al., 2004). This allowed to devise robust classification schemas, but the difficulties related to the implementation of the processing chain could limit their usage. Camps-Valls et al. (2008) coupled multitemporal SAR and optical data to set up a framework based on composite kernels. Davidson and Ouchi (2003) improved a segmentation scheme with the use of multitemporal data. Engdahl and Hyypä (2003) proposed to use multitemporal quantities to reduce the dimension of the original dataset, selecting the best-suited group of images for unsupervised classification. Quegan et al. (2000) analyzed the backscattering characteristics of a time series for forest mapping. Gaetano et al. (2014) applied a Markov random field framework for multitemporal data segmentation using an homomorphic transformation.

All these works used multitemporal data in order to improve the performances of previously introduced schema or to introduce new methods of data fusion. This allowed to better identify features which are assumed to be stable (i.e. the forest mapping presented by Quegan et al. (2000)), or which have a specific behavior along time with respect to a given multitemporal quantity (see Davidson and Ouchi, 2003). Here, we want to focus on the simplification of the classification schema, in which multitemporal data are exploited for building a product, i.e. the Level-1 α product, suitable to be treated with standard algorithms.

In general, given an application, different approaches lead to different solutions and answers to the following questions:

1. What is the best algorithm/technique to reach the goal?

2. Which parameters must be tuned for obtaining a more reliable output?

The answers to these questions dramatically change as a function of the expertise of the operator.

As for the first question, hundreds of algorithms/techniques with excellent results exist. An expert user could use his/her skills to select the best solution for the given application. A non-expert end-user looks toward solutions already implemented in some available software suite. Moreover, most of the existing methods suffer from the problem that both the end-user and the SAR expert have a limited control/vision of the operations that lead to the final product.

As for the second question, the great complexity of modern classification algorithms requires a non trivial expertise for tuning the required parameters.

However, the active participation of end-users in classification procedures represent an opportunity to improve the performances in terms of compliance to requirements of practical scenarios. This can be possible dulling the hurdles that they can encounter answering to the two aforementioned questions. Hence, it is necessary to include simple algorithms in a framework in which the knowledge/expertise required for classifying a scene is diluted and made accessible to the end-users. This is feasible via an extreme simplification of the classification schemas, starting from an adequate choice of the input data. This position is in countertrend with the past literature in which the development of new algorithms/techniques rather than the reasoning on input data is privileged.

Level-1 α products
can be exploited in
simple algorithms for
classification

In the following sections, we will show that Level-1 α products (Amitrano et al., 2015c), allows for the building of such a framework. It gives the opportunity to use simple, well-known user-oriented classification tools which, thanks to the characteristics Level-1 α imagery, provide excellent results on complex scenes (Amitrano et al., 2016). Thus, our framework satisfies the fundamental end-user requirements of simplicity, velocity, accuracy and repeatability.

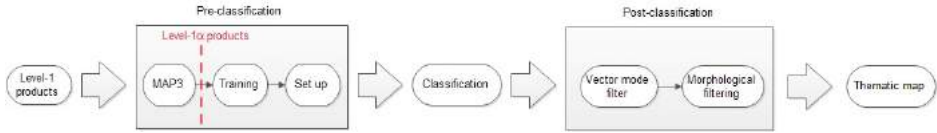


Figure 5.9. Block diagram of the adopted classification framework. Each block is constituted by a very simple operation. This is possible thanks to the characteristics of Level-1 α imagery.

Definition of the classification framework

In Figure 5.9 the block diagram of the proposed framework is shown (see also Amitrano et al., 2016).

It is articulated in three steps, pointing to the redistribution of the complexities from the classification phase to those of pre-classification and post-classification. This simple architecture is possible thanks to the usage of the Level-1 α products which are characterized by i) a high level of interpretability and ii) a good separability between the observable classes. The former characteristic makes these products very attractive for supervised classification procedures, since their interpretability makes easier the training phase. The latter characteristic allows to carry out the classification procedure with simple algorithms which require only few parameters to be set.

The architecture presented in Figure 5.9 is rather standard, at least in the first two stages (named as “pre-classification” and “classification” in the picture). Therefore, it is poorly suited for being applied to SAR data. However, the introduced novelties, concerning the input products for the classifier and the post-processing phase (in which two morphological operators are used) allows for obtaining performances fully comparable with those of definitely more complex algorithms. It is worthwhile to remark that Level-1 α imagery is obtained by multitemporal processing of N images, even if each product involves only two of them, as stated by Amitrano et al. (2015c). As all multitemporal processing, the quality of the products increases as the number of the available images increases.

The proposed approach looks towards temporal series of classification maps, introducing a novelty with respect to the past literature, which typically aimed at producing a single classification map by integrating the

information acquired by N acquisitions considered as stand-alone events. Nevertheless, when N acquisitions of the same scene are available, a series of $N - 1$ products and classification maps can be retrieved. This could be useful if a long-term environmental monitoring is required, since the changes between one acquisition and another can be recorded with a programmed revisit time. In the case of archive data the entire extent of the time series can be exploited to perform historic investigations and/or statistics. At semantic level, the detectable classes have a different meaning with respect to the classic SAR literature. In fact, our framework brings to a labeling which depends on the comparison of objects backscattering with respect to a pre-established reference situation, rather than on the temporal behavior of the scene with respect to some multitemporal quantities.

The block diagram presented in Figure 5.9 is quite generic. In the following subsections we will delineate a schema which allows to reduce drastically the pre-classification activities necessary for supervised classification procedure and, in general, introduces significant simplification in all the other blocks of the chain.

Pre-classification

Level-1 α products defined by MAP3 (Amitrano et al., 2015c) are particularly well-suited for supervised classification procedures. In fact, these products have the peculiarity of introducing a good separation between the classes, which can be identified using simple algorithms. Moreover, the training phase is simplified thanks to their high degree of interpretability, which allows to identify objects by simple and fast visual inspection. Conversely, this operation could be very complex and tedious when considering a single channel image, especially if the operator has no a priori knowledge about the scene.

The supervised approach requires the user to select representative training data for each of a predefined number of classes. Classification performance is highly dependent on users ability in modeling the class distribution. In this phase, users experience is crucial for identifying and locating the best training areas that, ideally, should be homogeneous and known, or, at least, recognizable (Townshend, 1981). The introduction of the Level-1 α products strongly mitigates these uncertainty factors since

they can be easily managed and interpreted also by nonexpert users.

Classification

In our framework the complexity of the classification algorithm is redistributed toward the phases of pre-classification and post-classification, allowing to avoid the usage of complicated decision rules. Therefore, in this work we used the maximum likelihood (ML) classifier. This choice was guided by the simplicity of the set up and its availability on commercial and open-source suites.

However, it is worthwhile to note that due to the strong inhomogeneity of the urban areas we exploit their high separability in Level-1 α products using an interferometric coherence-derived mask. In fact, as explained by Richards and Jia (2006), the ML classification used in this work assumes that the statistics for each class in each band are normally distributed and calculates the probability that a given pixel belongs to a specific class. In Level-1 α imagery, it is not possible to assume that the urban area has this type of distribution, being it very heterogeneous. Therefore, we used the interferometric coherence for identifying pixels belonging to this class. In fact, it is well known that urban structures have typically a high value of interferometric coherence, therefore they can be easily detected by thresholding this quantity.

Post-classification

The output of the classification procedure could be mottled, especially if complex scenes are analyzed. Hence, an adequate post classification step is required in order to obtain a more homogeneous output map. In this work, we propose a simple post classification phase composed of two steps: i) application of a mode filter and ii) of a morphological filter that generates the final output. These activities are described in detail in the following paragraphs.

The vector mode filter The mode filter is a reliable solution for filtering the output of the selected classifier, obtaining homogeneous classification maps with minimum alteration of its informative content.

The mode filter substitutes the central pixel of a sliding window with the mode computed within the window itself. However, the execution of that filter must take into account three requirements: simplicity, accuracy and velocity. In fact, the implementation of a sliding

The vector mode filter is a quick solution for compensating classification maps granularity

window running throughout the whole scene implies a high computational burden, which has to be carefully managed in order to reduce the execution times. As a matter of fact, for a $m \times n$ image and a square sliding window of dimension l (with l odd), the number of windows to be evaluated is $(n - l + 1)(m - l + 1)$. Hence, for example, if the scene is 5000×5000 pixels and the sliding window is a three pixels square side, the number of windows to be processed is 24980004.

A solution for implementing the sliding windows particularly suitable for vectorial languages is to exploit the Hankel indexing (Partington, 1989). To this end, assign to each element of the scene a linear index, for example proceeding along the rows of a $m \times n$ matrix, obtaining the matrix $\mathbf{I}: \mathbf{I}_{i,j} = j + (j - i)(i - 1) + n(i - 1)$, $i = [1, \dots, m]$, $j = [1, \dots, n]$. The purpose is to process the sliding window in vector form, i.e. to build an index matrix \mathbf{W} in which each row collects the elements of a window.

In order to get this matrix, consider the Hankel-like matrix of dimensions $k \times l$, where $k = n - l + 1$. In this matrix, the generic element a is such that $a_{i,j} = a_{i-1,j+1}$.

The matrix \mathbf{H} collects on each row the first l elements of the sliding windows involving the first row of the matrix \mathbf{I} . Therefore, we can place side by side to the \mathbf{H} matrix l columns with an index jump of n with respect to the elements $H_{i,l}$. This operation has to be repeated $l - 1$ times, i.e. once for each remaining row of the sliding window. Hence, the intermediate matrix \mathbf{T} has dimensions $k \times l^2$ and it is obtained as in the following pseudo-code:

```
FOR  $i = 1, l$  DO
     $\mathbf{T}(*, il\mathbf{u}_1 + \mathbf{v}_1) = \mathbf{H} + ni\mathbf{h}$ ,
END
```

where \mathbf{v}_1 is a vector, $\mathbf{v}_1 = [1, 2, \dots, l]$, \mathbf{u}_1 is a vector of dimensions $[1 \times l]$ whose elements are all equal to 1, and \mathbf{h} is a matrix whose elements are all equal to 1, i.e. $\mathbf{h}:\mathbf{h}_{i,j} = 1, \forall i \in [1, \dots, k], \forall j \in [1, \dots, l^2]$. We indicated with the asterisk $\mathbf{T}(*, \dots)$ that the operation involves all the rows of the matrix \mathbf{T} using a notation typical of many programming languages.

The matrix \mathbf{T} represents the sliding windows for the first l rows of the matrix \mathbf{W} . The last step is to replicate this matrix by rows k times inserting an index jump of l every k rows as follows:

```
FOR  $j = 1, k$  DO
     $\mathbf{W}(jk\mathbf{u}_2 + \mathbf{v}_2, *) = \mathbf{T} + nj\mathbf{t}$ ,
END
```

where \mathbf{v}_2 is a vector $\mathbf{v}_2 = [1, 2, \dots, k]$, \mathbf{u}_2 is a vector of dimensions $[1 \times k]$ whose elements are all equal to 1, and \mathbf{h} is a matrix whose elements are all equal to 1, i.e. $\mathbf{t}:\mathbf{t}_{i,j} = 1, \forall i \in [1, \dots, nw], \forall j \in [1, \dots, l^2]$.

Each row of the matrix \mathbf{W} is composed by the linear position index of the pixels belonging to a sliding window. The matrix dimensions are $nw \times l^2$, where $nw = n \times (l - 1)/2 \times m \times (l - 1)/2$, with l odd.

Applying the obtained index matrix to the input image, we get the sliding windows (in vector form) in which the mode has to be computed. We implemented an optimized procedure to accomplish this task, which is structured as below:

- Let N_C be the number of classes present within the classification map and build a three dimensional matrix \mathbf{C} of dimensions $nw \times l^2 \times N_C$ such that

$$\mathbf{C}(*, *, i) = \mathbf{Z} + i\mathbf{z}, \quad i = 1, 2, \dots, N_C, \quad (5.5)$$

where \mathbf{Z} is a matrix in which all the elements are equal to zero and \mathbf{z} is a matrix of the same dimensions of the matrix \mathbf{Z} whose elements are all equal to 1. The matrix \mathbf{C} is composed by a series of matrices in which all the elements are equal between them and equal to the position index of the element in the third dimension of \mathbf{C} . Hence, $\mathbf{C}(*, *, 1) = \mathbf{Z} + \mathbf{z}$, $\mathbf{C}(*, *, 2) = \mathbf{Z} + 2\mathbf{z}$ and so on;

- Let \mathbf{S} be the matrix of the sliding windows and replicate it in the third dimension N_C times. In such way, \mathbf{S} and \mathbf{C} are two matrices of dimensions $nw \times l^2 \times N_C$;
- Make the subtraction $\mathbf{A} = \mathbf{S} - \mathbf{C}$; in such way, along the third dimension of \mathbf{A} we find a zero when the position index coincides with the class indicated in the matrix \mathbf{S} ;
- Mark with one all the positions in which there is a zero in the matrix \mathbf{A} ;
- Introduce the matrix \mathbf{B} obtained by the summation by row of the matrix \mathbf{A} . This matrix have dimensions $nw \times N_C$. The values of \mathbf{B} along the columns indicates how many times each class appears in each window;
- Compute the maximum of the matrix \mathbf{B} by rows: the position index of the maximum of each rows indicates the mode of the window.

This vector approach allows to process a scene of size of the order of 4984×5831 in about one minute on a machine with 12 GB of RAM memory and 8 processors. The same operation carried out with no vectorization has been completed in about 8 minutes.

Morphological filtering Mathematical morphology is the most appropriate instrument for extracting image components that are useful in the representation and description of region shape, such as boundaries, skeletons and convex hull. Moreover, morphological techniques are also used for image preprocessing and/or postprocessing for filtering, thinning and pruning activities (Gonzalez and Woods, 2007). In this section, in order to optimize the classification map, we use a “closing” morphological operator with square structuring element. It is obtained through the application of the two basic morphologic operators of dilation and erosion. This filter tends to smooth sections of contours fusing narrow breaks and long, thin gulfs, eliminating small holes and filling gaps in the contour (Gonzalez and Woods, 2007).

Experimental results

In this Section, we applied the framework defined in Figure 5.9 to a Level-1 α product extracted from a set of COSMO-SkyMed images acquired over the city of Castel Volturno (Italy) in stripmap mode (3 m spatial resolution) and HH polarization between December 2009 and October 2011.

In particular, the product used for classification involves two images acquired on December 2009 (reference image, blue band) and August 2010 (test image, green band). Therefore, according with the rationale of Level-1 α imagery, this product is well-suited for detecting summer crops, since the enhancement of backscattering produced by the growth of plants and fruits (Fung, 1979) causes for this feature a dominant response of the green band in the RGB composite.

The validation of the classification procedure has been performed using an external ground truth derived from the CORINE land cover (Feranec et al., 2007). Seven classes (“Grassland”, “Summer crops”, “Urban areas”, “Woods”, “Winter crops”, “Water” and “Temporary water”) were extracted, as shown in Figure 5.10a, for a total extension in the order of 200000 pixels. However, it is remarkable that the class “Temporary water” does not exist in the CORINE land cover product. For this reason, a ground truth for such class has been built exploiting the SAR product.

The Level-1 α product of the study area is depicted in Figure 5.10b. The used training sets are depicted in Figure 5.11. The first experiment was the solution of the classical SAR classification problem, i.e. a four classes classification (Bruzzone et al., 2004) for the features “Water”, “Temporary water” (TW), “Urban areas” and “Woods/grasses/crops”. Results of this experiment are shown in Figure 5.12a and Table 5.2. The registered kappa coefficient κ and overall accuracy are 0.78 and 93.65%, respectively.

As a general comment, in Table 5.2, a moderate interclass confusion between the classes “Urban areas” and “Woods/Grasses/Crops” (WGC) is registered. This is due to the high resolution of the input maps which contrasts with the ground truth extracted from a land cover whose objective is to identify large homogeneous areas, which, in the case of urban areas, include other features (such as trees and grassland or gardens) detectable on the SAR product.

The model can be complicated separating “Woods” (prevalently lo-

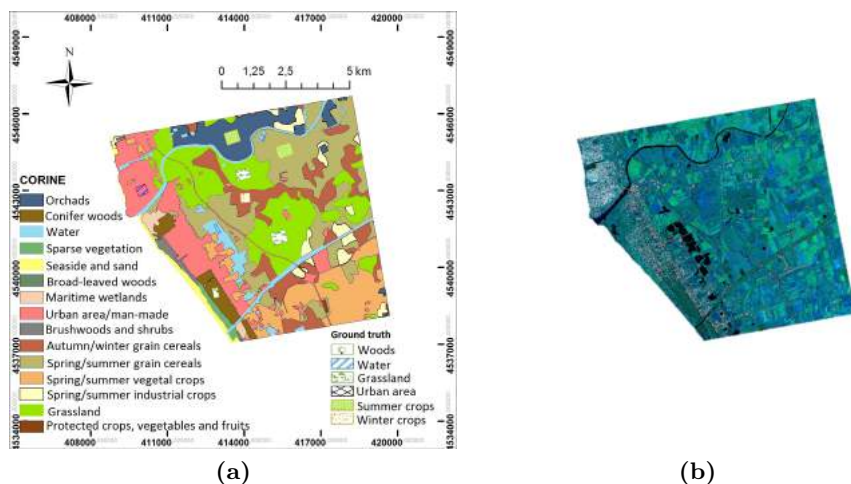


Figure 5.10. Castel Volturno, (a) CORINE land cover with the extracted ground truth and (b) the relevant Level-1 α product. Reference image: December 2009; Test image: August 2010.

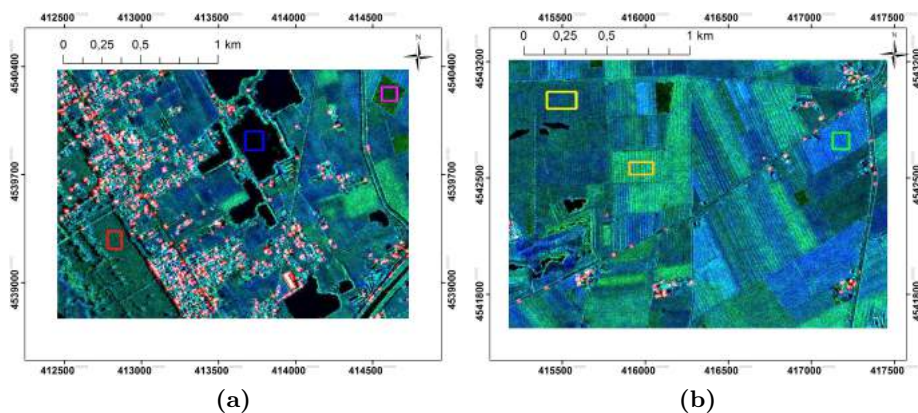


Figure 5.11. Castel Volturno, training sets: (a) woods (red), water (blue), temporary water (magenta) and (b) winter crops (green), grasses (yellow), summer crops (orange).

cated at south-west of our scene, as clearly observable in the Level-1 α product of Figure 5.10b) from “Grassland and Crops” (GC), as shown in Figure 5.12b and Table 5.3. The separation is quite good and in fact the

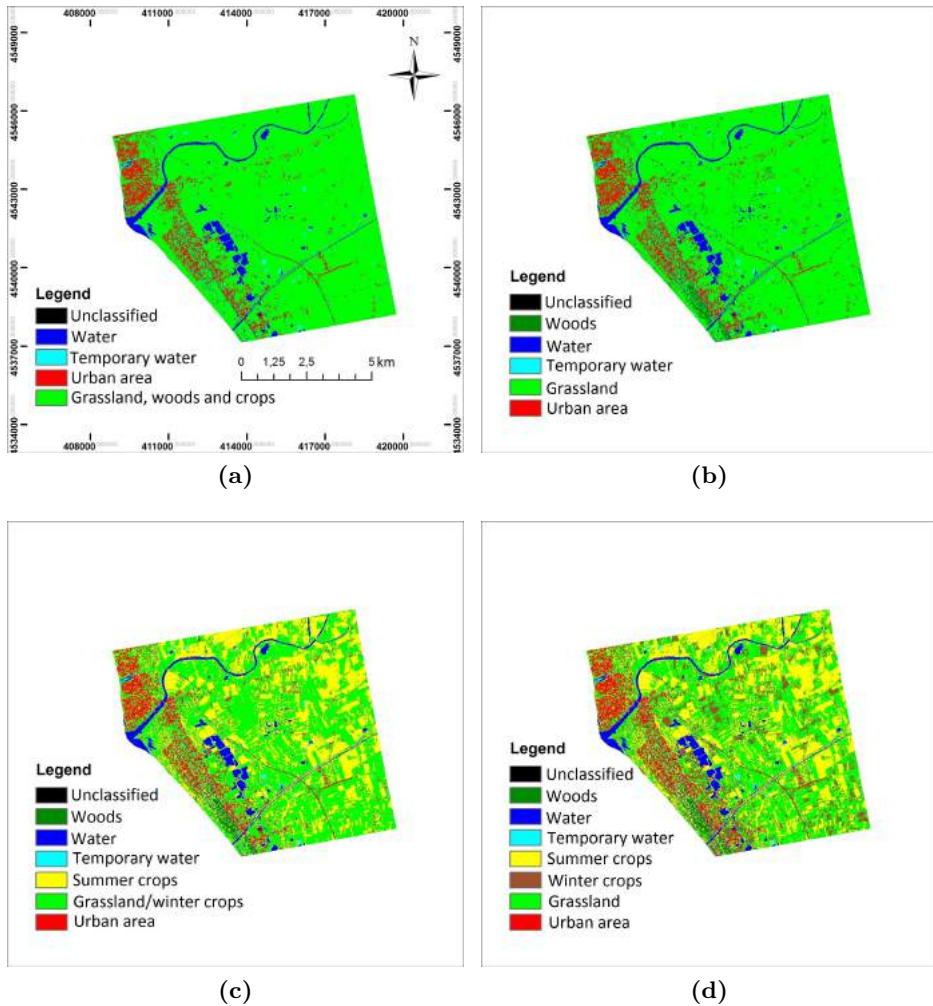


Figure 5.12. Castel Volturno, ML classification with (a) four, (b) five, (c) six and (d) seven classes.

registered κ and overall accuracy (which are 0.741 and 90.44%, respectively) are only slightly lower with respect to the four classes problem.

Results of the six classes experiment are shown in Figure 5.12c. In this case we separated “Summer crops” (Sc) from “Grassland and winter crops” (Gwc). This separation is possible thanks to the characteristics

Table 5.2. Castel Volturno, four classes classification confusion matrix. $\kappa = 0.78$, overall accuracy: 93.65%. PA: producer accuracy. UA: user accuracy.

	TW	WGC	Urban	Water	PA (%)	UA (%)
TW	89.75	1.30	0.52	0.47	89.75	59.84
WGC	10.25	96.61	33.46	0.19	96.61	95.97
Urban		1.13	65.44		65.44	86.97
Water		0.96	0.58	99.34	99.34	85.86

Table 5.3. Castel Volturno, five classes classification confusion matrix. $\kappa = 0.741$, overall accuracy: 90.44%.

	TW	GC	Urban	Woods	Water	PA (%)	UA (%)
TW	89.75	1.26	0.52	1.95	0.47	89.75	59.84
GC	10.25	94.72	33.18	30.85	0.19	94.72	93.64
Urban		1.20	65.44	0.01		65.44	86.87
Woods		1.99	0.28	64.31		64.31	68.67
Water		0.83	0.58	2.89	99.34	99.34	85.56

of the Level-1 α products. In fact, summer cultivations have a response which turns into green since the dominance of the test (summer) band with respect to the reference (winter) one. However, this splitting caused a fall of the κ and of the overall accuracy to 0.563 and 70.46%, respectively. This can be explained by the fact that if the harvesting has been already performed on some fields (in fact, the test band has been acquired at the end of August), their response loses the volumetric contribution given by the abundant presence of leafs and plants and turns toward that of a pasture. The complete confusion matrix for this experiment is reported in Table 5.4.

A further splitting is possible. In fact, winter crops should exhibit a higher backscattering in winter and, as a consequence, a response which turns into blue on the Level-1 α product. The results of this seven classes procedure are shown in Figure 5.12d. However this strongly depends on the type of cultivations and on the sowing time. In fact, in souther Italy, autumn/winter cereals are seeded usually in November due to the mild climate. The relevant fields not always reach a backscattering such as

Table 5.4. Caserta dataset, six classes classification confusion matrix. $\kappa = 0.563$, overall accuracy: 70.46%. Class abbreviations: TW: Temporary water, GWc: Grassland and winter crops, Sc: Summer crops. PA: producer accuracy, UA: user accuracy.

	TW	GWc	Sc	Urban	Woods	Water
TW	89.75	0.03	2.59	0.52	1.95	0.47
GWc		80.07	34.55	16.97	8.72	
Sc	10.25	17.36	57.22	16.21	22.12	0.19
Urban		0.86	1.58	65.44	0.01	
Woods		1.62	2.40	0.28	64.31	
Water		0.06	1.67	0.58	2.89	99.34
PA (%)	89.75	68.5	57.22	65.44	64.31	99.34
UA (%)	59.84	68.5	68.12	86.87	68.67	85.56

to be separated from grassland at the end of December. In fact, in the study area, this cultivations experienced their maximum growing stadium usually in May/June since the harvesting is performed at the beginning of summer. For this reason, the class “Winter crops” is highly absorbed by the “Grassland” one. The registered κ for this classification is 0.462 while the overall accuracy is 60.58%. The confusion matrix is omitted for brevity.

Summary

In this Section we introduced an end-user-oriented framework for the classification of multitemporal SAR data which exploits the characteristics of the Level-1 α products. The proposed framework makes use of simple tools well-known in literature and, fusing the concepts of classification and change detection, contextualizes its products in the field of the time series of classification maps. The extracted classes have a different semantic content with respect to the past literature since the detection is made by comparison of objects backscattering with respect to a reference situation rather by the analysis of the temporal behavior along the entire series.

The proposed approach exploits the peculiarity of the Level-1 α products: interpretability and class separability. The former property makes our framework particularly well-suited for supervised classifications, since the training step is simplified and fast. The latter property allows for using very simple classification algorithms. This allows for reduce the

complexity of the whole activity, making this schema particularly attractive for nonexpert SAR users. In fact, from the end-users' standpoint, the proposed framework fulfills the requirements of simplicity, repeatability, velocity and accuracy. In particular: i) the use of the Level-1 α optimize the the selection of the training sets; ii) the simplicity of the training step and of the decision rule ensures the classification procedure to be completed in a very short time. The post-classification phase also has a small computational burden, thanks to the vector form of the mode filter and to the mathematical morphology robustness; iii) the procedure is simply replicable either for different elements of the same time series or for other scenes with characteristics completely different; iv) the obtained results testify the potentiality of the proposed framework in terms of accuracy and reliability of the output maps.

Due to the simplicity of the schema, in many cases high level classes can be obtained. This is particularly useful in operative contexts when quick preliminary analysis are needed. In this optic, our framework acts as a pre-classifier whose results can be exploited for more detailed investigations.

Level-1 α products allows for designing end-user oriented classification frameworks, characterized by the usage of simple algorithms, giving excellent results on complex scenes

5.3.2 Crops monitoring using neural networks

In this Section, we explore the suitability of Level-1 β products with practical scenarios. In fact, visual data mining and interpretation are not enough to extract information at large scale from satellite images. As a consequence, it is highly desirable that such an easy to read data representation can be effectively processed with standard algorithms. It is quite clear that a more interpretable product is very well suited with supervised

procedures, as explained by Amitrano et al. (2016) and in Section 5.3.1.

However, the recent remote sensing literature paid great attention to neural network techniques because they can handle effectively large multidimensional datasets. Therefore, in the following we will linger on the usage of such techniques, with particular reference to one of the most consolidated and widespread, i.e. the Kohonen's Self-Organizing Map (Kohonen, 2001).

Objective and approach

The problem we consider is the evaluation of the agricultural production in two successive years. To this end, we used two calibrated Level-1 β products (obtained considering images belonging to the summer season of the years 2008 and 2009) concerning the city of Ciró Marina (see Section 3.5.1). These images are stacked in an unique six-bands data and used as input of a randomly initialized SOM to classify. Presenting the two year values together allows the network not only to classify the images based on the characteristics of the pixels, but also on their relative change during the two years. As a consequence, the changes in the two following years detected by the SOM can be interpreted as changes in the crops behavior or production.

However, this application is critical since it involves two different characteristic times concerning the building of the single multitemporal product and the two-years cropfields behavior classification. Therefore, we considered a strategy in which the SOM clustering was supported by hierarchical agglomeration and object-based reasoning. This allowed for reconstructing a more homogeneous clusters map and a better understanding of study area.

Level-1 β products can be used in combination with neural networks to classify

In particular, we used an agglomerative hierarchical clustering method to unify similar classes after classifying the input Level-1 β products in 64 categories through an overdimensioned 8×8 SOM (Gonçalves et al., 2011). Therefore, a top-down model was followed (Matsuyama and Hwang, 1990). In fact, our world model suggests that 64 categories are too much to

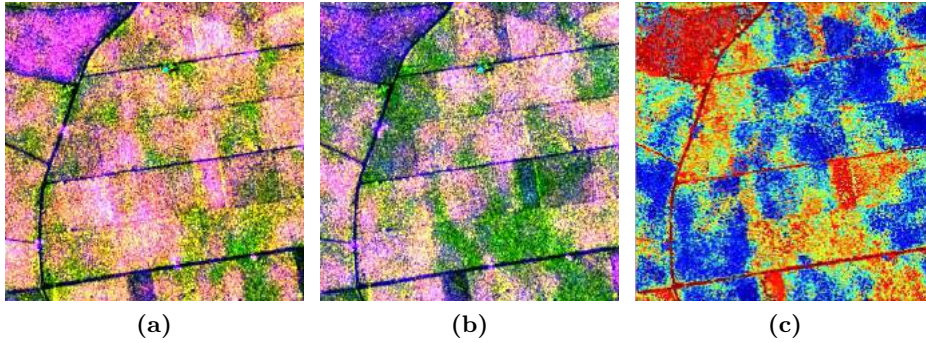


Figure 5.13. Ciró Marina (Italy): Level-1 β products relevant with years (a) 2008 and (b) 2009. (c) Clustered product. The 64 clusters identified by the SOM depicted in Figure 5.14 (counted from top to bottom and then from left to right) are associated to 64 different colors.

describe the dynamics of our scene. As a consequence, the initial number of clusters will be reduced according to the characteristics of the scene and to the application we are dealing with, reducing the clusters fragmentation at the same time.

In Figure 5.13a and Figure 5.13b the two subsets of the Level-1 β products relevant to the year 2008 and 2009 are depicted. SOM clusters are reported in Figure 5.13c. They have been obtained using 1000 training elements and 200 epochs. In Figure 5.14, a realization of the 8×8 Kohonen network is provided. Each element is represented by a plot showing the change of RGB values in the 2008 and 2009 years.

The clusters map depicted in Figure 5.13c, although oversegmented, has a smooth representation because at neighbouring pixels, having a similar behavior, are associated neighbouring classes, thus similar colors. In fact, observing the Kohonen network of Figure 5.14, the reader can see that adjacent nodes represent similar patterns in the pixel behavior between the two years.

Agglomerative hierarchical clustering and object-based reasoning

In order to reduce the map fragmentation, we performed an agglomeration of the obtained clusters to fuse those carrying similar information.

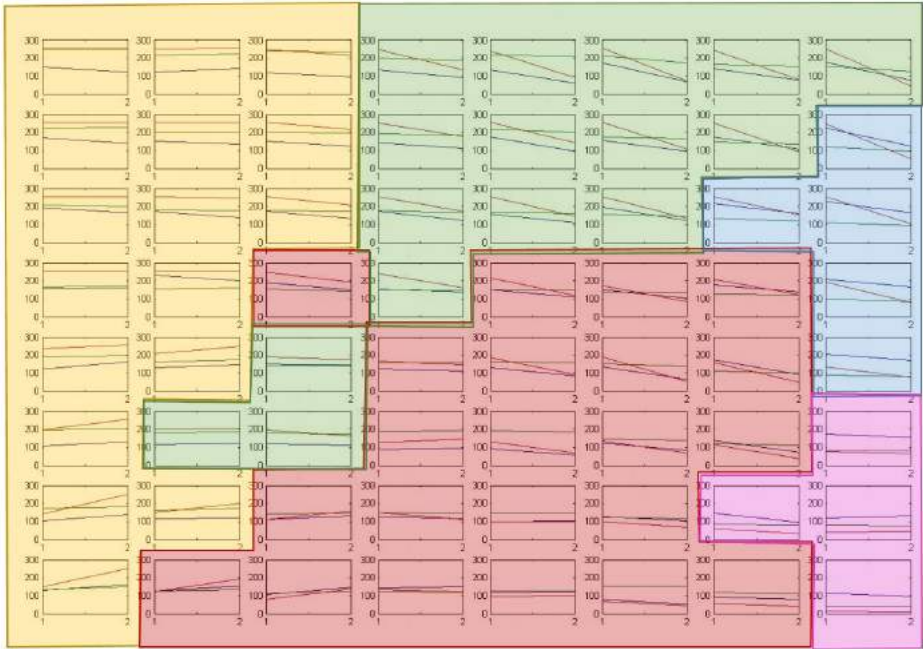


Figure 5.14. 8×8 SOM grid showing the RGB values for year 1 and their change to year 2. The colored rectangles represent the clusters after hierarchical agglomeration (see the dendrogram in Figure 5.15).

A dendrogram represents a quick solution for evaluating the relation between SOM clusters. In Figure 5.15, it has been built adopting as metric the Ward's distance J. H. J. Ward (1963). The red dashed line represents the distance under which dendrogram's branches are considered for clusters fusion. Clusters interested by fusion have been displayed in yellow, green, magenta, blue, and red. This representation is consistent with the SOM grid depicted in Figure 5.14, in which the colored rectangles enclose fused clusters, for a total of five categories. Looking at the dendrogram, the output of the fusion process is given by the intersection between the horizontal red dashed line and the black leaves above the colored groups of clusters.

The output of the hierarchical clustering is depicted in Figure 5.16a, in which an unpleasant granularity arises. It consists in small regions

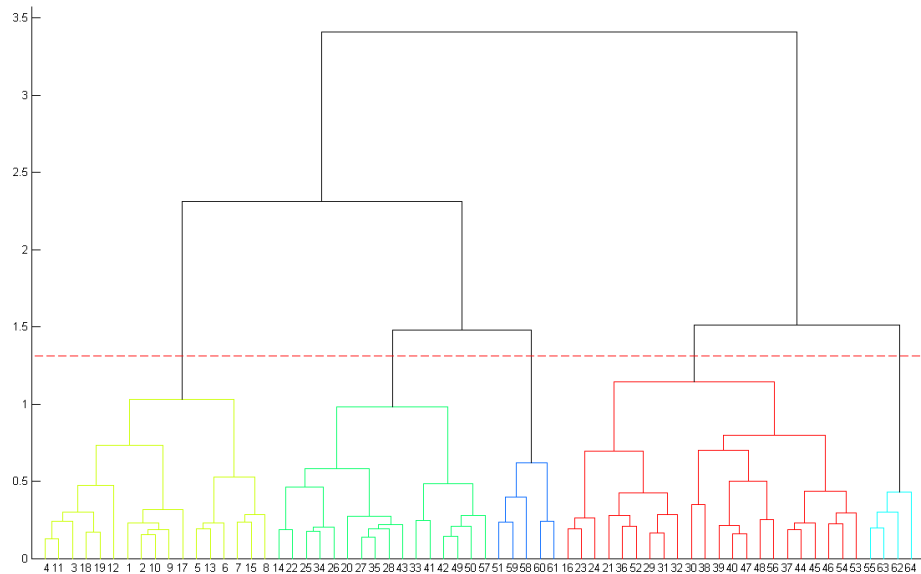


Figure 5.15. SOM clusters dendrogram. The red dashed line represents the distance under which diagram’s branches are considered for fusion. Clusters interested by the fusion are depicted in yellow, green, blue, red and magenta (see also the SOM grid depicted in Figure 5.14).

mainly sunk into an uniform background. In this picture, the cluster-color association is consistent with the dendrogram of Figure 5.15 and with the SOM grid of Figure 5.14, in which the semi-transparent colored rectangles enclose the agglomerated clusters.

Physical-based consideration and object-based reasoning can be used for improve the quality of the obtained clusters map.

We propose a simple processing based on connected components labeling (Shapiro and Stockman, 2002). This algorithm assigns an increasing numeric index to each connected region found within the image. For these objects, as suggested by Shackelford and Davies (2003), some shape parameters can be computed, as well as spatial relation between them. Here, we proposed to reason basically on few parameters, such as area, compactness, number of neighbours and percentage of shared borders between adjacent regions. These parameters have been used to generate an

appropriate physical-based and application-tailored set of rules. In this case, we want to reconstruct the homogeneity of the clusters representing the behavior of the cropfields. Therefore, the implemented rules basically aim at fuse small objects (also with irregular shape) to the background.

In Figure 5.16b, the result of the above described procedure application is depicted. It is quite evident that the fusion operation has definitely improved the quality of the clusters map. In fact, with respect to the map presented in Figure 5.16a, the number of regions is dramatically reduced, decreasing from 12529 to 207, also bringing undebatable benefits to the physical interpretation of the map.

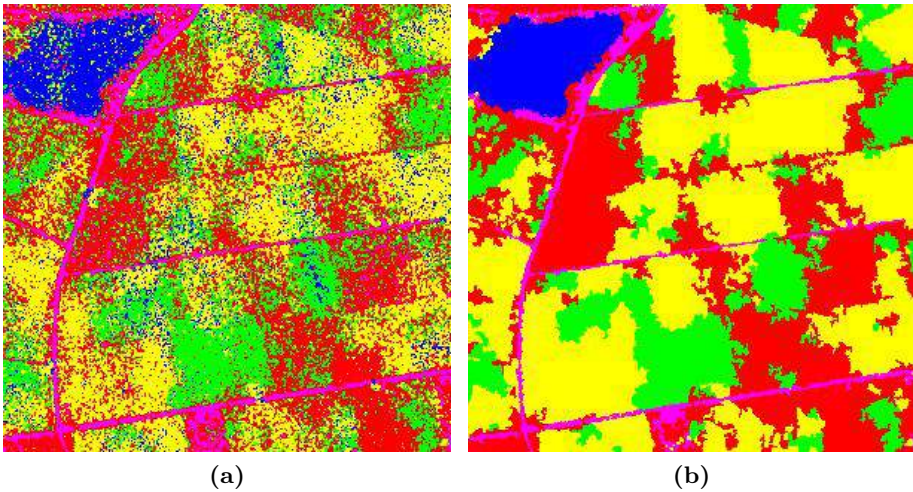


Figure 5.16. Clusters map (a) after the hierarchical agglomeration and (b) after the object-based region fusion. The cluster-color association is consistent the dendrogram of Figure 5.15 and with the SOM grid of Figure 5.14.

Interpretation

The physical interpretation of the retrieved clusters is now in order.

From the available data extracted from the CORINE land cover, it results that the study area is mainly destined to vineyards. Thus, observing the original Level-1 β subsets depicted in Figure 5.13a and Figure 5.13b, the clustered image reported in Figure 5.16b and the SOM grid of Fig-

ure 5.14, it is possible to infer the following:

- **Yellow cluster:** The RGB values in this cluster are medium to high, constant between the two years or with a slightly increasing red value. The red (variance) and the green (mean) values are generally higher than the blue (saturation index). This means that the yellow cluster is associated with agricultural areas that have a pink/yellow color in the Level-1 β product in 2008 and remained similar in 2009, or slightly turned to pink when they were yellow in the first year. These areas exhibit an almost constant behavior in the two examined years. Assuming that the study area is destined to vineyard, as stated in the CORINE land cover, we can argue that this cluster represent the fields in which the optimum behavior of the cultivation is reached. In fact, vineyards experience their full development between April and October, which is actually the period considered for the building of the Level-1 β products used for the classification (see Section 3.5.1). Thus, in this period the growth of the cultivation results in high values of the considered quantities;
- **Green cluster:** In this cluster the green value is medium and remains constant, but red and blue values are generally decreasing from high to low. This is associated with areas that were pink in the 2008 Level-1 β product and turned into green. Physically, the decreasing in variance and saturation index could be related to a smaller development of the cultivation, thus to a lacking of a volumetric contribution due to an increased foliage density and fruits growth. However, since the contribution of the mean remains high in both years, despite the decreasing in variance and saturation index, allows for arguing that the cultivation is structurally characterized by an almost dense foliage, such as a vineyard, whose structure and dynamics are fully compatible with the phenomenology inferable from the cluster analysis;
- **Red cluster:** The RGB values are constant or slightly decreasing, but generally lower than the values in the Yellow cluster. This cluster is mainly associated with green areas which remains almost constant in the two years. From the above considerations, we can infer that the behavior of the vineyards grouped in this cluster is not optimal

in both the considered years;

- Blue cluster: The areas in this cluster have high blue and red in the first year, decreasing in the second year, while the green remains constant. The red generally decreases more than the blue. This cluster is associated with areas that were pink in the 2008 Level-1 β image and turned into blue in 2009. It is mainly located in the field at the top left corner in Figure 5.13a, Figure 5.13a and Figure 5.16b. It represents an anomaly with respect to the behavior of the previously analyzed cluster. In fact, the abrupt fall in variance and the low values of the mean bring us to argue that this cluster is a fallow land on which some activities has been performed in the year 2008;
- Magenta cluster: these areas have very low values of RGB. They are almost constant in the two years. This is associated with dark blue areas in the Level-1 β products. The very small variance indicates that the pixels are quite stable. In fact, looking at the shapes, it is clear they are associated to roads.

The outcomes of clusters interpretation are summarized in Table 5.5. In this table, RGB attributes are referred to the 2008 Level-1 β product and their derivatives are qualitatively evaluated basing on the SOM grid depicted in Figure 5.14.

Table 5.5. Summary of clusters interpretation outcomes. RGB attributes are referred to the 2008 Level-1 β product. In table headings, t refers to time. In the table body, H stands for high, M for medium and L for low.

Cluster	R	G	B	R(t)	G(t)	B(t)	Interpretation
Yellow	H	H	H	=	=	=	Optimal crops behavior
Green	H	M/H	M	↓↓	=	↓	Non-optimal behavior (year 2)
Red	M	M	M/L	↓	=	=	Non-optimal behavior (both years)
Blu	H	L	H	↓↓	=	↓	Fallow lands
Magenta	L	L	L	=	=	=	Roads/Low scatterers

Summary

In this Section, we tested the reliability of Level-1 β products to be used in combination with neural networks. In particular, a complex two-years crops classification was performed by running a Kohonen's SOM on a stack of two calibrated Level-1 β products. Hierarchical agglomerative clustering and object-based processing was adopted to make the output of the SOM more homogeneous and to reconstruct the clusters meaning. In such way, five clusters was extracted, representative of the behavior of the cultivations in the two considered years with respect to the assumed reference situation.

5.3.3 Land cover mapping

In this Section, Level-1 β products will be used in combination with SOM for land cover mapping. In this case, as detailed in Chapter 4, SOMs offer the possibility of an immediate semantic transferring from the Level-1 β product to the classified map. In fact, during the training, the SOM (randomly initialized) nodes are updated to be representative of training elements, which are chosen within the RGB triplets constituting the input product. Thus, the resulting node colors will have the same semantic of the Level-1 β products.

Experimental results

We performed a land cover classification on a subset of the Sentinel-1 Level-1 β product presented in Section 3.5.2 relevant to the Dresden city area (Amitrano et al., 2015a). We used a 2×2 SOM grid for producing a 4-class land cover product. The original Level-1 β product and the SOM land cover classification are shown in Figure 5.17 and Figure 5.18, respectively.

As stated above, the use of a SOM allows for transferring immediately the semantic from one product to another. In particular, the interpretation of the class meaning, following the guidelines dictated in Section 3.5, is provided below:

- Blue Dodger class - Urban area: This class corresponds with cyan areas in the Level-1 β product due to the high contribution of the mean and of the interferometric coherence;

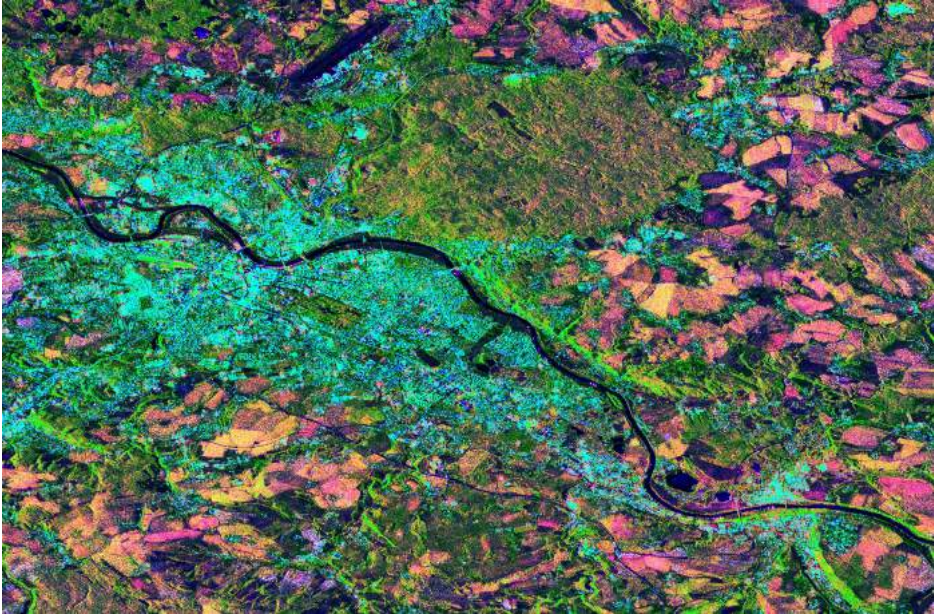


Figure 5.17. Dresden area: (a) Sentinel-1 Level-1 β product.

- Light green class - Grasslands: It corresponds to light green areas in the Level-1 β product, showing low variance and saturation index and a higher contribution of the mean due to a backscattering contribution of terrains;
- Orange class - Growing crops/vegetation: This category includes areas characterized by tones from yellow to pink in the Level-1 β product thanks to a medium/high contribution of all the considered indicators;
- Dark green class - Low scatterers: This is the more heterogeneous class, enclosing the darker objects of the scene (water bodies, bare soils and shadows).

The SOM land cover classification outcomes are summarized in Table 5.6.

This experiment output a very heterogeneous cluster (the Dark green one) enclosing objects of different nature. Thus, with respect to the scene

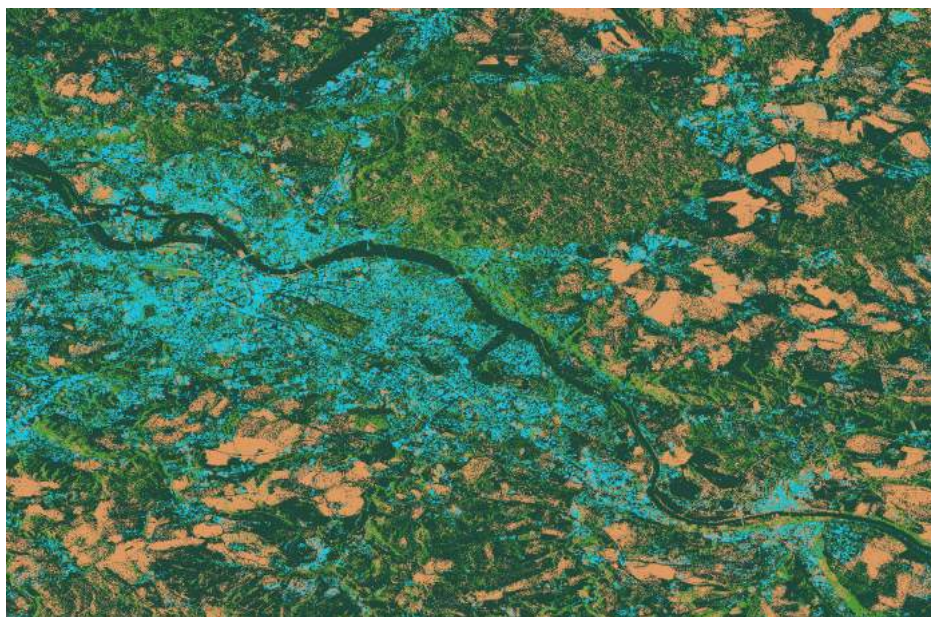


Figure 5.18. Dresden area, 4-class land cover map obtained through SOM clustering of the Level-1 β product depicted in Figure 5.17. The SOM allows for an automatic transferring of the Level-1 β product semantic to the classified product through the preservation of the chromatic content of the input RGB product. However, a very heterogeneous cluster has been obtained (the Dark green one); therefore it is desirable to split it into two more meaningful clusters.

Table 5.6. SOM 4-class land cover summary.

SOM/Map color	Level-1 β color	Interpretation
Blue dodger	Cyan	Urban areas
Marine green	Light green	Grasslands
Saffron	Pink/Yellow	Growing vegetation
Dark green	Dark green	Low scatterers

analyzed in Section 5.3.2, the problem is inverted, since we could desire to split this cluster in at least two more significant categories.

To this end, we analyze the reasons that led to the association of black and dark green objects in the Level-1 β product within a unique cluster.

As explained above, in the classic Kohonen's scheme, the SOM nodes are trained with RGB triplets randomly selected within the input product. However, looking at Figure 5.18, it arises that the "black object class" is the less represented in the datum. Therefore, it is highly uncertain that a SOM node can gain such color, since few training set relevant to this class are presented to the network.

The flexibility of a SOM, joined to a knowledge about the mapping of the real world into Level-1 β imagery, allows to easily solve this problem. In fact, in a Level-1 β product, an insufficient presence of an object class can occur mainly for water surfaces (as in the analyzed case) and urban areas, if an extended natural scene is considered. In the presented experiment, this caused substantially the aggregation of the class water bodies into the class woods, which is the closer for chromatic characteristics.

The SOM allow for an automatic transferring of the Level-1 β product semantic to the classified product through the preservation of the chromatic response of the input RGB product

In order to split this class into its two major features, it is sufficient to force a significant number of training sets (in the order of 15-20% of the total) to point toward the less represented class. In such way, we induced one of the SOM nodes to assume the corresponding color. In this case, we used a 2x1 SOM, acting the modification of the training sets toward the black color after masking out all the classes except the Dark green one. It is worthwhile to note that when the dimension of the SOM is reduced, the neighbor influenced by the winner neuron should be modified accordingly.

In Figure 5.19, we show the final 5-classes land cover map, after the application of the above described split of the Dark Green class into a Dark Green cluster (again) representative mainly of woods and a Black cluster relevant to water surfaces (see also Table 5.7). It is remarkable how the land cover map looks very similar to the input Level-1 β product.

However, as highlighted in Table 5.7, the Black cluster does not include only water surfaces, since other features, such as shadows and bare soils, are chromatically similar to them.

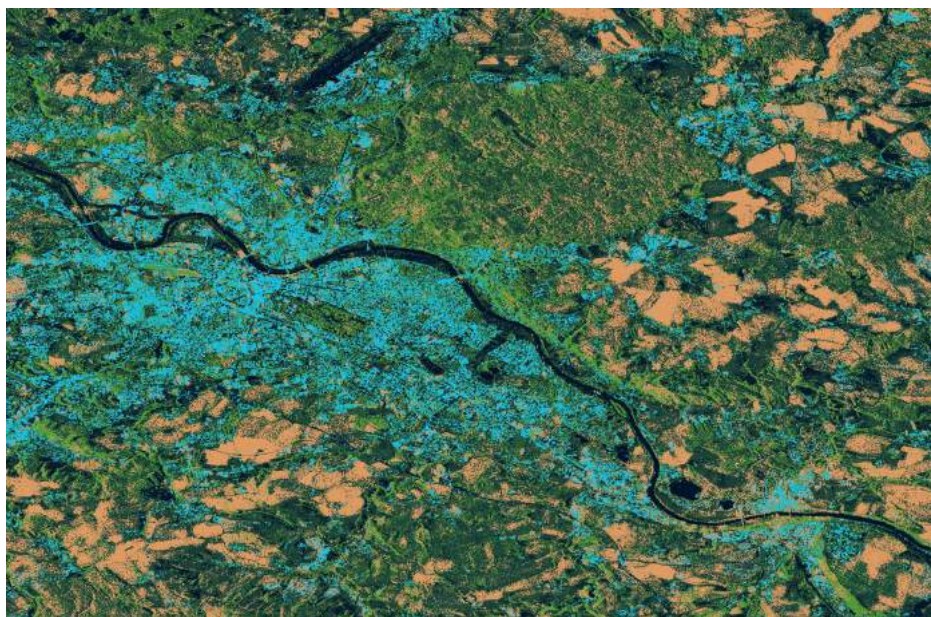


Figure 5.19. Dresden area, 5-class land cover map obtained after the splitting of the Dark green cluster visible in Figure 5.18. The retrieved classes are: “Urban areas” (Blue dodger cluster), “Grasslands” (Light green cluster), “Growing vegetation” (Saffron cluster), “Woods” (Dark green cluster) and “Water/shadows/bare soil” (Black cluster).

Table 5.7. SOM 5-class land cover summary after the split of the Dark green class.

SOM/Map color	Level-1 β color	Interpretation
Blue dodger	Cyan	Urban areas
Marine green	Light green	Grasslands
Saffron	Pink/Yellow	Growing vegetation
Dark green	Dark green	Woods
Black	Black	Water/shadows/bare soil

Summary

In this Section, we exploited Level-1 β products in combination with SOM for automatic land cover mapping. The most important outcome of this ac-

tivity is that the SOM allows for the preservation of the chromatic response of the input RGB product. Therefore, it makes possible an automatic and immediate transferring of the semantic content from the Level-1 β product to the classified product (see Chapter 4 for details).

5.4 Features extraction

5.4.1 A seasonal water index based on Level-1 α products

Introduction

Information about the extent of water surfaces is fundamental for water resource monitoring (Frazier and Page, 2000; Amitrano et al., 2014a) and flood prevention and resilience (Brivio et al., 2002; Martinez and Le Toan, 2007; Mason et al., 2012; Iervolino et al., 2015). Therefore, the past remote sensing literature devoted great attention to this topic, developing several methods to delineate water bodies and enhance their presence.

As stated by McFeeters (1996), these methods generally make use of reflected solar radiation or active microwave systems.

As for reflected solar radiation-based methods, Xu (2006) and McFeeters (1996) classified them in single-band methods and multi-band methods. In the first case, near-infrared radiation (NIR) is used to retrieve the water body surface since it is strongly absorbed by water and highly reflected by vegetation and terrain. A threshold is then chosen in order to separate the water class from the rest of the scene. However, the usage of a single band could induce the analyst in a wrong choice of the threshold, since moist and bare soils could exhibit low reflectance values at NIR frequencies. Hence, the coupling of NIR and visible frequencies has been proposed in order to enhance the presence of water surfaces at the expense of terrestrial features leading to the formulation of the Normalized Difference Water Index (NDWI) and its evolutions (McFeeters, 1996; Xu, 2006). These parameters are extremely popular in the end-users community and employed in a number of applications because they allow for a quick identification of water surfaces by mean of a simple bands ratio (Roshier et al., 2001; Davranche et al., 2010; Sakamoto et al., 2007; Langford et al., 2008; Ouma and Tateishi, 2006; McFeeters, 2013).

Multispectral data allows for the exploitation of simultaneous acquisitions made at different wavelengths which interacts differently with scene's features. This makes possible to define spectral indices, such as the NDWI in the case of water features, whose amplitude is related with some physical quantity linkable to the feature (for the NDWI this quantity is the turbidity McFeeters, 1996). These indices are extremely powerful since they are easily manageable and understandable even by non-expert remote sensing users, making them highly suitable for operative contexts.

Level-1 α products allows for the introduction of a bands-ratio methods, providing a simple and user-oriented tool for the detection of temporary water bodies in a change-detection framework

As for active microwave systems-based methods, in particular SAR, the problem of water body detection can be contextualized in the field of segmentation (Niedermeier et al., 2000; Deng and Clausi, 2005; Gaetano et al., 2014; Martinis et al., 2015). In this case, the main advantage is the independency from weather and illumination conditions. The new sensors generation added two fundamental characteristics, i.e. the high resolution and the short revisit time. All these features make SAR sensors very attractive for water resource management and flood monitoring, as testified by the wide literature on these topics. Anyway, SAR scattering mechanism makes the identification of water surfaces an open challenge which is still solved with application-oriented, time consuming and parametric techniques. For this reason, end-users facing a water-related application often prefer optical datasets, which are more easily interpretable and transformable in a Level-2 product carrying a physical information.

Level-1 α products Amitrano et al. (2015c) allows for introducing bands-combination-based methods also in the SAR world. In fact, the availability of multiple shots of the same scene allows the SAR-based methods to move towards change detection. This field has been already explored in the past literature which proposed valuable solutions, especially with regard to flood mapping Dellepiane and Angiati (2012). However, the for-

mulation of a water index derived from SAR data has not been reached yet, leaving a gap between the SAR theory and operative contexts which are principally populated by multidisciplinary users.

In this Section, we approach the problem of water bodies detection in a change detection framework. The proposed method is suitable to be applied in semi-arid environments for extracting small reservoirs contour (Amitrano et al., 2015b). Computational complexities and the required SAR expertise (which often discourage non-expert users) are moved in the product formation phase, making the feature extraction immediate and easy (Amitrano et al., 2013). In such way, SAR data can definitely encounter the necessities of end-users/costumers, which are principally interested in the possibility of operate with simple techniques, with few parameters to set.

Detection of seasonal water in semi-arid environment

Semi-arid regions are characterized by a particularly vulnerable environment due to the scarce rainfalls which, allowing the growth of vegetation coverage for limited period of the year, make terrains very inclined to erosion phenomena (Diamond, 2005). The combination of water scarcity and soil erosion has serious implications on terrains productivity and therefore on the available food quantity per capita in countries where the demographic pressure is the highest throughout the world.

The international community is well-aware of these problems and many initiatives for their monitoring, prevention and mitigation have been promoted. In particular, the TIGER initiative launched by the European Space Agency represents the most relevant program of the last decade concerning water resource management supported by remotely sensed (almost optical) data.

In a change-detection framework, the presence of surface water can be detected by comparing the electromagnetic response of a reference image with that of the test one. Reference situation can be identified during the dry season, when ponds are almost completely dry. Using multitemporal Level-1 α RGB composites, as explained in (Amitrano et al., 2015c), it is convenient to place this image on the blue band. In such way, placing the test acquisition on the green band, areas covered by seasonal water (i.e. covered by water only during the test image acquisition) will be displayed

in blue color due to the dominant response of the dry season image.

These above consideration lead us to define a seasonal water presence pseudo-probability as follows:

$$\text{SWPP} = \hat{G}^2 \left(\frac{B - G}{B + G} \right), \quad \text{SWPP} \in [-1, 1], \quad (5.6)$$

in which SWPP is the acronym for Seasonal Water Pseudo-Probability and

$$\hat{G} = \left(1 - \frac{G}{255} \right), \quad (5.7)$$

is the complement of the green band.

In (5.6) the first factor at the right hand side enhances the value of the pseudo-probability for those pixels in which there is a lack of response in the green band, i.e. pixels interested by surface water. The second factor at right hand side attenuates the value of the index for the pixels for which the response of the two bands is comparable. At the same time, pixels experiencing a dominant response in the green band (such as vegetated areas) will assume a negative value of the SWPP, while pixels for which the blue band is dominant are pushed toward positive values. Hence, the SWPP values of the pixels candidate to be classified as surface water are expected to be close to 1.

Physically, the proposed pseudo-probability is related with the state of the terrain during the acquisition of the reference scene. In fact, the higher the terrain response during this acquisition, the higher the value of the SWPP.

In Figure 5.20 we show some results of the application of (5.6) to some small reservoirs of the study area, located in Burkina Faso. In particular, in the first column of Figure 5.20, we reported the subsets of the Level-1 α product concerning the considered reservoirs. In the second column, the reservoir contour has been manually extracted. In the third column the SWPP map of the area (which covers in all cases approximately 1.5 km²) is shown. Finally, in the fourth column, a binary mask obtained via thresholding of the SWPP map at 0.3 is depicted. Mathematical morphology has been applied in order to make more homogeneous the final mask and remove islands (Amitrano et al., 2016).

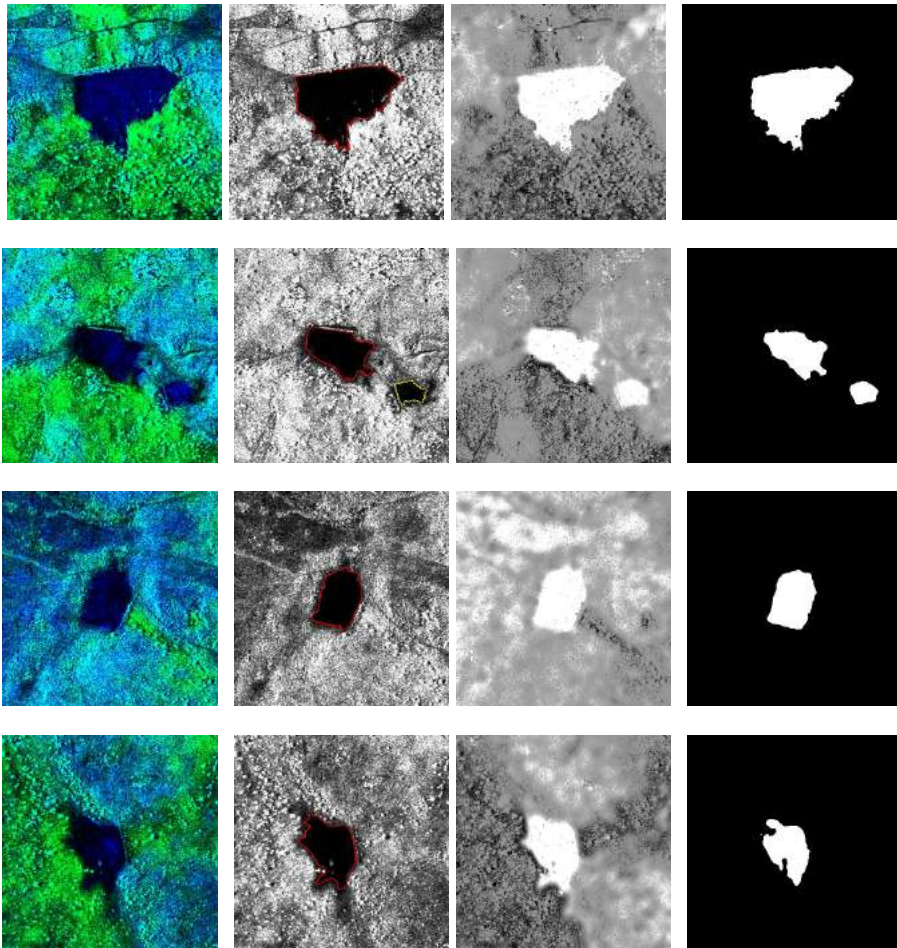


Figure 5.20. Burkina Faso: results of the application of the SWPP scene for five small reservoirs. First column: Level-1 α products (Reference image: 28 April 2011, Test image: 31 August 2010), second column: manually extracted water surface mask, third column: SWPP map, fourth column: binary mask obtained thresholding the SWPP map at 0.3.

This representations allows for qualitatively appreciating the performances of the proposed approach, whose results show a good agreement with the manually extracted ground truth, thus providing a quick solution for the individuation of water bodies of (almost) any dimension, compat-

ibly with the suitability of the input Level-1 α product.

A quantitative assessment of the proposed approach is now in order. The reader should easily understand that a ground truth is not available. Therefore, we exploited our *a priori* knowledge of the study area Amitrano et al. (2014a,c,b), for identifying and extract by visual inspection the contour of 19 reservoirs whose extent is between approximately 6000 and 300000 square meters (in general we did not consider objects with surface area smaller than 5000 m²). These reservoirs have been marked with a red dot in Figure 5.21, in which we report the Level-1 α product of the whole study area (reference image: 28 April 2010, test image: 31 August 2010). The numeric ID accompanying the red dots refers to the graph depicted in Figure 5.24.

As first, an object-based assessment is proposed. In fact, in the study area, where basins are not mapped and often built autonomously by local communities without a governmental coordination, the most important information the analyst needs is about the presence or not of a water resource in a certain area. To this end, we consider another Level-1 α product, whose green band has been acquired on 15 August 2010. Coming back to Figure 5.21, we already talk about red dots, which identify the ground truth. As for the other markers, yellow dots are related to structures detected by the SWPP but not associable to reservoirs through visual inspection on 15 August. Magenta dots are associated to structures detected by the SWPP but not associable to reservoirs through visual inspection on 31 August. White dots are related to structures not associable to reservoirs on both acquisitions. On 15 August we counted 13 of such structures; on 31 August the count was 17. The number of structures which persist in both dates is 6.

As for the object-based accuracy, the 19 basins constituting the ground truth were detected on both 15 and 31 August. As for false alarms, the analysis must take into account the climate conditions and the severe rainfalls that occur during the wet season in the study area, as well as in all West Africa. In other words, after a heavy rainfall, it is highly probable that in some areas the shape of the terrain favors water accumulation in natural hollows. This makes very difficult to establish whether a detected structure is a false alarm or a tank.

As a consequence, we define as a false alarm a structure which is not

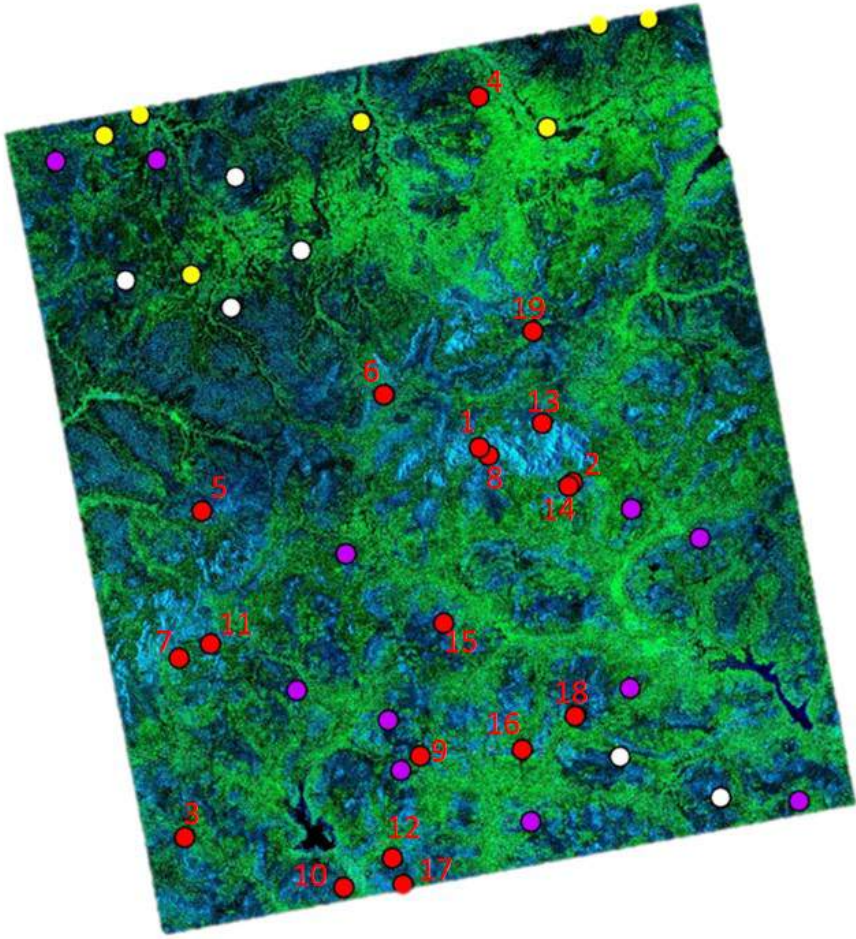


Figure 5.21. Level-1 α product of the study area. Reference image (blue band): 28 April 2010. Test image (green band: 31 August 2010. Red dots identify the basins assumed as ground truth, indexed with an ID which refers to the graph of Figure 5.24. Magenta dots indicate objects not recognized as basins by visual inspection in the image acquired on 31 August 2010. Yellow dots indicate objects not recognized as basins by visual inspection in the image acquired on 15 August 2010. White dots indicate objects not recognized as basins by visual inspection in both the aforementioned images.

stable on a short time scale, which, therefore, cannot be related to a stable reservoir. In fact, due to high evapotranspiration, a water accumulation in a natural hollow should disappear in few days. On the contrary, if the scene is such to cause a persistent anomaly (we are considering only objects whose area is greater than 5000 m^2), this can be considered as a false alarm or, at least, as a structure which is likely to be irrelevant from an hydrogeological viewpoint, although representing a pond.

The analysis of false alarms is supported by the the diagram of rainfall in the study area for the years 2010 and 2011 reported in Figure 5.22 Amitrano et al. (2014c). From this graph, it arises that the acquisition of 31 August is immediately preceded by a severe rainfall event. Therefore, it is more probable to detect structures due to water accumulation, probably irrelevant from an hydrogeological viewpoint.

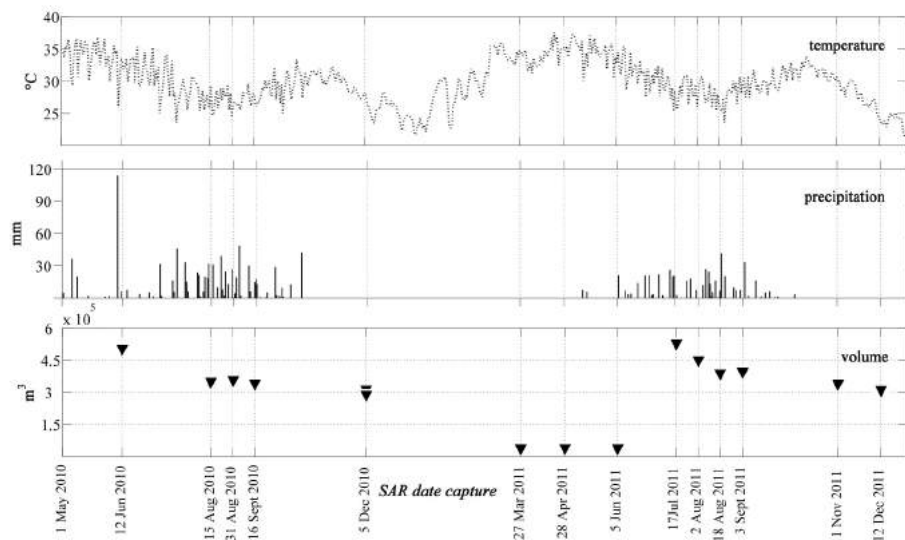


Figure 5.22. Rainfalls and temperature diagrams for the study area in the years 2010 and 2011.

If we consider the acquisition of 15 August, we can see that the closest rainfall occurred a couple of days before. Therefore, we can think that the reduction in the number of unidentifiable structures detected could be due to evapotranspiration.

Therefore, the behavior of the proposed technique with respect to objects not associable with reservoirs is dependent on the rainfalls, beyond scene characteristics. The more intense the rainfalls in the days preceding the acquisition, the highest the probability to detect such structures. However, if two acquisitions are considered, the probability to detect objects not relevant from an hydrological viewpoint is significantly reduced. In particular, as previously hinted, by considering the acquisitions of 15 and 31 August 2010 we found 6 structures which can be assumed as false alarms.

However, to remark the complexity of the false alarm evaluation, consider the situation depicted in Figure 5.23. We are talking about particular agricultural structures, typical of semi-arid environments, constituted by terracing built in counterslope using rudimental walls for collecting rain water. One of these structures as seen on a Google Earth view is shown in Figure 5.23a. This structure is one of the 6 we classified as false alarms (see white dots in Figure 5.21). In Figure 5.23b the same area is shown as it appears in the Level-1 α product of the 31 August 2010. Finally, in Figure 5.23c a 9 meters resolution DEM of the area Amitrano et al. (2014c) is shown, highlighting that it represent a natural hollow, whose walls are used for creating terracing. During the wet season, these structures can be covered by surface water as well as the hollow on which they are built. Thus they could be detected using the proposed methodology. The same situation occurs in at least 4 of the 6 cases we classified as false alarms. The question is therefore the following: can these 4 objects considered as false alarms?

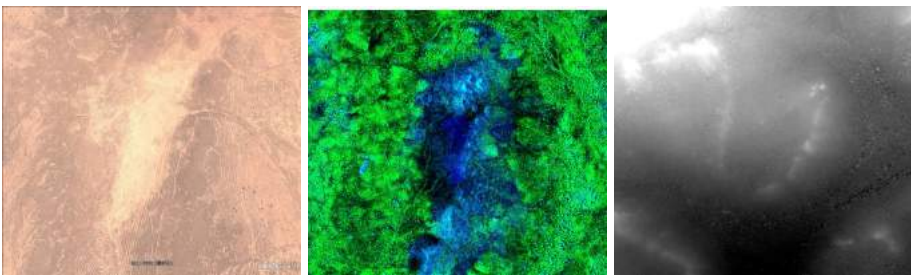


Figure 5.23. Terracing built exploiting a natural hollow as represented in (a) Google Earth, (b) Level-1 α imagery and (c) on a 9-meter resolution DEM.

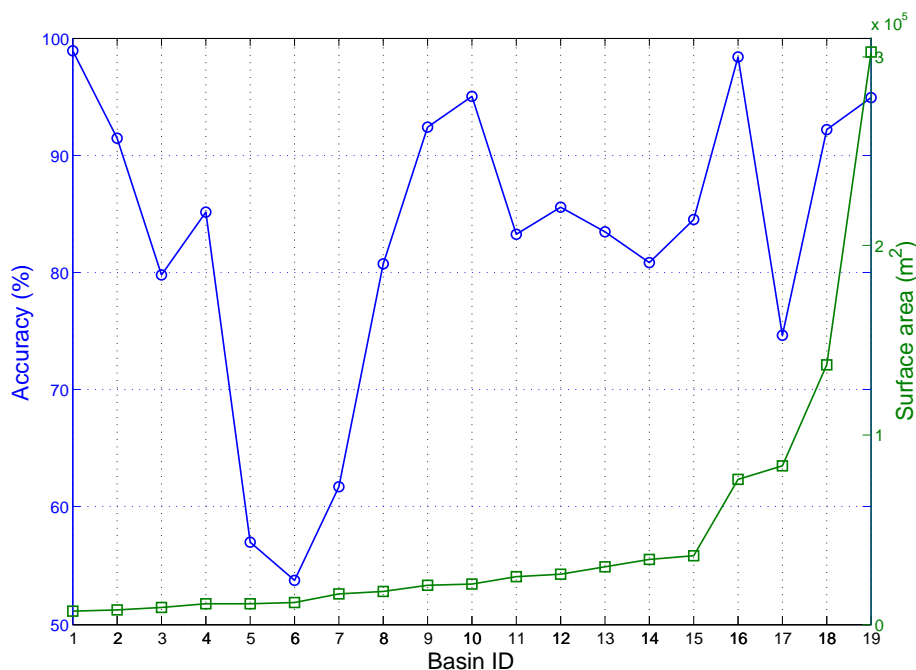


Figure 5.24. Detection accuracy in surface area estimation with respect to the extracted ground truth for the 19 analyzed reservoirs. Data refer to the acquisition of 31 August 2010.

From the point of view of the technique we propose, we can claim that they are not false alarms. In fact, the proposed method is designed for detecting temporary water coverage and, in this sense, it works. However, in the context of semi-arid regions, the application is fundamental. As an example, if the analyst wants to draw the water balance of the study area, these structures should be classified as false alarms, since they are probably not hydrologically relevant, such as a reservoir can be. On the contrary, if the application is more focused on agriculture, then terracing become important as well as reservoirs, since they represent areas in which cultivations are possible.

A pixel-based assessment was also performed. In Figure 5.24 we show the detection accuracy as a function of reservoirs' surface area for the 19 considered basins (see Figure 5.21) assumed as ground truth. Overall, the

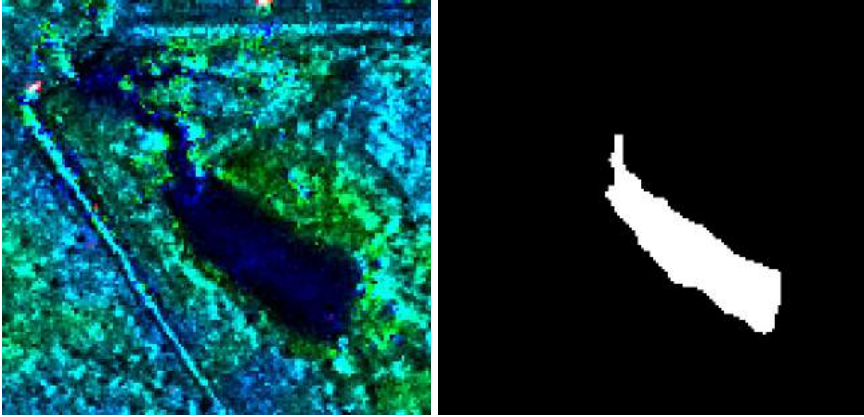


Figure 5.25. Basin 6: (a) close-up of the Level-1 α product and (b) its mask extracted after thresholding.

registered accuracy is of 88.6%.

As for false alarms, the above considerations hold. In general, it is expected to be negligible, since the probability to detect features that not represent water is scarce, due to Level-1 α products characteristics and SWPP formulation. Furthermore the application of morphological operators tends to erode slightly basins' extension.

As for missed detections, they are mainly related to the mixed response of boundaries in which the presence of mud and/or vegetation cause the reference image scattering to be not dominant. This also determines the accuracy with whom a reservoir is detected. As an example, consider the basin depicted in Figure 5.25a. This is the basin indexed with the number 6 in Figure 5.21 and that characterized by the worst performance of the detection (about 55%, see Figure 5.24). In this case, the causes of the missing detections can be attributed to the shape of the reservoir (particularly elongated and narrow in its upper part) and to the presence of vegetation at its boundary. This phenomenon represents the principal cause of missed detections.

Summarizing, the most critic parameter to assess for the proposed technique is the false alarm rate. The message we want to convey is that the its evaluation should be application oriented. In fact, a technique is designed to highlight a certain characteristic of the scene. In our case,

it is the a temporary water coverage. However, different structures can exhibit this characteristics and they are not all necessarily relevant for the application one is carrying out. The most important thing is to minimize the detection of objects which do not have the physical characteristic we are looking for. In this sense, the performance of the proposed method are satisfying, since only 2 of such objects have been detected. From an applicative viewpoint, the capability of the analyst is fundamental for recognizing the structures that could be interesting for his/her purposes. In this sense, Level-1 α imagery makes the interpretation of the images easier with respect to Level-1 products, representing a valid support even for users with limited expertise in SAR issues.

Summary

In this Section, we presented a change detection framework for temporary water bodies detection suitable to be applied to Level-1 α products. In particular, we introduced a Seasonal Water Pseudo-Probability, whose expression is given by a simple bands ratio between a reference image and a test one, representing the situation in which the abnormal water coverage occurs. This results in a quick, repeatable and user-oriented technique.

The proposed method is suitable to be applied to detect small reservoirs in a semi-arid environment, such as that of Burkina Faso. The proposed method allowed for detecting all the reservoirs in the study area with a negligible influence of false alarm. The performances are related with the selection of the reference band composing the input Level-1 α product. In fact, the seasonality of the water is related to its absence during the acquisition of this band. Therefore, a suitable choice for the reference image is an acquisition made at the peak of the dry season, when the environment is almost completely dry.

5.4.2 Buildings extraction

Introduction

In Chapter 2 we introduced the MAP3 framework for building of Level-1 α products. The main characteristic of this representation is that the association color-object, being physical-based, does not depend on the

scene. In this Section, we will focus, in particular, on the built-up feature. This class of objects is characterized by precise temporal properties: high (and stable) backscattering and high interferometric coherence, even when computed with a long temporal baseline. These properties make built-up features to appear in white in Level-1 α images, which are obtained combining the backscattering amplitude of two acquisitions and their interferometric coherence.

Coherence is estimated in a moving window, whose typical dimensions are of several meters in order to avoid bias (Hanssen, 2001). However, this choice is not optimal for all the scene targets. Moreover, decorrelation can occur due to several causes, such as orbital instability, baseline length, shadowing, feature shape and so on. The lack of the coherence contribution make the built-up class to appear in cyan (thanks to the contribution of the amplitude bands), thus introducing an alteration of the expected semantic.

Here, we present a technique which aims at improve the information content of the interferometric coherence band. Our technique exploits an input Level-1 α product for generating an *a priori* information which is used for i) adapting the coherence window dimension to the scene target and ii) introducing a texture measure which is used to identify the built-up feature class when it do not response to the variation of the coherence window dimension.

Already available Level-1 α products can be used to enhance the informative content of the interferometric coherence band through an expert feedback system

Enhanced MAP3 Framework

The calculation of the interferometric coherence (see Equation (2.2)) requires the selection of the dimension of the mean window, which determines the performances of the estimator with respect to the scene objects (Arrigoni et al., 2005). As an example, rough surfaces (like the sea) exhibit a stochastic and non-stationary backscattering, which is expected to produce incoherent signals. Thus, a small coherence window could produce a noisy coherence. Conversely, man-made structures, having a deterministic

stable scattering, typically exhibit high coherence values. Therefore, especially if the urban texture is not dense, a large computation window could include features like vegetation, shadows or roads together with buildings, thus causing a decrease in the resulting coherence beyond a poor resolution of the computed map.

In Figure 5.26a we show a three-meters resolution Level-1 α product of the city of Castel Volturno (Italy). The product is composed as follows: on the blue band, an image acquired on December 2009 (reference image) is loaded; the green band depicts an image acquired on August 2010 (test image); the red band is reserved to the interferometric coherence between the two images computed using a 11-pixels coherence window.

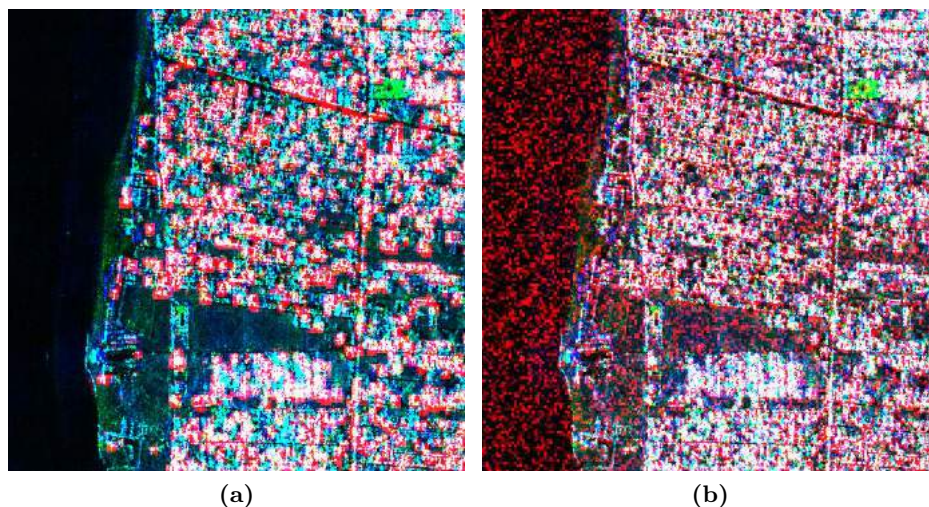


Figure 5.26. Castel Volturno (Italy): Level-1 α product (reference image December 2009, test image August 2010) computed setting the coherence window to (a) 11 pixels and (b) three pixels. The decreasing in the coherence window dimension causes a wrong estimate of the interferometric coherence on stochastic scatterers.

Since about eight months have passed between the two acquisitions we expect that only stable features, such as buildings, keep high values of the interferometric coherence. Being these features also highly reflective, urban structures should be represented in white color. However, since the interferometric coherence is computed using a mean window (i.e. the

interferometric coherence map has a coarser resolution than the intensity products), stable targets are surrounded by a red “crown” which helps the human photo-interpreter in their detection.

The product depicted in Figure 5.26a is consistent with the above considerations. Anyway, it is reasonable to think that a mean window of about 30 meters is too large for representing as best the detail of the urban areas. In fact, the reader should have the impression that white/red pixels are a bit sparse for being representative of a dense urban area. This means that the choice we made concerning the dimension of the coherence window was not optimal since it did not allow for fully exploit the characteristic of the estimator defined in Equation (2.2).

In Figure 5.26b we repeated the same experiment setting the coherence window to three pixels. In this case, it is clear that the image has an unacceptable granularity, which is physically inconsistent, beside being visually unpleasant. In fact, as an example, wide portions of the sea surface exhibit high coherence values. In this case the estimate on stochastic targets is affected by a bias which increases as the coherence window dimension decreases (Hanssen, 2001). Anyway, it is also true that the urban area is better represented using a smaller window, since it is possible to appreciate that more details arise with respect to the product depicted in Figure 5.26a. Therefore, a strategy for adapting the coherence window to the scene target is needed.

To this end, we propose a feedback system, whose rationale is depicted in Figure 5.27. The system is structured as follows:

- 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 estimated by Equation (2.2) is optimized for the considered target;
- 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 the following, we provide details about the implementation of the proposed feedback.

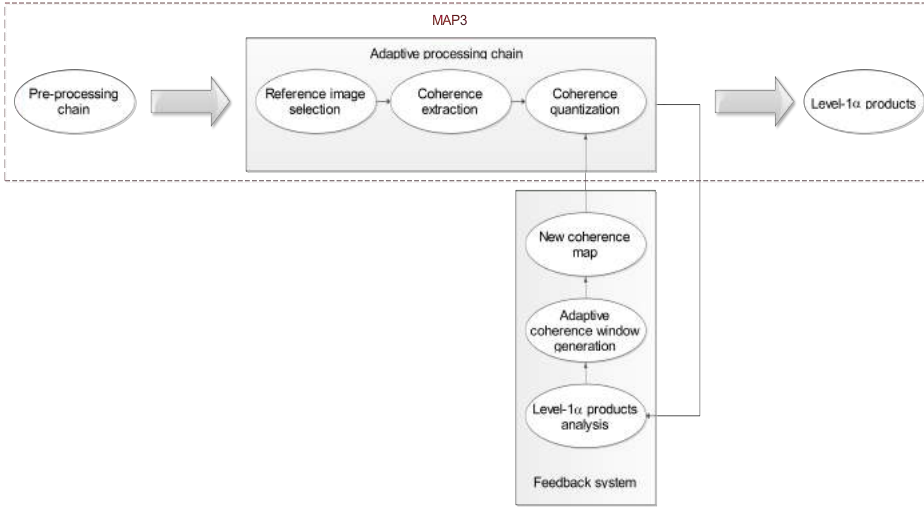


Figure 5.27. Coherence feedback system for the MAP3 framework. Level-1 α products are exploited for generating an *a priori* information which is used to adapt the coherence window dimension to the scene target and to introduce a texture measure which is used to identify the built-up feature class when it do not response to the variation of the coherence window dimension.

Fuzzification The coherence feedback system works through the fuzzification of Level-1 α 's coherence and intensity bands. We modeled these variables using three fuzzy sets with verbal attributes of “low” (Z-type), “medium” (π -type) and “high” (S-type). In particular, the following expressions was adopted Tobias and Seara (2002):

$$S(DN, a, b, c) = \begin{cases} 0, & DN \leq a \\ 2 \{(DN - a) / (c - a)\}^2, & a < DN \leq b \\ 1 - 2 \{(DN - c) / (c - a)\}^2, & b < DN \leq c \\ 1, & DN > c, \end{cases} \quad (5.8)$$

in which DN means digital number and a, c and $b = (a + c)/2$ are the parameters which model the S-function (see Table 5.8 for details) The

Z-type fuzzy set is obtained from the S-type one being $Z(DN, a, b, c) = 1 - S(DN, a, b, c)$. Finally, the π -type function is built by combining a S-type and a Z-type function sharing the parameters c and a , respectively, as shown in Table 5.8. The obtained fuzzy sets are plotted in Figure 5.28.

Table 5.8. Parameters used for modeling the adopted fuzzy set for Level-1 α 's coherence and intensity bands and for the coherence window.

Function type	Level-1 α			Coherence window		
	a	b	c	a	b	c
S-type	160	190	220	21	36	51
Z-type	0	60	120	0	2.5	5
π -type (S-part)	40	100	160	3	7	11
π -type (Z-part)	160	185	210	11	26	31

The purpose of this system is to adapt the dimension of the coherence window as a function of the phase-stability and reflectivity characteristics exhibited by the targets in the bi-temporal RGB composite. Essentially, we aim at reducing the mean window as the coherence and the intensity responses increase. To this end, the system output (i.e. the coherence window size) is fuzzified as well as the input Level-1 α product. In particular, we considered three type of windows, “small”, “medium” and “large”, which have been modeled with a S-rule a π -rule and a Z-rule, respectively.

The assignment of the coherence window category for each image target is now in order. The rationale is quite simple: the more stable and reflective the target, the smaller the coherence window. In Table 5.9, we reported the adopted rules for the assignment of the fuzzy categories “small” and “medium” to the coherence window. All the other cases are reserved to the “large” window attribute.

The selected fuzzy sets for the coherence window are depicted in Figure 5.29. They have been partitioned as a function of the membership degree. As an example, the fuzzy set “small” is subdivided in three parts. The first part is relevant to persistent scatterers (PSs), and is reserved to pixel with the highest membership within this fuzzy set, which are likely to be the more reflective and coherent in the input Level-1 α product. As the membership of “small” window decreases, the window dimension increases to three and five pixels, since the correspondent targets seem further away

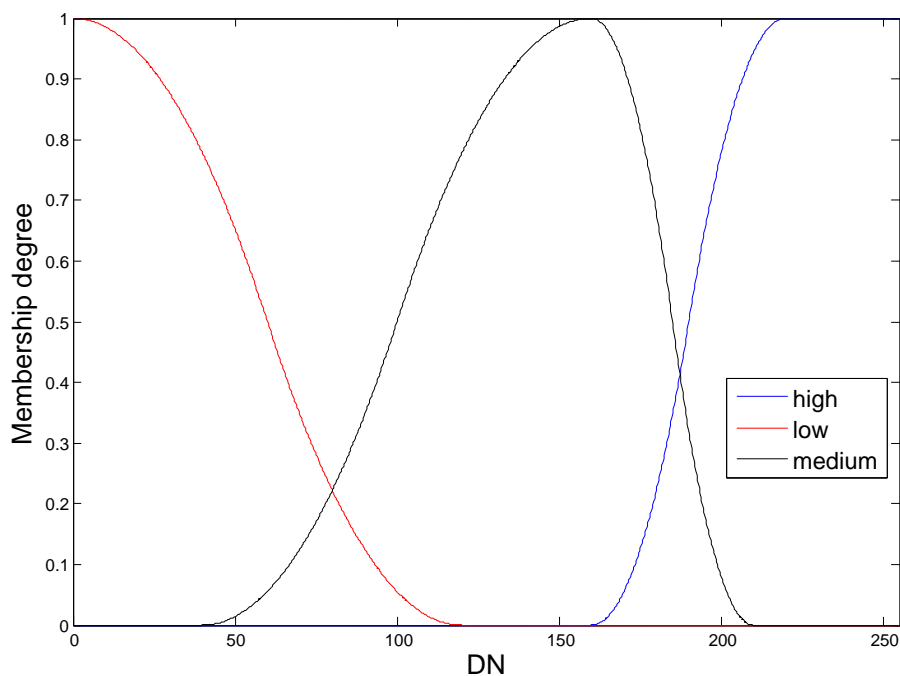


Figure 5.28. Fuzzy set adopted for the fuzzification of the input Level-1 α product.

to behave as PSs. Similar reasoning can be performed for the “medium” window fuzzy set. As for the “large” window fuzzy set, we do not define any partition, since all the scatterers belonging to it are expected to be stochastic.

In particular, we modeled “small”, “medium” and “large” windows with a S-rule a π -rule and a Z-rule, respectively. The parameters adopted for building these function are reported in Table 5.8.

The coherence window attribute (“small”, “medium” or “large”) and, as a consequence, the dimension of the coherence window according with the aforementioned partitions of the fuzzy sets, is assigned by computing the maximum membership degree after the application of the rules reported in Table 5.9.

It is remarkable that an ambiguity arises concerning the “medium”

Table 5.9. Adopted rules for the assignment of the coherence window attribute.

Window	Coherence	Intensity 1	Intensity 2
Small	High	High	High
	Medium	High	High
	High	Medium	High
	High	High	Medium
Medium	Medium	Medium	Medium
	Low	High	High
	Medium	Medium	High
	Medium	High	Medium
	High	Medium	Medium

coherence window. In fact, due to the bell-shape of this fuzzy-set, two windows dimensions correspond to each membership degree, one for the S-part of the π and one its Z-part. However, these windows correspond with targets with very different characteristics. In fact, on the S-part of the π , we expect to have targets with medium-high characteristics of reflectivity and coherence. On the contrary, on the the Z-part of the π , we expect to find targets more likely to be incoherent. Therefore, we require that targets which lie in the S-part of the “medium” window fuzzy set must have “high” reflectivity in both the images which compose the Level-1 α product. Otherwise, they are placed on the Z-part of the π , on the side of the largest windows associated to this fuzzy set.

Use of texture Coherence response of targets depends on several factors. Decorellation can occur, also on stable targets due to imaging geometry (incidence angle), shadowing, misalignment with respect to the flight direction or the shape of the building. As an example, a squared-shape building with smooth walls is more likely to have a coherent response with respect to one with an irregular shape or a pitched roof. Therefore, it is possible that some buildings do not exhibit a coherent response after the reduction of the coherence window.

In Level-1 α imagery, built-up features are represented in white color, due to the high contribution of both amplitude and interferometric channels. However, as explained above, decorrelation should cause this feature

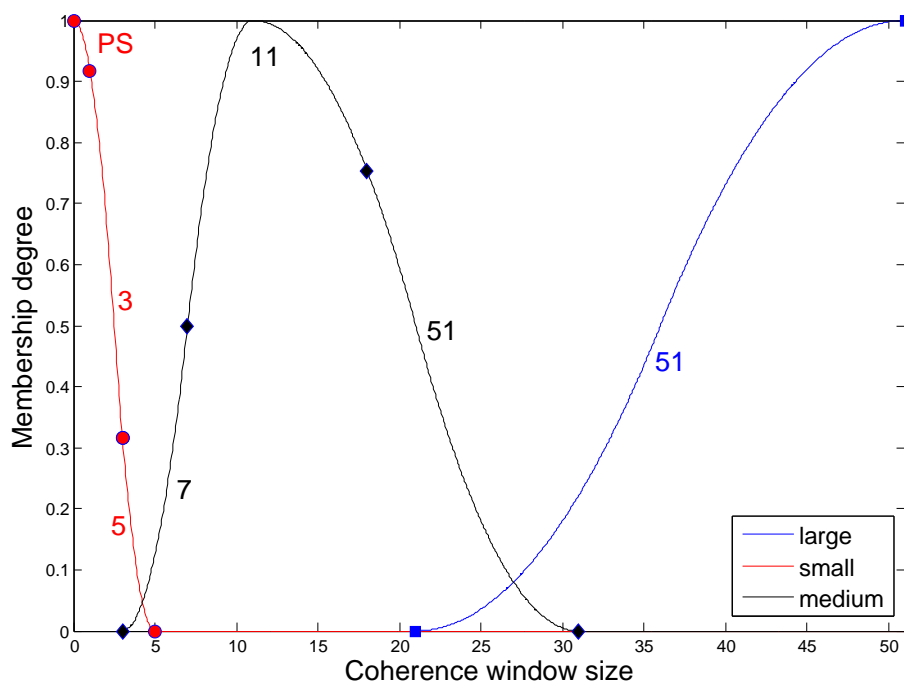


Figure 5.29. Fuzzy set adopted for the fuzzification of the coherence window.

to appear in cyan. Here, we want to introduce a texture measure for enhancing the red band informative content. In such way, it is possible to restore the semantic the user expects on built-up features, making this feature class to appear in white.

The principle we exploit is the convergence of evidences Matsuyama and Hwang (1990). Until now, we have assumed that the built-up class is characterized by high amplitude (in both the dates that constitute the Level-1 α bi-temporal composite) and high interferometric coherence. The introduction of a fourth evidence, i.e. the texture, allows for slacken the requirement on the interferometric coherence and considering as built-up also the objects that, even after the feedback application, do not exhibit a high response to the coherence estimator but are highly reflective (in both Level-1 α 's acquisitions) and located in areas characterized by high texture.

The texture measure we use was introduced by Nagao and Matsuyama in Nagao and Matsuyama (1980). To obtain it, first, we applied a k -means clustering to the input Level-1 α product (see Figure 5.30b and Figure 5.30a, respectively). A very coarse clustering is sufficient for our purposes. In fact, as shown in Figure 5.30b, we just separated the white, red and cyan color (grouped in the white class in the classified map) from all the others (green class). The white class can be roughly associated with built-up features.

The cluster map is used for contour extraction through the application of the second Laplacian operator Marr and Hildreth (1980) (see Figure 5.30c). Finally, the Nagao-Matsuyama texture can be computed. It is a very simple rule: a moving window of dimension N is applied to the contour map; if in a window fall at least $2N + 1$ border pixels, then the central pixel of the window is classified as high texture area. Otherwise, it is classified as a low texture area. The result of the application of this rule to the computed edge map is depicted in Figure 5.30d.

The texture evidence is activated (for pixels classified as high texture areas) on objects characterized by high backscattering in both Level-1 α acquisitions and low coherence. Obviously, the information conveyed by the red band for these targets does not concern anymore their phase stability.

Experimental results The proposed feedback process is based on the convergence of four evidences: the amplitudes of the two acquisition composing the input Level-1 α product, their interferometric coherence and the texture. The first three evidences can be considered strong, since high characteristic of reflectivity and coherence are almost an exclusive property of built-up features. Texture, instead, is assumed to be a weak evidence, because in SAR imagery high texture areas are not necessarily urban areas. However, the convergence of two strong evidences (i.e. the high backscattering in the two Level-1 α 's acquisitions) and one weak evidence (i.e. the texture) make us confident that the target we are considering is man-made.

In Figure 5.31a we show a Google Earth view of Macerata Campania, a small city in southern Italy. In Figure 5.31b and Figure 5.31c the correspondent Level-1 α products before (coherence window set to 11 pixels)

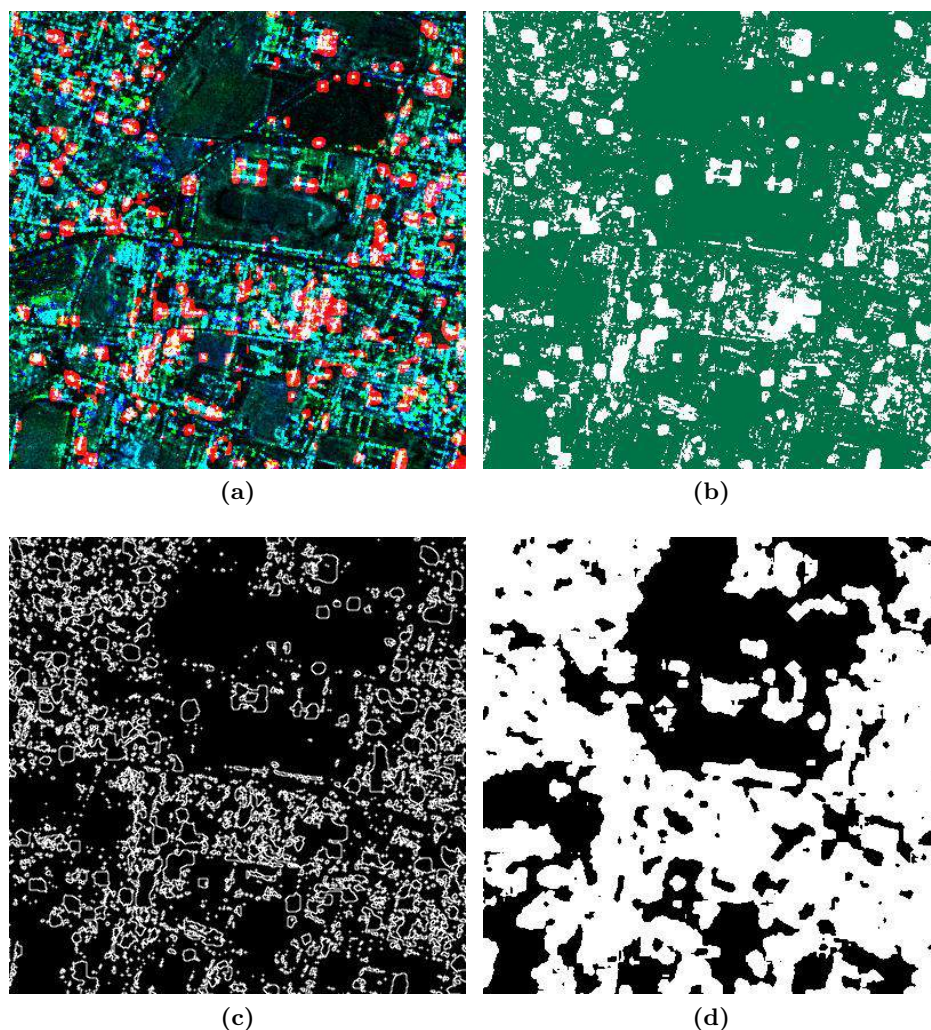


Figure 5.30. San Prisco (Italy): (a) Level-1 α product, (b) two categories k-means clustering, (c) edge map and (d) Nagao-Matsuyama texture.

and after the feedback application, respectively, are depicted. Qualitatively, the reader should appreciate as the feedback system improves the representation of the built-up feature class. In fact, in Figure 5.31c, more bright targets are visible with respect to Figure 5.31b. Moreover, the

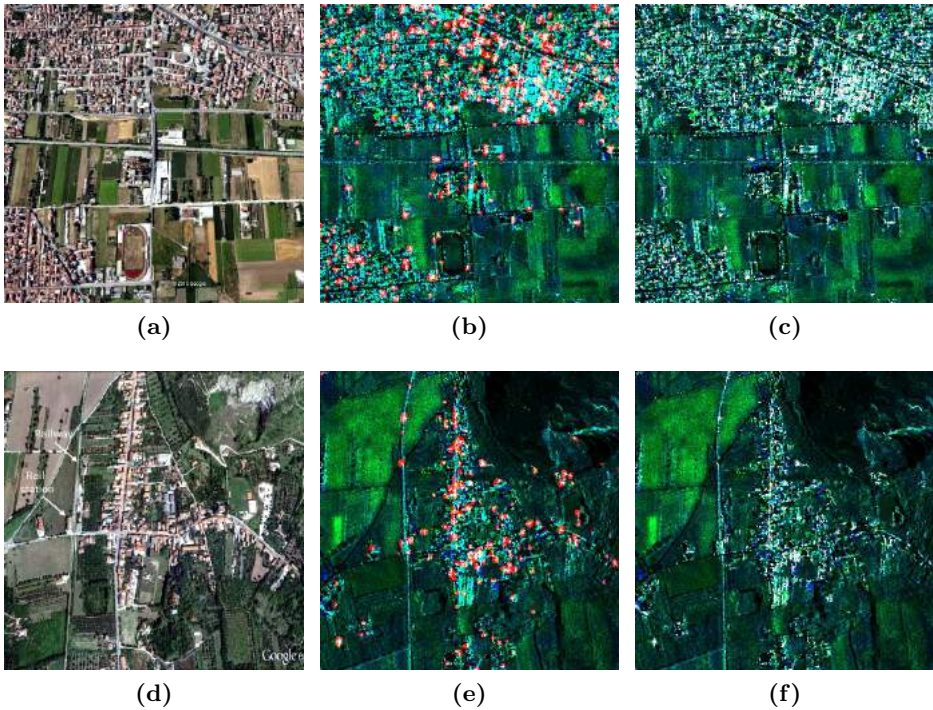


Figure 5.31. (a) Google Earth view of Macerata Campania stadium with its correspondent Level-1 α products (b) before and (c) after the feedback application. (d) Google Earth view of Sant’Angelo in Formis rail station with its correspondent Level-1 α products (e) before and (f) after the feedback application.

red “crown” surrounding coherent targets is practically disappeared. This means that the resolution of the red band is higher.

In Figure 5.31d we show a Google Earth view of the city of Sant’Angelo in Formis (Italy), with its rail station and railway at the left of the residential area (see annotations on the picture). In Figure 5.31b and Figure 5.31c the correspondent Level-1 α products before (coherence window set to 11 pixels) and after the feedback application, respectively, are depicted. The considerations made for the Macerata Campania scene hold: the feedback application allows for retrieving a number of bright targets, i.e. more buildings, in the residential area, more detail on the railway and improves the resolution of the red band.

These claims will be more and more evident if Figure 5.32 is considered. In particular, in Figure 5.32a and Figure 5.32b we reported the red band (i.e. the one representing the coherence/texture information) of the products displayed in Figure 5.31b and Figure 5.31c. In Figure 5.32c and Figure 5.32d the red band of the products displayed in Figure 5.31e and Figure 5.31f are depicted. All the maps in Figure 5.32 have been obtained through linear quantization of the input coherence product between the interval $[0.45, 0.6]$.

It is remarkable that more bright points appear in Figure 5.32b and Figure 5.32d (after the feedback application) with respect to the maps depicted in Figure 5.32a and Figure 5.32c obtained by applying Equation (2.2) with fixed 11 pixels window. This means that the urban area is better characterized. In fact, as an example, in Figure 5.32b, the shape of the stadium is recognizable (see the annotation on the picture) while in Figure 5.32a it is not.

In Figure 5.32d (after the feedback application) it is remarkable as more particular of the rail station arise with respect to Figure 5.32c. Moreover, the shape of the railway is now visible at the left of the residential area (see annotation on Figure 5.32d).

These experiments should qualitatively convince the reader of the effectiveness of the proposed algorithm. A quantitative assessment is provided in the following Section.

Assessment The quantitative assessment of the performances of the proposed algorithm is now in order.

To this end the urban atlas of the European Environmental Agency has been used. It is a land cover map in which the density of urban areas is measured with the respect to the soil sealing (SL), i.e. the covering of the ground by an impermeable material. In particular, here we are particularly interested in five urban categories of the urban atlas: “Continuous urban fabric” (SL $> 80\%$), “Discontinuous Dense Urban Fabric” (SL $50 - 80\%$), “Discontinuous Medium Density Urban Fabric” (SL $30 - 50\%$), “Discontinuous Low Density Urban Fabric” (SL $10 - 30\%$) and “Discontinuous Very Low Density Urban Fabric” (SL $< 10\%$).

In order to compare the results of the application of the proposed technique with the urban atlas database, we adopted the following pro-

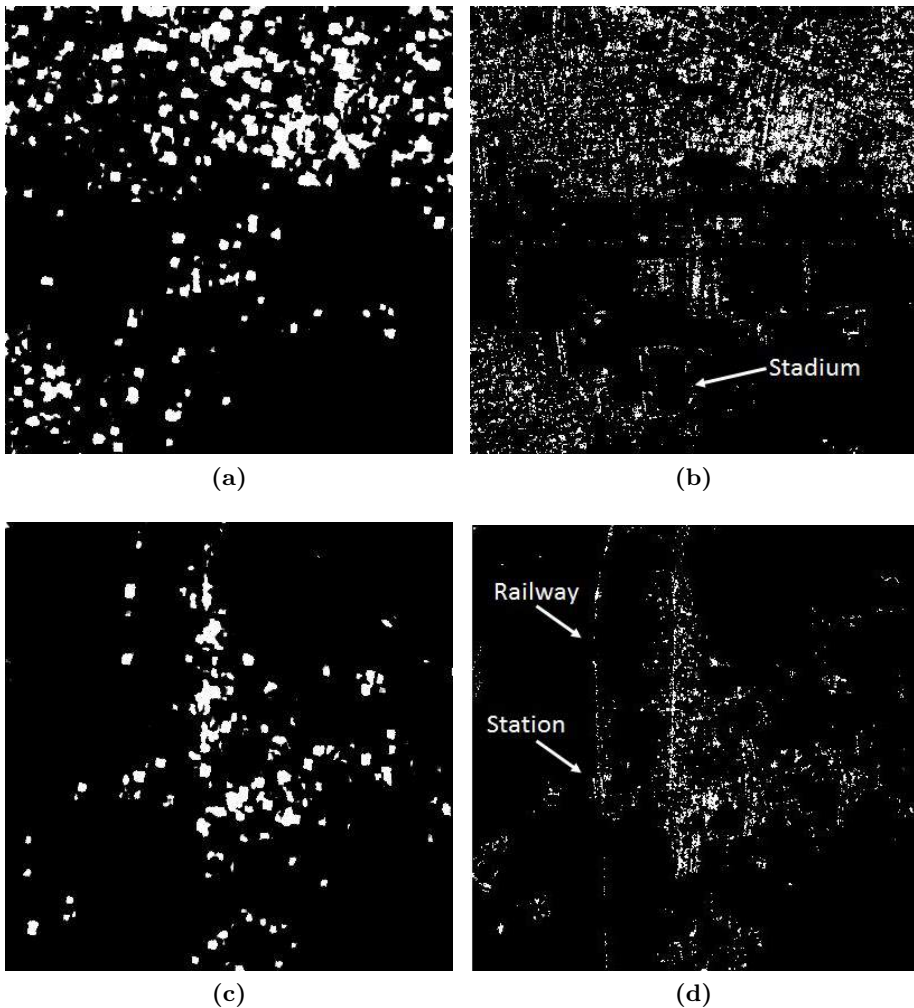


Figure 5.32. Macerata Campania (Italy): coherence map (a) before and (b) after the feedback application corresponding with Level-1 α products depicted in Figure 5.31b and Figure 5.31c, respectively. Sant'Angelo in Formis (Italy): coherence map (c) before and (d) after the feedback application corresponding with Level-1 α products depicted in Figure 5.31e and Figure 5.31f, respectively.

cedure. First, we applied the built-up index (BI) proposed in (Amitrano et al., 2015d) to the input Level-1 α product before and after the feedback

implementation. We report it here for the ease of the reader:

$$BI = \frac{RGB}{255^3}, \quad (5.9)$$

where R , G and B are the values of the red, green and blue bands of the input Level-1 α product, respectively.

Equation (5.9) has been applied to the input Level-1 α product before and after the feedback application. In both cases a threshold of 0.1 has been adopted for maps binarization.

The retrieved binary maps has been then compared with the polygons provided by the urban atlas. In particular, we compute the zonal statistics for each class, obtaining the total number of building pixels which fall in the polygons belonging to that class. Therefore, for each class, the following equation was computed

$$N^i = \sum_{k^i} n_k^i, \quad (5.10)$$

where N^i is the total number of building pixels for the i -th class, k^i is the index of the polygon belonging to the i -th class, and n_k^i is the number of building pixels which fall in the k -th polygon of the i -th class.

The urban atlas class list can be found in Table 5.10 together with the SL (when specified) and the total area of the classes relatively to the study area.

Once N^i has been computed for each class, the percentage of soil (on average) covered by built-up pixel for the i -th class is obtained by relation

$$\bar{SL}^i = 100r \frac{N^i}{A^i}, \quad (5.11)$$

where r is the map cell size and A^i the total area covered by polygons belonging to the i -th class.

In Figure 5.33 a picture relevant to the study area containing both the input Level-1 α product and the urban atlas layer is shown. The results of the application of the above described assessment procedure are reported in Table 5.10.

Table 5.10. Area covered by built-up pixels in relation with Urban Atlas categories before and after the feedback application.

Urban atlas categories	SL (%)	Area [km ²]	SL before (%)	SL after (%)
Continuous Urban	> 80	2.77	12.11	24.83
Dense Urban	50-80	5.22	9.6	20.54
Medium Density Urban	30-50	1.09	4.81	11.04
Low Density Urban	10-30	0.46	2.33	5.59
Very Low Density Urban	< 10	0.08	1.39	3.35
Agricultural/Wetlands		42.1	0.13	0.7
Isolated structures		0.6	1.82	3.96
Industrial/commercial		4.13	4.11	9.18
Other roads		25.6	0.54	1.32
Land without use		0.2	1.33	3.56
Fast transit roads		0.41	0.14	0.5
Railways		0.43	1.54	4.61
Dump sites		0.32	0.2	10.28
Construction sites		0.04	5.26	7.62
Green urban areas		0.09	4.03	10.21
Sports facilities		0.34	4.28	8.4
Forests		6.37	0.06	2.32
Water bodies		1.16	0.09	0.2

As expected, the percentage of soil covered by the built-up class significantly increased after the feedback application. In particular, the SL index passes: from 12.11% to 24.83% for the class “Continuous Urban Fabric”, from 9.6% to 20.54 for the class “Discontinuous Dense Urban Fabric”, from 4.81% to 11.04% for the class “Discontinuous Medium Density Urban Fabric”, from 2.33% to 5.59% for the class “Discontinuous Low Density Urban Fabric” and from 1.39% to 3.35% for the class “Discontinuous Very Low Density Urban Fabric”. It is remarkable that, even after the feedback, these values are very far from the percentages indicated in the urban atlas class description (see Table 5.10). This can be explained. In fact, consulting the urban atlas mapping guide, we know that the macro-class “Urban Fabric” (i.e. the one that contains all those above mentioned)

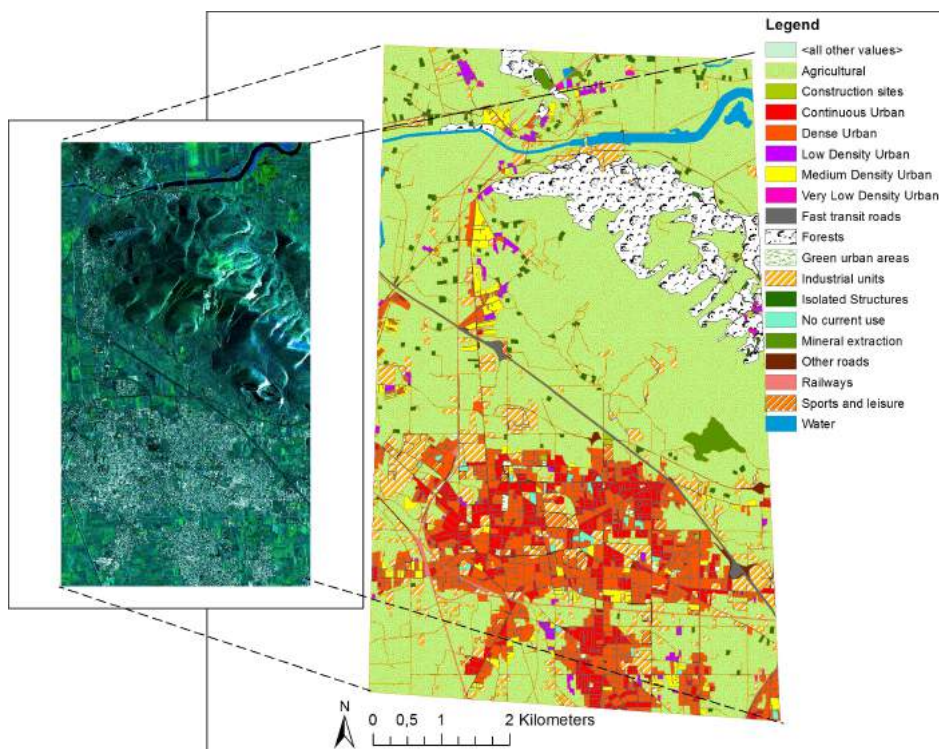


Figure 5.33. Level-1 α product of the study area with its relevant urban atlas layer.

is formed by “built-up areas and their associated land, such as gardens, parks, planted areas and non-surfaced public areas and the infrastructure” (European Environment Agency, 2011). Therefore: i) in the reference urban atlas class, other land covers are included beyond built-up; ii) some of this land cover (roads, sidewalks, car parks) do not have the fundamental property of high backscattering we use for the detection; iii) shadowing effects, influencing backscattering, partially prevent the feature detection using Equation (5.9). The last phenomenon is more severe as the density of the urban area increases. That’s why the more dense the urban area, the more distant the soil sealing index retrieved through Equation (5.11) on the Level-1 α product with respect to that indicated in the urban atlas.

The last observation concerns layover. Using a very coarse clustering, it is likely that the texture evidence brings to an incorrect decision on layover features. This should be more clear considering Figure 5.34. In Figure 5.34a we show the Level-1 α product before the feedback. Here, layover is correctly represented, since we have no response of the red band (i.e. of the interferometric coherence). The texture evidence action makes this feature to appear as a bright target, thus introducing a confusion with the built-up class see (Figure 5.34b). This problem can be solved using topographic corrections before building Level-1 α products or by applying a layover mask in post-processing. We opted for this choice. We used a 20 meters resolution DEM for generating the layover mask depicted in Figure 5.34c. Qualitatively, the reader should appreciate in Figure 5.34c as the masking operation applied to the Level-1 α product treated with the feedback algorithm mitigates the confusion between layover (which should be displayed in cyan) and the built-up class.

Quantitatively, the influence of layover in the soil sealing index calculated through Equation (5.11) on the SAR image can be found in the class “Forests” of Table 5.10. In the urban atlas layer, this class is mainly located in the upper left part of our study area (see Figure 5.33, where the presence of a relevant topography makes layover to occur. In this case, the application of the proposed feedback causes the soil sealing to increase from 0.06% (before the feedback) to 2.2% due to layover features included in the computation. The application of the layover mask allows for reducing this value to 1.14%. Better results can be obtained using a finer resolution DEM.

Summary In this Section, we introduced a technique for enhancing the information content of the red band of a Level-1 α product. This band is usually reserved to the interferometric coherence information and therefore it is aimed at the enhancement of targets characterized by high phase sta-

The Building Index is a quick and user-oriented solutions for the extraction of building features from Level-1 α products through bands product

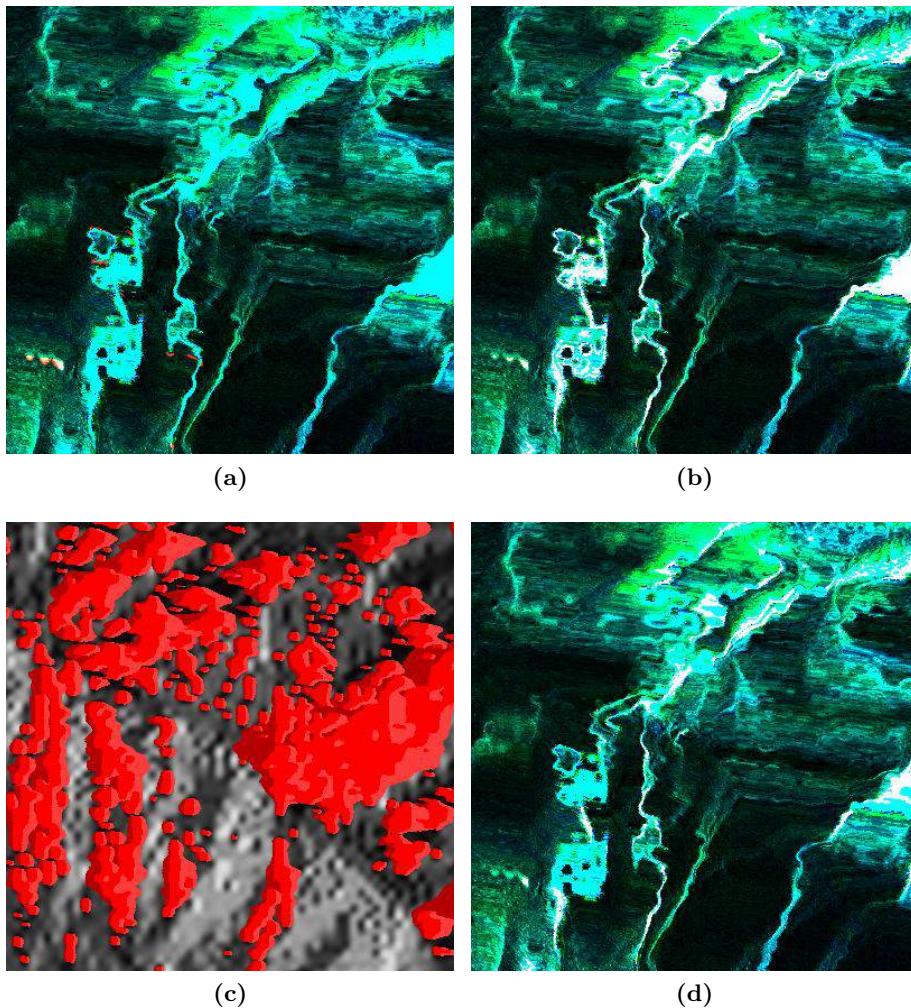


Figure 5.34. Layover treatment: Level-1 α product (a) before and (b) after the feedback application. (c) Layover mask. (d) Level-1 α product after the application of feedback and layover masking.

bility over the time, such as built-up features. The proposed technique is devoted at enhancing the performances of the coherence estimator through an adaptive selection of the moving window dimension. To this end, the input Level-1 α product is used to generate an *a priori* knowledge which

is used to build the expert system for the choice of the window dimension through fuzzy rules.

Texture evidence is used for supporting the enhancement of built-up features where the feedback fails due to decorrelation. We proposed the use of the simple Nagao-Matsuyama rule for improving the informative content of the red band of the refined Level-1 α product in areas characterized by high backscattering and texture.

We assessed the performances of the proposed algorithm qualitatively and quantitatively. Qualitatively, at the visual experience of the operator is enhanced by the feedback application. This makes easier the detection of the urban area through the enhancement of its built-up features. Quantitatively, we used the urban atlas layer for comparing the soil sealing measured on the Level-1 α product (before and after the feedback) with data provided in the reference layer description. We found that the feedback application significantly increased the number of detected pixels belonging to the built-up class for all the relevant classes. The obtained results in term of the measured soil sealing are congruent with data provided by urban atlas, compatibly with SAR imaging characteristics and classes definition.

5.4.3 Object-based image-analysis

Object-based image analysis (OBIA) has its roots in industrial image processing. In the remote sensing community, the most relevant research started in the 2000s, thus the topic is relatively new (see Blaschke, 2010, for an extended review). It is quite clear that object-based image understanding (Nagao and Matsuyama, 1980; Matsuyama and Hwang, 1990) is strictly related to the reliability of such objects, thus to an effective image segmentation.

Segmentation is one of the historical challenges of computer vision (Haralick and Shapiro, 1985; Pal and Pal, 1993) and one of the most addressed issues in modern SAR remote sensing (Gaetano et al., 2014; Wang et al., 2014). In fact, segmentation is an ill-posed problem in the sense of Hadamard since if a solution certainly exists, it is not unique (Marroquin et al., 1987). In Chapter 4, we showed how to use a Kohonen network to segment RGB SAR composites using color homogeneity and to attach to each region a semantic label indicating a color attribute. Here,

we want to enrich the characterization of the obtained objects through the introduction of parameters describing their shape, which can be then used to recognize a certain class present on the scene.

We consider an application related with precision agriculture, in particular with olive-trees monitoring. This culture is object of great attention in southern Italy due to plague of the *xylella fastidiosa* (Carlucci et al., 2013), a fatal parasite which is causing the destruction of a number of trees with severe damage to local economies.

The study area is located in Calabria (Southern Italy), in the nearby of the city of Ciró Marina (see Figure 5.35). This area renowned for the quality of its olives, whose cultivation represent one of the most important productive activity . The backscattering of this cultivation is not easy to be monitored on an aggregate product such as the Level-1 β due to irregular cycles of pruning and a strong variability of fruits production for different years. However we can exploit the proposed color composite for classifying different variety of olive-trees thanks to the trees-line geometry.

In Figure 5.36 we show an olive grove picture taken from the net. The red rectangle represents its planting layout, which is an ideal square or rectangle at which vertexes trees are positioned. This schema changes with trees variety and dimension and with the type of plantation (intensive or extensive). In any case, it should allow a satysfing illumination for each plant in order to ensure the quality of its fruits.

The planting layout is important also at a normative level. As an example, a local law of the Puglia region rules the safeguard of plurisecolar olive-trees and of plantations with landscape relevance. One of the criteria for establishing which trees fall in this categories is the irregularity of the planting, since in the past less attention was paid to the geometry of the plantation due to the absence of mechanized harvesting systems. The local law of the Calabria region authorizes the public administration to eradicates olive-trees from private terrains when the planting layout is such that it creates damages to the olive-grove. Many disciplinary of consortia for the production of extra-virgin olive oil dictate the minimum allowed planting layout, i.e. the maximum allowed number of trees per hectare.

Olive trees can be identified indirectly by the shadow they produce to the electromagnetic backscattering of terrains on the their backside with respect to sensor flight track. In Figure 5.37a and Figure 5.37b we show

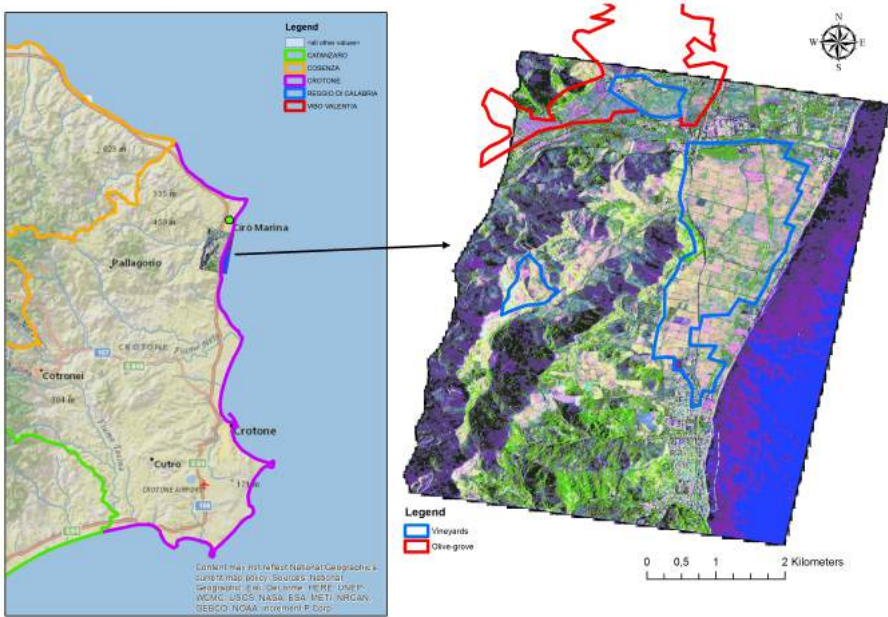


Figure 5.35. Study area, located in Southern Italy, for the proposed olive-trees monitoring application.

a Google Earth view and a Level- 1β view of our test area, respectively.

In order to identify a preliminary mask of the olive-tree feature, we first isolated a set of candidates by querying the color attribute of the semantic Level- 1β product. In this case, the colors corresponding to olive trees shadow are “Black”, “Dark slate blue”, and “Midnight blue”. In Figure 5.37c we show the mask obtained selecting only the pixels whose color attribute corresponds with those above enumerated. By comparison with the Google Earth view of Figure 5.37a, it is clear that other features (such as roads and buildings shadows) fall in the mask besides those we want to isolate. In the following, we use an approach derived from Marr’s computational theory of vision (Marr, 1982) for identifying olive-trees shadows within the extracted mask.

As first, we extracted objects from the binary image of Figure 5.37c through connected components labeling (Shapiro and Stockman, 2002)



Figure 5.36. An olive-grove with its planting layout (red rectangle).

obtaining a map in which each object is labeled with a progressive numeric identifier. In parallel, we computed an objects contour map by application the Laplacian operator to the mask image (Marr and Hildreth, 1980). In Figure 5.37d we show the 4-adjacency rule contour map which, together with the connected components map, constitute the instruments which allow for an objects shape analysis.

In the identification of olive-trees shadow, we used three shape parameters: area, circularity and elongation. The area parameter (A) is simply given by the area of the considered connected region. The circularity parameter is computed by relation Cox (1927):

$$f_C = \frac{A}{4\pi P^2}, \quad (5.12)$$

where P is the region perimeter defined through the contour map. Clearly, A can be obtained simply by counting the number of elements belonging to each numbered element displayed in Figure 5.37c (the seg-

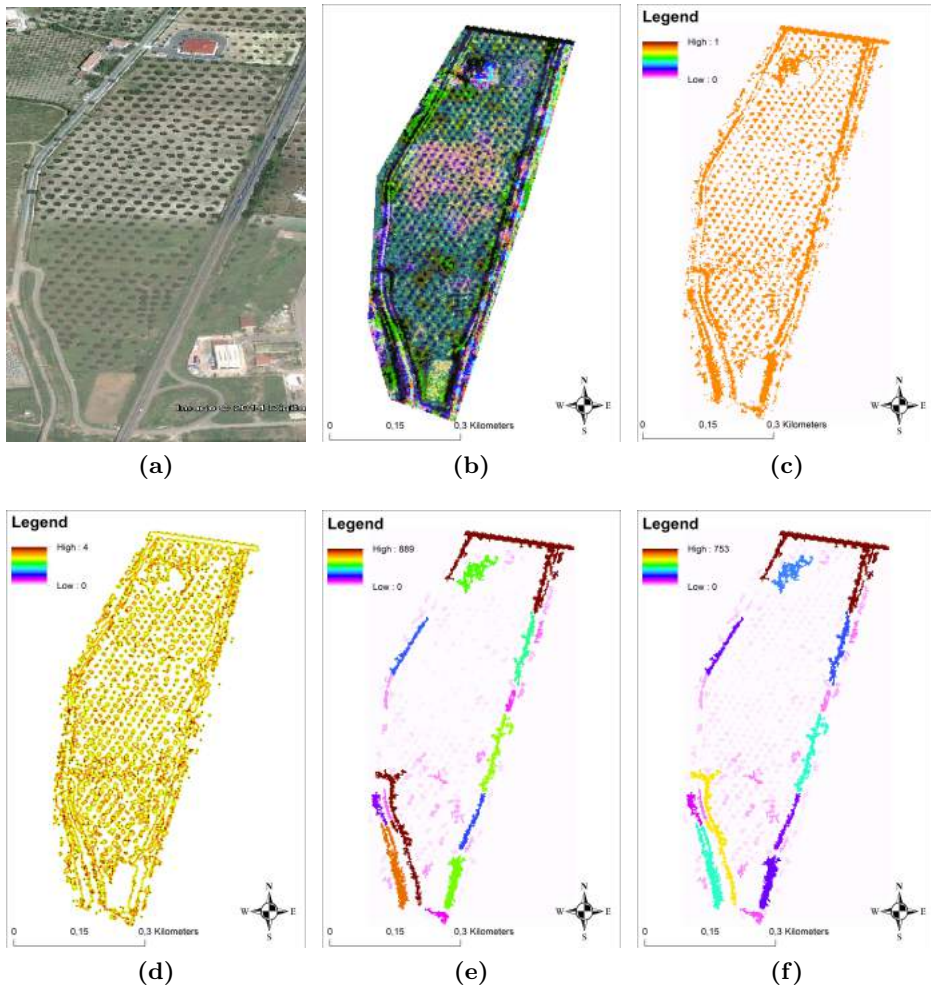


Figure 5.37. Ciró Marina (Italy), olive grove: (a) Google Earth view, (b) Level-1 β view, (c) preliminary olive-trees mask, (d) connected components image, (e) 4-adjacency rule contour map and (f) circularity map.

ments area map is reported in Figure 5.37e). The perimeter is obtained by superimposing the connected components map to the 4-adjacency map in order to isolate border pixels (see Figure 5.37f). In the digital world, the most round object is the isolated pixel (i.e. the square) for which $A = 1$

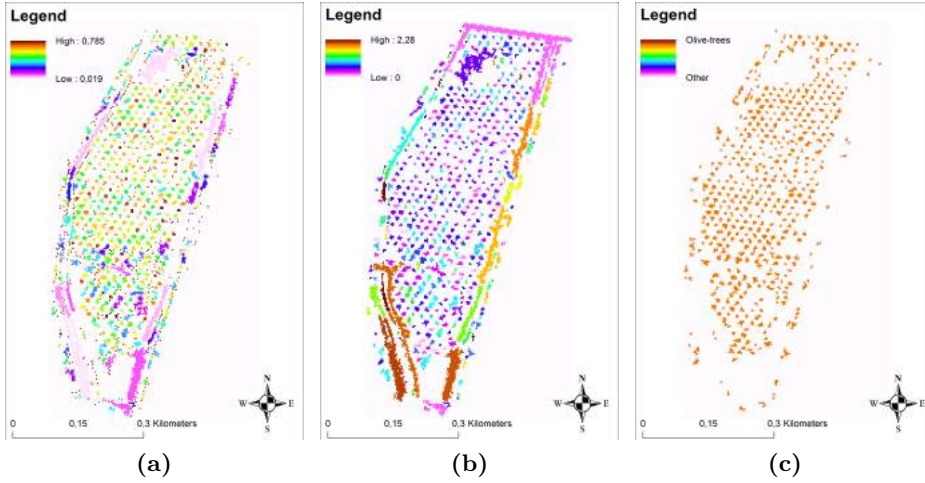


Figure 5.38. Ciró Marina (Italy), olive-grove: (a) circularity map, (b) elongatedness map and, (c) final olive-trees mask.

and $P = 4$ (if the 4-adjacency is considered). Therefore, the maximum value for the roundness is 0.785. The obtained roundness map is reported in Figure 5.38a.

The elongation parameter is defined by relation Matsuyama and Hwang (1990):

$$f_E = \frac{A}{W^2}, \quad (5.13)$$

where W is the number of iteration necessary to shrink completely the segment. In other words, at each shrinking iteration, pixels directly in contact with the background are eroded. The operation is repeated until the entire segment is fused into the background. The obtained elongatedness map is reported in Figure 5.38b.

By analyzing these maps it arises that the objects we are looking for are characterized by a precise shape. In fact, olive-trees area is low as well as their perimeter and elongatedness. Conversely, their circularity results high, since they have an almost square shape. Therefore, the request for extracting this feature starting from the preliminary mask of Figure 11c obtained by querying the color attribute should be:

Olive-trees have low area, low elongatedness and high circularity (5.14)

In order to query our shape maps with this statement, a crisp threshold is applied to the three considered parameters in order to establish if they are “low” or “high”. The result of the application of Equation (5.14) to the preliminary olive-trees mask is reported in Figure 5.38c. The retrieved number of trees is 494. A manual count performed by visual inspection of the Google Earth view restituted 517 trees and 26 false alarms of the obtained detections. Therefore, obtained accuracy net of false alarms is in the order of 88.8%. The false alarm rate is in the order of 5.3% of the total detections. Given the extension of the analyzed terrain (about 20 hectares) the retrieved datum is compatible with a planting layout in the order of 15 meters, which is typical of extensive olive growing in Calabria. This data was confirmed by a measure made thanks to Google Earth measurement tools and can be converted by an agronomist in an indication on the olive-trees class.

Low level semantic color attribute and OBIA can be exploited for extracting features otherwise scarcely identifiable

Obviously, this application is possible in the limits dictated by the resolution. In fact, treating intensive cultivations, with planting layouts in the order of 5-7 meters or less, is not possible with stripmap products (three meters spatial resolution). In these cases, spotlight products (about one meter resolution) must be used.

Obviously, this application is possible in the limits dictated by the resolution. In fact, treating intensive cultivations, with planting layouts in the order of 5-7 meters or less, is not possible with stripmap products (three meters spatial resolution). In these cases, spotlight products (about one meter resolution) must be used.

Summary In this Section, we exploited a semantic Level-1 β product and some basic shape parameters for an application of precision agriculture concerning with the extraction of the olive-trees contained in a field. In particular, the color attribute attached to the Level-1 β product through SOM clustering allowed for a scattering-based extraction of a preliminary feature mask. This mask have been then refined by calculating the shape of the identified objects. The obtained results, being highly congruent with the available ground truth, confirmed the reliability of our methodology.

5.5 Conclusions and discussions

In this Chapter, we discussed several applications exploiting, in general, multitemporal SAR images and, in particular, the new methodology defined by the MAP3 framework introduced in Chapter 2.

In Section 5.2, water resources management in semi-arid regions was implemented. We exploited multitemporal SAR images for implementing a simple basin shorelines extraction using a threshold on the intensity images. In this case, only the first part of the MAP3 chain was applied, dealing with geometric registration, radiometric calibration and despeckling (that is almost all the content of the “Pre-processing” block displayed in Figure 2.2). This information, in combination with a high-resolution digital elevation model allowed for the estimate of the retained volume in several reservoirs. Moreover, the obtained results was used for retrieving a semi-empirical relation between basins’ surface and retained volume, providing a quick and ready-to-use tool for the estimate of the available water volume in all regions whose morphological characteristics and climatic conditions are similar to that of the study area (i.e. the Sahel). Operatively, such relation represents a very useful investigation instrument since it allows for avoiding expensive bathymetric (in-situ) surveys.

In Section 5.3, classification activities exploiting using Level-1 α and Level-1 β products were discussed. In particular, in Section 5.3.1, a simple supervised ML classifier was used to classify a Level-1 α product. The obtained results, showing a good agreement with the considered ground truth, confirmed the reliability of MAP3 products to be processed with simple algorithms, widely available on commercial and/or open source software suites and very popular in the end-user community.

In Section 5.3.2, a complex application concerning two years crops monitoring using Level-1 β products in combination with neural networks was faced. In this case, due to the lack of a ground truth, our approach allowed for inferring useful information from the available data about the state of the cultivations in the study area. In Section 5.3.3, land cover mapping through neural networks starting from a Level-1 β products is performed. In this case, the most interesting outcome was that the use of Kohonen’s SOM allowed for the preservation of the input Level-1 β product chromatic response in the classified map. This allowed for an immediate and automatic transferring of the semantic from the RGB map to the

classified product.

Features extraction is discussed in Section 5.4. In particular, in Section 5.4.1, we introduced an index, based on a ratio between Level-1 α bands, for identifying temporary water bodies, i.e. surface water occurring only during the acquisition of the test image. We tested our approach for small reservoirs monitoring in semi-arid environment. The congruence of the obtained results with respect to the available ground truth confirmed the reliability and the effectiveness of the proposed index, which represents a quick and user-oriented solution for the considered problem.

In Section 5.4.2 we showed how to enhance the representation of urban areas in Level-1 α imagery exploiting a feedback system to adapt the dimension of the coherence window to scene's target and textural information. The enhanced Level-1 α product was then used for extracting the built-up feature using, again, a simple band product.

In Section 5.4.3 a precision agriculture application was faced. In this case, we exploited a semantic Level-1 β product and some shape parameters for the extraction of the number of olive-trees contained in a field. The coupling of the low level semantic (constituted by the color attribute attached to the RGB product by a Kohonen's SOM) and OBIA allowed for obtaining good performances in the features extraction process, which showed a good agreement with the considered ground truth.

MAP3 products showed to be a reliable support for all the considered scenarios, exhibiting the versatility, the interpretability and the usability, i.e. the possibility to operate with simple techniques, an end-user needs for his/her applications.

References

- Amitrano, D., Cecinati, F., Di Martino, G., Iodice, A., Riccio, D., and Ruello, G. (2015a). Sentinel-1 Multitemporal SAR Products. In *IEEE International Geoscience and Remote Sensing Symposium*, pages 3973–3976.
- Amitrano, D., Ciervo, F., Di Martino, G., Papa, M. N., Iodice, A., Koussoube, Y., Mitidieri, F., Riccio, D., and Ruello, G. (2014a). Modeling Watershed Response in Semiarid Regions with High Resolution Synthetic Aperture Radars. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 7(7):2732–2745.
- Amitrano, D., Di Martino, G., Iodice, A., Koussoube, Y., Mitidieri, F., Papa, M., Riccio, D., and Ruello, G. (2015b). Hydrological modeling in ungauged basins using SAR data. In *IEEE International Geoscience and Remote Sensing Symposium*, pages 2747–2750.
- Amitrano, D., Di Martino, G., Iodice, A., Mitidieri, F., Papa, M. N., Riccio, D., and Ruello, G. (2014b). Sentinel-1 for Monitoring Reservoirs: A Performance Analysis. *Remote Sensing*, 6:10676–10693.
- Amitrano, D., Di Martino, G., Iodice, A., Riccio, D., and Ruello, G. (2015c). A New Framework for SAR Multitemporal Data RGB Representation: Rationale and Products. *IEEE Transactions on Geoscience and Remote Sensing*, 53(1):117–133.
- Amitrano, D., Di Martino, G., Iodice, A., Riccio, D., and Ruello, G. (2016). An end-user-oriented framework for the classification of multitemporal SAR images. *International Journal of Remote Sensing*, 37(1):248–261.
- Amitrano, D., Di Martino, G., Iodice, A., Riccio, D., Ruello, G., Ciervo, F., Papa, M. N., and Koussoube, Y. (2013). Synthetic Aperture Radar for Humanitarian Purposes: Products and Opportunities. In *IEEE Global Humanitarian Technology Conference*, pages 546–551.
- Amitrano, D., Di Martino, G., Iodice, A., Riccio, D., Ruello, G., Papa, M. N., Ciervo, F., and Koussoube, Y. (2014c). Effectiveness of high-resolution SAR for water resource management in low-income semi-arid countries. *International Journal of Remote Sensing*, 35(1):70–88.
- Amitrano, D., Di Martino, G., Iodice, A., Riccio, D., and Ruello, G. (2015d). Urban Areas Enhancement in Multitemporal SAR RGB Im-

- ages Through a Feedback System. In *IEEE Joint Urban Remote Sensing Event*.
- Annor, F. O., van de Giesen, N., Liebe, J., van de Zaag, P., Tilmant, A., and Odai, S. N. (2009). Delineation of Small Reservoirs Using Radar Imagery in a Semi-Arid Environment: A Case Study in the Upper East Region of Ghana. *Physics and Chemistry of the Earth, Parts A/B/C*, 34(4-5):309–315.
- Arrigoni, M., D’Aria, D., and Monti Guarnieri, A. (2005). Space-adaptive coherence estimation. In *ESA FRINGE Workshop*.
- Blaschke, T. (2010). Object based image analysis for remote sensing. *ISPRS Journal of Photogrammetry and Remote Sensing*, 65(1):2–16.
- Boelee, E., Cecchi, P., and Kone, A. (2009). Health Impacts of Small Reservoirs in Burkina Faso. IWMI Working Paper 136, International Water Management Institute, Colombo, Sri Lanka.
- Brivio, P. A., Colombo, R., Maggi, M., and Tomasoni, R. (2002). Integration of remote sensing data and GIS for accurate mapping of flooded areas. *International Journal of Remote Sensing*, 23(3):429–441.
- Bruzzone, L., Marconcini, M., Wegmuller, U., and Wiesmann, A. (2004). An Advanced System for the Automatic Classification of Multitemporal SAR Images. *IEEE Transactions on Geoscience and Remote Sensing*, 42(6):1321–1334.
- Campbell, J. B. and Wynne, R. H. (2011). *Introduction to Remote Sensing*. The Guilford Press, New York, NY 10012, fifth edition.
- Camps-Valls, G., Gomez-Chova, L., Munoz-Mari, J., Rojo-Alvarez, J., and Martinez-Ramon, M. (2008). Kernel-Based Framework for Multitemporal and Multisource Remote Sensing Data Classification and Change Detection. *IEEE Transactions on Geoscience and Remote Sensing*, 46(6):1822–1835.
- Carlucci, A., Lops, F., Marchi, G., Mugnai, L., , and Surico, G. (2013). Has *Xylella fastidiosa* “chosen” olive trees to establish in the mediterranean basin? *Phytopathologia Mediterranea*, 52(3):541–544.
- Cecchi, P., Meunier-Nikiema, A., Moiroux, N., and Sanou, B. (2009). Towards an atlas of lakes and reservoirs in burkina. In Andreini, M., Schuetz, M., Harrington, T., and Battaramulla, L., editors, *Small reservoir toolkit*. International Water Management Institute, Colombo, Sri Lanka.

- Cox, E. P. (1927). A method of assigning numerical and percentage values to the degree of roundness. *Journal of Paleontology*, 1(3):179–183.
- Davidson, G. and Ouchi, K. (2003). Segmentation of SAR images using multitemporal information. *IEE Proceedings Radar, Sonar and Navigation*, 150(5):367–374.
- Davranche, A., Lefebvre, G., and Poulin, B. (2010). Wetland monitoring using classification trees and spot-5 seasonal time series. *Remote Sensing of Environment*, 114(3):552–562.
- Dellepiane, S. G. and Angiati, E. (2012). A New Method for Cross-Normalization and Multitemporal Visualization of SAR Images for the Detection of Flooded Areas. *IEEE Transactions on Geoscience and Remote Sensing*, 50(7):2765–2779.
- Deng, H. and Clausi, D. A. (2005). Unsupervised Segmentation of Synthetic Aperture Radar Sea Ice Imagery Using a Novel Markov Random Field Model. *IEEE Transactions on Geoscience and Remote Sensing*, 43(3):528–538.
- Diamond, J. (2005). *Collasso. Come le società scelgono di morire o vivere*. Einaudi.
- Engdahl, M. E. and Hyypä, J. M. (2003). Land-Cover Classification Using Multitemporal ERS-1/2 Insar Data. *IEEE Transactions on Geoscience and Remote Sensing*, 41(7):1620–1628.
- European Environment Agency (2011). Mapping Guide for a European Urban Atlas. Technical report, European Commission.
- Feranec, J., Hazeu, G., Christensen, S., and Jaffrain, G. (2007). CORINE land cover change detection in Europe (case studies of the Netherlands and Slovakia). *Land Use Policy*, 24(1):234–247.
- Frazier, P. S. and Page, K. J. (2000). Water Body Detection and Delineation with Landsat TM Data. *Photogrammetric Engineering & Remote Sensing*, 66(12):1461–1467.
- Fung, A. K. (1979). Scattering from a Vegetation Layer. *IEEE Transactions on Geoscience Electronics*, 17(1):1–6.
- Gaetano, R., Amitrano, D., Masi, G., Poggi, G., Verdoliva, A., Ruello, G., and Scarpa, G. (2014). Exploration of Multitemporal COSMO-SkyMed Data Via Tree-Structured MRF Segmentation. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 7(7):2763–2775.

- Gonçalves, M. L., Costa, J. A. F., and Netto, M. L. A. (2011). Land-Cover Classification Using Self-Organizing Maps Clustered with Spectral and Spatial Information. In Mwasiagi, J. I., editor, *Self Organizing Maps - Application and Novel Algorithm Design*, pages 299–322. InTech.
- Gonzalez, R. C. and Woods, R. E. (2007). *Digital Image Processing*. Prentice Hall, Upper Saddle River, NJ.
- Hanssen, R. F. (2001). *Radar Interferometry - Data Interpretation and Error Analysis*. Kluwer Academic Publishers, Dordrecht.
- Haralick, R. M. and Shapiro, L. G. (1985). Image segmentation Techniques. *Computer Vision, Graphics and Image Processing*, 29:100–132.
- Harger, R. O. (1973). Synthetic Aperture System Design for Random Field Classification. *IEEE Transactions on Aerospace and Electronic Systems*, 1(5):732–740.
- Iervolino, P., Guida, R., Iodice, A., and Riccio, D. (2015). Flooding Water Depth Estimation With High-Resolution SAR. *IEEE Transactions on Geoscience and Remote Sensing*, 53(5):2295–2307.
- J. H. J. Ward (1963). Hierarchical Grouping to Optimize an Objective function. *Journal of the American Statistical Association*, 58(30):236–244.
- Kohonen, T. (2001). *Self-Organizing Maps*. Springer-Verlag, Berlin, Heidelberg.
- Langford, M., Higgs, G., Radcliffe, J., and White, S. (2008). Urban population distribution models and service accessibility estimation. *Computers, Environment and Urban Systems*, 32(1):66 – 80.
- Liebe, J., van de Giesen, N., and Andreini, M. (2005). Estimation of small reservoir storage capacities in semi-arid environment: A case study in the Upper East Region of Ghana. *Physics and Chemistry of the Earth, Parts A/B/C*, 30:448–454.
- Liebe, J., van de Giesen, N., Andreini, M., Steenhuis, T., and Walter, M. (2009). Suitability and Limitations of ENVISAT ASAR for Monitoring Small Reservoirs in a Semiarid Area. *IEEE Transactions on Geoscience and Remote Sensing*, 47(5):1536–1547.
- Marr, D. (1982). *Vision*. W. H. Freeman, San Francisco.
- Marr, D. and Hildreth, E. (1980). Theory of edge detection. *Proceedings of the Royal Society B*, 207(1167):187–217.
- Marroquin, J., Mitter, S., and Poggio, T. (1987). Probabilistic solution

- of ill-posed problems in computational vision. *Journal of the American Statistical Association*, 82(397):76–89.
- Martinez, J. and Le Toan, T. (2007). Mapping of flood dynamics and spatial distribution of vegetation in the Amazon floodplain using multitemporal SAR data. *Remote Sensing of Environment*, 108(3):209–223.
- Martinis, S., Kuenzer, C., Wendleder, A., Huth, J., Twele, A., Roth, A., and Dech, S. (2015). Comparing four operational sar-based water and flood detection approaches. *International Journal of Remote Sensing*, 36(13):3519–3543.
- Mason, D. C., Schumann, G. J.-P., Neal, J. C., Garcia-Pintado, J., and Bates, P. D. (2012). Automatic near real-time selection of flood water levels from high resolution Synthetic Aperture Radar images for assimilation into hydraulic models: A case study. *Remote Sensing of Environment*, 124:705–716.
- Matsuyama, T. and Hwang, V. S.-H. (1990). *SIGMA - A Knowledge-Based Aerial Image Understanding System*. Plenum Press, New York.
- McFeeters, S. K. (1996). The use of the Normalized Difference Water Index (NDWI) in the delineation of open water features. *International Journal of Remote Sensing*, 17(7):1425–1432.
- McFeeters, S. K. (2013). Using the Normalized Difference Water Index (NDWI) within a Geographic Information System to Detect Swimming Pools for Mosquito Abatement: A Practical Approach. *Remote Sensing*, 5(7):3544–3561.
- Mutiti, S., Levy, J., Mutiti, C., and Gaturu, N. S. (2010). Assessing Ground Water Development Potential Using Landsat Imagery. *Ground water*, 48(2):295–305.
- Nagao, M. and Matsuyama, T. (1980). *A Structural Analysis of Complex Aerial Photographs*. Plenum Press, New York.
- Niedermeier, A., Romaneeß en, E., and Lehner, S. (2000). Detection of Coastlines in SAR Images using Wavelet Methods. *IEEE Transactions on Geoscience and Remote Sensing*, 38(5):2270–2281.
- Ouma, Y. O. and Tateishi, R. (2006). A water index for rapid mapping of shoreline changes of five East African Rift Valley lakes: an empirical analysis using Landsat TM and ETM+ data. *International Journal of Remote Sensing*, 27(15):3153–3181.

- Pal, N. R. and Pal, S. K. (1993). A Review on Image Segmentation Techniques. *Pattern Recognition*, 26(9):1277–1294.
- Partington, J. R. (1989). *An introduction to Hankel operators*. Cambridge University Press, Cambridge, UK.
- Quegan, S., Le Toan, T., Yu, J. J., Ribbes, F., and Floury, N. (2000). Multitemporal ERS SAR Analysis Applied to Forest Mapping. *IEEE Transactions on Geoscience and Remote Sensing*, 38(2):741–753.
- Richards, J. A. and Jia, X. (2006). *Remote Sensing Digital Image Analysis - An Introduction*. Springer-Verlag, Berlin, Germany, fourth edition.
- Ronse, C. and Serra, J. (2010). *Algebraic Foundations of Morphology*. John Wiley & Sons, Hoboken, NJ.
- Roshier, D. A., Whetton, P. H., Allan, R. J., and Robertson, A. I. (2001). Distribution and persistence of temporary wetland habitats in arid Australia in relation to climate. *Austral Ecology*, 26(4):371–384.
- Sakamoto, T., Nguyen, N. V., Kotera, A., Ohno, H., Ishitsuka, N., and Yokozawa, M. (2007). Detecting temporal changes in the extent of annual flooding within the cambodia and the vietnamese mekong delta from {MODIS} time-series imagery. *Remote Sensing of Environment*, 109(3):295–313.
- Shackelford, A. K. and Davies, C. H. (2003). A Combined Fuzzy Pixel-Based and Object-Based Approach for Classification of High-Resolution Multispectral Data Over Urban Areas. *IEEE Transactions on Geoscience and Remote Sensing*, 41(10):2354–2363.
- Shapiro, L. and Stockman, G. (2002). *Computer Vision*. Prentice Hall, Upper Saddle River, NJ.
- Shrivakshan, G. and Chandrasekar, C. (2012). A comparison of various edge detection techniques used in image processing. *International Journal of Computer Science Issues*, 9(5-1):269–276.
- Tobias, O. J. and Seara, R. (2002). Image segmentation by histogram thresholding using fuzzy sets. *IEEE Transactions on Image Processing*, 11(12):1457–1465.
- Townshend, J. R. G. (1981). *Terrain Analysis and Remote Sensing*. George Allen and Unwin, London.
- Wang, F., Wu, Y., Zhang, Q., Zhao, W., Li, M., , and Liao, G. (2014). Un-supervised SAR image segmentation using higher order neighborhood-

- based triplet markov fields model. *IEEE Transactions on Geoscience and Remote Sensing*, 52(8):5193–5205.
- Xu, H. (2006). Modification of normalised difference water index (NDWI) to enhance open water features in remotely sensed imagery. *International Journal of Remote Sensing*, 27(14):3025–3033.
- Zhang, A. and Jia, G. (2013). Monitoring meteorological drought in semi-arid regions using multi-sensor microwave remote sensing data. *Remote Sensing of Environment*, 134:12–23.

This page intentionally left blank.

Chapter 6

Conclusions

The limited usage of SAR data in the end-user community and in applicative contexts testified the failure of the recent literature, in which the research privileged the automatic extraction of information at the expense of users' experience with data. In this Book, we introduced several human-centered concepts for the understanding of SAR images. The necessity to restore the centrality of the user in remote sensing data processing has been widely expressed throughout the work and achieved thanks to the introduction of the MAP3 framework, leading to the definition of two classes of user-oriented RGB products we named as Level-1 α and Level-1 β . They have two principal characteristics:

- Ease of interpretation, thanks to an effective multitemporal processing and consistent rendering of suitable RGB channels;
- Possibility to be processed with simple, end-user-oriented algorithms.

These properties should definitely fill the gap between the academy and the applications. In fact, the rationale is to provide ready-to-use images, in which the technical expertise with electromagnetic models, SAR imaging and image processing has been absorbed in the products formation phase. In such way, the idea that SAR images are too complicated to be interpreted and processed in order to extract physical information is overcome.

To conclude the work, we want to linger on the following consideration: what is the role of the electromagnetic models in remote sensing data analysis?

As first, electromagnetic models constitute the basis of the interpretation of any SAR product. In fact, unlike the case of optical data, humans have no experience of radar imaging directly linked to the world they live, and from which they can take inspiration for understanding data. Therefore, Level-1 α and Level-1 β imagery gain their semantic from the knowledge of the phenomenology dictated by the interaction of the electromagnetic field with physical surfaces. The proposed products absorb the expertise required for understanding these complex mechanisms, re-elaborating the scene dynamic in a more user-friendly color display in which the non-expert user can more easily encounter his/her expectation about object appearance. However, models are often left aside in favour of the development of techniques, which are highly conditioned by the application and by the correct selection of free parameters.

We found a lack of integration between these two levels. Actually, we think that a data analysis technique, such as a neural-net-based method, can be a valid alternative to the application of electromagnetic models, which are probably out of reach for non-expert user, when high-level information are sought. Instead, understanding the scene at its lowest level requires the application of scattering models. As an example, the knowledge of the precise destination of a terrain belonging the class “Cropfields” requires the mastery of a vegetation scattering model. In the same way, the retrieval of the height of an object classified as “Urban structure” requires the knowledge of the complex scattering mechanism triggered by an urban environment.

At the end of the day, the message we convey is that a more effective integration of techniques and models is needed, especially looking toward automatic systems of image understanding which seem to become the essential core of remotely sensed data analysis in a big data scenario. We think that the full development of such systems can provide a complete answer to the challenges of modern remote sensing.