PhD Dissertation

International Doctorate School in Information and Communication Technologies

DISI - University of Trento

Advanced methods for change detection in VHR multitemporal SAR images

Carlo Marin

Advisor:

Prof. Lorenzo Bruzzone

University of Trento

Co-Advisor:

Dr. Francesca Bovolo

Bruno Kessler Foundation

March 2015

In girum imus nocte et consumimur igni

Abstract

Change detection aims at identifying possible changes in the state of an object or phenomenon by jointly observing data acquired at different times over the same geographical area. In this context, the repetitive coverage and high quality of remotely sensed images acquired by Earth-orbiting satellites make such kind of data an ideal information source for change detection. Among the different kinds of Earth-observation systems, here we focus on Synthetic Aperture Radar (SAR). Differently from optical sensors, SAR is able to regularly monitor the Earth surface independently from the presence of cloud cover or sunlight illumination, making SAR data very attractive from an operational point of view.

A new generation of SAR systems such as TerraSAR-X, TANDEM-X and COSMO-SkyMed, which are able to acquired data with a Very High geometrical Resolution (VHR), has opened new attractive opportunities to study dynamic phenomena that occur on the Earth surface. Nevertheless, the high amount of geometrical details has brought several challenging issues related to the data analysis that should be addressed. Indeed, even though in the literature several techniques have been developed for the automatic analysis of multitemporal low- and medium-resolution SAR data, they are poorly effective when dealing with VHR images. In detail, in this thesis we aim at developing advanced methods for change detection that are able to properly exploit the characteristics of VHR SAR images.

i) An approach to building change detection. The approach is based on a novel theoretical model of backscattering that describes the appearance of new or fully collapsed buildings. The use of a fuzzy rule set allows in real scenarios an efficient and effective detection of new/collapsed building among several other sources of changes.

ii) A change detection approach for the identification of damages in urban areas after catastrophic events such as earthquakes or tsunami. The approach is based on two steps: first the most damaged urban areas over a large territory are detected by analyzing high resolution stripmap SAR images. These areas drive the acquisition of new VHR spotlight images, which are used in the second step of the approach to accurately identify collapsed buildings.

iii) An approach for surveillance applications. The proposed strategy detects the changes of interest over important sites such as ports and airports by performing a hierarchical multiscale analysis of the multitemporal SAR images based on a Wavelet decomposition technique.

iv) An approach to multitemporal primitive detection. The approach, based on the Bayesian rule for compound classification integrated in a fuzzy inference system, takes advantage of the multitemporal correlation of images pairs in order to both improve the detection of the primitives and identify the changes in their state.

For each of the above mentioned topic an analysis of the state of the art is carried out, the limitations of existing methods are pointed out and the proposed solutions to the

considered problems are described in details. Experimental results conducted on simulated and real remote sensing data are provided in order to show and confirm the validity of each of the proposed methods.

Keywords:
Remote sensing, synthetic aperture radar, very high geometrical resolution images, change detection, surveillance, building change detection, multitemporal primitive detection, damage assessment.

Acknowledgments

I get to put my name on the cover of this thesis, but many are the people who contributed in part to the work and who supported me during these years. Without them, it wouldn't have been nearly as much fun.

First, I would like to thank Prof. Lorenzo Bruzzone for the privilege of having him as my supervisor during these three years. He showed me how to do research, and his clear-mind and forward thinking have been really stimulating points for me.

A special thank goes to Dr. Francesca Bovolo. Her influence and contributions were critical in many situations and are spread throughout this work.

I would like to thank all the staff of the Remote Sensing Laboratory at the Department of Information Engineering and Computer Science, University of Trento. In particular, I would like to thank Max for the extremely long discussions about new topics and for standing me as his co-supervisor, and Claudio who decided to stay in this open space for a while sharing in this way his valuable research experience with us.

During my career as PhD student I had the opportunity to meet some of the coolest people there are. I would like to thank the people I met here in Trento and to whom I shared a lot: Dani, the guy who is so patient to guide me around mountains, Ema with whom I enjoy to rope up in mountain, Luca Parma for his extraordinary positive attitude to life, Marco and Francesca (who will certainly find her way), Daniel for being so spontaneous, Calle for his cheerful personality, Teo and Anna (for the climbing training at the gym), the historical friends Bezzu and Gardu, Davide for sharing so much with me and Dr. Liu who arrived a little bit before me to the magical end. It is the calibre of these people and the warmth of their friendship that makes this a great place to live.

But foremost I would like to thank my family for their support. I was fortunate to have them by my side during these years. The accomplishments of this work are them as much as mine.

Contents

List of Tables

List of Figures

List of Symbols

B	constant for split selection
B_r	bandwidth of a chirp
$D_{\mathrm{KL}}(\cdot,\cdot)$	Kullback-Leibler (KL) distance
H	platform height
L	Equivalent Number of Looks (ENL)
L_{sa}	synthetic aperture length
$N-1$	optimal building level
Q^n	length of the wavelet filters at resolution level n
$RCM(t)$	range cell migration at time t
S_A	split length along azimuth
S_R	split length along range
T	pulse time
T_M	threshold on the aggregate membership
T_c	threshold for extracting the candidate regions Γ from the index image $\mathbf{C}$
Δ	set of regions of change that made up a building candidate
Γ	set of changed building candidates
Σ	response of linear detector
Θ_a	beamwidth of an antenna
Θ_{sa}	beamwidth of a synthetic antenna
$\bar{\sigma}^2$	average variance of splits
δ_a	azimuth resolution
δ_r	slant-range resolution
λ	wavelength
$\mathbf{C}$	index of the size of change
$\mathbf{M}$	change detection map
$\mathbf{M}^{\mathrm{opt}}$	map of changes at the optimal scale
$\mathbf{X}_{\mathrm{LR}}$	log-ratio image
$\mathbf{X}_{\mathrm{NR}}$	normalized-ratio image
$\mathbf{X}_1$	image acquired at time t_1
$\mathbf{X}_2$	image acquired at time t_2
$\mathbf{X}_D$	difference image
$\mathbf{X}_{LR}^{HH}$	high frequency log-ratio diagonal detail subband
$\mathbf{X}_{LR}^{HL}$	high frequency log-ratio vertical detail subband
$\mathbf{X}_{LR}^{LH}$	high frequency log-ratio horizontal detail subband

$\mathbf{X}_{LR}^{LL}$	log-ratio approximation subband
$\mathbf{X}_R$	ratio image
$\mathbf{X}_{\mathrm{LR}}$	log-ratio image
$\mathbf{X}_{\mathrm{LR}}^n$	multiscale representation of the log-ratio information at the n-th level
$\mathbf{X}_\tau$	VHR SAR images acquired at time τ
$\mathbf{X}_{\mathrm{MRD}}$	Mean Ratio Detector (MRD)
$\mathcal{H}(\cdot)$	homogeneity predicate
$\mathcal{P}(\cdot,\cdot)$	set of parcels associated with two images
$\mathcal{T}$	universe of discourse of changes
$\mathcal{U}$	universe of discourse of primitive
$\mathcal{X}_{\mathrm{MS}}$	set of images at different resolution levels
μ	mean
μ_p	membership function to test the degree of completeness
μ_{r_l}	membership function to test the degree of the equivalence of the lengths
μ_{r_s}	membership function to test the degree of the proportionality of the areas
ω_p	class of primitive
ω_{c1}	class of fully destroyed buildings
ω_{c2}	class of new buildings
ω_{c3}	class of changes that have size comparable to a building but do not present the typical pattern of full new/demolished buildings
ω_{c4}	class of changes that do not show a size comparable to a building
ω_{np}	class of non-primitive
ω_u	class of unchanged pixels
ϕ	aspect angle
ρ	response of linear detector D2
ρ_i	i-th parcel
σ^0	radar cross section
σ_s^2	split variance
σ_{σ^2}	standard deviation of the variance of splits
$\mathrm{CKLD}(\cdot,\cdot)$	cumulant-based Kullback-Leibler distance (CKLD)
CR	empirical contrast
LCV	local coefficient of variation
θ	incidence angle
φ^{scatt}	scattering phase
ξ^+	increase in backscattering class
ξ^-	decrease in backscattering class
ξ_u	unchanged class
a	parameters for tuning the slope of the sigmoid
b	parameters for tuning the center of the sigmoid
c_0	speed of light
d_a	length of an antenna

f, g, h	parameters for tuning the triangular membership function
f_i	instantaneous frequency
k_r	chirp rate
$n = 0, 1, \ldots, N-1$	resolution level
r	response of ship detector
$r(t)$	distance between SAR and a point on the ground at time t
r_0	shorter approach distance
r_E	ratio edge detector response
r_L	response of linear detector D1
r_{PT}	response of point detector
v	sensor velocity
$w_1 \times w_2 \times h$	building dimensions
w_b	building width limit
$x_{\tau,j}$	$j-th$ pixel of the monotemporal features that describe a primitive
$x_{c,j}$	$j-th$ pixel of the index of change
$z_1 \times z_2$	window size for the search of candidates
$z_{c,j}$	$j-th$ pixel of output of the multitemporal primitive detector

Chapter 1

Introduction

This chapter present an overview of the research proposed in the thesis. In detail, after introducing the problem of change detection in multitemporal very high geometrical resolution SAR images, the motivations, the objectives and the novel contributions of this thesis are highlighted and discussed. Finally, the structure and organization of the document is illustrated.

1.1 Introduction and Motivations of the Thesis

Remote sensing can be defined as a technology for acquiring measurements and information about a given phenomenon without being in physical contact with it. Due to its capability to provide information about the Earth surface through the use of passive and active sensors mounted on airborne or space-borne platforms, this technology recived a constantly growing interest from public and private institutions in the last decades. This interest stems from the fact that remote sensing represents a useful source of information in decision making for a wide range of applications [1, 2].

Due to the large amount of data to process, the extraction of useful information from remotely sensed data requires the definition of automatic and possibly unsupervised analysis techniques. Generally speaking, a remote sensing recognition scheme, which has the task to recognize in an automatic or semi-automatic way the physical reality of the phenomenon under study, is typically based on three major phases: (i) the data collection phase carried out by the remote sensor; (ii) the pre-processing phase, which aims at correcting all the possible signal distortions; (iii) the analysis phase for the extraction of the required information from the pre-processed remote sensing data. Regarding this last point, several approaches may be adopted depending on the focus of the final application. In this thesis, we focus on change detection (CD), which purpose is to identify possible changes in the state of an object or phenomenon by jointly observing signals acquired at different times.

CD is based on the analysis of two co-registered multitemporal remote-sensing images acquired at different times over the same geographical area. CD has of primary importance for a large number of applications, including urban planning [3], natural resources monitoring [4], agricultural surveys [5], natural hazard prevention and monitoring [6],

damage assessment [7]. In this context, optical sensors have been extensively exploited and different automatic and unsupervised change-detection methodologies have been developed. Synthetic Aperture Radar (SAR) has been less exploited than optical sensors in the context of change detection. This is due to the fact that SAR images, being acquired by a coherent side looking system, suffer from the presence of both an intrinsic speckle noise and geometry distortions, which make any automatic analysis difficult. Despite these problems, the use of SAR sensors in change detection is highly attractive from the operational viewpoint. In fact, SAR presents the advantage over optical sensors of being independent of sunlight illumination and almost insensitive to atmospheric conditions. This means for instance that it is capable of monitoring geographical areas regularly, even if covered by clouds or during night time, which is not the case of passive sensors, making it possible to ensure the data acquisition on an area of interest in advance according to end-user requirements[1] (e.g., seasonal and agricultural calendars [8, 9]) or during a crisis event (e.g., hurricanes, floods [10]).

In the past few years, we have observed an increased availability of very high geometrical resolution (VHR) SAR images. This is due to a new generation of satellites, such as TerraSAR-X [11], TANDEM-X [12], and COSMO-SkyMed [13], which can acquire images with a nominal resolution of up to 1 meter. This was impossible with traditional space-borne SAR systems, which are able to acquire only high and medium geometrical resolution SAR images (i.e., with resolution ranging from 10 to 150 meters). Therefore, the accessibility of VHR SAR images regularly acquired with satellites platforms opens new attractive opportunities for both the scientific and the user communities. Although the literature presents several works related to change-detection techniques applied to high- and medium-resolution images, only few techniques have been developed in the last years for the detection of changes in VHR data. Moreover, the techniques developed for low- and medium-resolution images are largely insufficient to exploit in a proper way the richness of information contained in multitemporal VHR data. In this context, we aim at developing effective and robust unsupervised change-detection methods for VHR multitemporal SAR images. In detail, we focus our attention on two challenging change detection scenarios. The former is related to detection of changed buildings, either after either catastrophic events or for urban expansion analysis. The latter is related to surveillance and monitoring problems of sites of general interest, such as ports, airports, industrial sites, and so on for which the short repetition interval guaranteed by the 4 satellites of the CSK constellation (i.e., 24 hours) can be exploited. Nevertheless, the combination of short repetition interval and very high geometrical resolution leads to some challenging issues that have to be addressed. These issues are described below.

1. Differently from the low and medium resolution SAR images, in VHR data only a small number of elementary scatterers are present inside the resolution cell and hence more features of the investigated ground object become visible. This increases the complexity of the electromagnetic backscattering mechanisms of the imaged objects. For instance a building is identified in VHR SAR images considering the composition of the backscattering contributions coming from its walls, its roof and its surrounding ground. Hence, in order to properly detect changes in the objects present in the scene

[1] In practice, the availability of the data is subjected also to the number of requests that the provider has to deliver.

it is necessary to take into account backscattering contributions from sub-objects of the investigated object.

2. In VHR SAR images changes may present an extension that varies from the single pixel to a relevant portion of the entire scene. Moreover, objects that are considered homogeneous from a semantic point of view may show a signature that is inhomogeneous because of the scattering contributions from sub-objects that compose them. This leads to a multiscale nature of the problem in which different scale levels are more appropriated to detect changes with that specific scale.

3. A large set of possible changes with different semantic meaning is detectable in multitemporal VHR SAR images. For instance it is possible to distinguish among changes due to the anthropogenic activity or the natural disaster effects. Depending on the application, some of these may be of interest to the end-users, whereas others may not.

4. External factors, such as different content of water on the ground due to different weather conditions may affect the local backscattering behaviors at the two considered dates also in absence of any change. Therefore, the same object may show different values of backscattering even though it is not affected by a relevant change, leading to a large amount of false alarms.

5. The very high geometrical resolution of new generation sensors and the progress in the technology result in a higher amount of data to analyze than in the case of previously available data. Therefore, for a proper exploitation of VHR images in real applications related to extended areas, it is mandatory to develop effective automatic techniques that can properly handle the huge amount of available data in addition to the high geometric content.

All these factors make the problem of the detection of changes in VHR SAR images complex. Indeed, they invalidate the usual assumption that two SAR images acquired on the same geographical area at different times are similar to each other except for the presence of changes occurred on the ground, which is often considered for medium-resolution SAR images. Thus the use of standard pixel based change detection techniques [14] is not possible as they would be affected by a large amount of false alarms. Given the complexity of the problem, also standard context-based techniques based on a local neighborhood analysis [15] would fail to solve the problem since they are not able to properly take into account the high geometrical detail of VHR data. To achieve change detection in VHR SAR images the analysis between the multitemporal images should be performed at a higher conceptual level that models the source of change from the prospective of interactions with the incidence electromagnetic wave and the available prior information about the scene.

1.2 Novel Contributions of the Thesis

On the basis of the analysis on the problems related to the automatic detection of changes in VHR SAR images carried out in the previous section, this thesis focuses on the devel-

opment of advanced techniques for unsupervised change detection that take into account the peculiar properties of VHR SAR images. In greater details, the following main novel contributions have been introduced in this work:

1. an approach to building change detection in multitemporal very high resolution SAR images;
2. an approach to rapid and accurate damage detection in built-up areas combining high (stripmap) and very high(spotlight) resolution SAR images;
3. a hierarchical approach to change detection in very high resolution SAR images for surveillance applications;
4. a multitemporal detector for the extraction of primitive from very high resolution SAR images.

In the next sub-sections the main objectives and novelties of these contributions are described.

Building change detection in multitemporal very high resolution SAR images

Monitoring of urban areas is of great importance for several applications such as urban planning, cadastral map updating, environmental monitoring, disaster assessment and so on. The increasing availability of images having a resolution of a meter or less allows the analysis of urban areas at a detail level never reached before resulting in the possibility of detecting buildings individually. Nonetheless, the high amount of geometrical details renders the detection of changes in buildings difficult for the reasons described in the previous section. In this thesis we introduce a novel approach that takes advantage from a theoretical modeling of the backscattering mechanisms of new and fully destroyed buildings in multitemporal VHR SAR images. In detail, new and destroyed buildings can be identified by a pattern made up of an area of both increase and decrease of backscattering with spatial properties and alignment depending on the proprieties of the buildings. In order to extract the changes associated with increase and decrease of backscattering, the proposed approach makes use of a multiscale representation of the multitemporal information allowing a detection of changes at the optimal building scale. This information is used to identify the candidates to be changed buildings. The building candidates are analyzed in order to properly detect the new or destroyed building by means of four fuzzy rules. The fuzzy rules are formulated by taking into account the proposed multitemporal changed building model. The aggregated membership resulting from the application of the fuzzy rules makes it possible to identify the class of each building candidate (i.e., new/destroyed building or general change with size comparable to the building size but not related to new or destroyed building).

Rapid and accurate damage detection in built-up areas combining high (stripmap) and very high (spotlight) resolution SAR images

Data coming from satellites equipped with SAR represent a useful tool to support decisions for initiating effective emergency response actions after catastrophic events. In particular

the new generation of SAR missions can perform acquisitions using two high resolution modes called StripMap (SM) and SpotLight (SL). SM is characterized by a wide coverage (around 1500 Km^2) with a high resolution (3 to 5 m), whereas SL is characterized by a moderate coverage (around 100 Km^2) with a very high resolution (1 m). This means that the two SAR modes are complementary to each other and can be used to perform different duties during the emergency response phase. In this thesis we propose to analyze the high resolution SM image to detect the urban changed areas (hot-spots) spread around the large observed territory. This is done using a multiscale strategy based on the Wavelet transform. Hot-spots are exploited to drive the selection of SL pre-crisis VHR SAR images from the archive and the acquisition of new post-crisis VHR SAR images. These images are used to perform change detection at high spatial resolution with the goal to detect collapsed buildings. This analysis is conducted by taking advantage of the theoretical model of new/destroyed building and the detection method illustrated in chapter 3 of the thesis, by exploiting the Curvelet transform (instead of Wavelet) in order to better preserve the linear details that characterize urban structures in VHR SAR images.

A hierarchical approach to change detection in very high resolution SAR images for surveillance applications

The main difficulty in a surveillance scenario is that the investigated scene can include a wide diversity of man-made structures (buildings, antennas, roads, lights, ground vehicles, etc), showing various shapes and sizes, different materials (metallic, asphalt, etc.) and varying orientations from the sensor. These wide diversity structures generate VHR SAR images showing different high-complexity backscattering behaviors. In order to deal with this scenario, the proposed approach identifies changed areas (hot-spots) at different resolution levels according to a multiscale technique based on the 2-dimensional stationary Wavelet transform. Each hot-spot is analyzed according to its spatial position within a specific area of interest by using the available prior information. The use of such information allows one to define the best change detectors to be used for extracting the expected changes at detailed scale. It is worth noting that the assumption on the availability of the prior information on the usage of different areas (and thus on the expected kinds of change) is reasonable since the method has been developed for high frequency surveillance/monitoring of sensitive areas such as maritime ports, airports, and so on.

A novel multitemporal detector for the extraction of primitive from very high resolution SAR images

As pointed out int the previous section external factors (e.g., different content of water on the ground due to different weather conditions) affect the backscattering behaviour altering the performance of object/primitive detection from an acquisition to another acquisition even though the same SAR set-up is used. In this thesis we propose a novel approach that exploits the multitemporal information in order to: i) perform the detection of primitives at a given time (e.g., lines, object with specific shape); and ii) identify the transitions in the state of primitives between two consecutive observations. This is done by introducing a compound detection approach based on the Bayesian framework. Nonetheless, due to the difficulty to estimate the statistical terms in an unsupervised way,

the proposed compound detection mechanism is integrated with the fuzzy logic. Fuzzy logic is used in order to test the semantic consistency among the state of primitives at the different temporal observations and infer the state of a primitive at a given time. This allows the compound detection to avoid the problems of the unsupervised statistical estimation. Even though the proposed approach is presented for primitive detection in VHR SAR images it has general validity and can be exploited for primitive detection in high- and medium-resolution SAR images.

1.3 Structure of the Thesis

The thesis is organized in seven chapters. The present chapter pointed out the motivations for this research, and highlighted the objectives as well as the main novel contributions.

Chapter 2 provides the basic knowledge about the SAR systems and the state of the art regarding the unsupervised change detection techniques for SAR images.

Chapter 3 investigates the appearance of changed buildings in multitemporal VHR SAR imagery. From this study a novel model that describes the backscattering behavior of the new and collapsed buildings is introduced and exploited in the developing of a novel building change detection approach. The effectiveness of the method has been validated on COSMO-SkyMed multitemporal spotlight images acquired in 2009 on the city of L'Aquila (Italy) before and after the earthquake that hit the region, and TerraSAR-X multitemporal spotlight images acquired on the urban area of the city of Trento (Italy).

Chapter 4 introduces the method for the rapid and accurate damage detection in built-up areas combining stripmap and spotlight SAR images. In particular, the quick identification of the hot spots after a catastrophic event is performed using stripmap data. Meanwhile the accurate detection of individually destroyed building relies on concepts presented in chapter 3. As test case of COSMOSky-Med images acquired after the earthquake event of L'Aquila (Italy) is considered.

Chapter 5 presents a hierarchical approach to change detection in very high resolution SAR images for surveillance applications. The method exploits both the multiscale proprieties of VHR SAR images and the prior information about the site to be monitored in order to extract the expected change. The characteristics of the method are highlighted using two complex data-sets acquired by the COSMOSky-Med constellation over the port and freight village in Livorno, Italy.

Chapter 6 illustrates the possibility to take advantage of the multitemporal information in order to increase the performance of primitive detectors. This is done by introducing a Bayesian compound detection mechanism integrated with the fuzzy logic, which avoid the problems of the unsupervised statistical estimation. Simulated and real TerraSAR-X images acquired over the city of Lüneburg in Germany are used to evaluate the performance of the proposed compound approach.

Finally, chapter 7 draws the conclusions of this work, and discusses future developments for the research conducted in the framework of this thesis.

All the chapters in this dissertation are to be considered as independent to each other and therefore result self-consistent. Readers interested to one of the above mentioned topics can read a single chapter without the need of reading the whole dissertation.

Chapter 2

Fundamentals on the analysis of multitemporal SAR images

The aim of this chapter is to provide the reader with the basic knowledge about synthetic aperture radars and the state of the art on unsupervised change detection for both low/medium and very high resolution multitemporal SAR images. The chapter is divided in three sections. In the first section the basic principles of SAR systems are reviewed. In detail, it provides the general aspect of radar remote sensing along with the detailed explanation of the SAR image formation. In the second section a general overview of the framework commonly used in unsupervised change-detection for low- and medium resolution SAR images is given. Moreover, for each step of the framework a survey on the techniques presented in the literature and exploited in this thesis is reported. A final section reviews the unsupervised techniques developed for very high resolution SAR data.

2.1 Background on SAR

SAR is an active radar system operating in the microwave region of the electromagnetic spectrum usually between P-band (i.e., 0.25-0.5 GHz) and Ka-band (i.e., 25-40 GHz). The inherent characteristics of SAR measurements render this sensor sensitive to dielectric and roughness properties of the investigated natural media. The active operating mode makes SAR independent of solar illumination allowing day and night acquisitions. Moreover, the effects of clouds, fog, rain, smokes, etc., are mostly avoided since the sensor operates in the microwave portion of the electromagnetic spectrum. SAR systems illuminate the scene using a side-looking geometry (see Fig. 2.1). The antenna of the radar system is mounted on a flying platform. Its horizontal and vertical axes are parallel and orthogonal to the azimuth direction, respectively. The angle between nadir and the radar beam direction is called incidence angle and it is usually denoted by θ. Such a system illuminates the Earth surface with microwave pulses and receives the electromagnetic signal backscattered from the illuminated scene. The amplitude and phase of the backscattered signal are related to the physical (e.g., geometry, roughness, dielectric proprieties) and electrical properties (e.g., permittivity) of the imaged object. Depending on the working frequency of the SAR, different levels of penetration can be achieved. More penetration

of the electromagnetic pulses in media occurs for radar systems using longer wavelengths. Besides these advantages, raw SAR acquisitions are far to be a conventional image; they rather resemble holograms wherein the useful information is hidden. A considerably amount of signal processing needs to be done to form the image. With the advent of digital signal processing (DSP) techniques and powerful computational resources, however, this is not a limiting factor for the development of SAR applications anymore. By means of signal processing techniques that exploit the Doppler shifts of the received electromagnetic echoes, SAR systems are able to synthesize a two-dimensional high spatial resolution image from all the received signals as we will see in the next section.

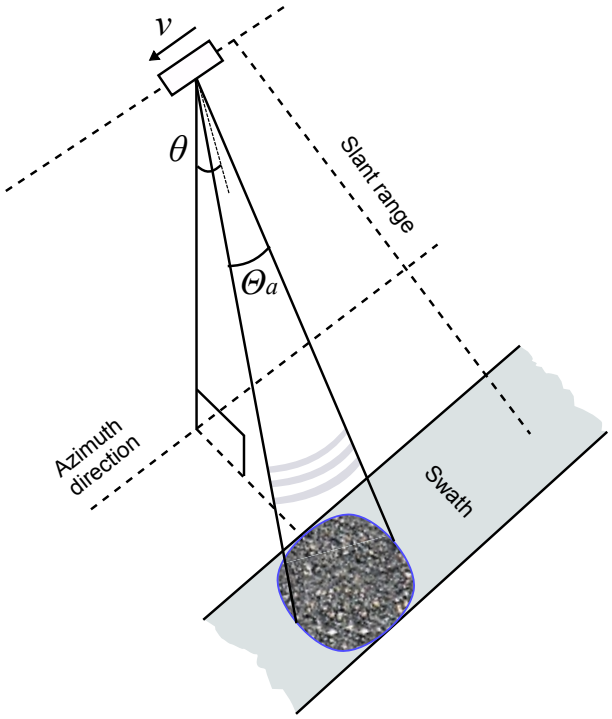


Figure 2.1: SAR imaging geometry.

2.1.1 Basic SAR Principles

The working principle of a Synthetic Aperture Radar (SAR) is similar to conventional radar for surveillance: electromagnetic waves are sequentially transmitted and the backscattered echoes are collected by the radar antenna. Since SAR is mounted on a moving platform the consecutive time of transmission and reception translates into different positions due to the platform movement. An appropriate coherent combination of the received signals allows the construction of a virtual antenna aperture that is much longer than the physical antenna length. This gives to the SAR the property of being an imaging radar.

SAR sensors commonly utilize frequency modulated pulsed waveforms for transmission, called chirp signals. The amplitude of the transmitted waveform is constant during the pulse time T, while the instantaneous frequency f_i is varied in a linear manner over time

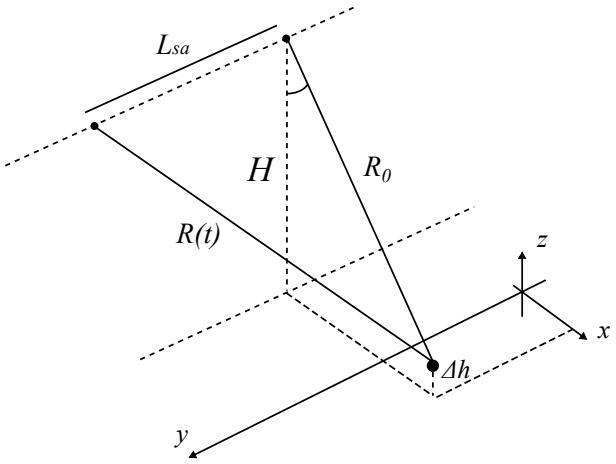


Figure 2.2: Range history of a point scatterer.

t according to

$$f_i = k_r t \tag{2.1}$$

where k_r is known as the chirp rate. The bandwidth B_r of a chirp is thus given by

$$B_r = k_r T \tag{2.2}$$

After sending the chirp signal the SAR listens to the scattered echoes and stores the received signals on-board. The transmission and listen procedure is repeated every PRI seconds, where the pulse repetition interval (PRI) is the reciprocal of the pulse repetition frequency $PRF = 1/PRI$. It is possible to prove that the slant-range resolution δ_r is inversely proportional to the system bandwidth according to

$$\delta_r = \frac{c_0}{2B_r} \tag{2.3}$$

where c_0 is the speed of light. The azimuth resolution δ_a is instead related to the length of the synthetic aperture. The beamwidth Θ_a of an antenna of length d_a can be approximated by

$$\Theta_a = \frac{\lambda}{d_a} \tag{2.4}$$

where λ is the wavelength. From Fig. 2.1 it can be seen that the corresponding synthetic aperture length L_{sa} is given by

$$L_{sa} = \Theta_a r_0 = \frac{\lambda r_0}{d_a} \tag{2.5}$$

where r_0 is the shorter approach distance. A long synthetic aperture is favorable since it results in a narrow virtual beamwidth

$$\Theta_{sa} = \frac{\lambda}{2L_{sa}} \tag{2.6}$$

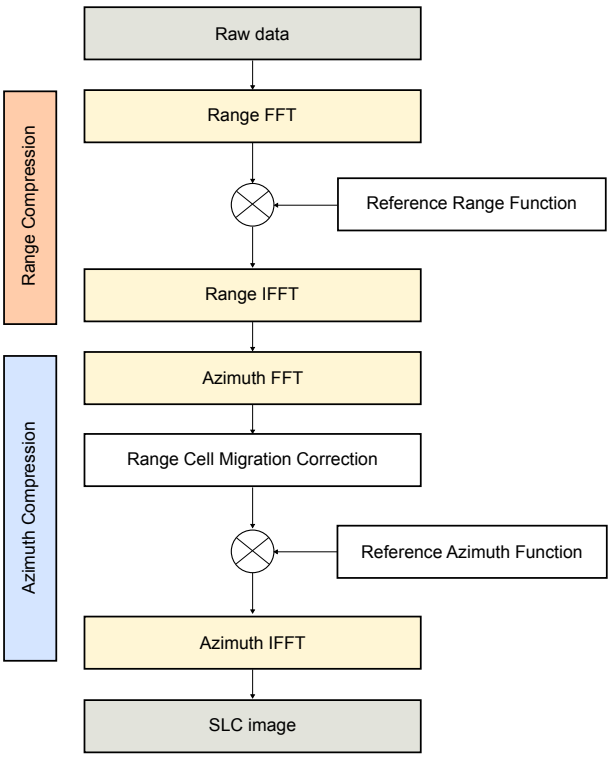


Figure 2.3: Summary of SAR image formation processing steps.

Where the multiplication by 2 appears because of the two-way path from transmission to reception. Moreover with a long synthetic aperture is possible to achieve a high azimuth resolution given by

$$\delta_a = r_0 \Theta_{sa} = r_0 \frac{\lambda}{2L_{sa}} = \frac{d_a}{2} \tag{2.7}$$

From (2.7) it seems that a short antenna yields a fine azimuth resolution. This appears surprising at first. However, it becomes clear considering that radar with a shorter antenna sees any point on the ground for a longer time. This means to have a longer virtual antenna length and thus a higher azimuth resolution.

2.1.2 SAR Image Formation

During the interval of time corresponding to the synthetic aperture, the backscattered energy is received from the target. The radar acquires a range line whenever it travels a distance $v \cdot PRI$, where v is the platform speed. A range line consists of the complex echo signal samples after being amplified, down converted to base band, digitized and stored in memory. By exploiting the moving of the platform the received echos form a two-dimensional data matrix of complex samples, where each element is given by quadrature and in-phase parts of the received echo (or by real and imaginary parts). As mentioned in the previous section and unlike optical sensors, visualizing raw SAR data does not give any useful information on the scene. In order to extract useful information from the raw data an articulated signal processing has to be performed. Fig. 2.3 summarizes the

basic SAR processing steps needed to obtain the image date. By simplifying two separate matched filter operations along the range and azimuth dimensions are needed for such a process. The first step is the compression of the transmitted chirp signals to a short pulse. Instead of performing a convolution in the time domain, a multiplication in the frequency domain is adopted due to the lower computational load. Thus, each range line is multiplied in the frequency domain by the complex conjugate of the spectrum of the transmitted chirp; the result is a range compressed image, which reveals only information about the relative distance between the radar and any point on the ground. In addition, the range matched filter frequency response typically includes an amplitude weighting to control sidelobes in the range impulse response.

At any time t the distance between the radar moving at constant velocity v and a point on the ground, described by its coordinates $(x, y, z) = (x_0, 0, \Delta h)$ is obtained by exploiting the Pythagoras theorem

$$r(t) = \sqrt{r_0^2 + (vt)^2} \tag{2.8}$$

where $t = t_0 = 0$ is the time when the distance is minimum, $r(t_0) = r_0 = \sqrt{(H - \Delta h)^2 + x_0^2}$ with the platform height H. In general the distance r_0 is larger than vt during the illumination time T_{ill} a point on the ground is observed. This allows expanding $r(t)$ into a Taylor series and by keeping the first two terms it is possible to write (2.8) as

$$r(t) = \sqrt{r_0^2 + (vt)^2} \approx r_0 + \frac{(vt)^2}{2r_0} \quad \text{for} \quad vt/r_0 \ll 1 \tag{2.9}$$

Note that the quadratic approximation in (2.9) is done for the sake of simplicity. Accurate SAR data processing takes into account the complete phase history without any approximation. The range variation of a point target over time is directly related to the azimuth phase $\varphi(t)$ by

$$\varphi(t) = \frac{-4\pi r(t)}{\lambda} \tag{2.10}$$

the phase variation has a parabolic behavior (the factor 4π is due to the two way range measurement of the SAR system). The distance between the radar and any fixed point on the ground is changing within the synthetic aperture time. This distance change is obtained from (2.9) by subtracting the constant r_0 and is given by

$$RCM(t) = \sqrt{r_0^2 + (vt)^2} - r_0 \approx \frac{(vt)^2}{2r_0} \quad \text{for} \quad vt/r_0 \ll 1 \tag{2.11}$$

RCM causes an azimuth defocusing because energy is distributed over several range cells. This can be observed as a curvature of the range compressed responses that depends on r_0. The fact that the range migration is range-variant makes SAR focusing a two-dimensional space-variant problem, and hence the data need to be correlated with a non-stationary two-dimensional reference function, making the accurate correction of RCM one of the most challenging aspect of SAR focusing. Detailed analyses and comparisons of these processors, can be found in [16].

After the range migration correction can follow the azimuth compression. The signal is convolved with its reference function, which is the complex conjugate of the response

expected from a point target on the ground. Considering an elemental scatterer at range $r(t)$ given in (2.9), the azimuth signal can be modeled by

$$s_a(t) = A\sqrt{\sigma^0}\exp(i\varphi^{\text{scatt}}) \tag{2.12}$$

where A accounts for the dependency of the received signal on system parameters such as transmit power and losses, and the antenna pattern weighting as a function of the azimuth and elevation angles. The radar cross section is given by σ^0 and φ^{scat} is the scattering phase.

2.1.3 SAR Image Properties

As discussed in the previous section SAR images are made up by two components: amplitude (or intensity) and phase. Amplitude (intensity) is related to the energy of the backscattered signal. Whereas, the phase is related to the sensor-to-target distance and it is the key element in any interferometric measurement. In this thesis we focus the attention on the amplitude (intensity) component of SAR images starting from the analysis of the proprieties of low-medium resolution SAR images and then introducing the VHR SAR data.

Being a coherent system, the amplitude (intensity) component of SAR images are corrupted by the so-called speckle, which causes the granular appearance of SAR imagery. Speckle is caused by the presence of several elemental scatterers within each resolution cell with a random distribution. The coherent sum of their amplitudes and phases results in strong fluctuations of the backscattering from resolution cell to resolution cell. Consequently, the intensity and the phase in the final image are no longer deterministic, but follow instead an exponential and uniform distribution, respectively [17]. The total complex reflectivity for each resolution cell is given by

$$\sum_n \sqrt{\sigma^n}\cdot\exp(i\varphi_n^{\text{scatt}}) \tag{2.13}$$

where i is the number of elementary scatterers within the resolution cell. Although the speckle is commonly referred as a form of noise, speckle is actually a physical measurement of the resolution cell structure at sub-resolution level. It is worth noting that it cannot be reduced by increasing the transmit signal power, since it has a multiplicative character, i.e., its variance increases with its intensity.

The side looking geometry of SAR together with non-flat terrain causes geometric distortions such as foreshortening, layover and shadows. In Fig. 2.4 an example of these phenomena is given considering a mountain area. Foreshortening arises when inclined surfaces are oriented towards the sensor. These surfaces appear shortened in SAR imagery. Layover and shadow effects, visible as relatively bright and dark regions in SAR imagery, arise when the inclination of the surface is larger than the incidence angle. In detail, since in the layover area parts of the ground surface, the slope facing the sensor, and parts of the slope turned away from the sensor are equidistant to the SAR antenna their backscattering return to the sensor at the same time, causing a high value of backscattering. On the other hand, shadows are areas where no backscattering is recorded at the sensor, because they are occluded from the radar beam resulting as dark area in the image.

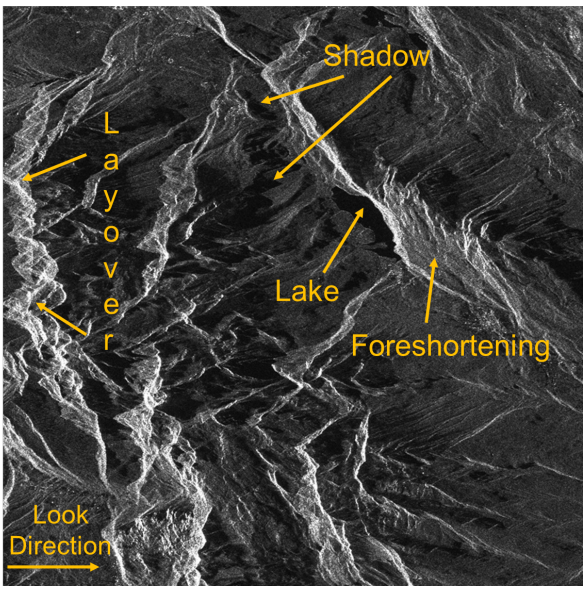


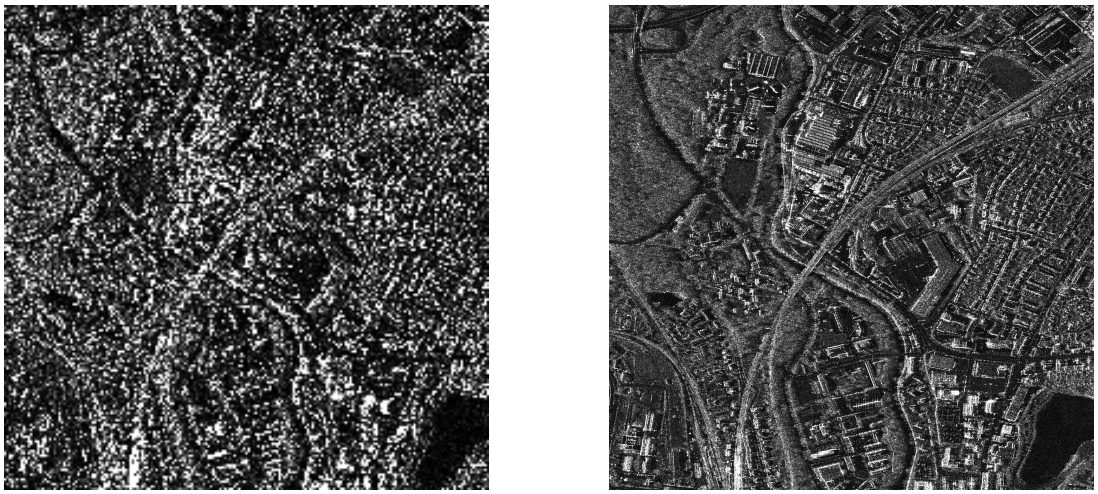
Figure 2.4: Geometrical distortions in Trentino region (Italy) due to the later vision of SAR (copyright ESA).

The scattering of microwaves from a surface is composed by a mix of specular and Lambertian scattering, depending on the surface roughness σ_s with respect to the wavelength. For a perfectly smooth ($\sigma_s \ll \lambda_c$) area the electromagnetic wave is entirely scattered in specular direction. Hence, no backscattering is recorded by the sensor. This is the case of the water body in Fig. 2.4. A perfectly rough surface ($\sigma_s \gg \lambda_c$) instead scatters according to the Lambertian cosine law. For a slightly rough surface ($\sigma_s < \lambda_c$) the scattering is characterized by a large specular component, and a Lambertian component with less power scattered in all directions. The rougher the surface, the weaker are the specular and the stronger are the Lambertian components.

Besides being dependent on the surface roughness, the scattering is also influenced by the relative dielectric constant ϵ_r. ϵ_r is a complex number that depends on the dielectric properties of the material and varies with respect to the material moisture content and the SAR frequency. Materials with low ϵ_r have less reflectivity and hence a higher penetration into the medium.

When VHR resolution images are considered the prospective is changing with respect medium-law resolution SAR data: the images are more heterogeneous and the speckle is not always fully developed since may happen that a small number of elementary scatterers are present inside the resolution cell, the geometrical distortions and the multiple reflections are dominant effects in urban areas, and the very high resolution allows the identification of more geometrical features of the ground objects. In order to give to the reader an idea of the differences between medium and very high geometrical resolution images, Fig. 2.5 shows two images acquired over an area of the city of Lüneburg (Germany). Fig. 2.5a shows the SAR image with a resolution of 10 m, while Fig. 2.5b shows the SAR image acquired at 1 m resolution. Observing the two images it is simple to imagine the potentiality of new applications arisen from the high geometrical content of

VHR data and the need of developing new and proper techniques for effectively manage such data.



(a) (b)

Figure 2.5: Example of TerraSAR-X images acquired over the city of Lüneburg (Germany) at the resolution of: (a) 10 meters (simulated); and (b) 1 meter.

For more detail we refer to the rich selection of publications and excellent textbooks on the theoretic aspects of the acquisition, image formation and image understanding [16,18].

2.1.4 Overview of Spaceborne SAR Missions

Since the launch of the first spaceborne SAR (SEASAT, L-band) by the Jet Propulsion Laboratory in 1978, several SAR missions have been operating and many are planned in the next future [19]. Among those, notables are the ERS-1/2, the C-band twin European Remote Sensing satellites launched by ESA in 1991 and 1995 respectively, that allowed the demonstration of repeat-pass interferometry; J-ERS, the L-band Japanese Earth Remote Sensing Satellite launched in 1992; RADARSAT-1/2, the C-band satellites launched by the Canadian Space Agency in 1995 and 2007 respectively, now with full-polarimetric capabilities; ALOS/PALSAR, the Japanese Phased Array L-band SAR operating since the beginning of 2006; the SIR-C/X-SAR in 1994 and SRTM in 2000, NASA SAR systems carried on-board of the Space Shuttle, which were able to operate in multi-polarization and multi-frequency; ENVISAT/ASAR, launched in 2002 by ESA, which was the first SAR satellite using an antenna able to operate in different imaging modes; the X-band SAR TerraSAR-X and TanDEM-X launched in 2007 and 2010, respectively, by the German Aerospace Center (DLR) and the Italian COSMO-skyMed-1/4 Constellation that comprises four X-band SARs, launched starting from 2007 and nowadays fully operational, which are a new class of SAR satellites able to provide images with resolution of a meter or less; Sentinel-1a, launched in 2014, which is the first of two satellites of the ESA sentinel constellation operating in C band; ALOS-2, launched in 2014 is a L-band full-polarimetric SAR system operated by JAXA. New SAR missions are planned in the future: the L-band full-polarimetric SAOCOM-1/2, which launch is scheduled in 2015 by

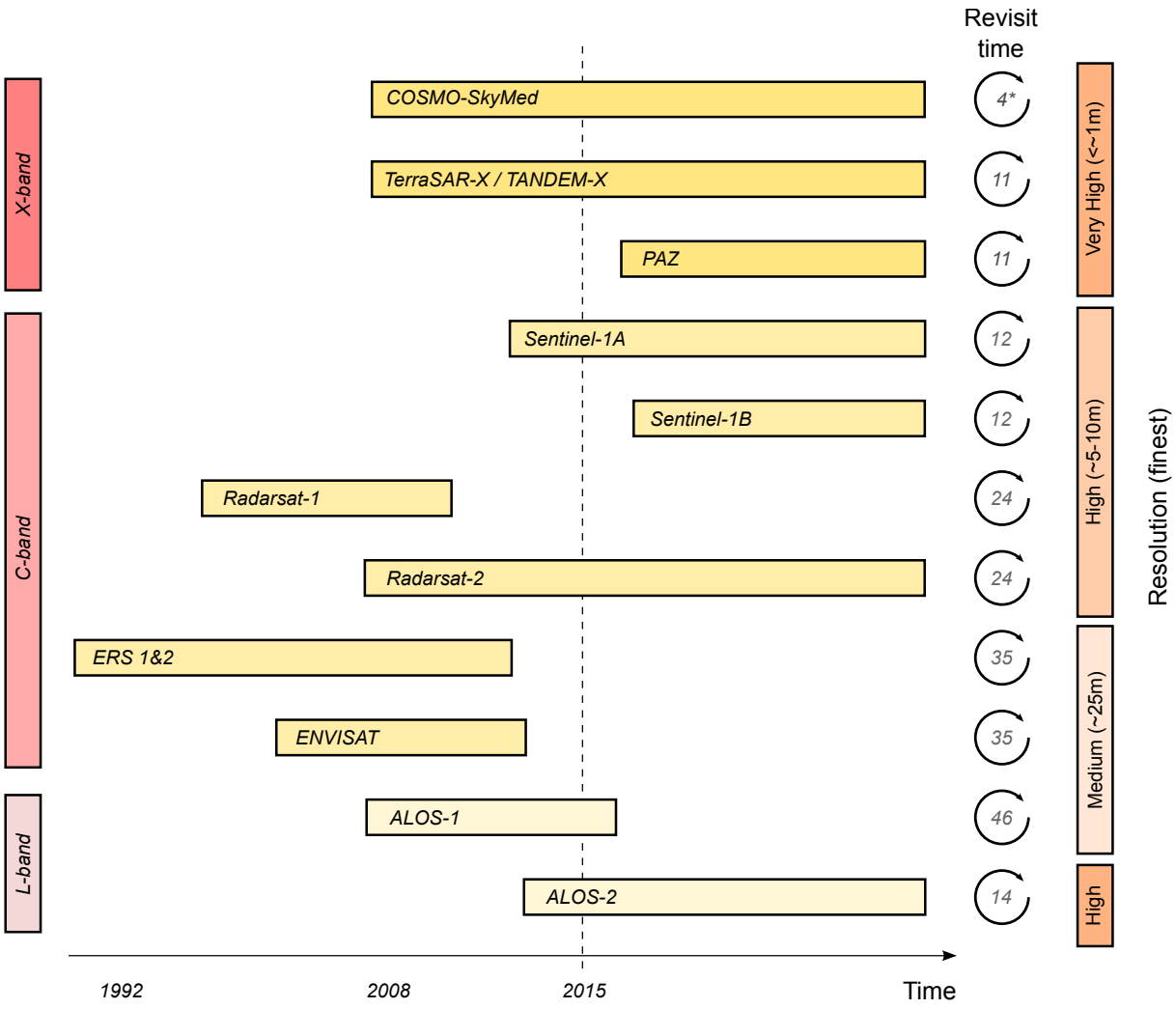


Figure 2.6: Main commercial SAR satellite missions and their characteristics.

CONAE; Radarsat Constellation-1/2/3 scheduled for launch in 2017 by Canadian Space Agency. In Fig. 2.6 we report the main SAR missions and their characteristics. It is indeed expected that more than 10 SAR satellites will be launched before the end of 2020.

2.2 Unsupervised Change Detection with SAR Images

As pointed out in chapter 1, images acquired at different time over the same geographical areas can be used in order to identify changes happened on the Earth surface. This process, called change detection, has become of primary importance for a large number of applications. In this section, by taking into account the proprieties of the SAR images given in the previous section, we give a general overview of the framework commonly used in unsupervised change-detection for low- and medium resolution SAR images.

In the literature, change detection approaches are divided in two main categories:

supervised and unsupervised [20–22]. The former requires a priori information on the investigated area and allows one to determine both the presence/absence and the kind of change [23], whereas the latter does not assume availability of ground truth data and usually identifies only the presence/absence of changes [14,20,21,24,25]. Even if the large part of unsupervised change detection does not produce information about the nature of the class of changes, it is often preferred in real-data applications. This is due to the difficulties in collecting proper ground-truth information (necessary for supervised techniques), which is a complex, time consuming and expensive process (in many cases this process is not consistent with the application constraints).

Unsupervised change-detection techniques are based on the comparison of multitemporal images and a subsequent analysis of the comparison output aim at deriving a change detection map with possible labels "change" or "no-change". We can give a formal definition as follows:

Definition 2.1 *Let us consider two remote sensing images,* $\mathbf{X}_1$ *and* $\mathbf{X}_2$ *acquired over the same area at different times* τ_1 *and* τ_2*. Let us assume that no ground-truth information is available for the design of the change-detection algorithm. We define the unsupervised change detection as the process aimed at generating a change-detection map that represents changes occurred on the ground between the two considered acquisition dates.*

In the literature, the most widely used unsupervised change-detection techniques for low/medium and very high resolution multitemporal SAR images are based on a 3-step procedure [14, 20, 21]. This framework, which is reported in Fig. 2.7 has general validity and can be seen as the starting point for the development of the novel techniques in the context of VHR SAR images. Unsupervised change detection steps are as follows

1. **Pre-processing** The aim of pre-processing is to make the two considered images as more comparable as possible. In general, this step includes: co-registration, noise reduction, geometric and radiometric corrections. From the practical point of view, co-registration and radiometric corrections are essential steps as they allow the derivation of a stable pair of images where corresponding pixels are associated to the same position on the ground. It is worth noting that the use of noise filtering should be carefully evaluated because it may introduce both geometrical and radiometric artifacts in the images. When VHR SAR images are considered, the rationale behind the pre-processing step does not change and the same flow order can be used.

2. **Comparison** The comparison step aims at producing a further image where differences between the two considered acquisitions are highlighted. Different mathematical operators can be adopted to perform image comparison and multiscale strategies can be applied; this choice gives rise to different kinds of techniques [20, 26–31]. Some of them are illustrated in the following since they are exploited in the thesis. It is worth noting that most of these techniques were developed for low- or medium resolution SAR images.

3. **Image analysis** In order to extract the change information, a proper unsupervised image analysis technique should be adopted. Among unsupervised techniques, the most widely used is based on the selection of decision thresholds that separates

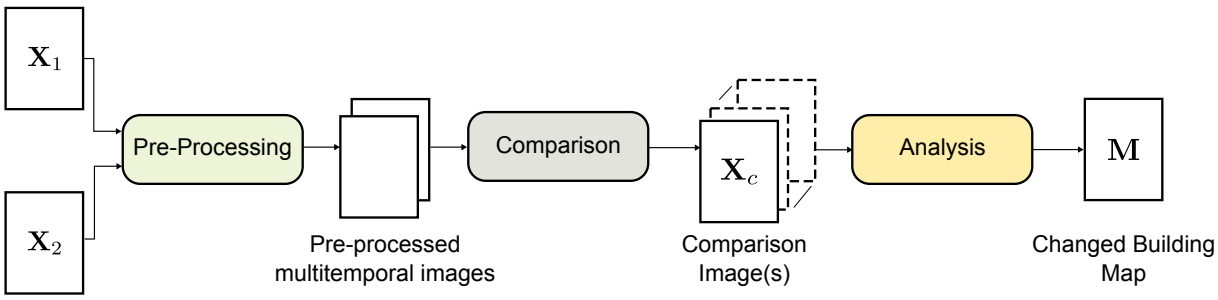


Figure 2.7: Block scheme of a standard unsupervised change-detection approach.

changed from unchanged pixels. The decision thresholds can be selected either with a manual trial-and-error procedure (according to the desired trade-off between false and missed alarms) or with automatic techniques e.g., by analyzing the statistical distribution of the image obtained after comparison, by fixing the desired false alarm probability [8], or following a Bayesian minimum-error decision rule [14, 24, 25].

In the following each of these steps are discussed in detail introducing the state of the art techniques that have been evaluated and exploited in some of the methods presented in this thesis.

2.2.1 Pre-processing

The first step for properly performing change-detection is image pre-processing. This procedure aims at generating two images that are as much similar as possible unless in changed areas. Staring from the focused SAR images (see previous section) the pre-processing consists in four steps:

co-registration It allows the alignment of the temporal images in order to have corresponding pixels in the spatial domain associated to the same geographical position on the ground. This requires spatial registration and potentially resampling (in cases where pixel sizes differ) to correct for relative translational shift, rotational and scale differences. In the literature two kinds of methods to image registration can be found [32]: (i) semi-manual registration methods and (ii) automatic methods.

noise reduction It aims at reducing the noise in the images. In SAR images the presence of the multiplicative speckle noise makes the task more complex. A first step to reduce speckle is usually performed during the multi-looking, where range and/or azimuth resolution cells are averaged. This allows a mitigation of the speckle at the expense of spatial resolution. Beside the multi-looking many techniques have been developed in the literature for reducing the speckle [33, 34]. Among the mono-temporal filtering usually the adaptive despeckling procedures are exploited such as Frost, Lee, Kuan, Gamma Map [35], [36]. Within the multitemporal filtering the Quegan filter is the most used [34, 37, 38]. However, the real need for filtering should be carefully evaluated on the basis of the specific application considered, as filtering degrades the spatial resolution.

geometric calibration Geocoding, Georeferencing, Geometric Calibration, and Orthorectification are synonyms. All these words mean the conversion of SAR images into a map coordinate system (e.g., cartographic reference system) in order to reduce the distortions that are due to the ground topography. This process compensates for the different response to illumination due to the irregular shape of the terrain. In particular, it allows the reduction of specific geometric distortion effects like layover and foreshortening.

radiometric calibration Calibration of the backscatter values is necessary for intercomparison of radar images acquired with different sensors, or data coming from the same sensor but processed with different image formation algorithm. The radiometric calibration of the SAR images is done by considering the radar equation law. The radiometric calibration corrects: i) the scattering area. Each output pixel is normalized for the actual illuminated area of each resolution cell, which may be different due to varying topography; ii) the antenna gain pattern. The effects of the variation of the antenna gain in range are corrected, taking into account topography or a reference height and iii) the range spread loss. The received power for the range distance changes from near to far range. It is worth recalling that in practical cases three calibrated values are used: i) the radar brightness β^0, which is the reflectivity per unit area in slant range; ii) the backscattering coefficient σ^0, which is a normalized dimensionless number (usually in dB) that expresses the power per area of one square meter. σ^0 is defined with respect to the nominally horizontal plane, and in general has a significant variation with incidence angle, wavelength, and polarization, as well as with properties of the scattering surface itself; and iii) the γ, which is the backscattering coefficient normalized by the cosine of the incidence angle. This is done to correct the dependency of the backscatter from the incidence angle.

2.2.2 Comparison

The second step is the image comparison. This step aims at generating an image that highlights the changes between the two (or more) pre-processed multitemporal SAR images. The comparison is done by exploiting pixel- or context-based change detection operators as discussed in the following.

Pixel-based comparison

Pixel-by-pixel comparison can be performed by means of different mathematical operators. In SAR image analysis the ratio operator demonstrated to be more effective than the difference one, due to the multiplicative noise model commonly adopted for this kind of data [14, 26, 36]. In the following, we describe the reasons motivating the preference of image rationing over image differencing in SAR imagery basing on [26]. Let us define $\mathbf{X}_c$ the image resulting from the comparison of the two SAR images $\mathbf{X}_1$ and $\mathbf{X}_2$ acquired at time τ_i with $i = 1, 2$, respectively. The indexes $c = D$ and $c = R$ refer to the difference and ratio images, respectively. Each pixel of the difference image is obtained in the following way

$$\mathbf{X}_D = \mathbf{X}_2 - \mathbf{X}_1 \tag{2.14}$$

while pixels of the ratio image are given by:

$$\mathbf{X}_R = \frac{\mathbf{X}_2}{\mathbf{X}_1} \tag{2.15}$$

Let us assume that pixels of the two intensity images $\mathbf{X}_1$ and $\mathbf{X}_2$ follow a Gamma distribution, i.e.,

$$p(\mathbf{X}_i) = \frac{L^L \mathbf{X}_i^{L-1}}{\mu_i^L (L-1)!} \exp\left(-\frac{L\mathbf{X}_i}{\mu_i}\right), \quad i = 1, 2. \tag{2.16}$$

where μ_i is the mean intensity of a homogeneous region at time τ_i, and L is the Equivalent Number of Looks (ENL). Another assumption is that the intensity images $\mathbf{X}_1$ and $\mathbf{X}_2$ are supposed to be statistically independent. The purpose of this assumption (though not entirely realistic) is to simplify the analytical derivation of the statistical models for the difference and ratio images. Under these conditions, the distribution of the difference image $\mathbf{X}_\mathrm{D}$ is given by:

$$p(\mathbf{X}_\mathrm{D}) = \frac{L^L \exp\left(-L\frac{\mathbf{X}_\mathrm{D}}{\mu_2}\right)}{(L-1)!(\mu_1+\mu_2)^L} \sum_{j=0}^{L-1} \frac{j!(L-1+j)!}{(L-1-j)!} \mathbf{X}_\mathrm{D}^{L-1-j} \left[\frac{\mu_1\mu_2}{L(\mu_1+\mu_2)}\right]^j \tag{2.17}$$

In [26] the authors showed that the variance of difference image distribution increases with the intensity level. In terms of the probability of error in detecting changes, this means that the difference operator will produce more errors in high intensity regions of the image than in low intensity regions. This represents an undesired effect that makes the difference operator not suitable to the statistics of SAR data.

Similarly, it is possible to prove that the distribution of the ratio image $\mathbf{X}_\mathrm{R}$ can be written as follows:

$$p(\mathbf{X}_\mathrm{R}) = \frac{(2L-1)!(\overline{\mathbf{X}}_\mathrm{R}^L \mathbf{X}_\mathrm{R}^{L-1})}{(L-1)!^2(\overline{\mathbf{X}}_\mathrm{R} + \mathbf{X}_\mathrm{R})^{2L}} \tag{2.18}$$

where $\overline{\mathbf{X}}_\mathrm{R} = \mu_2/\mu_1$. The ratio operator shows two main advantages over the difference operator. The first one is that the ratio image distribution depends only on the relative change $\overline{\mathbf{X}}_\mathrm{R}$ in average intensity between two dates and not on the intensity level of the pixels, in contrast with the distribution of the difference image. This means that changes will be detected in the same manner both in high and low intensity regions. The second advantage is that the rationing allows the reduction of common multiplicative error components, as far as these components are the same for images acquired with the same geometry. It is worth noting that, in the literature, the ratio image is usually expressed in a logarithmic scale. With this operation the distribution of the two classes of interest in the ratio image can be made more symmetrical and the residual multiplicative speckle noise can be transformed in an additive noise component. Thus the log-ratio operator is typically preferred when dealing with SAR images and change-detection is performed analyzing the log-ratio image $\mathbf{X}_\mathrm{LR}$ defined as:

$$\mathbf{X}_\mathrm{LR} = \log \mathbf{X}_\mathrm{R} = \log \frac{\mathbf{X}_2}{\mathbf{X}_1} = \log \mathbf{X}_2 - \log \mathbf{X}_1 \tag{2.19}$$

It is worth noting that for keeping the changed class on one side of the histogram of the ratio (or log-ratio) image, a normalized ratio can be computed pixel-by-pixel as follows:

$$\mathbf{X}_{\mathrm{NR}} = \min\left\{\frac{\mathbf{X}_1}{\mathbf{X}_2}, \frac{\mathbf{X}_2}{\mathbf{X}_1}\right\} \tag{2.20}$$

This operator allows all changed areas (independently of the increasing or decreasing value of the backscattering coefficient) to play a similar role in the change-detection problem.

Context-based comparison

The context insensitive analysis based only on the values of single pixel may still suffer from the presence of speckle. In order to address this problem in the literature some works have been presented that take into account the context information. This is usually done by considering the local neighborhood information of each pixel. In [15] the Mean Ratio Detector (MRD) is implemented as the normalized local means in the neighborhood of each pixel as follows

$$\mathbf{X}_{\mathrm{MRD}}(i,j) = 1 - \min\left\{\frac{\mu_1^W}{\mu_2^W}, \frac{\mu_2^W}{\mu_1^W}\right\} \tag{2.21}$$

where μ_1^W and μ_2^W are the mean values compute on a moving window over the images at time τ_1 and τ_2, respectively. Even though the MRD is more robust against the speckle noise than pixel-based techniques there is no a systematic procedure in order to select the optimal windows size for a certain data-set. Indeed, from one hand by choosing a small size for the moving window one gets a good sensitivity to geometrical details but a low mitigation of the speckle noise. On the other hand by choosing a large size for the moving window one gets a low sensitivity to geometrical details but a high mitigation of speckle. Generally the size that gives the best trade-off between details preservation and smoothing effect is selected after a trial-and-error procedure.

Because of the limitations in properly model the spatial-context information in possibly complex changes, change-detection maps obtained with this technique generally does not achieve a high overall accuracy. In order to overcome this limitation, in [39] an advanced context-sensitive change-detection methods has been proposed. This is done by modeling the spatial-context information of pixels by considering multitemporal homogeneous regions called parcels. In detail, let $\mathcal{P}(\mathbf{X}_1, \mathbf{X}_2) = \{\rho_1, \rho_2, \cdots, \rho_N\}$ be the set of parcels associated with the images $\mathbf{X}_1$ and $\mathbf{X}_2$. Let $\mathcal{H}(\mathbf{Z})$ the homogeneity predicate and $\mathbf{Z}$ the set of spatial adjacent pixels involved in the evaluation of the homogeneity. The homogeneity predicate $\mathcal{H}(\cdot)$ is defined considering different spatial, temporal, and backscattering attributes. Let finally denote with $\mathbf{X}_i(\rho_s)$ the portion of image $\mathbf{X}_i$ $(i = 1, 2)$ belonging to

the parcel ρ_s. According, $\mathcal{P}(\mathbf{X}_1, \mathbf{X}_2)$ should satisfy the following conditions:

$$\text{Pixels in } \rho_s \text{ are connected;} \tag{2.22}$$

$$\bigcup_{s=1}^{N} \mathbf{X}_1(\rho_s) = \mathbf{X}_1 \text{ and } \bigcup_{s=1}^{N} \mathbf{X}_2(\rho_s) = \mathbf{X}_2; \tag{2.23}$$

$$\rho_s \cap \rho_r \ = \ \varnothing; \tag{2.24}$$

$$\mathcal{H}\left\{\mathbf{X}_1(\rho_s)\right\} = true \ \wedge \ \mathcal{H}\left\{\mathbf{X}_2(\rho_s)\right\} = true \tag{2.25}$$

$$\mathcal{H}\left\{\mathbf{X}_1(\rho_s) \cup \mathbf{X}_1(\rho_r)\right\} = false \ \vee \ \mathcal{H}\left\{\mathbf{X}_2(\rho_s) \cup \mathbf{X}_2(\rho_r)\right\} = false, \quad \text{with } s \neq r \tag{2.26}$$

Conditions (2.22)-(2.24) guarantee that all pixels in $\mathbf{X}_1$ and $\mathbf{X}_2$ are distributed into N connected and non-overlapping regions. Condition (2.25) determines the homogeneity properties of the parcels in both original images (i.e. the geometrical homogeneity in the temporal domain) and condition (2.26) expresses the maximality of each parcel (i.e., pairs of adjacent parcels cannot be merged without relaxing the homogeneity criterion).

In conclusion, the adaptive nature of parcels allowed a notable reduction in false alarms and an accurate location of the border of the changed area. On the other hand, the main difficulty is represented by the definition of the homogeneity (or heterogeneity) of the neighborhoods around pixels for the definition of the parcels. Nonetheless, this concept will be used in chapter 5 for extracting multitemporal primitives.

Statical similarity measure for change-detection

The ratio, log-ratio and MRD assume that a change in the scene will appear as a modification of the local mean value of the image. If the change preserves the mean value but modifies the local texture, it will not be detected. Moreover, the speckle is different from one image to the other (even though they are acquired at the very same area), and it can induce a high number of false alarms in the change-detection procedure. Recently in the literature promising statistical similarity measures such as Kullback-Leibler Divergence (KLD) and Mutual Information (MI) have been successfully applied to highlight change information in multitemporal images. In the following we introduce these information similarity measurements in detail. Let $p_{X_1}(x_1)$ and $p_{X_2}(x_2)$ be the two marginal Probability Density Functions (PDF), and let $p_{X_1,X_2}(x_1, x_2)$ be the joint probability density function of two random variables X_1 and X_2 related to the images acquired at time τ_1 and τ_2. The mutual information [40] is defined as

$$I(X_1, X_2) = \int_{X_2} \int_{X_1} p_{X_1,X_2}(x_1, x_2) \log \frac{p_{X_1,X_2}(x_1, x_2)}{p_{X_1}(x_1) p_{X_2}(x_2)} \, \mathrm{d}x_1 \mathrm{d}x_2 \tag{2.27}$$

mutual information determines how similar the joint distribution $p(X_1, X_2)$ is to the products of the factored marginal distribution $p(X_1)p(X_2)$. Mutual information is always positive and it is equal to zero if and only if X_1 and X_2 are independent.

The Kullback-Leibler divergence [41] from X_2 to X_1 is defined as

$$K(X_2|X_1) = \int \log \frac{p_{X_1}(x_1)}{p_{X_2}(x_1)} p_{X_1}(x_1) \, \mathrm{d}x_1 \tag{2.28}$$

The measure $\log \frac{f_{X_1}(x_1)}{f_{X_2}(x_1)}$ can be thought of as the information on x_1 for the discrimination between the hypothesis $\mathcal{H}_{X_1}$ and $\mathcal{H}_{X_2}$, if hypothesis $\mathcal{H}_{X_1}$ is associated with the PDF $f_{X_1}(x_1)$ and $\mathcal{H}_{X_2}$ with $f_{X_2}(x_1)$. Therefore, the KL divergence $K(X_2|X_1)$ can be understood as the mean information for the discrimination between $\mathcal{H}_{X_1}$ and $\mathcal{H}_{X_2}$ per observation. This divergence appears to be an appropriate tool to detect changes when we consider that changes on the ground induce different shapes on the local PDF. Since the KL divergence can be understood as the entropy of P_{X_1} relative to P_{X_2} , it is also called information gain. It can easily be proven that $K(X_2|X_1) \geq 0$; $K(X_2|X_1)$ vanishes only when the two laws are identical. $K(X_2|X_1)$ can be used as a measure of the divergence from P_{X_1} to P_{X_2}. This measure is not symmetric as it stands: $K(X_2|X_1) \neq K(X_1|X_2)$, but a symmetric version may be defined by writing

$$D_{\mathrm{KL}}(X_1, X_2) = D_{\mathrm{KL}}(X_2, X_1) = K(X_2|X_1) + K(X_1|X_2) \tag{2.29}$$

that will be called the KL distance.

In order to estimate the KL distance or the MI, the PDFs of the two variables to be compared have to be known. The estimation problem can be addressed in several ways [42]. One can use parametric approaches such as the maximum likelihood or the maximum a posteriori estimators implemented with refined algorithm to reach the convergence of the estimation e.g. the expectation-maximization algorithm. However, parametric approaches are applicable only if the underlying distributions are known and sufficiently separated. In case the distributions are not known, one can try to solve the estimation problem by using nonparametric approaches such as the kernel density estimation. The nonparametric methods need the definition of a bandwidth or tuning parameter, which controls both the degree of complexity and the goodness of the results. The choice of bandwidth is thus critical. In order to avoid the errors of both parametric and nonparametric approaches, an interesting method that takes into account some descriptors of the shape of the distributions is presented in [15]. In detail, the authors proved that an effective way to model in a nonparametric way the shape of a statistical distribution is to use the infinite Edgeworth series expansion of cumulants truncated at a given order. The cumulants themselves do not provide such a PDF estimation directly but are necessary to describe its shape: e.g., third-order (κ_3) is linked to the symmetry (i.e., skewness), while the fourth-order (κ_4) is linked to the flatness (i.e., kurtosis). The density is then estimated through the Edgeworth series expansion by assuming that the density to be approximated is not too far from a Gaussian PDF. It is worth noting that, according to the results provided by the authors, the methods presents a more robust behavior with respect other parametric indicators (i.e., Gaussian or Pearson based), even though the hypothesis that the PDF to be approximated is Gaussian does not fit perfectly the statistical model of SAR images. It yields an approximation of the KL divergence by Edgeworth series, truncated at a given order. In [15], such an approximation was truncated to order of four and the symmetric version of the cumulant-based KL divergence between two observations X_1 and X_2 has been derived as

$$\mathrm{CKLD}(X_1, X_2) = KL_{\mathrm{Ew}}(X_1, X_2) + KL_{\mathrm{Ew}}(X_2, X_1) \tag{2.30}$$

The results obtained with this method will be used in Chapter 5.

Multiscale change analysis

A scale-dependent representation can be obtained by applying different methods to the data e.g., Laplacian/Gaussian pyramid decomposition, wavelet transform, recursively up-sampled bi-cubic filter, etc. In the SAR literature for instance, image multiresolution representation has been applied extensively to image denoising [43], [44]. In [39] the authors proposed a multiscale adaptive approach to change-detection in multitemporal SAR images. This is based: (i) on a multiscale decomposition of the log-ratio image; (ii) on a selection of the reliable scales for each pixel (i.e., the scales at which the considered pixel can be represented without border problems) according to an adaptive analysis of its local statistics; and (iii) on a scale-driven combination of the selected scales. In this way the method exploits only high-resolution levels in the analysis of the expected edge (or detail) pixels and uses low-resolution levels in the processing of pixels in homogeneous areas. This allows the method to exhibits both a high sensitivity to geometrical details and a high robustness to noisy speckle components in homogeneous areas.

Generally speaking, the nice propriety of multiscale strategies is that there should be an optimal representation level for each imaged object. For example, in the multiscale analysis at finer scale we can identify cars. At coarser scales we can identify groups of cars or larger size objects such as buildings. At the coarse scale we can identify city blocks. In order to properly model objects at different scales it is necessary to take into account: i) the logical connection of the objects at the same level; and ii) the hierarchical connection of the object represented at a generic level with the objects at the finer and coarser scales [45]. This concept can be exploited in VHR SAR multitemporal analysis where changes may present different optimal scales.

In the following we give the theoretical background of the Wavelet decomposition method introduced in [39] and used in this thesis work. From $\mathbf{X}_{\mathrm{LR}}$ a set of multi-level images $\mathcal{X}_{\mathrm{MS}} = \left\{\mathbf{X}_{\mathrm{LR}}^{0}, \ldots, \mathbf{X}_{\mathrm{LR}}^{n}, \ldots, \mathbf{X}_{\mathrm{LR}}^{N-1}\right\}$ is computed, where the superscript n, $n = 0, \ldots, N$ indicates the resolution level of images. The output at resolution level 0 corresponds to the original image, i.e., $\mathbf{X}_{\mathrm{LR}}^{0} \equiv \mathbf{X}_{\mathrm{LR}}$. For n ranging from 0 to $N-1$, the images are characterized by different trade-offs between spatial-detail preservation and speckle reduction. The decomposition is based on the two-dimensional discrete stationary wavelet transform (2D-SWT). As the log operation transforms the multiplicative residual speckle noise of the ratio image into an additive noise, SWT can be applied to $\mathbf{X}_{\mathrm{LR}}$ without any additional processing. Moreover, differently from the DWT, SWT avoids down-sampling the filtered signals after each convolution step. 2D-SWT applies to the considered signal at each resolution level n appropriate level-dependent high- and low-pass filters with impulse response $h^{n}(\cdot)$ and $l^{n}(\cdot)$, $n = 0, 1, \ldots, N-1$, respectively. Filter impulse response depends on the selected Wavelet family and on the desired length of filters. $h^{n}(\cdot)$ and $l^{n}(\cdot)$ are applied first along rows and then along columns in order to produce N different images at the next scale. Thus, the image $\mathbf{X}_{\mathrm{LR}}$ is decomposed into N images of the same size as the original one. In detail, the decomposition phase produces: i) the so called approximation subband $\mathbf{X}_{\mathrm{LR}}^{\mathrm{LL}_n}$, which is a lower resolution version of image $\mathbf{X}_{\mathrm{LR}}$, and contains low spatial frequencies both in the horizontal and the vertical directions at resolution level n; and ii) the so called detail subbands, which are the three high-frequency images $\mathbf{X}_{\mathrm{LR}}^{\mathrm{LH}_n}$, $\mathbf{X}_{\mathrm{LR}}^{\mathrm{HL}_n}$, and $\mathbf{X}_{\mathrm{LR}}^{\mathrm{HH}_n}$ corresponding to the three im-

ages in horizontal, vertical, and diagonal directions at resolution level n, respectively. Note that, superscripts LL, LH, HL, and HH specify the order in which $h^n(\cdot)$ and $l^n(\cdot)$, $n = 0, 1, \ldots, N-1$, have been applied to obtain the considered sub-band.

Multiresolution decomposition is obtained by recursively applying the described procedure to the approximation subband $\mathbf{X}_{\mathrm{LR}}^{\mathrm{LL}_n}$ at each scale. Thus, the outputs at a generic resolution level can be expressed analytically as follows:

$$\begin{aligned}
\mathbf{X}_{\mathrm{LR}}^{\mathrm{LL}_{n+1}}(i,j) &= \sum_{p=0}^{Q^n-1}\sum_{q=0}^{Q^n-1} l^n[p]l^n[q]\mathbf{X}_{\mathrm{LR}}^{\mathrm{LL}_n}(i+p, j+q) \\
\mathbf{X}_{\mathrm{LR}}^{\mathrm{LH}_{n+1}}(i,j) &= \sum_{p=0}^{Q^n-1}\sum_{q=0}^{Q^n-1} l^n[p]h^n[q]\mathbf{X}_{\mathrm{LR}}^{\mathrm{LL}_n}(i+p, j+q) \\
\mathbf{X}_{\mathrm{LR}}^{\mathrm{HL}_{n+1}}(i,j) &= \sum_{p=0}^{Q^n-1}\sum_{q=0}^{Q^n-1} h^n[p]l^n[q]\mathbf{X}_{\mathrm{LR}}^{\mathrm{LL}_n}(i+p, j+q) \\
\mathbf{X}_{\mathrm{LR}}^{\mathrm{HH}_{n+1}}(i,j) &= \sum_{p=0}^{Q^n-1}\sum_{q=0}^{Q^n-1} h^n[p]h^n[q]\mathbf{X}_{\mathrm{LR}}^{\mathrm{LL}_n}(i+p, j+q)
\end{aligned} \tag{2.31}$$

where Q^n is the length of the wavelet filters at resolution levels n and (i, j) are the spatial coordinates of the pixels in the image. Filter coefficients of the first decomposition step for $n = 0$ depends on the selected wavelet family and on the length of the chosen wavelet filter. According to an analysis of the literature, selected the Daubechies wavelet family and set the filter length to 8. Dubechies of order 4 low-pass filter prototype impulse response is given by the following coefficient set:

$$\begin{array}{llll}
\{0.230378, & 0.714847, & 0.630881, & -0,0279838, \\
-0.187035, & 0.0308414, & 0.0328830, & -0.0105974\}
\end{array} \tag{2.32}$$

The finite impulse response of the high-pass filter for the decomposition step is obtained by satisfying the properties of the quadrature mirror filters. This is done by reversing the order of the low-pass decomposition filter coefficient and by changing the sign of the even indexed coefficients. Finally, in order to obtain the image set $\mathcal{X}_{\mathrm{MS}}$ (where each image contains information at a different resolution level) for each approximation subband $\mathbf{X}_{\mathrm{LR}}^{\mathrm{LL}_{n+1}}$ the inverse stationary wavelet transform (2D-ISWT) is applied $n+1$ times.

Multiscale decomposition is at the base of the proposed analysis and methods presented in Chapters 3 and 5.

2.2.3 Image analysis

The most widely used approach to extract change information from the image $\mathbf{X}_c$ obtained after comparison is based on histogram thresholding. This procedure allows the discrimination between the changed and unchanged regions obtaining the so called change detection map. In detail from the log-ratio image, which is widely used in the presented

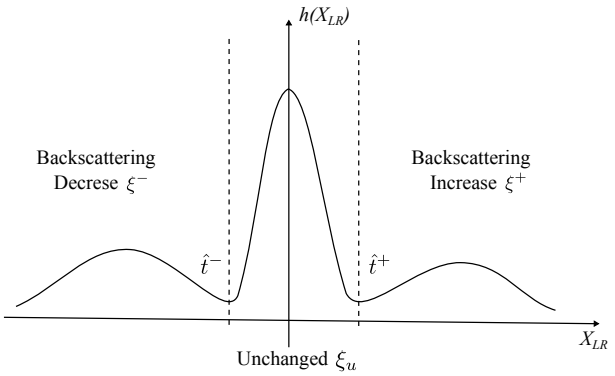


Figure 2.8: Log-ratio image thresholding.

work, a change detection map made up of three classes is obtained as follow

$$\mathbf{M} \in \begin{cases} \xi^-, & \text{if} \quad \mathbf{X}_c < \hat{t}^- \\ \xi_u, & \text{if} \quad \hat{t}^- \leq \mathbf{X}_c \leq \hat{t}^+ \\ \xi^+, & \text{if} \quad \mathbf{X}_c > \hat{t}^+ \end{cases} \tag{2.33}$$

where ξ_u represent the class of unchanged pixels, ξ^+ the class of pixels that increase their backscattering value and ξ^- the class of pixels that decrease their value (see Fig. 2.8).

The derivation of appropriated thresholding values (e.g., $\hat{t}^-$ and $\hat{t}^+$) may be done using empirical strategies [46] or manual trial-and-error procedures. In detail, according to the assumption that changed pixels are few and show values significantly different from the unchanged ones, a reasonable solution is to label as changed those pixels that are far from the mean of the density function associated to the image to be thresholded. A more sophisticated technique is to select the value of the thresholds according to a desired probability of correct detection P_d (which is the probability to be over the threshold if a change occurred) or false alarm P_{fa} (which is the probability to be over the threshold if no change occurred). It has been shown that the thresholding values can be analytical defined as a function of the true change in the radar backscattering σ^0 and of the equivalent number of looks L [26], once P_d and P_{fa} are fixed. The dependence on subjective and empirical criteria represents a critical limitation of the aforementioned approaches to image thresholding. An interesting alternative consists in formulating the change-detection problem in the framework of the Bayesian decision theory in order to optimize the separation between changed and unchanged pixels in an unsupervised way. The main problem to be solved for the application of the Bayes decision theory consists in the explicit estimation of the statistical terms associated to the classes of change and no-change i.e., the prior probabilities $P(\xi^-)$, $P(\xi^+)$ and $P(\xi_u)$, and the probability density functions $p(x_{c,j}|\xi^-)$, $p(x_{c,j}|\xi^+)$ and $p(x_{c,j}|\xi_u)$ without any ground-truth information (i.e. without any training set). The starting point of methodologies based on the Bayesian decision theory is the hypothesis that the statistical distributions of pixels in the analyzed image can be modeled as a mixture of two densities associated with the classes of changed and unchanged pixels, i.e.,

$$p(x_{c,j}) = P(\xi^-)p(x_{c,j}|\xi^-) + P(\xi_u)p(x_{c,j}|\xi_u) + P(\xi^+)p(x_{c,j}|\xi^+) \tag{2.34}$$

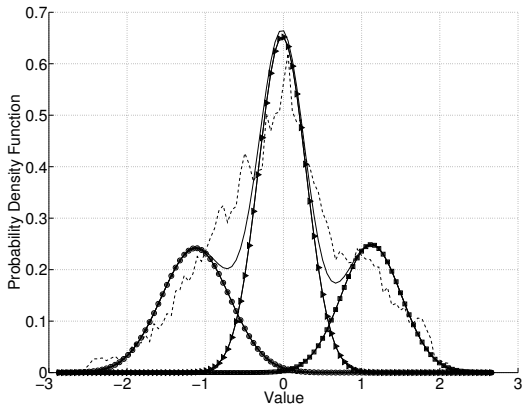


Figure 2.9: Probability density function estimated from the pixels of the selected splits $\mathbf{X}_s$ using a mixture of three Gaussians (solid line) compared with rescaled histogram of the pixels of the selected splits $\mathbf{X}_s$ (dashed line). Estimated distributions of the classes of no-change, decrease in backscattering and increase in backscattering are reported with line marked with triangles, circles and squares, respectively.

In the literature, explicit estimation of class statistical parameters has been addressed with the Expectation-Maximization (EM) algorithm which is an iterative approach to maximum-likelihood (ML) estimation for incomplete data problems [47]. The iterative equations that characterize the EM algorithm are different according to the statistical model adopted for the distributions of the change and no-change classes. The more suitable statistical model varies according to the kind of data to be analyzed. Examples of distribution successfully used for log-ratio image are the Generalized Gaussian distribution and the Gaussian distribution [48]. Once the statistical parameters are computed, pixel-based or context-based decision rules can be applied. In the former group, we find: the Bayes rule for minimum error, the Bayes rule for minimum cost, the Neyman-Pearson criterion, etc. [42]. In the latter group, we find the contextual Bayes rule for minimum error formulated in the Markov Random Field (MRF) framework [49, 50].

Statistical parameters estimation under hypothesis in (2.34) may be also performed adopting an implicit approach [14, 24]. As for the EM algorithm, the mathematical formulation changes according to the statistical model adopted for the change and no-change class distribution. As an example, we mention here the well-known Kittler and Illingworth (KI) thresholding technique which can be used under both the Gaussian and the Generalized Gaussian assumption for the statistical distributions of the change and no-change classes. Despite its simplicity (the change-detection map is computed in a one-step procedure), the KI technique produces satisfactory change-detection results.

In the analysis of large multitemporal images, which is the case of VHR SAR images as discussed in chapter 1, it is possible that the population of changed pixels is in sharp minority in the full $\mathbf{X}_c$. This may affect the accuracy of the threshold selection technique. The authors in [25] presented a method that addresses this issues. In detail, the method divides the considered image into sub-images of a given size (splits) and performs the threshold selection considering only the splits that present the highest probabilities to contain changed pixels. This selection allows the definition of a subset of pixels in which

the class of change shows a higher prior probability than in the whole image. $\mathbf{X}_c$ is split into a set of S sub-images of a user-defined size. The choice of the split size $S_R \times S_A$ is driven by the average extension of the expected changes. Once the sizes S_R and S_A are defined, the splits are identified and selected according to their probability to contain a significant amount of changed pixels. The selection is done according to the value of the variance σ_s^2, $s = 1, \ldots, S$, computed on the pixels of each split. This is a reasonable index for predicting the presence of changes in the log-ratio image since the residual multiplicative noise of the ratio image becomes additive due to the log operator. The desired set $\mathcal{P}_S$ of splits with the highest probabilities to contain changes is defined by selecting the splits that satisfy the following inequality:

$$\sigma_s^2 \geq \bar{\sigma}^2 + B\sigma_{\sigma^2}, \quad s = 1, \ldots, S \tag{2.35}$$

where $\bar{\sigma}^2$ denotes the average variance of splits, σ_{σ^2} is the standard deviation of the variance of splits, and $B > 0$ is a constant. High values of B results in the selection of only those splits with high variance (i.e., the ones with the highest probability to contain changes); vice-versa, low values of B results in the selection of more splits. The split-based selection allows defining a subset of pixels $\mathbf{X}_\mathrm{s} = \{\mathbf{X}_c | \mathbf{X}_c \in \mathcal{P}_S\}$ in which the classes of change show a higher prior probability than in the whole image. Consequently, the statistical estimation of the parameters related to the three probability density functions associated with no-change, increase, and decrease of backscattering between the two dates (i.e., ξ_u, ξ^+, ξ^-), can be correctly derived and used to separate the three classes using the methods described before.

Fig. 2.9 depicts an example of the gaussian mixture probability density function (pdf) estimated from $\mathbf{X}_\mathrm{s}$, which are extracted from real SAR data. As one can notice the three modes associated to three classes ξ_u, ξ^+, ξ^- are discernible by means of two thresholds $\hat{t}^+$ and $\hat{t}^-$. Here the thresholding method described in [48] is adopted and the Bayes decision rule for minimum error is applied to separate the 3 classes. To this end a statistical model for class distributions is required together with an approach for class statistical parameters estimation. The Gaussian model and the well-known Expectation-Maximization (EM) algorithm [47] are employed to derive $\mathbf{M}^{\mathrm{opt}}$.

Split-based analysis is at the base of the proposed analysis and methods presented in Chapters 3 and 5.

2.3 Unsupervised change-detection techniques for multitemporal VHR SAR images

Even though the techniques developed for low- and medium-resolution data are not always effective on VHR images, the framework presented in the previous section for unsupervised change detection remains valid and has been used as starting point to develop novel methodologies able to properly deal with VHR SAR data. Indeed during this research work, in the literature some methods have been published [38, 51–54] that tackle the problem of unsupervised change detection in VHR SAR images. In the following a brief description of these works is reported.

In [51] the authors presents a method called MIMOSA (Method for generalIzed Means Ordered Series Analysis) that exploit a new comparison mechanism based on the analysis of the scatterplot of the two images to investigate. In detail, a statistical model is used to predict the area of the scatterplot corresponding to changes, which is then isolated and used to derive the final change detection map. This representation allows the detection of changes with a low false alarm rate and with a good preservation of the geometrical details. The method, which has been applied in the context of monitoring/surveillance problems, points out all the changes despite some may be not of interest for the final user. The method has been integrated in a framework for the analysis of long time series which requires a low computational low rendering the operational use of the system possible.

In [52] the scatterplot of two SAR images is considered as base for the comparison step. In detail, the mean-shift algorithm is applied to the scatterplot of the amplitude levels in the two images to find the modes of the bivariate distribution. If some of the founded modes lay outside the main diagonal of the scatterplot they are associated to changes. This representation allows a better discrimination capability with respect the classical comparison method (i.e., log-ratio and KL divergence) even though a certain amount of changes has to be present in a scene in order to be effective.

The authors in [53] proposed a comparison method based on Markov random field hyperparameters. In detail, the hyperparameter maps are estimated starting from the real and imaginary part of the multitemporal SAR images. Their similarity is then measured using the two-sample Kolmogorov–Smirnov (KS) statistical test. The maximum distance calculated by the two-sample KS test provides the change map. The method is able to effectively discriminating the changes producing a low false alarm rate. This at the expense of the geometrical details, which are not always well preserved in the final change detection map.

The authors in [54] exploit the Curvelet transformation in order to preserve structural details while filtering the noise. In detail, the image comparison is performed in the Curvelet coefficient domain of the two multitemporal images. A soft thresholding of the compared Curvelet coefficients allows the derivation of the change detection map in the image domain. The approach, which takes advantage of the Curvelet transformation, has demonstrated its effectiveness in detecting structural changes related to man-made environments.

In [38] a two-step multitemporal nonlocal means denoising in presented. The method is able to detect the changes by means of a similarity measurement performed among patches. In detail, the similarity between patches is calculated considering both the generalized likelihood ratio and the KL divergence. This allows in real scenarios to obtain a reduction of the false alarm rate.

Even though these techniques have been successfully applied to VHR multitemporal images they do not solve all the problems presented in chapter 1. Filling this gap on the state of the art is one of the main objectives of the presented thesis.

Chapter 3

Building change detection in multitemporal SAR images

The increasing availability of very high resolution (VHR) images regularly acquired over urban areas opens new attractive opportunities for monitoring human settlements at the level of individual buildings. This chapter[1] *presents a novel approach to building change detection in multitemporal VHR Synthetic Aperture Radar (SAR) images. The proposed approach is based on two concepts: i) the extraction of information on changes associated with increase and decrease of backscattering at the optimal building scale; and ii) the exploitation of the expected backscattering properties of buildings to detect either new or fully demolished buildings. Each detected change is associated with a grade of reliability. The approach is validated on a) COSMO-SkyMed multitemporal spotlight images acquired in 2009 on the city of L'Aquila (Italy) before and after the earthquake that hit the region, and b) TerraSAR-X multitemporal spotlight images acquired on the urban area of the city of Trento (Italy). Results demonstrate that the proposed approach allows an accurate identification of new and demolished buildings while presents a low false alarm rate and a high reliability.*

3.1 Introduction

Monitoring of urban areas is of great importance for several applications such as urban planning, cadastral map updating, environmental monitoring, disaster assessment and so on. The uncensored synoptic view and the repeat-pass nature of satellites render them an ideal platform from where acquiring information about human settlements. Nonetheless, the huge amount of data acquired from the satellite sensors requires the development of automatic algorithms that can process the data and extract the desired information without any manual processing or ground truth information. As discussed in the introduction chapter of this thesis, in the last decades a new generation of satellite sensors has been operated, which can regularly acquire very high geometrical resolution (VHR) images i.e., images having a resolution of a meter or less. The increasing availability of

[1] Part of this chapter appears in:
[55] C. Marin, F. Bovolo, and L. Bruzzone, "Building Change Detection in Multitemporal Very High Resolution SAR Images," in *Geoscience and Remote Sensing, IEEE Transactions on,*, v. 51, n. 4 part I (2013), p. 2042-2054.

such data allows the analysis of urban areas at a detail level never reached before resulting in the possibility of detecting buildings individually. In this context, several automatic techniques for the study of urban areas have been developed by exploiting both active Synthetic Aperture Radar (SAR) and passive sensors [30,56–59]. Between these two technologies, the use of SAR systems for addressing the problem of monitoring urban area changes is very attractive from an operational point of view since, different from optical sensors, SAR is independent from the sun illumination and it is relatively insensitive to atmospheric weather conditions. This makes it possible to plan either the acquisition of data in advance (e.g., according to end-user requirements without unpredictable intervention of atmospheric effects) and to ensure data availability during crisis events (e.g., floods, earthquakes). Nowadays several SAR missions are operating that can acquire regularly VHR SAR images. Among them we recall TerraSAR-X, Tandem-X and the COSMO-SkyMed constellation. Data acquired from these missions can be exploited to detect changes in urban areas at the level of each single building. Nevertheless, the combination of high resolution and multitemporal analysis leads to some challenging issues that should be addressed as pointed out in chapter 1.

VHR SAR images are more heterogeneous than high or medium resolution data [60,61]. Objects that are considered homogeneous from a semantic point of view (e.g., buildings) show a signature that is inhomogeneous at high spatial resolution because of the scattering contributions from sub-objects (e.g., facade and roof in a building). Furthermore, on the one hand the side-looking illumination required by SAR systems leads to phenomena such as layover, shadow and multi-path signals [62, 63], which are very pronounced in urban areas. On the other hand, the appearance of a ground object depends on radar system parameters (i.e., wavelength, polarization, pulse length, incidence angle, look direction, etc.), surface feature properties (e.g., dielectric constant) and environmental variables (e.g., ground water content) [64]. On top of these aspects, SAR images are corrupted by speckle noise.

All these factors, propagated to the multitemporal analysis, make the problem of the detection of changes complex. Due to the high resolution a large set of possible changes with different semantic meaning and scale are detectable in multitemporal VHR SAR images. In general, each change may be associated with the cause of the change itself [22] and may present an extension that varies from the single pixel to a relevant portion of the entire scene. For instance it is possible to distinguish among changes due to the anthropogenic activity, the phenological evolution of the vegetation and the natural disasters. Depending on the application, some of these may be of interest to the end-users, whereas others may not [65]. In addition, external factors, such as different content of water on the ground due to different weather conditions may affect the local backscattering behaviors at two dates also in absence of any other change. Therefore, the same object may show different value of backscattering even though it is not affected by a relevant change. This invalidates the assumption that two SAR images acquired on the same geographical area at different times are similar to each other except for the presence of changes occurred on the ground, which is often considered for high- and medium-resolution SAR images. Thus the use of standard pixel based change detection (CD) techniques is not applicable as they would be affected by a large amount of false alarms.

Given the complexity of the problem, also standard context-based techniques based on a local neighborhood analysis would fail to solve the problem since they are not able to properly take into account the high geometrical detail of VHR data. To achieve CD in VHR SAR images the comparison between the multitemporal images should be performed at a higher conceptual level that models the source of change from the prospective of interactions with the incidence electromagnetic wave.

Numerous studies have been recently presented in the literature that deal with the problems of recognition of changes in urban areas using VHR SAR images [6, 31, 66–76]. In detail, [6,31,66–73] are focused on the detection of earthquake damages, [74] is focused on the building databases updating, and [75] addresses the detection of changes due to the urban evolution. The strategies that are exploited in these papers include post-event supervised analysis [67, 68, 71–73], joint use of optical and SAR data [6, 66], joint use of LiDAR and SAR data [75], unsupervised detection of damages at the level of aggregated blocks [31, 69] or GIS polygons [70]. Despite the great interest, the only work that addresses the problem of building change detection using multitemporal VHR SAR data in an unsupervised way is presented in [76]. In detail, the authors present an approach to the detection of damaged structures in urban areas from VHR SAR images, which is based on the multitemporal detection of double-bounce line generated by the multiple backscattering between the wall of the building and the ground. If a double-bounce line appears (disappears) between two acquisitions a new (destroyed) building is recognized. The method uses only one of the salient feature used for detecting a building. This partial modeling of the source of change may generate a relevant number of missed alarms. Furthermore, the double-bounce line of a building alone is not a reliable feature for the identification of buildings because in several cases it may be not visible [77].

In this chapter, we propose a novel approach to building change detection in VHR SAR images that: i) is unsupervised; ii) extends the studies on the backscattering properties of buildings presented in the literature for single date images [78–84] to bi-temporal images; iii) takes advantage of the multitemporal correlation between images; iv) considers the intrinsic multiscale nature of objects present in VHR images; and v) is flexible. In greater detail, the approach is based on two concepts: i) the extraction of information on changes associated with increase and decrease of backscattering at the optimal building scale; and ii) the exploitation of the expected backscattering properties of buildings to detect new and fully demolished buildings with their grade of reliability. The effectiveness of the proposed approach is demonstrated in experiments carried out on two data-sets: the former is acquired by COSMO-SkyMed satellites over L'Aquila, Italy, which was heavily damaged during the 2009 earthquake. The latter is acquired by TerraSAR-X and Tandem-X over the city of Trento, Italy and it represents the urban evolution of the city from 2011 to 2013.

This chapter is organized into five sections. Section 3.2 presents and reviews the fundamentals of the backscattering mechanism of buildings in monotemporal images. Moreover it proposes an analysis of this mechanism for bi-temporal SAR images. This introduces the concept used in the proposed approach. The proposed approach to building change detection is described in Section 3.3. Section 3.4 presents the data set and show the experimental results. Section 3.5 draws the conclusions of the work.

3.2 Fundamentals on the Building Backscattering Mechanism

In this section we analyze the backscattering mechanisms with respect to the variation of the acquisition geometry when isolated buildings with rectangular layout are sensed by VHR SAR (Section 3.2.1). This analysis is used to interpret multitemporal VHR SAR images in order to study and model the effects associated with new/demolished building (Section 3.2.2). This study gives the base for the development of the proposed approach.

3.2.1 Building Backscattering Mechanisms in Single Detected VHR SAR Images

SAR is an active system that measures the backscatter of a transmitted signal in the microwaves portion of the spectrum. The backscattering value is mainly determined by the geometry of acquisition, the dielectric and geometric properties of the target, and the transmission configuration of the antenna [18, 64]. In the literature several works have been presented that analyzed the scattering mechanism for different building models in order to derive their appearance in SAR images [6, 80, 86]. These analyses make use of the geometrical optics (ray theory) approximation in order to model the electromagnetic scattering. This has the power to be intuitive and geometrically accurate even though the electromagnetic interactions and the transmission configurations of the antenna are not taken into account. In geometrical optics the wave propagation is described by rays, which are modeled as lines perpendicular to the wavefronts, that may be reflected, absorbed or split at the interface between two media. By knowing the geometry of acquisition of the SAR system and by exploiting trigonometric functions it is possible to simulate the appearance of a building in VHR SAR images [6, 80, 86].

Let us consider an isolated flat-roof building with dimensions $w_1 \times w_2 \times h$ illuminated by a SAR sensor that is moving along the azimuth direction and is illuminating the building from the left (Fig. 3.1a). Let us consider a building model that does not take into account the building features such as windows, eaves, ridges, railings and so on. This is equivalent to analyze the image at the scale level comparable with the building, and it allows us to derive the appearance of a building without losing generality. The acquisition geometry of a SAR system (Fig. 3.1a) is characterized by two parameters: the incidence angle θ (i.e., the angle defined by the incident radar beam and the normal to the intercepting surface) that is generally included between 20° to 55°, and the aspect angle $0° \leq \phi < 90°$ (i.e., the angle between the azimuth direction and the orientation of an object in the horizontal plane). For the sake of argument, let us first assume that the SAR sensor illuminates the section A-A (light blue area in Fig. 3.1a) of the building from a fixed position in the azimuth. Fig. 3.1b shows in a qualitative way the amplitude of the backscattering projected in ground range and slant range of A-A. Following [6] the main contributions that can be identified are: the return a from the ground, the double bounce effect b caused by the dihedral reflector formed by the building wall and the ground, the backscattering c from the front wall, the returns d from the building roof and the shadow area e (see Fig. 3.1b). As one can observe the contributions from the ground, the wall and the roof are summed up into an area in front of the building ($a + c + d$), which appears brighter than each of these contributions taken singularly because of their superposition. This

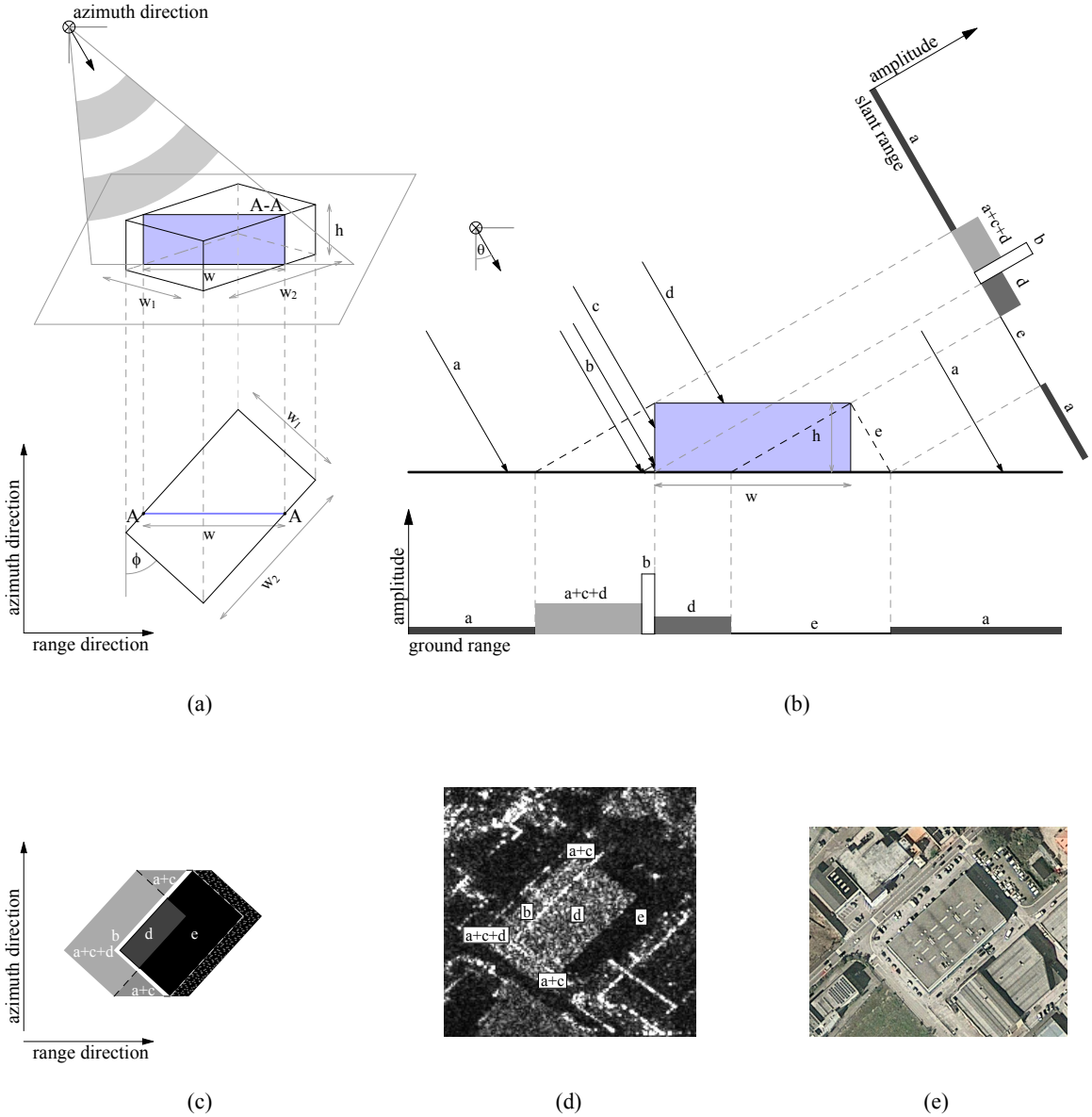


Figure 3.1: Example of backscattering mechanism for a flat roof building. (a) Simplified model of a flat roof building oriented with a given aspect angle ϕ. The SAR sensor is moving on the left of the building. Theoretical scattering model for a given incidence angle θ in the case the sensor is illuminating: (b) only the section A-A of the building from a fixed position in the azimuth (the backscattering is reported in both slant and ground range); (c) the whole extension of the building (the backscattering is in slant range geometry). Different gray levels represent different amplitudes. Example of: (d) a real flat-roof building in Trento (Italy) acquired at 1 m resolution by TerraSAR-X in slant range (SAR illumination is from the left); (e) the same building in an optical image [85].

phenomenon, called layover, occurs when SAR sensors are imaging a surface with a slope steeper than the incidence angle θ, such as the wall of a building. It is worth noting that the length of the layover, the roof contribution and the shadow depends on the width (w) and the height (h) of the section A-A, and the incidence angle θ.

In order to derive the appearance of the whole building in a VHR SAR image it is necessary to perform the aforementioned analysis for all the sections parallel to A-A that form the building. Nevertheless, as the width of each parallel section changes, the length of each mentioned contribution changes accordingly. Three cases can be observed on the basis of the relationship between the width w of the section A-A and the limit value w_b defined as [78, 86]:

$$w_b = \frac{h}{\tan\theta} \tag{3.1}$$

where h and θ are the height and the incidence angle of the building, respectively. Fig 3.1a shows the case in which $w > w_b$. As the width of the section w decreases, the layover and shadow result larger whereas the backscattering from the roof will be absorbed in the layover. This behavior is valid up to the limit condition $w = w_b$ for which the whole roof contribution is sensed in the layover area before the double-bounce. In the case $w < w_b$ two contributions to layover areas can be distinguished: one due to the ground, the front wall and the roof ($a + c + d$), and one due to the backscattering from the ground and the front wall of the building ($a + c$). Fig. 3.1c shows the radar building footprint[2] generated from the building of Fig. 3.1a. It is a convex polygon made up of: i) a bright L-shaped region due to the layover; ii) a dark L-shaped region due to the shadow; iii) a bright L-shaped line due to the double bounce; iv) a rectangular region due to the direct return from the roof; and v) two bright triangular regions due to the layover of only ground and wall ($a+c$). It is worth noting that, since in the acquisition phase the radar is moving, the shadow casted by a ground object is moving as well resulting in a blurring of the border of the shadow [87,88] (see the dotted area in Fig 3.1c). By comparing Fig. 3.1c and Fig. 3.1d it is possible to observe that the geometrical optics approximation can effectively describe the real behavior of scattering in VHR SAR images.

The same building may appear differently in SAR images according to the value of the aspect angle ϕ. When ϕ is approaching the limit values $\phi = 0°$ and $\phi = 90°$ the layover, the double bounce and the shadow change by approaching a rectangular shape. A special attention has to be given to the double bounce line with respect to the variation of ϕ [62, 63]. In [77] an empirical study on the relationship between the strength of the double bounce and the aspect angle highlighted that the double-bounce contribution drops off significantly if the aspect angle increases from $\phi = 0°$ up to $10°$, whereas it decays moderately for higher angles. θ also affects the appearance of a building in VHR SAR images. By reverting Eq. (3.1), it is possible to derive a limit value θ_b that is equivalent to w_b. Nonetheless, in urban areas the choice of θ can be critical for two reasons: i) the value of backscattering is related to θ i.e., small value of θ generates higher value of backscattering and vice-versa; and ii) a large value of θ generates long shadow areas and small layover areas, and vice-versa. Thus, the backscatter has a dependence on the incidence angle, and there is potential for choosing optimum configurations for different

[2]Differently from [83] here with the term footprint we always refer to the radar footprint of a building.

applications.

The same analysis described in this section for flat-roof building can be conducted for other building models e.g., gable-roof buildings [6]. The outcomes of such a study can be summarized as follows: the footprint of any type of isolated buildings with a rectangular base is given by a specific convex pattern made up of a bright area (due to layover and double bounce effects) followed by a dark area (due to the shadow effect). These features may have different thickness, shape and internal variability of backscattering on the basis of the considered building structure and material. Nevertheless they will systematically arise when SAR systems are sensing an isolated building with an adequate resolution.

3.2.2 Building Backscattering Mechanisms in Multitemporal VHR SAR Images

The aim of this section is to analyze the behavior of the radar backscattering in multitemporal images when a building changes by taking into account the single data analysis carried out in the previous section. This analysis introduces the basic concept on which the proposed approach is based. In order to properly illustrate the problem, let us focus on a building that fully disappears between two acquisitions. Let us assume to sample this situation acquiring one image when: i) the building is standing; and ii) the building is totally dismissed. Let us assume to use a VHR SAR system configured with the same geometrical parameters (i.e., same incidence angle θ and same azimuth path) for the two acquisitions. As in Section 3.2.1, let us start the analysis by comparing the backscattering considering the illumination source fixed at a given point along the azimuth direction corresponding to section A-A (Fig. 1.a) in the case the building is present. The backscattering profile obtained for the standing building is reported in Fig. 3.2a. As one can notice, the backscattering behavior is the same as the one obtained in the previous section and reported in Fig. 3.1b. Whereas for the case of totally dismissed building, by assuming bare ground as depicted in Fig. 3.2b, the value of backscattering is approximately constant. By comparing Fig. 3.2a and Fig. 3.2b one can observe that the region with a high value of backscattering due to the layover $a+c+d$, the double bounce b, and the roof contribution d decreases its value when the building disappears, whereas the region hidden by the shadow e becomes visible to the line-of-sight of the radar and therefore increases its value (the dashed line in Fig. 3.2a represents the envelope of backscattering of bare ground). In other words, in the case of a new building we are likely to observe a structured pattern made up of two regions having increase and decrease of the backscattering oriented from near-to-the-far-range. Vice-versa if we consider the case in which a building disappears between two acquisitions, we expect that a structured pattern made up of two regions having decrease and increase in the backscattering values arises in near-to-the-far-range (the dashed line in Fig. 3.2b represents the envelope of the backscattering amplitude of a building).

This specific multitemporal behavior obtained when the illumination source is fixed at a given position in azimuth can be used to retrieve the appearance of new/demolished buildings in multitemporal VHR SAR images by repeating the same analysis for all the sections parallel to A-A that form the scene. The obtained new/destroyed radar building footprint is made up of a pattern included in a convex polygon. The pattern includes two

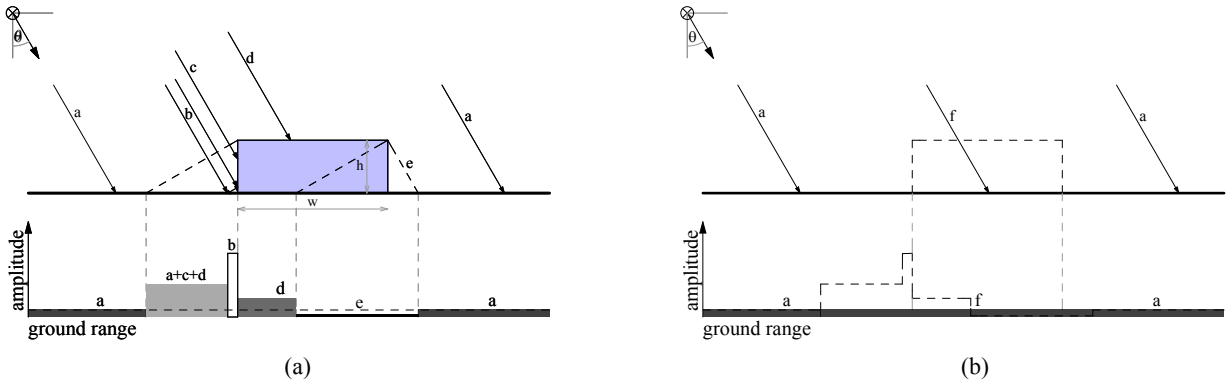


Figure 3.2: Building scattering mechanisms for a fixed illumination source corresponding to: (a) a section of a building; (b) bare ground. The comparison of Figures (a) and (b) describes the multitemporal scattering mechanism in the case (a) a new building is built up (the envelope of backscattering of bare soil is reported for comparison in dotted line); and (b) a building is dismissed (the envelope of backscattering of the building in (a) is reported in dotted line).

regions that can be classified as: i) area of increase of the value of backscattering and, ii) area of decrease in the value of backscattering. The order of appearance of these two regions along the near-to-far-range direction defines if the pattern is due to new buildings (see Fig. 3.3a) i.e., the increase area (in this work depicted conventionally in magenta) is closer to the sensor than the area of decrease (in this work depicted conventionally in green), or demolished building (see Fig. 3.3b) i.e., the decrease area is closer to the sensor than the increase area.

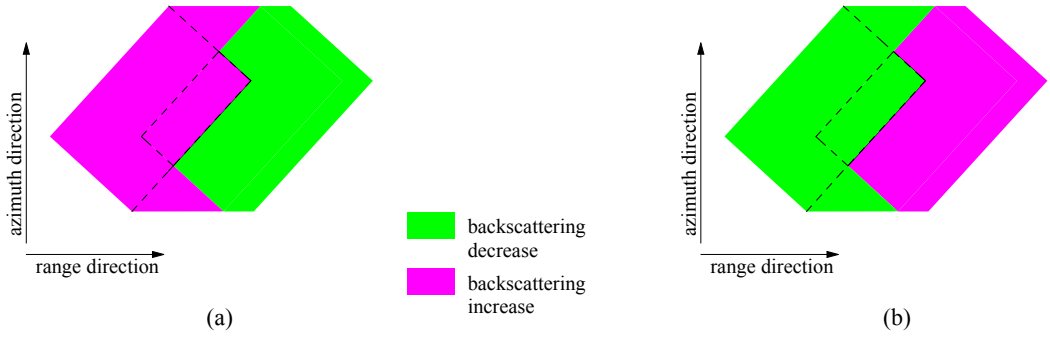


Figure 3.3: Building scattering mechanisms in multitemporal VHR SAR images (illumination source on the left). Idealized map highlighting decrease and increase in the value of backscattering (unchanged pixels are in white) obtained in the case of (a) new and (b) fully demolished building.

By taking into account the analysis done in section 3.2.1 on the radar signature of a general isolated building it is possible to derive the appearance of the radar footprint of any type of newly built up or destroyed buildings in VHR SAR multitemporal images. In detail, the change radar building footprint can be identified checking: i) the presence of both the regions of increase and decrease in backscattering; ii) the proportion between the areas (in pixels) of the regions of increase and decrease in backscattering. The proportion depends on the incidence angle of the acquisition; iii) the equality between the lengths

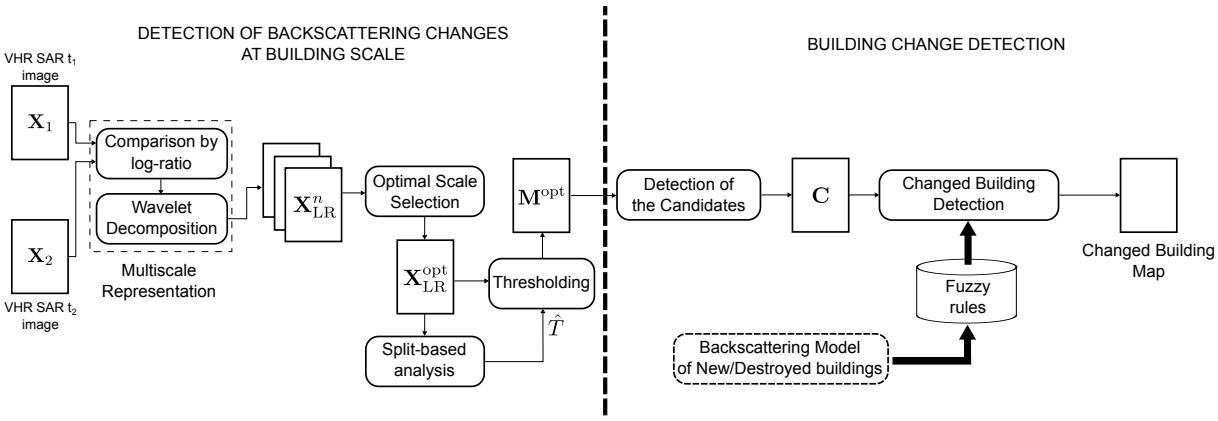


Figure 3.4: Architecture of the proposed approach to building change detection.

of the regions of increase and decrease in backscattering in the azimuth direction; iv) the alignment of the regions of increase and decrease in backscattering with respect to the range direction i.e., the barycenters of the two regions lay on the line with the range (As the regions in the model are not regularized the barycenters were calculated by means of the geometric decomposition method). The order of appearance of the two regions of increase and decrease in backscattering value determines if the changed radar building footprint is due to a new or a destroyed building.

It is worth noting that the regions of change may have different thickness, shape and variability of backscattering change on the basis of the considered building structure and acquisition geometry. Despite this variability, the pattern will systematically arise when SAR systems are sensing an isolated building with an adequate resolution [89]. Hence, a method for the detection of the changed building based on this concept results to be robust to both the noise and the uncertainty (i.e., the impossibly of modeling the reality precisely as described in Section 3.3.4) of multitemporal VHR SAR data.

3.3 Proposed Building Change Detection Approach

Let us consider two amplitude VHR SAR images $\mathbf{X}_1$ and $\mathbf{X}_2$ of size $I \times J$ acquired with the same incidence angle on the same geographical area at different times τ_1 and τ_2, respectively. Let $\Omega = \{\omega_u, \Omega_c\}$ be the set of classes of changes to be identified: ω_u represents the class of pixels having unchanged backscattering value, whereas $\Omega_c = \{\omega_{c1}, \omega_{c2}, \ldots, \omega_{cK}\}$ is a meta-class that gathers all the K possible classes (kinds) of change that may arise on the ground. One of the most critical issues dealing with this kind of scenario is related to the presence of many kinds of changes on the ground. Nevertheless in this work we are interested to investigate an urban area with the goal to only detect changed buildings. In detail, we consider $K = 4$ classes of change: i) fully destroyed buildings (ω_{c1}); ii) new buildings (ω_{c2}); iii) changes that have a size comparable to a building but do not present the typical pattern of full new/demolished buildings (ω_{c3}); iv) all the other changes that do not show a size comparable to a building and are therefore not related to changed buildings (ω_{c4}). In order to achieve this classification we introduce

an approach made up of two stages: i) identification of the areas affected by changes in the backscattering at the scale of buildings; and ii) exploitation of the backscattering models presented in Section 3.2.2 in order to detect the classes $\Omega_c = \{\omega_{c1}, \omega_{c2}, \ldots, \omega_{c4}\}$. Each stage of the proposed method is explained in detail in the next subsections. Fig. 3.4 shows the block scheme of the proposed approach.

3.3.1 Detection of Backscattering Changes at Building Scale

As described in Section 3.2.2, under the hypothesis of using a SAR sensor with a resolution comparable to the building size, the signature of isolated new/demolished buildings in multitemporal data is given by a pattern made up of increase and decrease of backscattering regions. The first stage of the proposed approach is thus devoted to extracting the areas that present whether a significant increase or decrease in the backscattering value. Nonetheless, at a resolution of a meter or less the small objects that form the buildings such as windowsills, or rain drains are visible. This results in an inhomogeneous signature of the building, which may affect also the appearance of new/destroyed buildings in multitemporal images rendering the regions of increase and decrease far from being homogeneous. Therefore, an optimum scale level for representing buildings, and not subparts of them, has to be derived from the VHR images and used to identify the regions of increase and decrease of backscattering. Working at the scale level of buildings has the additional advantages of: i) reducing the impact on the detection of small changes and thus reducing the false alarm rate; ii) making it possible a mitigation of the speckle effect on the detection.

In order to work at the scale of a building we propose to build a multiscale representation of the multitemporal information made up of N scale level. The $(N-1)^{\text{th}} = \text{opt}$ resolution level, which represents the optimum scale level, is select according to the minimum size of the building in the investigated scene. For n ranging from 0 to $N-1$, the images are characterized by resolution that is degraded approximately by a factor of 2^n. Therefore, as expected by decreasing the resolution, small changes tend to disappear and only changes of a given size are fully preserved. More details on this strategy are given in chapter 2.2.2.

Once the optimum scale level (opt) has been selected, a change detection (CD) map $\mathbf{M}^{\text{opt}}$ is derived from $\mathbf{X}_{\text{LR}}^{\text{opt}}$ according to an unsupervised thresholding procedure. $\mathbf{M}^{\text{opt}}$ presents three classes: i) no-change (ξ_u); ii) increase (ξ^+); and iii) decrease (ξ^-) of backscattering. As the analysis of large urban areas with VHR SAR sensors leads to the generation of large multitemporal images it is possible that the population of changed pixels is in sharp minority in the full $\mathbf{M}^{\text{opt}}$. This may affect the accuracy of the threshold selection technique. In order to address this issues we adopt the thresholding method presented in chapter 2.2.3. This method divides the considered image into sub-images of a given size (splits) and performs the threshold selection considering only the splits that present the highest probabilities to contain changed pixels. This selection allows the definition of a subset of pixels in which the class of change shows a higher prior probability than in the whole image. For our purpose, $\mathbf{X}_{\text{LR}}^{\text{opt}}$ is split into a set of S sub-images of a user-defined size. The choice of the split size $S_R \times S_A$ is driven by the average extension of the expected changes. Since we are considering changes related to buildings, we

can determine S_R and S_A taking into account the average size of the radar footprint of the changed buildings in the considered radar image. This information can be inferred by considering the actual average size of the buildings i.e., w_1, w_2 and h (see Fig. 3.1a) in the considered area and converting it into SAR geometry by applying the following equations[3] [86]:

$$S_R = X + Y = \begin{cases} w_1 \sin\theta + \frac{h}{\cos\theta}, & \text{if } \theta \geq \tan^{-1}(\frac{h}{w_1}) \\ h\cos\theta + (w_1 + \frac{h}{\tan^{-1}\theta})\sin\theta & \text{Otherwise.} \end{cases} \tag{3.2}$$

$$S_A = w_2 \tag{3.3}$$

Where X is the length of the return from the building (i.e., the sum of the lengths of contributions ($a + c + d$), b and d of Fig. 2.a) and Y the length of the shadow (i.e., the length of contribution e in Fig. 2a). Consequently, the statistical estimation of the parameters related to the three probability density functions associated with no-change, increase, and decrease of backscattering between the two dates (i.e., ξ_u, ξ^+, ξ^-), can be correctly derived and used to separate the three classes.

In Fig. 3.5 an illustrative example that depicts the detection process of the proposed approach is reported for the case of a demolished building. This example will be used in the chapter to better illustrate the proposed approach. As one can notice, in this example the CD map $\mathbf{M}^{\text{opt}}$ reported in Fig. 3.5a verifies the assumptions defined for the model presented in Section 3.2.1: i) the detected regions of increase or decrease in backscattering at the building scale are homogeneous; and ii) the change detection map presents a reduced number of changes smaller than the size of buildings i.e., changes that are not related to changed buildings (ω_{c4}) are filtered out. From $\mathbf{M}^{\text{opt}}$ it is possible to identify the building radar footprints by locating the pattern of increase and decrease of backscattering (building candidates) and by evaluating the matching between the properties of the building candidates with respect to the properties of new and destroyed building models described in Section 3.2.2. This task is preformed by the building detection stage of the proposed approach and it will be described in detail in the following section.

3.3.2 Building Change Detection

This stage of the proposed approach represents the most novel contribution of this work and it is devoted to the detection of demolished building (ω_{c1}), new buildings (ω_{c2}), changes that have a size comparable to a building but do not present the typical pattern of full new/demolished buildings (ω_{c3}) and all the other changes that do not show a size comparable to a building and are therefore not related to changed buildings (ω_{c4}). To this end a procedure based on two steps is developed. First the areas of change candidate to be associate to one of the classes ω_{c1}, ω_{c3} or ω_{c3} (changed building candidate) are detected among all the backscattering changes highlighted in $\mathbf{M}^{\text{opt}}$. The set of changed building candidates is denoted by $\Gamma = \{\gamma_1, \gamma_2, \ldots, \gamma_H\}$. Then the matching between

[3]Eq. (3.2) and Eq. (3.3) hold assuming flat-roof buildings with rectangular layout and SAR images in slant range geometry.

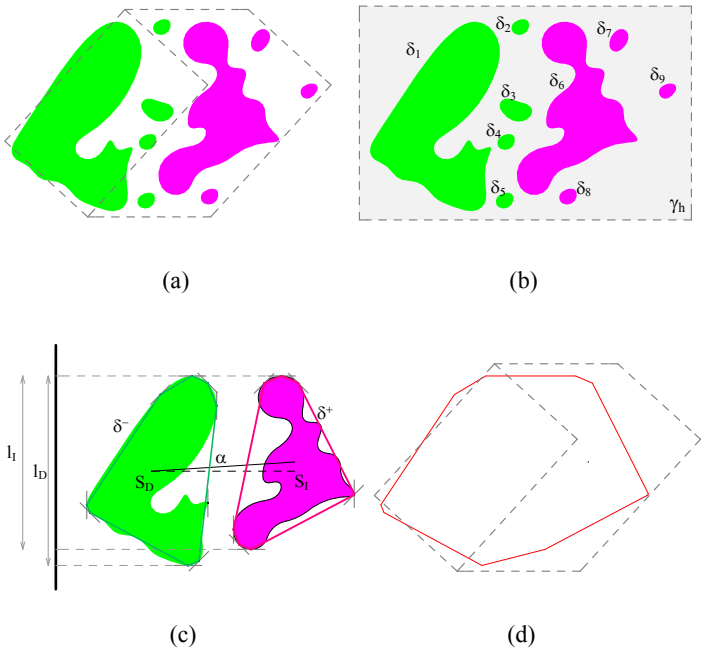


Figure 3.5: Conceptual example of detection of a destroyed building (illumination source on the left). (a) $\mathbf{M}^{\text{opt}}$. The pixels associated to the classes of backscattering decrease (ξ^-) and increase (ξ^+) are reported in green and magenta, respectively. The idealized dismissed radar building footprint is drawn in dashed line; (b) in gray the bounding box of changed building candidate γ_h containing the regions of change $\delta_1, \delta_2, \ldots, \delta_8$ that form the candidate; (c) illustration of the parameters S_I, S_D, l_I, l_D and α, which are evaluated by the fuzzy rules described in Table 3.1 in order to check the matching between the candidate and the expected pattern; (d) identified demolished radar building footprint (red continuous-line) compared with the actual radar footprint (gray dashed-line).

the expected backscattering behavior of changed buildings (presented in Section 3.2.2) and the characteristics of the changed building candidates is evaluated. This is done by considering the physical properties and the relation of the regions of change inside each candidate i.e., $\Delta = \{\delta_1, \delta_2, \ldots, \delta_K\}$. In order to properly model the uncertainty inherent in the process (that can be due for example to cluttered objects placed in front of buildings) fuzzy theory is used [90]. Fig. 3.5b-d illustrates the process of the proposed building change detection stage for the case of a demolished building. In the following each step is described in detail.

The first step of the change detection phase aims at detecting the set of changed building candidates $\Gamma = \{\gamma_1, \gamma_2, \ldots, \gamma_H\}$. Ideally each changed building in $\mathbf{M}^{\text{opt}}$ is made up of regions of both increase and decrease in backscattering, whose total extension is comparable with the expected size of buildings. One simple and effective approach to identify the changed building candidates $\gamma_1, \ldots, \gamma_H$ is to use a sliding window algorithm for each pixel in the image. Pixels within a window around it are taken and used to compute the detector output. In this work, the window is moved in $\mathbf{M}^{\text{opt}}$ from left to the right by one pixel and the number of changed pixels (labeled both as increase ξ^+ or decrease ξ^-) is counted. The process is repeated over 5 moving windows: four

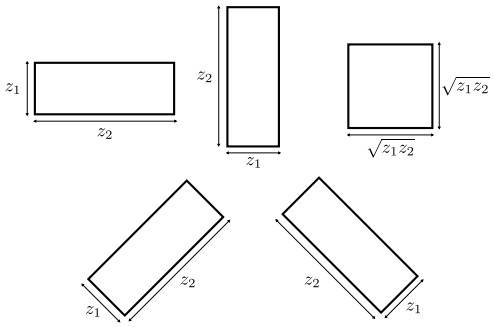


Figure 3.6: Windows used to derive the changed building candidates.

rectangular moving window showing different directions and a square window having the same area of the rectangular ones. This choice allows us to properly capture the most of the building orientations. The detector output is given by the window that results in the maximum value. The 5 windows are depicted in Fig. 3.6. The size of the windows $z_1 \times z_2$ is chosen according to the minimum size of the buildings in the considered scene in order to minimize missed alarms. Hence, let $\mathbf{M}^{\text{opt}}_{W^{\beta}_{i,j}} = \{\mathbf{M}^{\text{opt}} | \mathbf{M}^{\text{opt}} \subset \mathbf{W}^{\beta}_{i,j}\}$ be the set of pixels of $\mathbf{M}^{\text{opt}}$ included in the windows $\mathbf{W}^{\beta}_{i,j}$ centered at the pixel (i, j), with $\beta = 1, \ldots, 5$ indicating the different windows. Let $\mathbf{C}$ be the image with size $I \times J$ that reports for each pixel $C_{i,j}$ an index of the size of the changes. $C_{i,j}$ is computed as the maximum on the five windows as follows:

$$C_{i,j} = \underset{\beta \in 1,2,\ldots,5}{\arg\max} \left\{ \left| \mathbf{M}^{\text{opt}}_{\mathbf{w}^{\beta}_{i,j}} \in \xi^{+} \vee \mathbf{M}^{\text{opt}}_{\mathbf{w}^{\beta}_{i,j}} \in \xi^{-} \right| \right\} \quad \forall \quad i = 1, \ldots, I; \quad j = 1, \ldots, J. \tag{3.4}$$

where $|\cdot|$ indicates the cardinality of a set. The obtained index image $\mathbf{C}$ exhibits relatively high values when the sliding window contains a large amount of changes oriented in the same direction of the moving window. $\mathbf{C}$ exhibits relatively low values when the sliding window contains small changes and when the windows does not match the orientation of the change. Therefore, it is possible to detect the changed areas that can be associated in terms of size to the expected radar footprint of changed buildings by selecting the areas of $\mathbf{C}$ exceeding a certain threshold T_C, thus obtaining the image $\bar{\mathbf{C}}$ as follows:

$$\bar{\mathbf{C}} \in \begin{cases} 1, & \text{if} \quad \mathbf{C} \geq T_C \\ 0, & \text{otherwise.} \end{cases} \tag{3.5}$$

T_C is selected according to the expected minimum size of buildings $z_1 \times z_2$ in the investigated scene. The map $\bar{\mathbf{C}}$ represents the areas containing changes with size comparable or bigger than the minimum building size and therefore they belong to one of the classes $\omega_{c1}, \omega_{c2}, \omega_{c3}$. As one can notice, the pixels belonging to ω_{c4} are implicitly identified in this operation. From the map $\bar{\mathbf{C}}$ the set of changed building candidates $\Gamma = \{\gamma_1, \gamma_2, \ldots, \gamma_H\}$ is extracted by calculating the connected components considering a 8-connected neighborhood. This is done by means of a flood-fill algorithm [91]. Each candidate γ_h $(h = 1, \ldots, H)$ contains a subset of regions $\Delta = \{\delta_1, \delta_2, \ldots, \delta_K\}$ labeled as

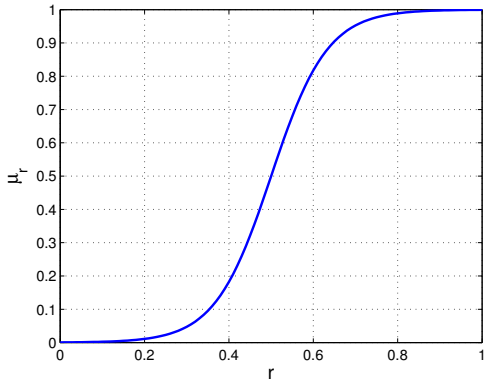


Figure 3.7: Example of sigmoid function $\mu_r(r)$ defined according to (3.6). The slope of the sigmoid is tuned by the parameter a_i. b_i defines instead the center of the sigmoid i.e., $\mu_r(b_i) = 0.5$. In the represented case $a_i = 15$ and $b_i = 0.5$.

ξ_u, ξ^+ or ξ^-. The information associated to the kind of change is derived considering $\mathbf{M}^{\text{opt}}$ in the areas delimited by the region γ_h. Fig. 3.5b shows an example of candidate γ_h that contains δ_k $(k = 1, \ldots, 8)$ regions for the case of a demolished building. In general the number of regions, k may vary from 1 to K depending on the size and proximity of the changes. For the sake of clarity, in the next step of the building change detection procedure we assume that only two regions belonging to increase (i.e., $\delta^+ = \delta_k \in \xi^+$) or decrease i.e., $(\delta^- = \delta_k \in \xi^-)$ occur inside a candidate. If more than 2 regions are present, all the combinations among the regions of increase and decrease are automatically analyzed by the proposed approach and the most reliable(s) selected.

In order to properly classify each changed building candidate according to the classes ω_{c1}, ω_{c2}, and ω_{c3}, the physical characteristics and the spatial arrangement of the pattern formed by the regions δ_k $(h = 1, \ldots, K)$ have to match with the four characteristics of the models of new or dismissed building discussed in Section 3.2.2. The matching is tested exploiting four fuzzy rules called here: completeness, proportionality of areas, equivalence of lengths and alignment. These rules aim at associating a grade of membership to the changed building candidate for each of specific performed test. On the basis of the aggregate membership the final classification decision is taken. In the following each rule is presented in detail.

Completeness

A pair of regions of increase and decrease is strictly needed to identify a new (ω_{c2}) or destroyed building (ω_{c1}). The first rule evaluates the simultaneous presence of δ^+ and δ^- inside the candidate γ_h. This is done by means of a crisp membership function μ_p that takes value 0 or 1 $(\mu_p(p) = \{0, 1\})$ depending on whether regions of increase and decrease are simultaneously present $\mu_p(p) = 1$ or not $\mu_p(p) = 0$.

Proportionality of Areas

This rule aims at verifying that the area δ^+ is not prevailing with respect to the area of δ^-, and vice-versa. The mathematical modeling of this rule depends on the parameters of acquisition and the size of the building as discussed in Section 3.2.1. It can be effectively represented by a sigmoid membership function that evaluates the attribute $r_s = \min\{s_I/s_D, s_D/s_I\}$, where s_I and s_D indicate the areas of δ^+ i.e. $S_I = |\delta^+|$ and δ^- i.e., $S_D = |\delta^-|$, respectively. In general, a sigmoid function is defined as follows:

$$\mu_r(r) = \frac{1}{1 + e^{-a_i(r-b_i)}} \tag{3.6}$$

The constant a_i tunes the slope of function and the constant b_i locates the center of the function. Eq. 3.6 returns values in [0,1]. Fig. 6.3a shows an example of sigmoid function. In accordance with the model it is expected that the smaller is r_s the smaller is the membership grade of the candidate hence for this rule $a_i = a_1 > 0$ and $b_i = b_1 > 0$.

Equivalence of Lengths

This rule aims at checking that the lengths of the regions δ^+ and δ^- in the azimuth direction are equivalent. This is done by identifying the extrema of the regions δ^+ and δ^-. To test the reliability of a candidate with respect to this criterion, the attribute $r_l = \min\{l_I/l_D, l_D/l_I\}$ (where l_I and l_D are the lengths in azimuth direction of δ^+ and δ^-, respectively) is used in the sigmoid membership function defined in eq. (3.6) with parameters $a_i = a_2 > 0$ and $a_i = b_2 > 0$. Since the farther is r_l from 1 the smaller is the membership grade of the candidate to be associated to a new/destroyed building, the slope of the sigmoid is steep and the center moved toward 1. Fig. 3.5.c illustrates an example of correct alignment between δ^+ and δ^-, whereas Fig. 3.8c depicts a case in which this rule is not satisfied.

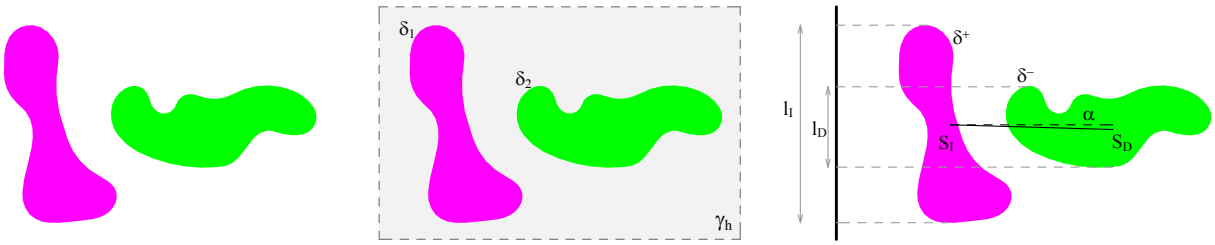


Figure 3.8: Example of a candidate γ_h that does not respect the equivalence of lengths. (a) $\mathbf{M}^{\text{opt}}$. The pixels associated to backscattering decrease (ξ^-) and increase (ξ^+) classes are reported in green and magenta, respectively. (b) In gray the changed building candidate γ_h containing the regions δ_1 and δ_2, which form the candidate; (c) Analysis of psychical properties and relations of the regions δ^+ and δ^-.

Alignment

This rule is devoted to test the alignment of the entire pattern increase/decrease (which is made up of regions δ^+ and δ^-) with respect to the range direction. The alignment of the

centroids of the regions of increase and decrease are considered, which are expected to lay on the same line. In detail, the absolute value of the angle α defined as the angle included by the line that connects the centroids of the regions δ^+ and δ^- and the SAR range direction is used in the equation of the sigmoid eq. (3.6). α ranges from $[-\pi/2, \pi/2)$. As expected, by imposing $a_i = a_3 < 0$ the larger is $|\alpha|$ the smaller is the membership grade of the candidate to be associated to a new/destroyed building. Fig. 3.5c illustrates an example of correct orientation of the pattern increase/decrease, whereas Fig. 3.8c depicts a case in which this rule is not satisfied.

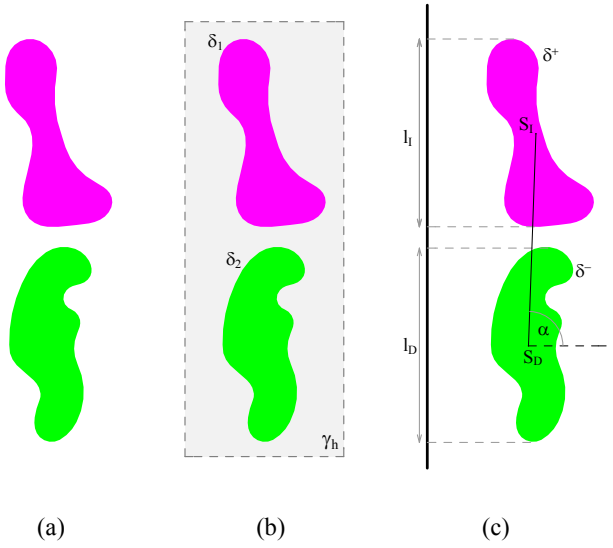


Figure 3.9: Example of a candidate γ_h that does not respect the fourth rule. (a) $\mathbf{M}^{\text{opt}}$. The pixels associated to backscattering decrease (ξ^-) and increase (ξ^+) classes are reported in green and magenta, respectively. (b) In gray the changed building candidate γ_h containing the regions δ_1 and δ_2, which form the candidate; (c) Analysis of psychical properties and relations of the regions δ^+ and δ^-.

A summary reporting the description of each rule, the parameters and membership functions used to represent the rules is reported in Table 3.1.

The final decision on the label of the candidate is done according to the aggregated membership grade M, which is calculated by combining the 4 membership functions. Among the fuzzy aggregation methods presented in the literature in this work the Larsen product implication [92] is used i.e.,

$$\mathrm{M} = \mu_p \mu_{rs} \mu_{rl} \mu_\alpha \tag{3.7}$$

In this way we apply a conservative rule for which one small value among the individual membership grades yields the final result to be small regardless of the other values (other rules can be selected [90]). The candidates that present a value of M greater than a user-defined threshold T_{M} are selected as fully demolished (ω_{c1}) or new buildings (ω_{c2}) considering the order of appearance of the pair increased/decreased with respect to the range direction as described in Sec. 3.2.2. The others are classified as ω_{c3} (i.e., they are not fully demolished or new buildings). It is worth noting that T_{M} will be a small value since the algebraic product of value smaller than 0 result in value smaller than the factors

Table 3.1: Fuzzy rules used to classify γ_h as new building (ω_{c1}), demolished buildings (ω_{c2}) or change comparable to the size of a building but not related to fully new/demolished buildings (ω_{c3}).

Rule	Parameter	Membership Function
1. Presence of both regions δ^+ and δ^- (completeness)	p	$\mu_p(p) = \{0, 1\}$
2. Proportionality of areas of δ^+ and δ^- (proportionality of the areas)	$r_s = \min\left(\frac{s_I}{s_D}, \frac{s_D}{s_I}\right)$	$\mu_{r_s}(r_s) = \frac{1}{1+e^{-a_1(r_s-b_1)}}$
3. Equivalence of the lengths of δ^+ and δ^- along the azimuth direction (equivalence of the lengths)	$r_l = \min\left(\frac{l_I}{l_D}, \frac{l_D}{l_I}\right)$	$\mu_{r_l}(r_l) = \frac{1}{1+e^{-a_2(r_l-b_2)}}$
4. Alignment of the barycenters of the regions δ^+ and δ^- (alignment)	$\lvert\alpha\rvert$	$\mu_{\lvert\alpha\rvert}(\lvert\alpha\rvert) = \frac{1}{1+e^{-a_3(\lvert\alpha\rvert-b_3)}}$

in (3.7). Nonetheless, the value of M obtained for each of these classes is representing a grade of reliability of the detection that can be used by the final users.

For each changed building candidate recognized as new or demolished building the 8-extrema points (i.e., from top-left to bottom-right) of both the regions δ^+ and δ^- are calculated and used to find the smallest convex polygon that contains the two regions i.e., the convex hull (see Fig. 3.5d). The computation of the convex hull is done using the Quickhull algorithm [93]. This allows the generation of an envelope of a new or dismissed building footprint as much as similar to the ideal expected one. Once the building footprint candidate is classified as new or demolished building, reverting the equations (3.2) and (3.3) it is possible to estimate the actual size of the building affected by the change, which is expressed in terms of $w_1[m] \times w_2[m] \times h[m]$ (see Fig. 3.1a).

3.3.3 Selection of the Parameters

In this section the role of the parameters used in the proposed approach is discussed in detail. The parameters can be divided into 3 groups: the first two groups are related to the building size and the third to the fuzzy rules. The first group includes S_R and S_A, which represent the size of the split used for the automatic thresholding procedure of the log-ratio image at the optimal building scale. As discussed in the chapter, the split size is selected according to the average size of building in the considered area. It is worth noting that the split-based thresholding approach has demonstrated to be tolerant to relative large variations of the selected split size with respect to the expect average size of buildings (see Sec. 4.4). The second group of parameters includes $N-1, z_1,\ z_2$ and T_C are related to the minimum size of the buildings in the investigated site. In general, the minimum size of buildings is driven by the typology of urban area (industrial or residential) and the zone of the world under analysis (e.g., typically buildings in US metropolis are larger than buildings in European towns). Nonetheless, by selecting large values of $N-1, z_1,\ z_2$ and T_C building candidates with size smaller than the chosen values can be excluded from the analysis. The third group of parameters is associated to the membership functions of the fuzzy rules. In detail the center of the sigmoids (regulated by b_1, b_2, b_3) represents the boundary situation that has half membership with respect to the theoretical model described in Section 3.2.2, while the slope (regulated by a_1, a_2, a_3) represents the flexibility of the rule to cope with clutter and noise. In the considered case

the centers are fixed from the model, whereas the slope can be chosen considering real situations e.g., urban density.

On the basis of this analysis and of our experimental results we can conclude that the tuning of the parameters to give as input to the proposed approach requires very basic prior information and is not critical.

3.3.4 Critical Analysis

It is worth noting that, as discussed in Section 3.2 the appearance of a building in VHR SAR imagery can vary with respect to the model of the building. For instance, in the case discussed in Section 3.2.1 the area of the regions of increase and decrease of backscattering are similar to each other. This is not always the case as it depends on both the size of the building and the incidence angle θ used for the acquisition. Moreover, objects placed in front of the building, such as trees or other buildings, or the high urban building density results in the occlusion of neighboring of the building facade and the corner reflector so that the double bounce and the layover might not be visible for the full extent of the building front wall. As a consequence, important primitives that identify a building may be partially not visible reducing the validity of the aforementioned conditions. This challenging situation is taken into account in the proposed method by using the fuzzy theory. Moreover, note that the comparison of multitemporal images according to the log-ratio operator also mitigates the negative effects of possible occlusions common to both images. Accordingly, the proposed approach is expected to be more robust to non ideal conditions than methods based on the separate detection of buildings in single data images. Thus the proposed approach is robust against false and missed alarms as it will be discussed in the experimental results section.

As a matter of fact, the multitemporal analysis becomes complex if we consider the case of building partially damaged or under construction. In these cases the backscattering characteristics are dependent on the specific kind and the extension of the change. In [6, 94, 95] it was proven that not all kinds of building damages are discernible in meter resolution SAR imagery. This is valid also in the case of partially built buildings. At present this is still an open topic that will be further investigated as future development of this work.

3.4 Data Sets Description and Experimental Results

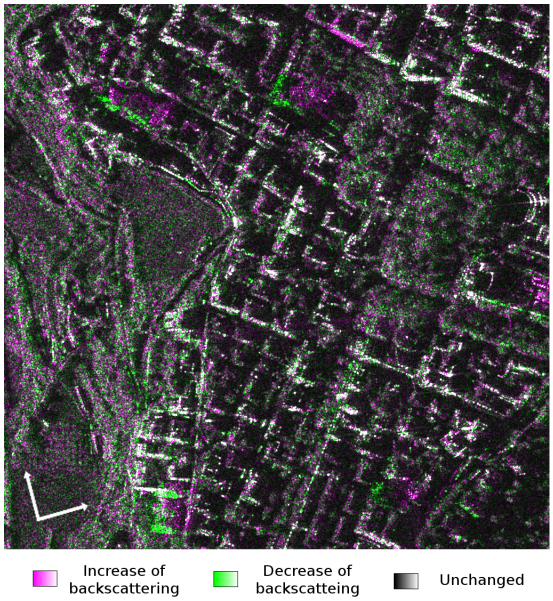
In order to assess the effectiveness of the proposed approach, different experiments were carried out on two data sets. The first one is related to the 2009 L'Aquila earthquake occurred in the region of Abruzzo, in central Italy. The second one documents the urban evolution of the city of Trento, in north Italy, between 2011 and 2013.

3.4.1 2009 L'Aquila Earthquake: Detection of Destroyed Buildings

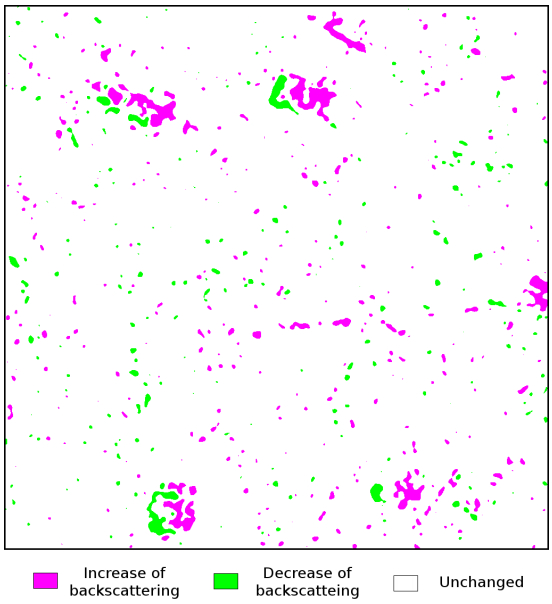
The first data set is made up of two spotlight mode X-band CSK 1-look amplitude images processed according to the standard processing level 1C (Geo-coded ellipsoid corrected, 1m resolution, 0.5m×0.5m pixel spacing). They were acquired in HH-polarization on


N

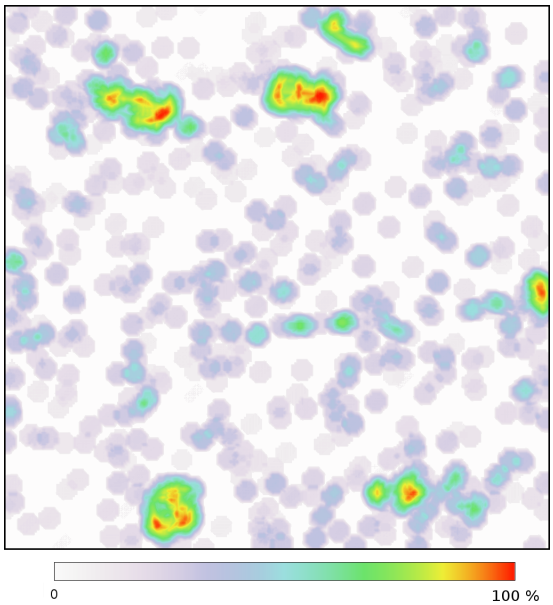
(a)


Increase of backscattering
Decrease of backscatteing
Unchanged

(b)


Increase of backscattering
Decrease of backscatteing
Unchanged

(c)


0
100 %

(d)

(e) (f)

Figure 3.10: L'Aquila (Italy) data set: (a) optical image. (September 4, 2006 [85]); and (b) RGB multitemporal composition of spotlight COSMOSky-Med images (R:09/12/2009, G:04/05/2009, B:09/12/2009). The images are geo-projected according to the reference ellipsoid. Actual range and the azimuth directions are reported in white arrows. (c) Change detection map $\mathbf{M}^{\text{opt}}$ obtained at the level $N-1=3$. (d) Map showing the index of the size of changes $\mathbf{C}$. (e) Changed building map overlapped to the RGB multitemporal composition of CSK data. (f) Cadastral map of the area (destroyed buildings are reported in red).

April 5, 2009 and on September 12, 2009, over the city of L'Aquila (42°21′N 13°24′E) in ascending orbit with 57-58 degree incidence angle. On April 6, 2009, an earthquake of 6.3 Moment Magnitude Scale (MMS) struck central Italy with its epicentre near L'Aquila.

A test site of 1024×1024 pixels of the full spotlight scene was selected in the city of L'Aquila in order to test quantitatively the proposed method. Fig. 4.6a shows the optical image corresponding to the area of interest. Fig. 4.6b shows a false color composition of the two CSK images (red channel:09/12/2009, green channel:04/05/2009, blue:09/12/2009) in which pixels with an increase in the value of backscattering appears in magenta tone, pixels with a decrease in the value of backscattering appear in green tone and unchanged pixel in gray-scale. From the cadestrial map of the considered site 200 buildings were counted and 6 of them were identified as totally collapsed because of the earthquake. This information was derived by analyzing a couple of orthophotos with a resolution of about 20 cm acquired April 8, 2009 and October, 2009 [96]. It is worth noting that a relevant number of changes are present between the two acquisitions. Some of those are not relevant from the application viewpoint because they are due to the activities occurred after the earthquake (e.g., shoring up of damaged structures or deployment of emergency services). This increases the complexity of the problem of building change detection.

As described in the methodological part, the first step of the proposed approach is the calculation of the log-ratio image $\mathbf{X}_{\text{LR}}$. It was computed from the two calibrated and

Table 3.2: Performance of the proposed approach for L'Aquila Dataset in terms of number of detected destroyed buildings and missed destroyed buildings. The total number of buildings present in the site is also reported in the table.

Total Number of Buildings	Number of Destroyed Buildings (reference)	Detected Destroyed Buildings (ω_{c1})	Missed Destroyed
200	6	7	0

co-registered CSK images. In detail, the co-registration was performed with sub-pixel accuracy and the radar brightness calibration was applied in order to render the two images comparable. From $\mathbf{X}_{\mathrm{LR}}$ the set $\mathcal{X}_{\mathrm{MS}} = \left\{\mathbf{X}^0_{\mathrm{LR}}, \ldots, \mathbf{X}^n_{\mathrm{LR}}, \ldots, \mathbf{X}^{N-1}_{\mathrm{LR}}\right\}$ was computed by applying sequentially the 2D-SWT and the 2D-ISWT with an 8-length *Daubechies* filter. The level $(N-1)^{\mathrm{th}} = 3^{\mathrm{rd}}$, which gives a resolution of approximately 8 meters, was selected as the optimum one. This allows to observe radar footprints of buildings as small as 8 by 8 meters, which is compatible with the minimum size of building footprints in the considered scene. $\mathbf{X}^3_{\mathrm{LR}}$ was thresholded according to the unsupervised split-based procedure described in section 3.3.1. The split size was calculated starting from the average building size, which was estimated to be $w_1 \times w_2 \times h = 25 \times 20 \times 15$ m^3. By substituting these values in eq. (3.2) S_{SR} was obtained to be equal to 49.50 m in slant-range geometry. Since the CSK images are ground-projected, S_{SR} was projected in ground range considering the reference incidence angle i.e., 58 degree obtaining in this way $S_{GR} =$ 58.36 m. From the calculation and taking into account a pixel spacing in range of 0.5 m, S_R results to be equal to 120 pixels. By applying eq. (3.3) and considering a pixel spacing in azimuth of 0.5 m, S_A was estimated to be equal to 40 pixels. It is worth noting that the split-based thresholding technique has a relative high tolerance to the selection of the values of S_R and S_A. This was proven by a sensitivity analysis in which S_A was ranged from 20 to 60 pixels and S_R from 80 to 170 pixels. Despite the large range (about $\pm 1/3$ of the optimal selected size) the error on the estimation of the thresholds is small. That is of 0.054 in average with a peak value of 0.14 for $\hat{t}^-$, and of 0.045 in average with a peak value of 0.15 for $\hat{t}^+$. Moreover, in all the cases the generated maps did not present any critical behavior and resulted in a very similar final detection of changed buildings.

The obtained CD map $\mathbf{M}^{\mathrm{opt}}$ highlighting the three classes ξ_u, ξ^+, ξ^- is reported in Fig. 4.6c. The detector described by (3.4) was then applied to $\mathbf{M}^{\mathrm{opt}}$ in order to obtain the image $\mathbf{C}$. With regard to the considered scene, it is expected that the minimum size of building footprint is 40×20 pixels. Therefore, the set of 5 windows reported in Fig. 3.6 were set to size $z_1 \times z_2 = 40 \times 20$ pixels in order to derive $\mathbf{C}$ (see Fig. 4.6d). A set of changed areas $\Gamma = \{\gamma_1, \ldots, \gamma_{49}\}$ compatible with building footprint changes was derived by threhsolding $\mathbf{C}$ with $T_C = 160$ (which corresponds to 20% of maximum value of $\mathbf{C}$) and by applying a flood-fill algorithm. It is worth noting that both the window size and the threshold value were selected according to minimum size of buildings in the considered area in order to limit the missed alarms. Each of the candidate areas was analyzed in order to detect the changed buildings. To this end the four fuzzy rules were automatically evaluated for each candidate γ_h, $(h = 1, \ldots, 49)$. Finally, for each building candidate that presents an aggregated membership T_{M} greater than 0.125 the radar building footprint was approximated by the convex-hull containing the candidate. It is worth noting that

T_{M} was selected in order to minimize the missed alarms. This is done by taking into account the definition of the membership functions and the aggregation strategy (i.e., algebraic product). In the present experiment $\mu_{rs} = \mu_{rl} = \mu_{\alpha} = 0.5$ were selected as limit case, which results in an aggregate membership function $\mathrm{M} = 0.125$. Fig. 4.6e shows the final result. In the specific case, all the reconstructed radar building footprints have an aggregate membership function greater than 0.615. Table 4.1a reports the settings of the parameters used for the dataset of L'Aquila.

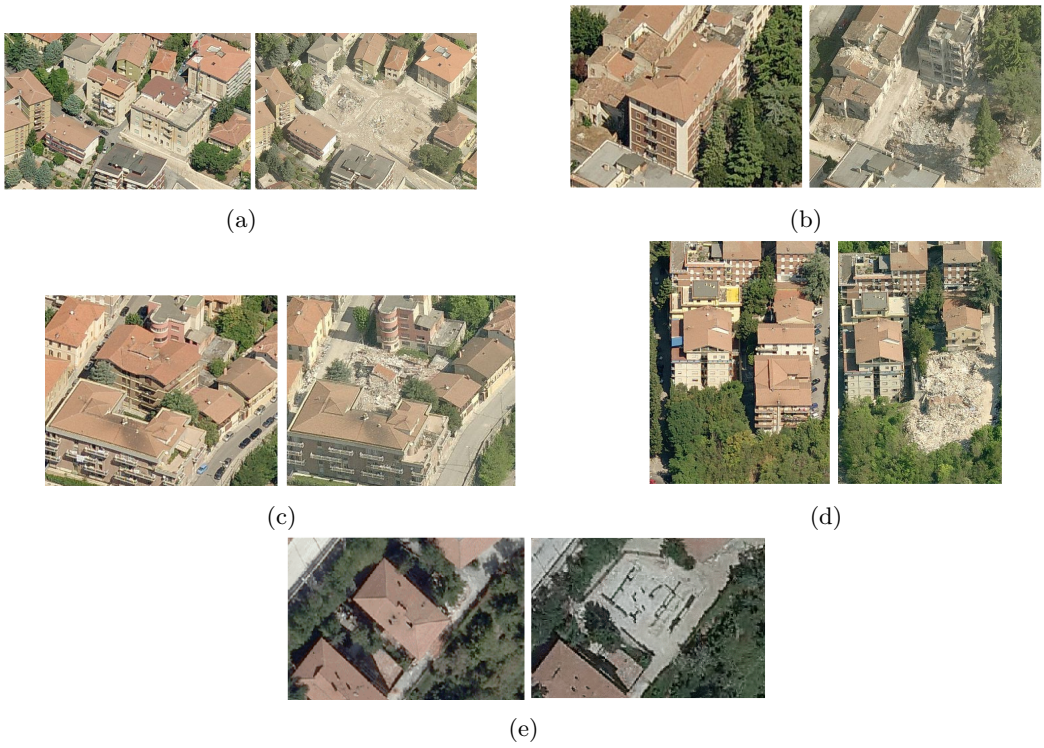


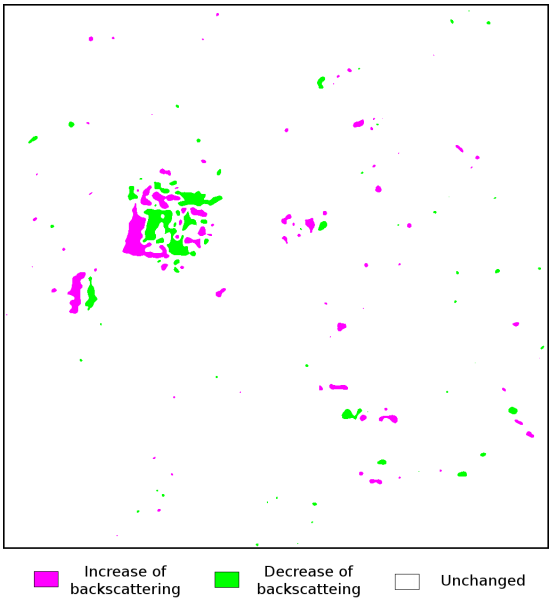
Figure 3.11: (a)-(d) Zoom of the 6 buildings destroyed and correctly detected by the proposed method [97]. See Fig. 4.6f for the correspondence of the buildings on the map. (e) Zoom of the orthophotos on the building wrongly detected as destroyed. As one can notice building was totally demolished after the earthquake.

By comparing the changed building map with the false color composition of the multitemporal images (Fig. 4.6e) one can notice that the proposed approach can effectively detect the backscattering changes associated to disappeared buildings as described in Section 3.2.2, which patterns present a decrease and increase of backscattering oriented near-to-far-range (see Fig. 4.6e). From a quantitative point of view, Table 4.2 reports the number of total buildings present in the scene and the total number of demolished radar building footprints correctly detected and missed by the proposed approach. Even though the dataset presents a relevant number of changes in the backscattering value, none of the actual demolished buildings (see Figg. 4.7a-d) was miss-detected and only one building was wrongly identified as destroyed (see Fig. 4.7.e). As a matter of fact, from the orthophotos and the optical images [85] acquired after the CSK acquisition of September 2009, it was observed that this building was seriously damaged from the earthquake and

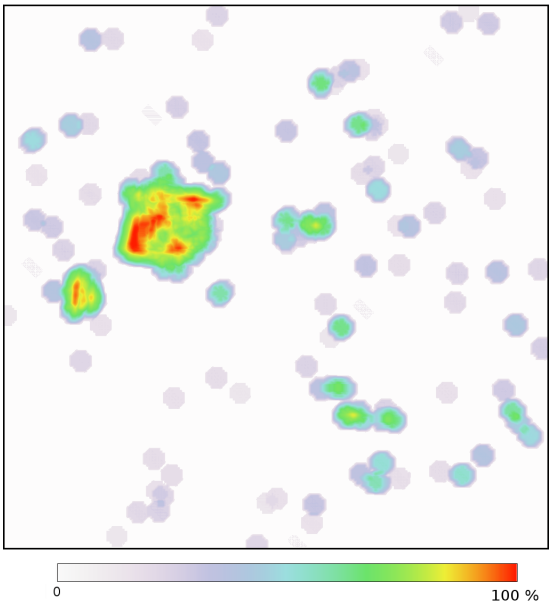
(a)


Increase of backscattering
Decrease of backscatteing
Unchanged

(b)


Increase of backscattering
Decrease of backscatteing
Unchanged

(c)


0
100 %

(d)

(e)

Figure 3.12: Trento (Italy) data set: (a) optical image (2011 [85]); and (b) RGB multitemporal composition of spotlight TerraSAR-X and TanDEM-X images (R:03/04/2013, G:21/01/2011, B:03/04/2013). (c) Change detection map $\mathbf{M}^{\text{opt}}$ obtained at the level $N-1=3$. (d) Map showing the index of the size of changes $\mathbf{C}$. (e) Changed building map overlapped to the RGB multitemporal composition of CSK data.

demolished by the authorities after the acquisition. Fig. 4.6f shows the position of the destroyed buildings in the cadestrial map obtained by transforming the geometry of SAR images to the geo-referenced cadastral map. The analysis confirms the high accuracy of the proposed technique. It is worth noting that the detected demolished buildings has different aspect angle ϕ. Moreover, even though the density of buildings in the considered site is high, the proposed approach can effectively discriminate between demolished and standing buildings. This points out the high robustness of the proposed technique to the possible deviations from the adopted ideal model of changes. A similar analysis has been performed for changes that are not associated with buildings, i.e., ω_{c3} (yellow areas in Fig. 4.6d). The orthophotos point out that these changes are not associated to fully destroyed buildings, but to the post earthquake activities e.g., emergency housing units or to partially damaged buildings. This confirms the effectiveness of the proposed method also in the detection of general changes.

3.4.2 Trento Data Set: Detection of New Buildings

The second data set is made up of two high resolution spotlight single look slant range complex images acquired by TerraSAR-X and TanDEM-X with resolution 0.58 m in range by 1.1 m in azimuth (0.454 by 0.855 meters of pixel spacing) in X-band HH-polarization. The images have been taken on January 21, 2011 and April 3, 2013 over the city of Trento (46°04′N 11°07′E), with a perpendicular baseline of 489.29 m in ascending orbit with 53° incidence angle. The selected test site is a section (1024×1024 pixels) of the full scene

where the department of engineering and computer science of the University of Trento and the Fondazione Bruno Kessler (FBK) are represented. Fig. 3.12a shows the optical image corresponding to the same area [85]. Fig. 3.12b shows a false color composition of the two images (R:03/04/2013, G:21/01/2011, B:03/04/2013) in which pixels with an increase in the value of backscattering appear in magenta tone, pixels with a decrease in the value of backscattering appear in green tone and unchanged pixel in gray-scale. From a ground field survey it is know that three new buildings were built up in the considered site. Moreover, several buildings were subject to restoration works. As one can notice, in this data-set the size of the changed buildings is not homogeneous: we have a large building (still partially under construction during the second acquisition) in the center left of the image, a medium size building in the left part of the image and a small building in the center of the image. Therefore, we aim to test the robustness of the proposed approach to this heterogeneous condition.

Table 3.3: Performance of the proposed approach for Trento Dataset in terms of number of detected new buildings and missed new buildings. The total number of buildings present in the site is also reported in the table.

Total Number of Buildings	Number of New Buildings (reference)	Detected New Buildings (ω_{c2})	Missed New Buildings
187	3	3	0

In order to apply the proposed method the two VHR SAR images were calibrated (radar brightness) and coregistered (with sub-pixel accuracy) and the log-ratio image $\mathbf{X}_{\mathrm{LR}}$ calculated. In the considered city, it is expected that the smallest radar footprint of a building is about 20 by 10 meters. Considering the proposed multiscale approach the resolution of $\mathbf{X}_{\mathrm{LR}}^{\mathrm{opt}}$ is approximately given by $1.1 \cdot 2^{(N-1)}$ in the worst case i.e., in azimuth. Therefore the level $(N-1)^{\mathrm{th}} = 3^{\mathrm{rd}}$, which gives a resolution of 8.8 m is suitable for the purpose. The automatic split-based thresholding approach was applied to $\mathbf{X}_{\mathrm{LR}}^{3}$ in order to derived $\mathbf{M}^{\mathrm{opt}}$. The split size was calculated starting from the average building size, which was estimated to be $w_1 \times w_2 \times h = 30 \times 80 \times 13$ m^3. By substituting these values in eq. (3.2) S_R was obtained to be equal to 45.56 m in slant-range geometry. By taking into account the pixel spacing in range of 0.454 m S_R results to be equal to about 100 pixels. By applying eq. (3.3) and considering a pixel spacing in azimuth of 0.855 m, S_A was estimated to be equal to about 90 pixels. Fig. 3.12c shows the CD map $\mathbf{M}^{\mathrm{opt}}$ generated with the split-based thresholding method from $\mathbf{X}_{\mathrm{LR}}^{3}$. As done for the dataset of L'Aquila a sensitivity analysis was performed. By ranging S_R and S_A from 60 to 120 the error on the estimation of the thresholds is small. That is of 0.058 in average with a peak of 0.14 for $\hat{t}^{-}$, and of 0.041 in average with a peak of 0.1 for $\hat{t}^{+}$. In all the cases the generated maps did not present any critical behavior and resulted in a very similar final detection of changed buildings. The index image $\mathbf{C}$ (Fig. 3.12d) was derived by applying the set of filters described by (3.4) to $\mathbf{M}^{\mathrm{opt}}$ with size $z_1 \times z_2 = 20 \times 10$ pixels selected according to the expected minimum radar building footprint size. $\mathbf{C}$ was thresholded and 25 candidates were detected. They were analyzed in order to discriminate the appropriate class of change by testing the proposed fuzzy rules. Table 4.1b reports the settings of the parameters used for the dataset of Trento.

Figure 3.13: New large building in the Trento dataset: some month before the 2011 TerraSAR-X acquisition (on the left) and on the same day of the second acquisition (on the right). As one can notice the building is still under construction during the second acquisition with several materials and vehicles present in the building site.

Table 3.4: Parameters used in the experiments carried out on: a) L'Aquila dataset; and b) Trento dataset.

Parameter	Value
$N-1$	3
S_R	120 pixels
S_A	40 pixels
z_1	40 pixels
z_2	20 pixels
T_C	160 pixels
a_1, b_1	10, 0.3
a_2, b_2	10, 0.5
a_3, b_3	-10, $\frac{\pi}{3}$
T_{M}	> 0.125

(a)

Parameter	Value
$N-1$	3
S_R	100 pixels
S_A	90 pixels
z_1	20 pixels
z_2	10 pixels
T_C	40 pixels
a_1, b_1	10, 0.3
a_2, b_2	10, 0.5
a_3, b_3	-10, $\frac{\pi}{3}$
T_{M}	> 0.125

(b)

As one can observe from the changed building map (Fig. 3.12e) the proposed approach allows the detection of the building footprint of the three new buildings. The aggregated membership values for all buildings is always grater than 0.74. In detail, for the small and the medium size buildings the reconstructed building footprints result to be accurate. The footprint for the large building in the center left of the image was reconstructed considering the several regions of increase and decrease in backscattering that arise because of both the complexity of the structure of the building (i.e., a inner court and three front wings) and the construction works, which were still in progress during the second TerraSAR-X acquisition (see Fig. 3.13a). As consequence the reconstructed footprint results to be divided in different parts. Nevertheless, the building footprint is completely detected even if in this case we are far from the ideal reference condition of the considered model.

Table 3.3 reports the total number of buildings present in the scene and the number of new buildings detected and missed by the proposed approach. Even though the size of the new buildings present in the scene is not homogeneous, none of the actual new buildings were missed from the detection and only one building was wrongly identified

as destroyed with a relatively low membership of 0.39. As a matter of fact, from the ground field survey it is known that this building was completely renovated between the two acquisitions. Thus the building footprint really changed in the considered time interval. As for L'Aquila dataset, the proposed approach can effectively discriminate between demolished and standing buildings despite the high density of buildings in the area. A further analysis has been performed on changes that are not associated to fully destroyed or new buildings to buildings i.e., ω_{c3} (yellow areas in Fig. 3.12e), which result to be buildings partially renovated. This confirms the effectiveness of the proposed method in the detection of changes in buildings.

3.5 Conclusion

In this chapter an approach to building CD in multitemporal VHR SAR images has been proposed that detects changes by distinguishing between new and demolished buildings. The approach is based on two concepts: i) the extraction of information on changes associated with increase and decrease of backscattering at the optimal building scale; and ii) the exploitation of the expected backscattering properties of buildings to detect either new and fully demolished buildings with their grade of reliability.

The method takes advantage of the theoretical modeling of the backscattering mechanisms in multitemporal VHR SAR images. From this modeling new and destroyed buildings can be identified by a pattern made up of an area of both increase and decrease of backscattering with specific spatial properties and a specific alignment. In order to extract the changes associated with increase and decrease of backscattering, the proposed approach makes use of a multiscale representation of the multitemporal information. This allows a detection of changes at the optimal building scale. This means that the changes smaller than the selected building scale are rejected while the changes related to building size are made homogeneous. This information is used to identify the candidates to be changed buildings. The building candidates are analyzed in order to properly detect the new or destroyed building by means of four fuzzy rules. The fuzzy rules are formulated by taking into account the ideal backscattering mechanisms that arise when a building appears or disappears between two acquisitions. The aggregated membership resulting from the application of the fuzzy rules makes it possible to identify the class of each building candidate (i.e., new/destroyed building or general change with size comparable to the building size but not related to new or destroyed building). Moreover, the membership value gives an indication of the reliability of the detected class. The proposed approach requires the tuning of some parameters that depend on the approximated size of the buildings in the investigated area. This is a prior information usually available and easy to include in the processing. It is worth remarking that after this tuning the method is completely automatic.

The proposed approach was tested on two VHR SAR dataset acquired by two different X-band systems i.e., COSMOSky-Med and TerraSAR-X/Tandem-X. The first dataset is related to the earthquake of L'Aquila (Italy) in 2009. Here the changes are mainly associated to demolished buildings. The second data set is related to a condition of urban expansion with the construction of new buildings in the city of Trento (Italy). In

both cases the proposed method demonstrated to be effective in detecting both new and demolished buildings with a high accuracy. In detail, only two buildings were misclassified over 387 present in the two datasets with no missed detection. This was possible also because of the high robustness of the proposed method to deviations from the ideal model considered for changes. This robustness is due to both the use of adequate fuzzy rules and the intrinsic reliability obtained when working on the comparison of two images rather than on the modeling of building on each image.

As future development of this work, we plan to improve the building detector by better modeling the geometrical behaviors of building primitives. Moreover, we plan to investigate the possibility to discriminate among several building construction stages and/or building damage levels.

Acknowledgment

The authors would like to thank the Italian Space Agency (ASI) for providing the CSK images of L'Aquila within the agreement no. I/041/09/0, project ID-2181. The authors would also like to thank the German Aerospace Center (DLR) for providing the TerraSAR-X images of Trento which were acquired in the framework of the ISPRS WG VII/2 and TerraSAR Project LAN0634.

Chapter 4

Rapid and accurate damage detection in built-up areas using SAR images

In the last years Earth Observation (EO) has been proven to be a reliable tool to support the operations after crisis events such as earthquakes and floods. In this context, the use of SAR systems is very attractive from an operational point of view since, differently from optical sensors, SAR is independent from the sun illumination and it is relatively insensitive to atmospheric weather conditions. This makes possible to ensure data availability during crisis events. Since 2008 a new generation of SAR missions has been daily acquiring and archiving images of the Earth surface with a short repetition interval and a high geometric resolution. This allows the analysis of damages caused from earthquakes on urban areas in an effective way. SAR systems on-board of these missions are able to acquire images with variable geometrical resolution and swath. By increasing the resolution the maximum ground coverage guaranteed (swath) decreases accordingly. In this chapter[1] *we propose to exploit and combine the acquisition modes offered by commercial SAR sensors in order to: i) quickly and automatically identify the areas severely affected by the earthquake (i.e., hot-spots) by analyzing images of large areas using data characterized by a high geometrical resolution (3-5 m); and ii) accurately detect collapsed buildings inside each hot-spot by acquiring and analyzing images of small areas using data with very high geometrical resolution (1 m). Experimental results obtained on COSMO-SkyMed multitemporal images acquired in 2009 on the city of L'Aquila (Italy) before and after the earthquake that hit the region demonstrate the effectiveness of the proposed approach.*

4.1 Introduction

Earthquakes have struck cities around the globe causing injury, loss of life and damage to human settlement since the ancient times. In the last years large earthquakes have

[1] Part of this chapter appears in:

[22] L. Bruzzone, C. Marin, F. Bovolo, "Damage Detection in Built-up Areas Using SAR Images." in Beer M., Patelli E., Kougioumtzoglou I., Au I. (Ed.) Encyclopedia of Earthquake Engineering: SpringerReference (www.springerreference.com). Springer-Verlag Berlin Heidelberg, 2013.

been reported globally affecting wide geographical areas and a relevant number of people. A remarkable example was the 2011 Tohoku earthquake with moment magnitude scale (MMS) of 9.0. The earthquake caused 15,884 deaths, 6,147 injured, and 2,636 people missing across twenty prefectures, as well as 129,316 buildings totally collapsed, 263,845 buildings half collapsed, and 725,760 buildings partially damaged [98]. The earthquake triggered tsunami waves that caused the nuclear accidents at three reactors in the Fukushima Daiichi Nuclear Power Plant complex, and the associated evacuation zones affecting hundreds of thousands of residents [98]. Moreover, the earthquake and tsunami also caused extensive and severe structural damage in north-eastern Japan, including heavy damage to roads and railways as well as fires in many areas, and a dam collapse. The World Bank has estimated economic damage of US 210 billion, making it the costliest natural disaster in the world history [98]. Another relevant example is the 2009 L'Aquila earthquake (Italy) with 6.3 MMS. The earthquake caused damages to a relevant number of medieval buildings in the city of L'Aquila: several buildings collapsed, about 300 people died and about 1500 people were injured making this event the deadliest earthquake to hit Italy since 1980. The list of earthquakes and their damages occurred globally can be continued. Indeed, according to the United Stated Geological Survey [99] records, the frequency of such catastrophic events is high: 332 earthquake with a MMS 7 or higher have been reported in the last 20 years.

In order to understand and possibly mitigate the impact of such events on the citizenry, several governmental initiatives and research activities have been carried out in the past (e.g., the international charter on space and major disasters [100], UNOSAT [101], or the geohazard supersites and natural laboratories [102]) to tackle each of the phases of such events. In detail a catastrophic event can be partitioned in three characteristic phases: 1) before the event in which a system of risk assessment for early warning is essential; 2) the moment the event occurs in which a disaster-alerting systems is the main priority; and 3) after the event in which a emergency response with impact assessment has to be initiate.

Focusing on the third phase, among all the possible damages caused by natural disasters, the building damages contribute mostly to the casualties [7]. Nonetheless, the identification of struck built-up areas may be particularly complex because of both the possible large extension of the affected region and the difficulties to access remote villages. In this scenario Earth Observation (EO) data acquired by remote-sensing satellites represent a useful tool to support decisions of civil protection authorities after the event. Indeed, satellites equipped with optical or synthetic aperture radar (SAR) sensors can monitor large regions in a quick, synoptic and uncensored manner. In particular, SAR systems presents the advantage over optical sensors to be independent from sun illumination and to be relatively insensitive to atmospheric weather conditions. This ensures the data availability during the crisis event even though the site affected by the crisis event is covered by clouds or if the satellite is passing during night.

As we discussed in the introduction chapter of this thesis, since 2008 a new generation of SAR missions (i.e., TerraSAR-X, TanDEM-X and COSMO-SkyMed) have regularly acquired data over Earth. These systems can perform acquisitions using two high resolution modes called StripMap (SM) and SpotLight (SL). SM is characterized by a wide coverage (around 1500 Km^2) with a high resolution (3 to 5 m), whereas SL is characterized by a

moderate coverage (around 100 Km^2) with a very high resolution (1 m). This means that the two SAR modes are complementary to each other and can be used for performing different duties during the emergency response phase. SM mode can be used for identifying the blocks of the city more hit by the earthquake whereas the SL images can be used to identify detailed information about collapsed buildings. It is worth noting that, these SAR missions are daily populating a huge archive of multitemporal. This archive can be employed to detect changes in the state of an object or phenomenon by jointly observing the data acquired at different times over the same geographical area. Nonetheless, SAR images with a resolution ranging from the 5 m (referred in this work as high resolution images) of the SM images up to 1 m (referred in this work as very high resolution images) of the SL images are more heterogeneous than old generation sensor images, which have resolution generally ranging from 10 to 30m [60, 61]. In high resolution (HR) and very high resolution (VHR) images objects that are considered homogeneous from a semantic point of view (e.g., buildings) show a signature that is inhomogeneous at high spatial resolution because of the scattering contributions from sub-objects (e.g., facade and roof in a building). Furthermore, on the one hand the side-looking illumination required by SAR systems leads to phenomena such as layover, shadow and multi-path signals [62,63], which are very marked in urban areas. On the other hand, the appearance of a ground object depends on radar system parameters (i.e., wavelength, polarization, pulse length, incidence angle, look direction, etc.), surface feature proprieties (e.g., dielectric constant) and environmental variables (e.g., ground water content) [64]. On top of these aspects, SAR images are corrupted by speckle noise.

All these factors, within the context of the multitemporal analysis, make the problem of the detection of changes complex because of the issues presented in chapter 1. Indeed they invalidate the assumption considered for medium-resolution SAR images that two images acquired on the same geographical area at different times are similar to each other except for the presence of changes occurred on the ground. Thus the use of standard pixel and context-based techniques would fail to solve the problem since they are not able to properly take into account the high geometrical detail.

Numerous studies have been recently presented in the literature that deal with the problem of recognition of changes in urban areas using HR and VHR SAR images [6, 31, 66–76]. In detail, [6,31,66–73] are focused on the detection of earthquake damages, [74] is focused on the building databases updating, and [75] addresses the detection of changes due to the urban evolution. The strategies that are exploited in these works include post-event supervised analysis [67, 68, 71–73], joint use of optical and SAR data [6, 66], joint use of LiDAR and SAR data [75], unsupervised detection of damages at the level of aggregated blocks [31, 69] or GIS polygons [70]. Despite the great interest, the only work that addresses the problem of building change detection using multitemporal VHR SAR data in an unsupervised way is presented in [76].

In this work we present an approach that exploits the characteristics of the two complementary acquisition modes (i.e., stripmap and spotlight) to: i) quickly and automatically identifies the areas severely affected by a catastrophic event (i.e., hot-spots), such as an earthquake by analyzing images characterized by a large coverage and a medium to high geometrical resolution; and ii) analyzes images characterized by very high geometrical res-

olution acquired over hot-spots in order to detect collapsed buildings. The effectiveness of the proposed approach are demonstrated in experimental results obtained on COSMO-SkyMed multitemporal images acquired in 2009 on the city of L'Aquila (Italy) before and after the earthquake that hit the region.

This chapter is organized into five sections. Section 4.2 reviews the main operational modes of SAR systems. The proposed approach to building change detection is described in Section 4.3. Section 4.4 presents the data set and shows the experimental results. Section 4.5 draws the conclusions of the work.

4.2 Synthetic Aperture Radar modes

As we discussed in the introduction of this thesis SAR is an active radar system operating in the microwave region of the electromagnetic spectrum generally between P-band (i.e., 0.1-0.39 GHz) and Ka-band (i.e., 26-40 GHz). SAR systems illuminate the scene using a side-looking geometry. The antenna of the radar system is mounted on a flying platform. Its horizontal and vertical axes are parallel and orthogonal to the azimuth direction, respectively. The angle between nadir and the radar beam direction is called incidence angle and it is usually denoted by θ. Such a system illuminates the Earth surface with microwave pulses and receives the electromagnetic signal back-scattered from the illuminated scene. By means of signal processing techniques that exploit the Doppler shifts of the received electromagnetic echoes, SAR systems are able to synthesize a two-dimensional high spatial resolution image from all the received signals.

Regarding the plane perpendicular to the flight direction, two reference systems are usually used to define the position of a point. These are the slant range and the ground range. The slant range is the direction identified by the conjunction between a point target and the SAR system. The ground range is the projection of the slant range on the ground (corresponding to the across-track direction of the RS case). Therefore, the ground range resolution δ_{gr} depends on the incidence angle and on the surface topography.

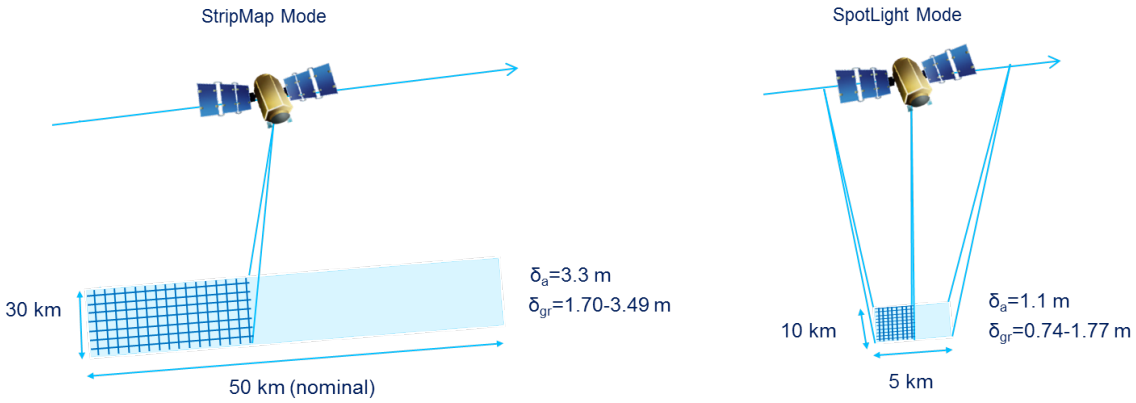


Figure 4.1: SAR modes. The reported parameters are refereed to TerraSAR-X imaging modes.

Several sensor operation modes for acquiring SAR data were developed in the past. The

most common modes which are implemented in space-borne SAR missions are StripMap (SM) and SpotLight (SL). In the SM (see Fig. 4.1 left side) the radar antenna points along a fixed direction with respect to the flight path. As the platform moves the resulting image covers a strip of the ground surface. The extension of the imaged area in azimuth direction is limited only by the power and free memory available on-board of the satellite, whereas the resolution in range direction is constrained by half of the actual antenna length. Generally speaking for the new SAR generation the standard coverage is of about 1500 Km^2 with a resolution that varies between 1.5 to 5 m in ground range and about 3 m in azimuth depending on the platform and acquisition parameters. In the SL mode (see Fig. 4.1 right side) the azimuth resolution is improved by increasing the length of the synthetic antenna. This is done by steering the antenna beam so that to illuminate the same area on the ground during the acquisition. Nonetheless, the improvement in the azimuth resolution is at the expense of the spatial coverage. The standard coverage is in fact of about 100 Km^2 (instead of 1500 Km^2) with a resolution that varies between 1 to 3.5 m (instead of 1.5 to 5 m) in ground range and 1 m (instead of 3 m) in azimuth depending on the platform.

4.3 Proposed Approach to Damage Detection

Let us assume to analyze two HR amplitude SAR images $\mathbf{I}_1$ and $\mathbf{I}_2$ acquired in Strip-Map mode before (pre-event) and after the event (post-event) respectively, which have the same incidence angle and centered over the epicenter of the event. The goal is to detect the areas severely damaged by the event (i.e., hot-spots) in order to quickly organize the rescue operations. The most critical issue for the identification of the damages is related to the presence of many possible kinds of changes on the ground. A single site may include many kinds of change that show significantly different characteristics in terms of size, shape and semantic meaning. Note that not all the changes are associated to the built-up site and thus they are not relevant for the rescue operation. The analysis of SM image allows the detection of large portions of changed areas. Here we refer them as hot-spots. The obtained hot-spots are then exploited to drive the selection of a pre-crisis VHR SAR $\mathbf{X}_1^1, \ldots, \mathbf{X}_1^Q$ and the acquisition of post-crisis VHR SAR images $\mathbf{X}_2^1, \ldots, \mathbf{X}_2^Q$ taken with the same incidence angle. Note that the acquisition of the post crisis images $\mathbf{X}_2^1, \ldots, \mathbf{X}_2^Q$ should be scheduled on the basis of the detected hot-spots. These images are used to perform change detection at high spatial resolution with the goal to detect collapsed buildings with a high level of detail.

Hence, the proposed approach is made up of two main phases: i) a coarse detection of the areas of change (hot-spots); and ii) a fine detection of the fully collapsed buildings inside each hot-spot. Fig. 4.2 shows the block schemes of the two phases of the proposed approach. Each phase of the method is explained in detail in the next subsections.

4.3.1 Detection of Damages (Hot-Spots)

The first phase of the proposed approach is devoted to rapidly analyze a large scene and detect the changes due to the crisis event (see Fig. 4.3). This is done by comparing pre- and

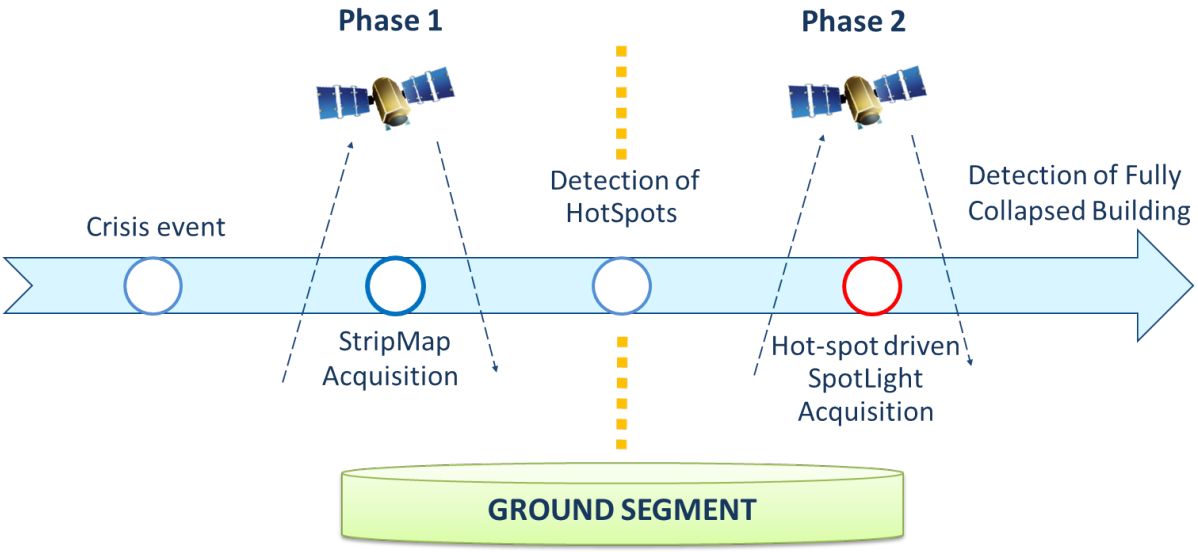


Figure 4.2: Temporal scenario for a rapid and accurate damage detection.

post-event SAR images $\mathbf{I}_1$ and $\mathbf{I}_2$ acquired in SM mode and by identifying the areas that present either a significant increase or decrease in the backscattering value as described in chapter 2.2.2. In detail, this is done by the log operator i.e., $\mathbf{I}_{\mathrm{LR}} = \log \frac{\mathbf{I}_2}{\mathbf{I}_1} = \log \mathbf{I}_2 - \log \mathbf{I}_1$. Given the nature of the HR images, the application of this strategy will result in a high number of false alarms. Several works have been presented in the literature in order to overcome this problem and identify the changes due to crisis events [7, 25, 69]. In this chapter a pixel-by-pixel comparison followed by a Wavelet de-noising procedure is used in order to mitigate the effect of speckle [39]. In detail, a two-dimensional discrete stationary Wavelet transform (2D-SWT) is applied recursively $N+1$ times to the log-ratio image. The de-noised image $\mathbf{I}_{\mathrm{LR}}^{\mathrm{fil}}$ is obtained by applying the inverse two-dimensional discrete stationary Wavelet transform (2D-ISWT) considering only the $N+1$ sub-bands of approximation. From $\mathbf{I}_{\mathrm{LR}}^{\mathrm{fil}}$ it is possible to isolate the areas affected by a significant backscattering change and that present an extension equal or larger than the resolution of the selected Wavelet level $N+1$. Therefore, in order to detect the hot-spots the selected level has to be comparable to the scale of aggregated buildings or city blocks size.

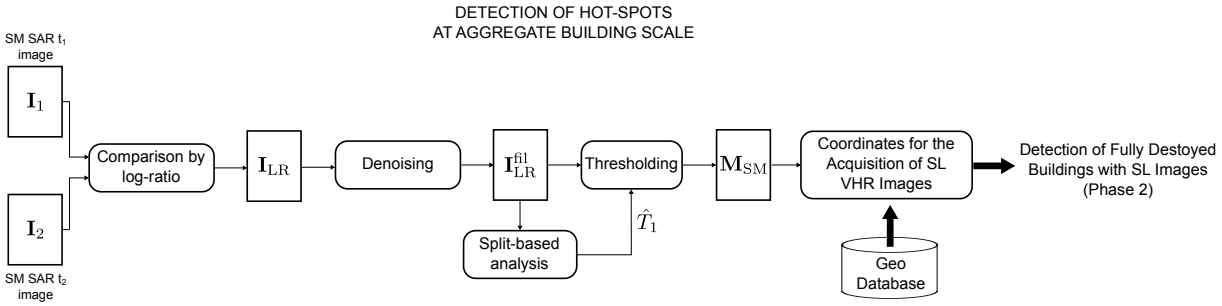


Figure 4.3: Logical flow for the detection of hot-spots using SM images (first phase).

As the analysis of SAR images acquired in SM mode involves the processing of large multitemporal images it is possible that the population of changed pixels is in sharp minority when considering $\mathbf{I}_{\mathrm{LR}}^{\mathrm{fil}}$. This may affect the accuracy of the adopted automatic threshold selection technique. In order to address this issue, we use the split-based thresholding method presented in chapter 2.2.3. It is worth noting that the obtained change detection map $\mathbf{M}_{\mathrm{SM}}$ does not report any information about the source of the changes. Indeed the changes may be associated to the crisis event as well to other source of change such as phenological changes. By exploiting prior information (e.g., cadastral map) and focusing on changes related to build up areas, the hot-spots possibly related to the crisis event that are in urban areas can be isolated. Nevertheless, among these changes some may be more critical than others. The severity of the change can be inferred considering the prior information available to civil protection or exploiting the intensity and texture information of the SAR images as proposed in [7, 69]. Once the most relevant hot-spots are detected, an algorithm for solving the "packing problem" of acquiring the minimum number of SL acquisitions $\mathbf{X}_2^1, \ldots, \mathbf{X}_2^Q$ by covering the highest number of critical hot-spots can be applied. Under the assumption the pre-event $\mathbf{X}_1^1, \ldots, \mathbf{X}_1^Q$ images have been already acquired, when a crisis occurs the acquisition of $\mathbf{X}_2^1, \ldots, \mathbf{X}_2^Q$ can be effectively planned (if not already done) by the outcome of this phase.

4.3.2 Detection of Fully Collapsed Buildings

The aim of the second phase is to detect collapsed buildings considering $\mathbf{X}_1^1, \ldots, \mathbf{X}_1^Q$ and $\mathbf{X}_2^1, \ldots, \mathbf{X}_2^Q$ identified in the previous phase. This is done by extending the method presented in chapter 3, which takes explicitly into account the backscattering mechanism of changed buildings in SAR images. In detail, we set the proposed approach for this specif task in order to produce a map made up of 3 classes: i) fully destroyed buildings (ω_{c1}); ii) changes that have a size comparable to a building but do not present the typical pattern of full new/demolished buildings (ω_{c2}); iii) all the other changes that do not show a size comparable to a building and are therefore not related to changed buildings (ω_{c3}). For the sake of notation, and without any loss of validity in the following we assume that $Q = 1$ and thus that critical hot-spots can be fully covered by means of a single couple of images $\mathbf{X_1}$ and $\mathbf{X_2}$ acquired with the same incidence angle[2] on the same geographical area at different times t_1 and t_2, respectively.

As depicted in Fig. 4.4 the proposed method firstly detect the backscattering changes at the level of building i.e., the level that better represents buildings and not sub-parts of them. Differently from our original work, this representation is derived according to a Curvelet transform [103]. Curvelet is a generalization of the concept of Wavelet that allows the representation of an image at different scales and different angles. Curvelets are an appropriate basis for representing images which are smooth apart from singularities along smooth curves. The Curvelet transform is applied to the log-ratio image $\mathbf{X}_{\mathrm{LR}}$ derived from $\mathbf{X_1}$ and $\mathbf{X_2}$ and the coefficients are selected according to a hard thresholding rule [103]. In this way Curvelets results sensitive to structural changes due to anthropogenic changes or damages in human made structures. The residual Curvelet coefficients are used to

[2]It is worth noting that the incidence angle can be different from the incidence angle of SM images.

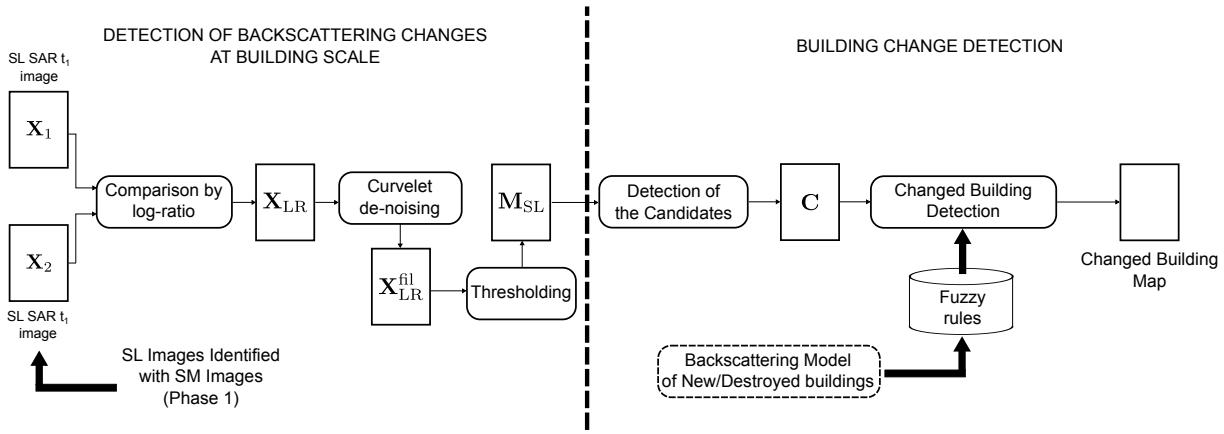


Figure 4.4: Logical flow for the detection of fully destroyed buildings using SL images (second phase).

reconstruct the filtered version of the log-ratio image by means of an inverse Curvelet transform. According to the proposed method, is used to extract the areas of changes related to increase or decrease in the value of backscattering. This is done by applying a thresholding procedure as describe in chapter 2.2.3. The obtained map MSL is employed to identify the candidate to be changed building. A changed building candidate is an area of change that has size equal or larger than the minimum size of the building present in the investigated site. The candidates are derived by counting the changed pixels inside a set of moving windows, with size equal to the minimum size of the buildings present in the scene and covering the most of the possible building orientations.

Once the candidates are identified the proposed method associates the proper class of change ω_{c1}, ω_{c2} and ω_{c3} to each candidate. This is done by evaluating five fuzzy rules. These rules have been derived by considering the expected backscattering mechanisms of changed building. Each rule returns a grade of membership of the candidate to one specific expected propriety related to the physical characteristics and the spatial arrangement of the candidate with respect to the model of destroyed building. The rules aim at verifying: i) the presence of both increase and decrease of backscattering contributions inside the candidate; ii) the grade of membership of the building candidate area, in terms of number of pixels, covered by the regions of increase and decrease of backscattering; iii) the orientation of the pattern increase-decrease; and iv) the alignment of the region of increase with respect to the region of decrease in backscattering; v) the overall size of the radar footprint w.r.t. the size of the areas of increase and decrease of backscattering. If the aggregate membership of the five fuzzy rules exceed a certain threshold TM, the estimated convex-hull of each candidate is reported in the final change detection map as result of the detection. We refer the reader to chapter 3 for more details on this part of the method.

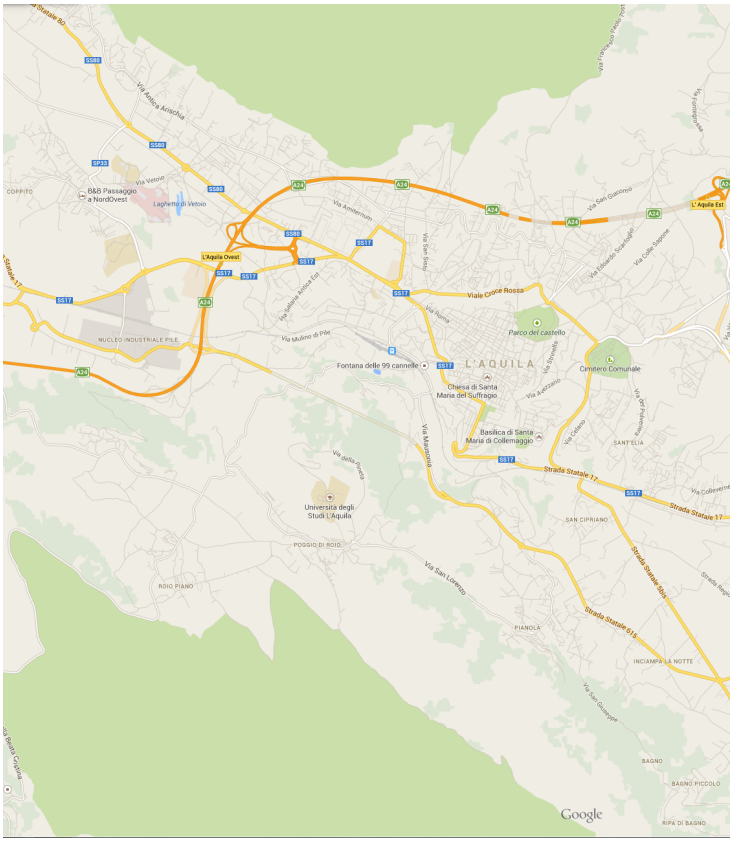
4.4 Data Set Description and Experimental Results

On April 6, 2009 an earthquake of 6.3 Moment Magnitude Scale (MMS) struck central Italy with its epicenter near L'Aquila (42°21′N 13°24′E). The earthquake caused damages

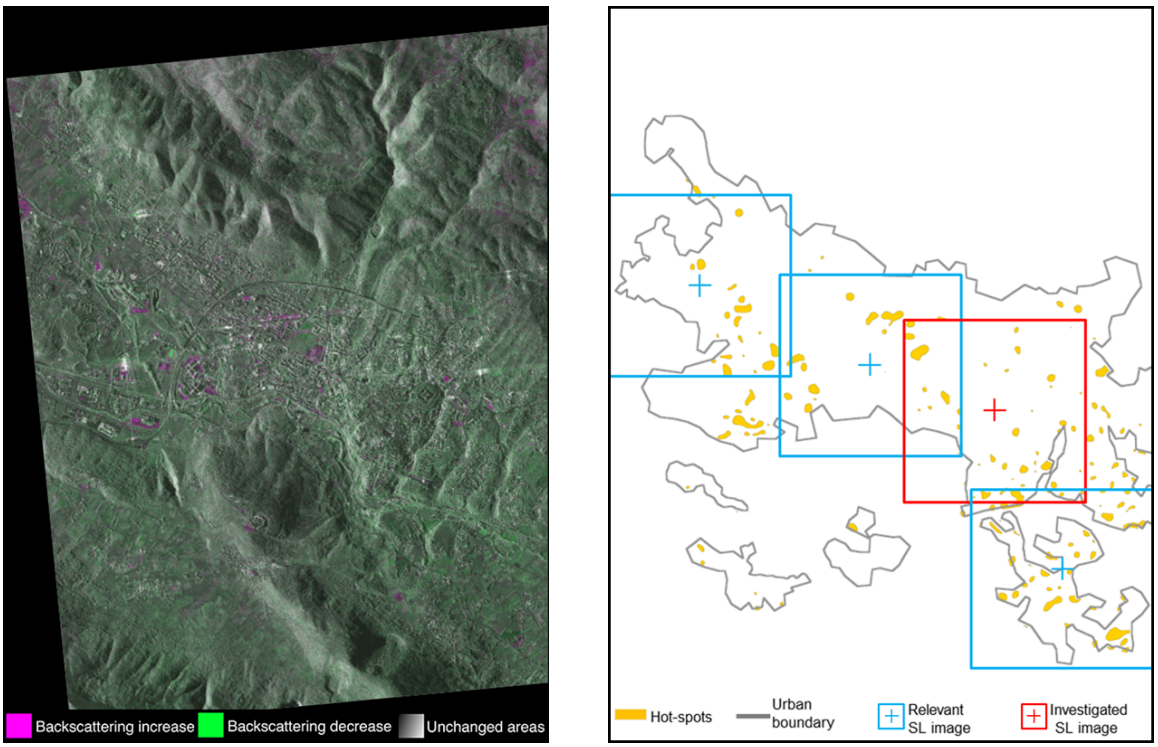
to a relevant number of buildings in the city of L'Aquila: several buildings collapsed and about 1500 people were injured and about 300 people died, making this event the deadliest earthquake to hit Italy since the 1980 Irpinia earthquake.

At the time of the earthquake no SM images were available on the area of the event. Nonetheless, two SL (1-look, 0.5m×0.5m pixel spacing, X-band) amplitude images ($\mathbf{X}_1$ and $\mathbf{X}_2$) acquired in HH-polarization in ascending orbit with 57-58 degree incidence angle and processed according to the standard processing level 1C (Geo-coded ellipsoid corrected) acquired on April 5, 2009 and April 21, 2009, respectively. This is due to the fact that the COSMO-Skymed constellation was operational for few years at the time and that only 2 out of the 4 satellites were fully operational. Therefore, the COSMO-Skymed archive was just started to be populated with images. In order to assess the effectiveness of the proposed method the SL images were used to generate a couple of images with a degraded resolution (pixel spacing of 2.5 m) that can be assimilated to SM images ($\mathbf{I}_1$ and $\mathbf{I}_2$). This was done by using a nearest neighbor interpolation without multilooking operations, and thus obtaining in images with size 5000 × 5000. In the following, we refer to these images as SM images. Moreover, in order to be consistent with the real relationship between the size of SM and SL images, in the following we considered SL images with size 5 times smaller than the obtained SM images i.e., 5000 × 5000.

Fig. 4.5a shows the map of the area covered by the two SM images. Fig. 4.5b shows



(a)



(b) (c)

Figure 4.5: L'Aquila: (a) Map of the area covered by the SM COSMOSky-Med images (Google). (b) RGB multitemporal composition of SM COSMOSky-Med images (red channel:21/04/2009, green channel:04/05/2009, blue:21/04/2009). (c) Hot-spots map $\mathbf{M}_{\mathrm{SM}}$. In red and blue the most damaged regions. The size of these regions corresponds to the coverage guarantee by SL images.

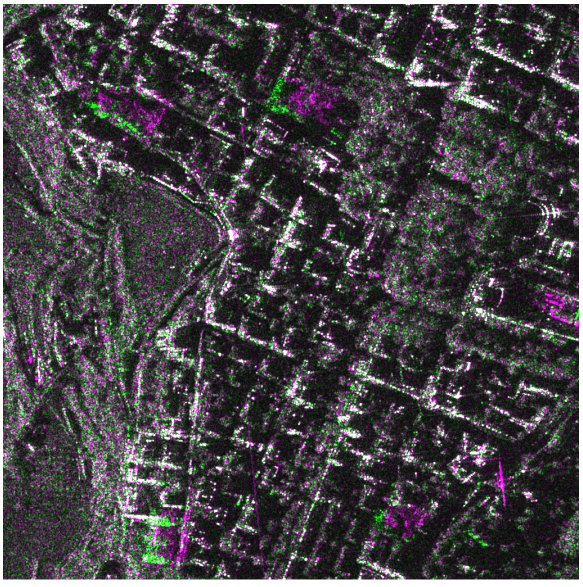
a false color composition of the two SM images (red channel:21/04/2009, green channel:05/04/2009, blue:21/04/2009) in which pixels with an increase in the value of backscattering appears in magenta tone, pixels with a decrease in the value of backscattering appear in green tone and unchanged pixel in gray-scale. It is worth noting that a relevant number of changes are present between the two acquisitions (see Fig. 4.5b). This increases the complexity of the detection of collapsed buildings.

In order to apply the proposed method the two SM images were calibrated and coregistered (with sub-pixel accuracy) and the log-ratio image $\mathbf{I}_{\mathrm{LR}}$ calculated. $\mathbf{I}_{\mathrm{LR}}$ was filtered by applying recursively a 2D-SWT transformation and reconstructing considering only the low-pass information at the 4th level. From $\mathbf{I}_{\mathrm{LR}}^{\mathrm{fil}}$ the hot-spots of changes were derived by applying the split-based thresholding. The size of splits $S_R^{\mathrm{SM}} \times S_A^{\mathrm{SM}} = 64 \times 64$ pixels (corresponding to 160×160 m^2 on the ground) were selected according to average size of expected changes in the test site. The obtained map is reported in Fig. 4.5c. From $\mathbf{M}_{\mathrm{SM}}$ and the cadestrial map $Q = 4$ areas with extension comparable to SL images were selected to be further analyzed in order to detect the presence of collapsed buildings. Among these areas we selected a representative case (red box in Fig. 4.5c) to be used in the following in order to effectively illustrate how the proposed method for the detection of collapsed buildings works. Indeed, the considered area present a relevant number

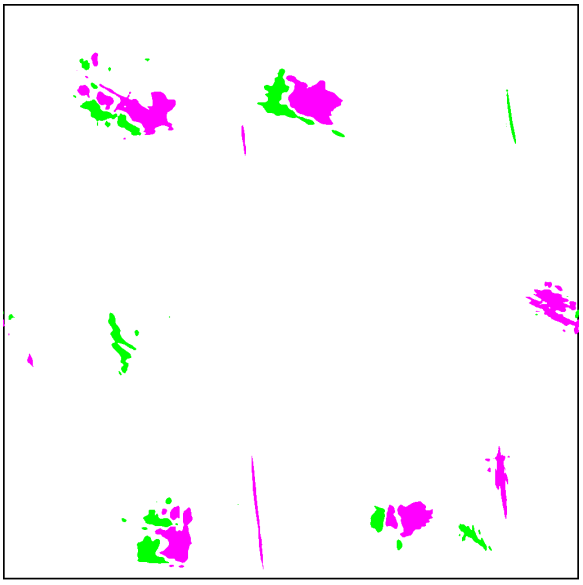
of collapsed buildings, differently from the other areas that present changes due to the construction of refugees camps.

As described in the methodological part, the proposed approach for the identification of collapsed buildings firstly calculates the log-ratio image $\mathbf{X}_{\mathrm{LR}}$. This was computed from the two calibrated and co-registered CSK SL images $\mathbf{X}_1$ and $\mathbf{X}_2$. In detail, the co-registration was performed with sub-pixel accuracy and the radar brightness calibration was applied in order to render the two images comparable. From $\mathbf{X}_{\mathrm{LR}}$ the set $\mathcal{X}_{\mathrm{MS}} =$

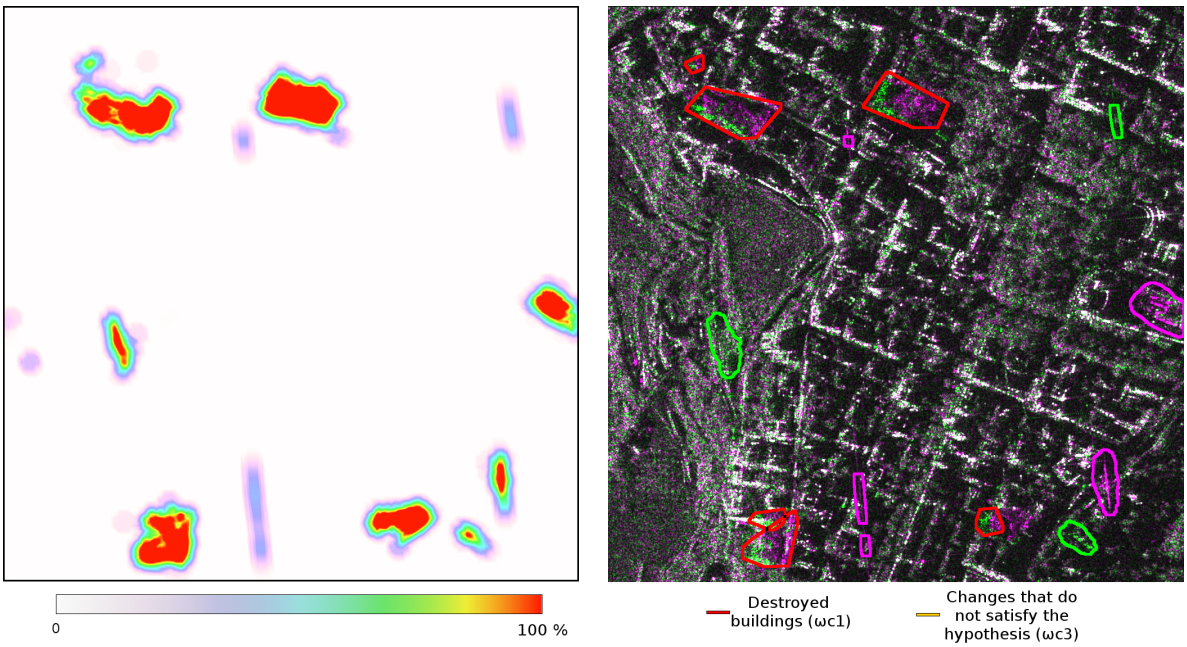
(a)



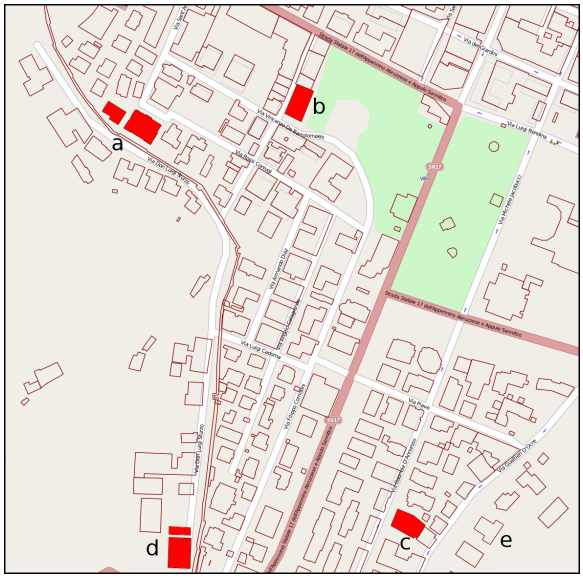
(b)



(c)



(d) (e)



(f)

Figure 4.6: L'Aquila (Italy) data set: (a) orthophoto acquired few days after the earthquake on April 2009 (Geoportale Abruzzo); and (b) RGB multitemporal composition of spotlight COSMOSky-Med images (R:21/04/2009, G:05/04/2009, B:21/04/2009). The images are geo-projected according to the reference ellipsoid. Actual range and the azimuth directions are reported in white arrows. (c) Change detection map $\mathbf{M}_{\mathrm{SL}}$ obtained after the curvelet de-noising. (d) Map showing the index of the size of changes $\mathbf{C}$. (e) Changed building map overlapped to the RGB multitemporal composition of CSK data. (f) Cadastral map of the area (destroyed buildings are reported in red).

$\left\{\mathbf{X}_{\mathrm{LR}}^{0},\ldots,\mathbf{X}_{\mathrm{LR}}^{n},\ldots,\mathbf{X}_{\mathrm{LR}}^{N-1}\right\}$ was computed by applying sequentially the 2D-SWT and the

Table 4.1: Parameters used in the experiments carried out on L'Aquila dataset.

Parameter	Value
$N-1$	3
S_R	120 pixels
S_A	40 pixels
z_1	40 pixels
z_2	20 pixels
T_C	160 pixels
a_1, b_1	10, 0.3
a_2, b_2	10, 0.5
a_3, b_3	-10, $\frac{\pi}{3}$
T_{M}	> 0.125

2D-ISWT with an 8-length *Daubechies* filter. The level $(N-1)^{\text{th}} = 3^{\text{rd}}$, which gives a resolution of approximately 8 meters, was selected as the optimum one. This allows to observe radar footprints of buildings as small as 8 by 8 meters, which is compatible with the minimum size of building footprints in the considered scene. $\mathbf{X}^3_{\mathrm{LR}}$ was thresholded according to the unsupervised split-based procedure described in section 4.3.1. The split size was calculated starting from the average building size, which was estimated to be $w_1 \times w_2 \times h = 25 \times 20 \times 15$ m^3. By substituting these values in eq. (3.2) S_{SR} was obtained to be equal to 49.50 m in slant-range geometry. Since the CSK images are ground-projected, S_{SR} was projected in ground range considering the reference incidence angle i.e., 58 degree obtaining in this way $S_{GR} = 58.36$ m. From the calculation and taking into account a pixel spacing in range of 0.5 m, S_R results to be equal to 120 pixels. By applying eq. (3.3) and considering a pixel spacing in azimuth of 0.5 m, S_A was estimated to be equal to 40 pixels. It is worth noting that the split-based thresholding technique has a relative high tolerance to the selection of the values of S_R and S_A. This was proven by a sensitivity analysis in which S_A was ranged from 20 to 60 pixels and S_R from 80 to 170 pixels. Despite the large range (about $\pm 1/3$ of the optimal selected size) the error on the estimation of the thresholds is small. Moreover, in all the cases the generated maps did not present any critical behavior and resulted in a very similar final detection of changed buildings. The obtained CD map $\mathbf{M}^{\text{opt}}$ highlighting the three classes ξ_u, ξ^+, ξ^- is reported in Fig. 4.6c. The detector described by (3.4) was then applied to $\mathbf{M}^{\text{opt}}$ in order to obtain the image $\mathbf{C}$. With regard to the considered scene, it is expected that the minimum size of building footprint is 40×20 pixels. Therefore, the set of 5 windows reported in Fig. 3.6 were set to size $z_1 \times z_2 = 40 \times 20$ pixels in order to derive $\mathbf{C}$ (see Fig. 4.6d). A set of changed areas $\Gamma = \{\gamma_1, \ldots, \gamma_{38}\}$ compatible with building footprint changes was derived by threhsolding $\mathbf{C}$ with $T_C = 160$ (which corresponds to 20% of maximum value of $\mathbf{C}$) and by applying a flood-fill algorithm. It is worth noting that both the window size and the threshold value were selected according to minimum size of buildings in the considered area in order to limit the missed alarms. Each of the candidate areas was analyzed in order to detect the changed buildings. To this end the four fuzzy rules were automatically

evaluated for each candidate γ_h, $(h = 1, \ldots, 38)$. Finally, for each building candidate that presents an aggregated membership T_{M} greater than 0.125 the radar building footprint was approximated by the convex-hull containing the candidate. It is worth noting that T_{M} was selected in order to minimize the missed alarms. This is done by taking into account the definition of the membership functions and the aggregation strategy (i.e., algebraic product). In the present experiment $\mu_{rs} = \mu_{rl} = \mu_{\alpha} = 0.5$ were selected as limit case, which results in an aggregate membership function $\mathrm{M} = 0.125$. Fig. 4.6e shows the final result. In the specific case, all the reconstructed radar building footprints have an aggregate membership function greater than 0.615. Table 4.1a reports the settings of the parameters used for the dataset of L'Aquila.

Figure 4.7: (a)-(d) Zoom of the 6 buildings destroyed and correctly detected by the proposed method (Bing Maps, Microsoft Corporation). See Fig. 4.6f for the correspondence of the buildings on the map. As one can notice building was totally demolished after the earthquake.

By comparing the changed building map with the false color composition of the multitemporal images (Fig. 4.6e) one can notice that the proposed approach can effectively detect the backscattering changes associated to disappeared buildings as described in Section 3.2.2, which patterns present a decrease and increase in backscattering oriented near-to-far-range (see Fig. 4.6e). From a quantitative point of view, Table 4.2 reports the number of radar building footprints of demolished building detected by the proposed approach. Even though the dataset presents a relevant number of changes in the backscattering value, none of the actual demolished buildings (see Figg. 4.7.a-d) was miss-detected and one of the building was wrongly identified as totally destroyed (see Fig. 4.7.e) even though it was half collapsed on itself.

Fig. 4.6f shows the positions of the destroyed buildings in the cadestrial map obtained by transforming the geometry of SAR images to the geo-referenced cadastral map. It is worth noting that differently from the results obtain in chapter 3, the building denoted

Table 4.2: Performance of the proposed approach for L'Aquila Dataset in terms of number of detected destroyed buildings, missed alarms and false alarms.

Total Number of Buildings	Detected Destroyed (ω_{c1})	Missed Destroyed	False Destroyed
200	6	0	1

with the letter e in Fig. 4.6f, was demolished by the authorities after the acquisition of the 21th of April (considered in this chapter) but before the acquisition of 12th September (used in chapter 3). The analysis confirms the high accuracy of the proposed technique.

4.5 Conclusions

In this chapter an approach for the rapid and accurate detection of damages in build-up areas after a catastrophic event using SAR data has been presented. SAR systems on-board of new generation SAR missions are able to acquire images with a geometrical resolution that varies form 5 meters up to 1 meter using two different operational modes called StripMap and SpotLight. By increasing the resolution the maximum ground coverage guaranteed decrease accordingly. In this chapter we have proposed to jointly exploit these two acquisition modes in order to: i) quickly identify the areas severely affected by the crisis event (i.e., hot-spots) by analyzing large areas using data characterized by a high geometrical resolution (3-5 m); and ii) accurately detect collapsed buildings inside each hot-spot by analyzing small areas using data with very high geometrical resolution (1 m).

In the first phase of the proposed approach the hot-spots, which are areas with both a significant increase or decrease in the backscattering value are detected from the SM images. Since the goal during this phase is to quickly identify the areas severely affected by the crisis event and not perform detection at a fine resolution, a comparison between the SM images with a log-ratio operator and a de-noising operation based on the Wavelet transformation was applied in order to obtained the hot-spots. By exploiting the available prior information about the scene i.e., cadastral maps it was possible to define the regions to be further analysed in order to detect the collapsed buildings.

The analysis at the level of single building is conducted in the second phase of the proposed approach by taking advantage of the theoretical model of new/destroyed building and the detection method illustrated in chapter 3 of the presented thesis but exploiting the curvelet transform in order to better preserve the linear details that characterize urban structures in VHR SAR images.

The proposed approach was tested using SAR data related to the earthquake of L'Aquila (Italy) in 2009 acquired by COSMOSky-Med constellation. The proposed method demonstrated to be effective in detecting all the demolished buildings with a high accuracy. In detail, only one buildings was misclassified over 200 present in the two datasets with no missed detection. This was possible because the high robustness of the proposed method to deviations from the ideal model considered for changed buildings. This robustness is due to both the use of adequate fuzzy rules and the intrinsic reliability obtained when working on the comparison of two images rather than on the modeling of building

in each image.

As future development of this work, we plan to test the proposed approach considering SAR dataset related to other earthquakes. Moreover, we plan to engineer the proposed approach to be compliant with the cogent requirements of real emergency scenarios.

Chapter 5

A hierarchical approach to change detection in multitemporal SAR images for surveillance applications

The availability of Very High Resolution (VHR) Synthetic Aperture Radar (SAR) images, which can be acquired by satellites over the same geographical area with short repetition interval, makes the development of effective unsupervised change-detection techniques very important. This chapter[1] *proposes a hierarchical approach to change detection in VHR SAR images for addressing surveillance applications, where VHR data are acquired with high temporal resolution (e.g., one image every few days). The proposed approach is based on two concepts: i) exploitation of a multiscale technique for a preliminary detection of areas containing changes in backscattering at different scales (hot-spots); and ii) explicit modeling of the semantic meaning of changes by using both the intrinsic SAR image properties (e.g., acquisition geometry and scattering mechanisms) and the available prior information. In order to illustrate the effectiveness of the proposed approach, a problem of freight traffic surveillance is addressed considering two data sets. Each of them is made up of a pair of multitemporal VHR SAR images acquired by the COSMO-SkyMed constellation in spotlight mode. Each data set defines a complex change-detection problem due to both the presence of a variety of changes on the ground and the complexity of object backscattering. Experimental results point out the effectiveness of the proposed approach.*

5.1 Introduction

The availability of very high geometrical and temporal resolution images makes it possible to address new challenging problems that were behind the possibilities of remote sensing before. Here we concentrate the attention on surveillance of sites of interest, such as maritime ports, airports, logistic centers, industrial areas and so on. As described in the chapter 1, the combination of short repetition interval and very high geometrical

[1] Part of this chapter appears in:

[65] F. Bovolo, M. Marin and L. Bruzzone, "A hierarchical approach to change detection in very high resolution sar images for surveillance applications," in *Geoscience and Remote Sensing, IEEE Transactions on,*, vol. 51, no. 4, pp. 20422054, April 2013.

(a) (b)

Figure 5.1: "Calata Neghelli — Porto Nuovo", Livorno, Italy. (a) Aerial image of the terminal showing warehouses, containers, forklifts, trucks, etc. ©Microsoft Corporation ©NAVTEQ — 2011. (b) RGB multitemporal composition of spotlight COSMO-SkyMed images (R:04/24/2010, G:04/23/2010, B:04/24/2010).

resolution leads to some challenging issues that should be addressed. In this chapter we focus the attention on the problem of freights surveillance. In order to introduce the problem, the case of a maritime port is considered in the following. This will allow a detailed description of the issues associated with this kind of problem. Nonetheless, the method is general and can be used also in other applications. Ports are complex systems often developed over wide surface. They are made up of several terminals used to properly handle cargoes passing through them [104–106]. In such a scenario, one of the main problems is the management of the large movement of goods for increasing the productivity of the port [104, 107]. In order to monitor the port operations and define new effective strategies for decreasing the handling time (and thus the overall cost of port management) surveillance activities should be carried out. A synoptic spatial view of a port can be obtained by analyzing available charts. However, given the large area to be monitored and its dynamic behaviours, it is desirable to exploit multitemporal space-borne sensor acquisitions.

In order to illustrate the scenario described above, let us consider a zone of the seaport of Livorno (Italy), and an aerial image taken over it (see Figure 5.1a). The main objects in the scene (i.e., buildings, containers, cargo ships, etc.) have dimensions in the order of some meters. Thus 1 meter resolution SAR images (e.g., acquired by COSMO-Skymed and TerraSAR-X/TanDEM-X) are a very important information source for addressing this problem. Dealing with VHR SAR images become very complex when multitemporal data are considered. Let us consider for example Figure 5.1b which was acquired on the same geographical area of Figure 5.1a. In this image, which is a false color composition of two VHR SAR images acquired at only one-day distance (i.e., April 23rd and April 24th 2010, respectively), changes due to a backscattering decrease appear in green, whereas the ones due to backscattering increase appear in magenta; unchanged pixels appear in gray levels. As one can notice, there is a large number of changes. However most of them are false alarms with respect to the goal of the detection container movements. This depends on the fact that many changes occurred in the backscattering are either related to the different conditions on the ground (an heavy rain occurred between the two acquisitions)

or to changes that are not related to the scenario under analysis.

Thus there is a need of distinguishing changes relevant from non-relevant changes. In order to accomplish this task we can exploit the prior information available in the considered problem. In the considered case of maritime ports, zones that differ because of their operational purpose can be identified. In detail it is possible to distinguish loading zones, storage tanks zones, docking facilities and so on. For each of these zones one can expect specific changes depending on the typical operations carried out. As an example, one can expect movements of containers in the container terminal, which are not expected in the chemical terminal. This means that, given the expected kind of change(s), for each zone it is possible to envisage the most suitable primitive(s) to enhance relevant changes and reduce the effects of non-relevant ones. These primitives may differ zone by zone. In order to render the detection of the expected change(s) more robust, it is possible to drive this procedure according to a preliminary identification of the areas affected by changes in backscattering (hot-spot) based on a multiscale analysis of the images. Large changed areas at a coarse scale are identified (hot-spots). According to them one can guess the expected kind of change. If hot-spots are localized in the car terminal, the change can be associated with the movement of cars. Therefore, we can extract primitives for detecting cars in the identified hot-spots on the two images in order to generate a change-detection map of car movements. A similar analysis can be carried out in order to describe any kind of change in maritime ports and extended to other sites of interest such as airports, industrial areas, logistic centers and so on.

In the literature few unsupervised approaches to change detection in VHR SAR images have been presented that exploit the semantic meaning of the backscattering changes in order to effectively separate the changes of interest from those that are not interesting [6, 10]. However, they do not take explicitly into account the multiscale proprieties of the expected changes. As we saw in chapter 2 some works have been presented in the literature that use multiscale strategies [15, 39, 108]. Despite their nice property of being multiscale, these methods do not model the change in backscattering of complex objects and are not able to exploit the semantic meaning of changes in backscattering (i.e., each kind of change is treated in the same way) since they were proposed for data having resolution in the range 10-30 meters. Therefore, they mainly address change-detection problems showing a single kind of change and are not optimized for separating multiple changes.

In order to overcome the limitations of state-of-the-art methods, in this work we propose a hierarchical approach to change detection in VHR multitemporal SAR images, which is based on two concepts: i) exploitation of a multiscale technique for a preliminary detection of areas containing changes in backscattering at different scales (hot-spots); and ii) explicit modeling of the semantic meaning of changes by using both the intrinsic SAR image properties (e.g., acquisition geometry and scattering mechanisms) and the available prior information. Due to the mentioned properties and the specific requirements on the availability of prior information, the proposed method is particularly suitable for surveillance and monitoring applications.

Two data sets, each made up of a pair of VHR CSK SAR images acquired in spotlight mode, were considered in the validation of the proposed approach. For both of them a

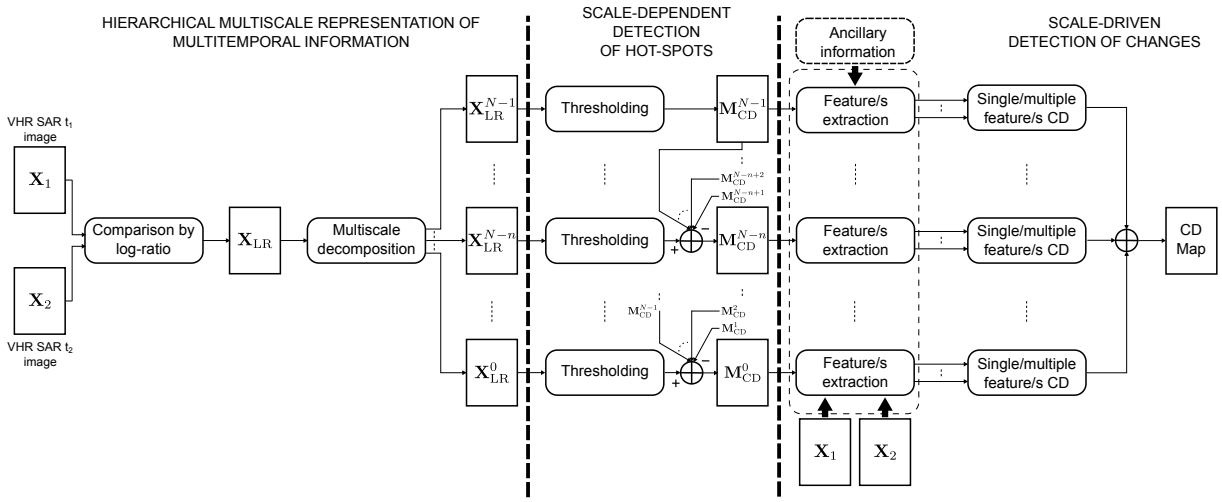


Figure 5.2: Conceptual flow of the proposed approach for monitoring and surveillance applications.

priori knowledge is available on the investigated scenario. The first data set was acquired over the freight village *"A. Vespucci"*, Livorno (Italy). Changes that occurred on the ground are mainly due to a single reason: movements of cargo (i.e., only one kind of change is present). The second data set was acquired over the port of the city of Livorno (Italy). This scenario is much more complex and between the two acquisitions several kinds of changes occurred on the ground. Changes are mainly due to cargo ship, truck and cargo movements. Moreover non interesting changes were observed due to the variation in the water content of the ground due to heavy rain between the two acquisitions. Results obtained on both data sets confirmed the effectiveness of the proposed approach.

The rest of the chapter is organized into four sections. Section 5.2 describes the proposed approach to change detection. Section 5.3 presents the data sets, illustrates the tuning of the proposed method on the considered application, and shows the experimental results. Section 5.4 draws the conclusions of this work.

5.2 Proposed Hierarchical Approach to Change Detection

Let us consider two amplitude VHR SAR images $\mathbf{X}_1$ and $\mathbf{X}_2$ acquired on the same geographical area at different times τ_1 and τ_2, respectively. Let us assume that the area of interest is associated to a surveillance problem, i.e., it is an airport, a port, a logistic center or an industrial area, and that the goal is to detect changes for monitoring commercial traffic. The most critical issue dealing with this kind of applications is related to the presence of many possible kinds of changes on the ground. A single test site may include many kinds of change that show significantly different characteristics in terms of size, shape and semantic meaning. Nevertheless, these kinds of change may exhibit similarities in the backscattering values even though they have a different semantic meaning. Here we propose to deal with this problem by exploiting all the available prior information about the scene. Given the considered kind of application, it is reasonable to assume that prior information about the test site is available, such as the position of buildings, docks,

landing strips, or the usage of specific zones (e.g., loading zones, storage tanks areas, docking facilities). The kind of available information depends on each specific problem.

Under this assumption we introduce an approach made up of three stages: i) a hierarchical multiscale representation of the multitemporal information; ii) a preliminary identification of the areas affected by changes in backscattering (hot-spots); and iii) a scale-driven detection of the changes, which uses the prior information. Figure 5.2 shows the block scheme of the proposed approach.

5.2.1 Hierarchical Multiscale Representation of Multitemporal Information

In order to detect hot-spots and to reduce the noise impact, a multilevel representation of the multitemporal information is computed. The representation is achieved according to a wavelet-based procedure [39]. Here the multiscale representation is used to derive the hot-spots rather than to directly detect changes. The detected hot-spots will drive the next steps of the proposed method (see Sec. 5.2.3).

The multiscale representation is obtained as described in chapter 2.2.2. Differences in backscattering are highlighted by means of the log-ratio image $\mathbf{X}_{\mathrm{LR}}$ i.e., $\mathbf{X}_{\mathrm{LR}} = \log \mathbf{X}_2/\mathbf{X}_1$. As we saw in chapter 2.2.2 this is the most common operator for highlighting changes in multitemporal SAR data. Indeed, the ratio operator is used in order to reduce the effects of speckle in the resulting image, and log operator is used to transform the residual multiplicative noise (which is expected to be high in portions of the ratio image associated with changed areas on the ground) in an additive noise component. $\mathbf{X}_{\mathrm{LR}}$ includes information about changes associated to both increase and decrease of backscattering. Unchanged pixels assume values close to zero, whereas positive and negative changes assume positive and negative values far from zero, respectively. Nevertheless, no information can be retrieved from $\mathbf{X}_{\mathrm{LR}}$ about the semantic label of such changes. From $\mathbf{X}_{\mathrm{LR}}$ a set of multilevel images $\mathcal{X}_{\mathrm{MS}} = \left\{\mathbf{X}_{\mathrm{LR}}^0, \ldots, \mathbf{X}_{\mathrm{LR}}^n, \ldots, \mathbf{X}_{\mathrm{LR}}^{N-1}\right\}$ is computed, where the superscript n, $n = 0, \ldots, N$ indicates the resolution level of images. The output at resolution level 0 corresponds to the original image, i.e., $\mathbf{X}_{\mathrm{LR}}^0 \equiv \mathbf{X}_{\mathrm{LR}}$. For n ranging from 0 to $N-1$, the images are characterized by different trade-offs between spatial-detail preservation and speckle reduction.

5.2.2 Scale-Dependent Preliminary Detection of Changes in Backscattering (Hot-Spots)

Once all the resolution levels have been brought back to the image domain, for each element $\mathbf{X}_{\mathrm{LR}}^n$ $(n = 0, 1, \ldots, N-1)$ of the set $\mathcal{X}_{\mathrm{MS}}$, a CD map is computed according to the split-based unsupervised thresholding approach proposed in 2.2.3. This approach was adopted since it can effectively detect changes in images of large size even when the prior probability of the class of change is small. This is due to the ability of the method to analyze only the sub-parts of the images that have the highest probabilities to contain changed pixels. In greater detail, each image $\mathbf{X}_{\mathrm{LR}}^n$ is automatically split into a set of S sub-images $\mathbf{X}_{\mathrm{LR}}^{n,s}$, $s = 1, \ldots, S$ of user defined size. The choice of the split size depends on the extension of the expected changes, and thus on the level n considered. Nevertheless, a minimum size should be guaranteed so that the estimation of statistical

parameters is reliable. At each resolution level the defined pixel sub-set is used to compute the change-detection map. Here the thresholding method described in [48] is adopted. Each sub-set of pixels is modeled as a sum of 3 probability density functions associated to no-change, positive change, and negative change classes. Under this assumption, the Bayes decision rule for minimum error is applied to separate the 3 classes. To this end a statistical model for class distributions is required together with an approach for class statistical parameters estimation. Following [48], the generalized-Gaussian model and the well known Expectation-Maximization (EM) algorithm [47] were employed. In this way, a CD map that shows three different classes is obtained at each resolution level n ($n = 0, \ldots, N-1$). The generated CD maps exhibit different trade-offs between details preservation and homogeneity i.e., the higher is n the larger and more homogeneous are the detected areas of change and viceversa (see Figure 5.6). Since these maps partially share changed areas, in order not to process the same changed areas multiple times at different resolution levels, a cancellation is performed. Starting from the coarsest level ($n = N-1$) and moving toward the finest one ($n = 0$), from each level n ($n = 0, \ldots, N-1$) changed areas that were detected (and therefore processed) at coarser levels m, $m = n+1, \ldots, N-1$, are removed. The result is a set of $\mathcal{M}_{\text{CD}} = \left\{\mathbf{M}_{\text{CD}}^0, \ldots, \mathbf{M}_{\text{CD}}^n, \ldots, \mathbf{M}_{\text{CD}}^{N-1}\right\}$ change-detection maps complementary to each other. Thus each map in $\mathcal{M}_{\text{CD}}$ contains a set of complementary changed areas $\mathcal{C}_h^n$, ($h = 1, \ldots, H_n$). In the following we will refer to these areas as hot-spots. It is worth noting, that differently from [25] these change-detection maps are not the final result of the proposed approach, they will be instead used to drive the detection of changes at finest scale during the next stage of the proposed approach.

5.2.3 Scale-Driven Detection of Changes Based on Prior Knowledge

In the last stage, changes of interest are extracted with their semantic meaning. This is done by taking advantage of the prior knowledge about the considered application, which is mainly associated to the typical usage of zones to be controlled and of the detected hot-spots. Thus the scene can be divided into different zones of interest in which different kinds of change are expected. Each expected change can be modeled and extracted by using specific features and change detectors that take into account the radiometric and geometric properties of these expected changes. The spatial context can also be taken into account. Which kind of features and change detectors should be involved in the process depends on the specific considered application. Feature extraction (and thus change detection) is performed by only considering hot spots $\mathcal{C}_h^n$, ($h = 1, \ldots, H_n$) in $\mathcal{M}_{\text{CD}}$ defined in the previous stage. Starting from the lowest resolution level $n = N-1$, and on the basis of the position of hot-spots within the zones of interest, the strategy for the detection of the specific expected changes is applied. Once all the $\mathcal{C}_h^n$ are analyzed, a finer scale is considered. This iterative process stops at the resolution level most suitable to properly detect the expected kinds of change. It is worth noting that one can expect to consider only few scales associated with the effects of expected hot-spots of change. The final change-detection map is built by combining in a single map the results achieved within the different hot-spots. In the following, examples of possible features and feature-dependent change detectors are presented, which are inspired by a problem of freight

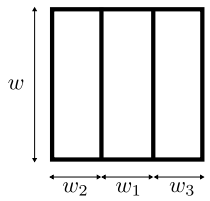


Figure 5.3: Window used by the line detector.

traffic surveillance. Despite features and change detectors may be applied to different hot-spots $\mathcal{C}_h^n$, in our notation we will omit this dependency in order to keep the notation as simple as possible.

Let us consider the example of Livorno maritime port surveillance described in Section 2.1 (see Figure 5.1a). In this scenario the main changes that characterize the cargo area of a commercial port are due to: 1) container movements; 2) cars movements; and 3) cargo ships movements. In the following they will be analyzed in more detail presenting their properties, the features and the change detectors more suitable to identify them.

Detection of Changes Associated with Movement of Containers

Shipping containers are reusable transport and storage metal units for moving products and raw materials between locations. They are made in different size according to ISO 6346 standard for containers. Common lengths are 20 and 40 feet. In addition, they are produced in two heights: "standard" (8.6 feet) and "high" (9.6 feet). Given the geometry, a single isolated container may be identified exploiting the strong response coming from the dihedral reflector, generated by the wall of the container and the ground where it sets down, which results in a double bounce effect that involves bright lines in SAR images. It is worth noting that the double-bounce Radar Cross Section (RCS) may depend on radar parameters (i.e., frequency and polarization), container parameters (e.g., shape, material), geometry parameters (i.e., incidence angle and container aspect angle i.e., the angle between the container wall facing the sensor and the azimuth direction), and background parameters (e.g., ground type, weather conditions). As presented in [77], the strength of the double-bounce varies as the aspect angle and background parameters vary according to a non-liner relation that can be empirically derived. These parameters have to be taken into account during the phase of detection. Thus, a container can be detected by extracting its double-bounce effect generated by the long side of the container, i.e., a bright line with a predefined minimum dimension of 20 feet with respect to the aspect angle.

In the proposed technique the extraction of bright lines is carried out on each single-date image by means of the line detector proposed by Tupin *et al.* in [109] but modified in order to extract the bright lines instead of the dark ones. This detector was successfully used in [110] for the detection of the bright lines associated with the presence of a building considering VHR TerraSAR-X images. It consists of the fusion of two-line detectors, namely D1 and D2, which are based on the geometry reported in Figure 6.5 considering 16 different directions. In greater detail, the response of D1 is the minimum response of

a ratio edge detector applied to both sides of the central linear structure of Figure 6.5:

$$r_L = \min(r_E^{(1,2)}, r_E^{(1,3)}) \tag{5.1}$$

where superscript 1 denotes the central region, superscripts 2 and 3 the two lateral regions, and $r_E^{(t,c)}$, with $(t,c) \in \{(1,2),(1,3)\}$, is defined as follows:

$$r_E^{(t,c)} = \begin{cases} 1 - \min\left(\frac{\mu_t}{\mu_c}, \frac{\mu_c}{\mu_t}\right) & \text{if } \mu_1 \geq \mu_2, \mu_3 \\ 0 & \text{Otherwise.} \end{cases} \tag{5.2}$$

where μ_t is the empirical mean of the region $t = 1$ having $M_t = \omega \times \omega_t$ pixels with amplitude A_p^t, i.e., $\mu_t = (\frac{1}{M_t}) \sum_{p \in t} A_p^t$. The same definition is applied to μ_c (with $c = 2, 3$) i.e., $\mu_c = (\frac{1}{M_c}) \sum_{p \in c} A_p^c$. The response of D2 is the minimum response of the cross-correlation edge detector applied to both sides of the central linear structure of Figure 6.5:

$$\rho = \min(\rho_{(1,2)}, \rho_{(1,3)}) \tag{5.3}$$

where subscript 1 denotes the central region, subscripts 2 and 3 the two lateral regions, and $\rho^2_{(t,c)}$ with $(t,c) \in \{(1,2),(1,3)\}$ is defined as follows:

$$\rho^2_{(t,c)} = \begin{cases} \frac{1}{1+(M_t+M_c)\frac{M_t \mathrm{LCV}_t^2 \mathrm{CR}^2_{(t,c)} + n_t \mathrm{LCV}_c^2}{M_t M_c (\mathrm{CR}_{(t,c)} - 1)^2}} & \text{if } \mu_1 \geq \mu_2, \mu_3 \\ 0 & \text{Otherwise.} \end{cases} \tag{5.4}$$

where $M_t = \omega \times \omega_t$ and $M_c = \omega \times \omega_c$ are the number of pixels, $\mathrm{CR}_{(t,c)} = \mu_t / \mu_c$ is the empirical contrast, LCV_t and LCV_c are the local coefficient of variation for regions $t = 1$ and $c = 2, 3$, respectively. The results of the two detectors, D1 and D2, are finally merged for each direction, by using an associative symmetrical sum $\Sigma(r_L, \rho)$ as follows:

$$\Sigma(r_L, \rho) = \frac{r_L \rho}{1 - r_L - \rho + 2 r_L \rho}, \quad \text{with } r_L, \rho \in [0, 1] \tag{5.5}$$

A line is detected when $\Sigma(r_L, \rho)$ is higher than the decision threshold $\Sigma(r_L, \rho)_{\min}$. The threshold can be identified manually or automatically, and its application results in a map of detected lines. Nevertheless, when VHR SAR data are considered, the hypothesis of fully developed speckle is not anymore satisfied. Thus the statistical properties of speckle studied for medium and low resolution SAR images cannot be used to automatically derive thresholds with a constant false alarm ratio as was done in [109]. In this work a fixed threshold equal to 0.7 has given a good detection of the linear bright features.

Once the containers are identified in $\mathbf{X}_1$ and $\mathbf{X}_2$ by means of their double-bounce response, they can be compared in order to detect possible changes. The comparison can be carried out according to the ratio of the means or to statistical similarity measures (e.g., *Kullback-Leibler* divergence, Mutual Information). In this chapter we use the *Kullback-Leibler* (KL) divergence. In detail we use the KL approximation given by the *Edgeworth* series expansion of cumulats truncated at a given order described in chapter 2.2.2. The obtained KL distance is then thresholded in order to identify changes in each analyzed double-bounce line of containers.

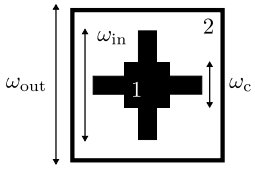


Figure 5.4: Window used by the isolated scatterer detector.

It is worth noting that, containers are normally stored together in large numbers and they may be piled up in stacks forming densely packed clusters with height, orientation and size. Even thought, the container clusters can be viewed as a problem where each element has similar characteristics, the dense packing and large variation in stacking configuration make it difficult the detection of single containers. Indeed, due to both the occlusion problems and different orientations of containers relative to the radar antenna (i.e., different aspect angles) the appearance of containers is not always the same, and a detailed quantitative information of the number of containers may be difficult or even impossible to extract. Nevertheless, a further processing of close double-bounce lines based on the method presented in [111] can mitigate the ambiguity on possible staking configurations of containers.

Detection of Changes Associated with Movement of Cars

Cars are often one of the freights handled by ports. In order to detect cars in SAR images it is necessary to identify their radar cross section. As described in [112] and like for the RCS of double-bounce lines, RCS of cars may depend on radar parameters (i.e., frequency and polarization), car parameters (e.g., model, shape, material), geometry parameters (i.e., incidence angle and car aspect angle), and background parameters (e.g., ground type, weather conditions). In [112] it was shown that there is a high probability of detection of cars if they are in front, lateral or back view with respect to the sensor. Such orientations produce a large RCS registered in the SAR image. Given their dimension and assuming front or back orientation, cars can be modeled as isolated scatterers with a given size. This is not true anymore if the cars are in lateral view with respect to the sensor. The extraction of isolated scatterers for the detection of vehicles is performed by means of the single-data detector proposed by Lopes *et al.* in [35]. Because the SAR impulse response is generally a separable function of range and azimuth directions, the main part of the point target response, i.e., the main lobe and the first side lobe, is spread on a neighborhood similar to a cross. Figure 5.4 shows the geometry of the isolated scatterer detector proposed in [35]. On the basis of this geometry, let us denote with subscript 1 the inner region and with subscript 2 the outer region, and let A_p be the amplitude of pixel, so that the radiometric empirical mean μ_Ω of a given region Ω having M_Ω pixels is $\mu_\Omega = (\frac{1}{M_\Omega}) \sum_{p \in \Omega} A_p$. The response of the detector is defined as:

$$r_{\text{PT}} = 1 - \min\left(\frac{\mu_1}{\mu_2}, \frac{\mu_2}{\mu_1}\right) \tag{5.6}$$

Thus, a pixel is considered as belonging to an isolated scatterer when its response r_{PT} is higher than a manually or automatically chosen threshold $r_{\text{PT}\,\text{min}}$. This detector can

be applied directly to the images $\mathbf{X}_1$ and $\mathbf{X}_2$ and the two results compared in order to detect the changes. Given the small number of samples that make up a car signature, statistical comparison methods like similarity measures (e.g., *Kullback-Leibler* divergence or Mutual Information) are not suitable. We perform the comparison of the isolated scatterers extracted as described before through a logical XOR operator. Then we associate a change in a car with regions with a predefined minimum size resulting for the XOR operation. Nevertheless, other techniques that operate in the multitemporal data extracting differential features can be applied to the detection of changes in cars.

Detection of Changes Associated with Movement of Cargo Ships

Cargo ships are one of the main vectors for the transportation of goods in a port. They are categorized according to the task they have to accomplish (e.g., container ships or tankers) and they can exhibit different size and shape. Nevertheless, due to the expected large size, changes associated with their positions can be detected using coarser levels of the multiresolution representation. Since cargo ships float in docks, changes due to cargo ships movements are often in between the ones associated to the natural movement of the water surface. Thus to remove this undesired effect the normalized ratio of means can be used. This is because changes in water backscattering do not strongly affect the mean value as the appearance/disappearance of ships do [113]. Hence, the normalized ratio of means of the images $\mathbf{X}_1$ and $\mathbf{X}_2$, computed inside the hot-spots $\mathcal{C}_h^n$ present in the wet dock, can be used. It is worth noting that other indicators could have been used to detect the movement of ships. Nevertheless, in order to maintain the computation load low, the hot-spots identified in the second step of the proposed approach were used. This image is the one used for hot-spots detection in the second step of the proposed approach. Let A_p^τ be the amplitude of pixel at time $\tau = \tau_1, \tau_2$, so that the radiometric empirical mean $\mu_{C_h^n}^\tau$ of a given hot-spot $\mathcal{C}_h^n$, $h = 1, \ldots, H_N$ having $M_{\mathcal{C}_h^n}^\tau$ pixels is $\mu_{\mathcal{C}_h^n}^\tau = (\frac{1}{M_{\mathcal{C}_h^n}^\tau}) \sum_{p \in \mathcal{C}_h^n} A_p^\tau$. The response of the detector is defined as:

$$r = 1 - \min\left(\frac{\mu_{\mathcal{C}_h^n}^{\tau_1}}{\mu_{\mathcal{C}_h^n}^{\tau_2}}, \frac{\mu_{\mathcal{C}_h^n}^{\tau_2}}{\mu_{\mathcal{C}_h^n}^{\tau_1}}\right) \tag{5.7}$$

In this work we choose to only consider low resolution levels since we are interested only in detecting cargo ships movements. More detailed information about the cargo ships (e.g., the kind of ship) could be retrieved by processing higher resolution levels (e.g., analyzing the histogram of the ship at the finest scale and comparing it with a database of ship histograms). However this kind of analysis is out of the aim of this work. As a final remark, it is worth noting that using low resolution levels the results suffer of a smoothing effect due to the low-pass filter applied during the multiscale decomposition. This may lead to under/over estimation of the size of the change and thus of the size of the appeared/disappeared ship.

5.2.4 Parallel Processing Architecture

Given the nature of the problem, the requirements that have to be fulfilled from the computational viewpoint are two: i) the analysis has to be performed in near-real time;

and ii) the scene can be associated to a large area. These two constrains are in contrast: increasing the area of the scene increases the computational burden and thus the overall time required to process the data. In order to face this problem, we developed the proposed approach in such a way as to be run on a computer cluster infrastructure. To this end, we defined a parallelization strategy based on the concept *divide et impera*. The main idea is that the same job can be performed by K different nodes on a small subset of the scene. The maximum speedup achievable using this strategy will depend on both the portion of code that can be parallelized (i.e., *Amdahl's law* [114]), and the computation resource available (taking into account the law of diminishing returns).

In greater detail, the proposed approach can be divided into two computational load phases. In the first phase the entire scene has to be analyzed in order to derive the multiscale representation of the changes in backscattering (see Sec. 5.2.1 and Sec. 5.2.2). This phase generates a computational burden that is proportional to the size of the multitemporal images. Thus the concept of *divide et impera* can be applied as follows. First, the VHR SAR images are split into tiles of a given size. In order to avoid borders problem (i.e., to detect changes located at the borders between two tiles) every tile overlaps with its neighbors. Second, the tiles are distributed among the computational nodes, which independently execute the multiscale decomposition and the detection of hot-spots.

In the second phase (Sec. 5.2.3), the change detectors are applied to the hot-spots found in the previous step of the proposed approach. It is worth noting that the selection of the more suitable detector is driven from the prior information available about the scene to be analyzed. Therefore, the computational load of this phase will depend on their extension and on the kind of change to be detected (i.e., on the complexity of the detector).

The results for each tile are finally merged in the full size change-detection map.

5.3 Data Description and Experimental Results

To assess the effectiveness of the proposed approach, experiments were carried out on two different data sets both describing a problem of freight traffic surveillance. The first data set represents a problem of monitoring of the movement of cars and it is related to the logistic center "*A. Vespucci*", Livorno (Italy), whereas the second data set represents a complex maritime port scene where different kinds of transport operations are carried out, and it is related to the port of Livorno (Italy).

5.3.1 Logistic Center "A. Vespucci": Cars Handling Surveillance

The first data set is made up of two spotlight mode (1m×1m resolution, with 0.5m×0.5m pixel spacing, X-band) CSK® 1-look amplitude images. They were acquired in HH-polarization on the 4th and 20th May 2011, on the freight village "'*A. Vespucci*" in ascending orbit with 23–24 degree incidence angle. The logistic center (43°36'14" N, 10°23'39" E) was set up in order to allow the exchange among the different modes of transportation (i.e., rail and truck) in such a way as to facilitate the traffic of freights to the final destination. One of the main kinds of freight handled by the center are cars.

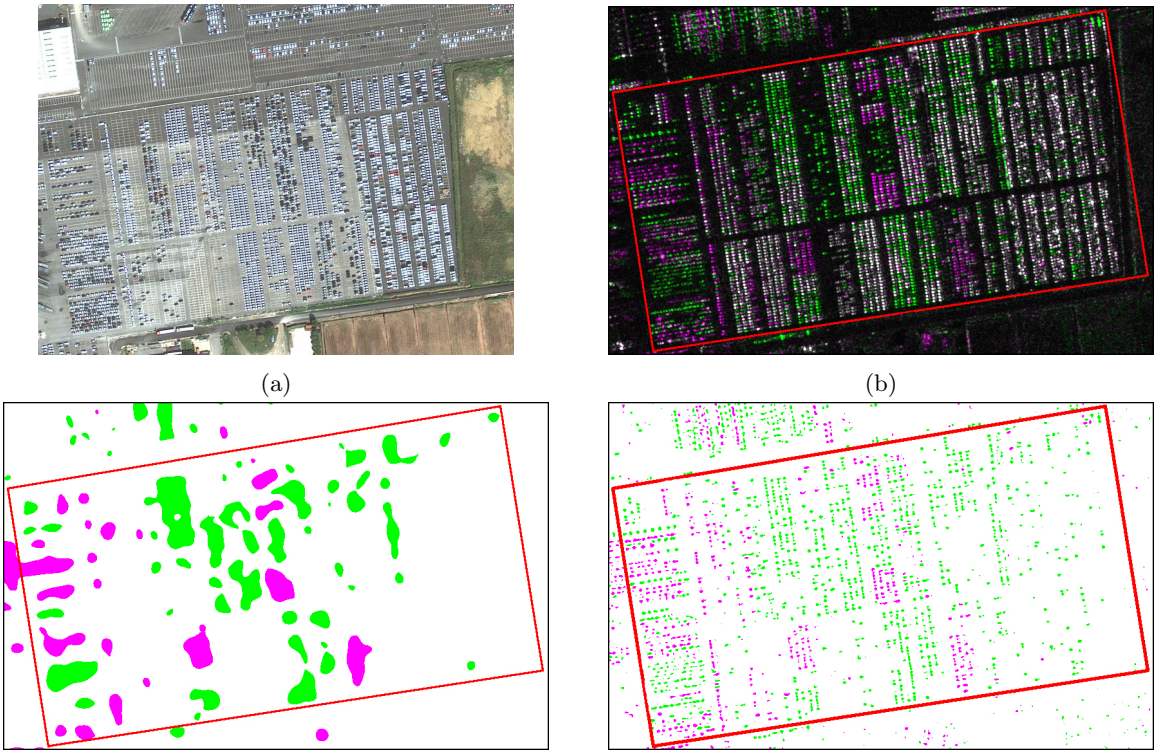


Figure 5.5: Logistic center "*A. Vespucci*" of Livorno (Italy) data set: (a) optical image — GeoEye, Tele Atlas — ©Google — 2011; (b) RGB multitemporal composition of spotlight COSMOSky-Med images (R:20/05/2011, G:04/05/2011, B:20/04/2011) ©ASI — Agenzia Spaziale Italiana — 2011. All Rights Reserved; (c) thresholding of $\mathbf{X}_{\mathrm{LR}}^{4}$; and (d) the proposed technique. New cars appear in magenta, and removed once are in green color.

Available prior information about the site tells us where car parking lots are positioned and how they are oriented. Thus, a test site of 883×693 pixels of the full scene was selected in which all the cars face the sensor.

In order to apply the proposed method, the log-ratio image $\mathbf{X}_{\mathrm{LR}}$ was computed from the two calibrated and co-registered CSK® images. The co-registration was performed with sub-pixel accuracy. From $\mathbf{X}_{\mathrm{LR}}$, the set $\mathcal{M}_{\mathrm{CD}} = \left\{\mathbf{M}_{\mathrm{CD}}^{0}, \ldots, \mathbf{M}_{\mathrm{CD}}^{n}, \ldots, \mathbf{M}_{\mathrm{CD}}^{N-1}\right\}$, with $N = 5$ resolution levels, was computed by applying firstly the 2D-SWT and the 2D-ISWT with an 8-length *Daubechies* filter. The impulse response of low-pass decomposition *Daubechies* filter of order 4 is given by the following coefficient set:

$$\{-0.0105974, 0.0328830, 0.0308414, -0.187035, -0.0279838, 0.630881, 0.714847, 0.230378.\}$$

The finite impulse response of the high-pass filter for the decomposition step can be computed by satisfying the properties of the quadrature mirror filters. From the multiresolution image representation, the complementary set of hot-spots $\mathcal{C}_{h}^{n}$, $(h = 1, \ldots, H_{n})$ have been extracted according to the unsupervised split-based method described in Section 5.2.2. In particular, we set the value of B equal to 3. This value gives for the considered data set the better trade-off between selecting high variance splits and suffi-

Table 5.1: Quantitative parameters associated with changed areas retrieved with the proposed approach in the logistic center "*A. Vespucci*".

Parameter	Retreived	Reference
# of added cars	440	302
# of removed cars	928	808

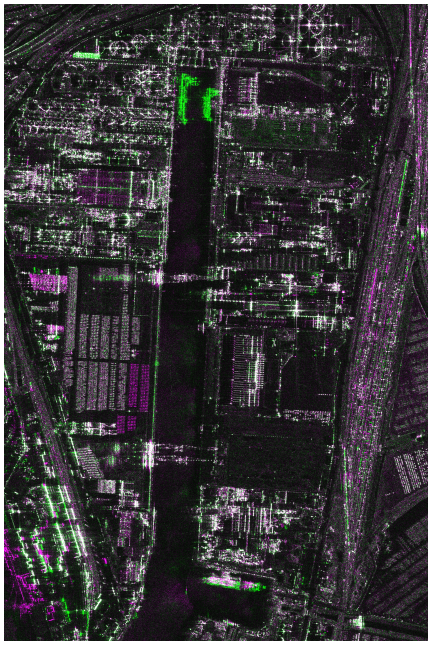
cient number of samples. Since from the available prior information the expected change in this scene is related only to the movement of cars, for each hot-spot $\mathcal{C}_h^n$ found at the levels n equal to 4 and 3, isolated scatterers are detected in $\mathbf{X}_1$ and $\mathbf{X}_2$ using the detector with response given by (5.6), with $\omega_{\text{c}} = 3$, $\omega_{\text{in}} = 5$ and $\omega_{\text{out}} = 11$. Isolated scatterers obtained for each image are compared according to the logical XOR operator in order to detect differences between them: i.e., appeared or disappeared cars. Considering the spatial resolution of the CSK® images, this feature may result in more than one scatterer for each vehicle. Therefore, a slight overestimation of the number of cars is expected.

Since no ground truth is available for the scene considered, the CD map obtained with the proposed method was compared from a qualitative point of view with: (a) an RGB multitemporal false color composition; (b) the change-detection map obtained by standard pixel-based thresholding of $\mathbf{X}_{\text{LR}}$ (i.e., $\mathbf{M}_{\text{CD}}^0$) and (c) the change-detection map obtained by thresholding $\mathbf{X}_{\text{LR}}^4$ (i.e., $\mathbf{M}_{\text{CD}}^4$). As expected, the change-detection map obtained with a standard pixel-based thresholding of $\mathbf{X}_{\text{LR}}$ is affected by a high number of false alarms due to noisy components. The change-detection map obtained by thresholding $\mathbf{X}_{\text{LR}}^4$ (Figure 5.5c) is less affected by isolated errors but it shows poor geometrical details. At this resolution level, cars cannot be counted. Differently, the proposed approach (Figure 5.5d) results in a change-detection map with a lower impact of false alarms, while at the same time it preserves changes associated to single cars. Quantitatively, we can automatically count 440 new cars and 928 cars removed from the logistic center (green and magenta colors in Figure 5.5d, respectively), whereas a visual inspection on $\mathbf{X}_1$ and $\mathbf{X}_2$ resulted in 302 new cars and 808 removed (see Table 5.1). This operation was not possible considering the CD map obtained with standard pixel-based thresholding of $\mathbf{X}_{\text{LR}}$.

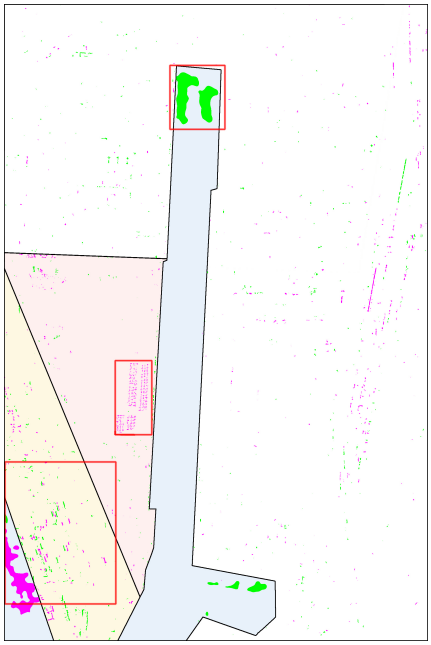
5.3.2 Port of Livorno Data Set: Port Surveillance

The second data set is made up of two spotlight mode CSK® images acquired on the port of Livorno (Italy). The port of Livorno is one of the largest Italian commercial seaports with an annual traffic capacity of around 50 million tonnes of cargo [105]. It is located on the Tyrrhenian Sea in the north-western part of Tuscany (43°32'6" N, 10°17'8" E). The two spotlight CSK® 1-look amplitude images (1m×1m resolution, with 0.5m×0.5m pixel spacing, X-band) were acquired in VV-polarization the 23rd and 24th April 2010 in descending orbit with 25–26 degree incidence angle. The selected test site is a section (2880×1920 pixels) of the full scene. Fig. 6.9a shows the optical image corresponding to the same area taken from Google Maps. This is a GeoEye RGB true color composition. Fig. 5.6b shows a false color composition of the two CSK® images. The available prior information about the scene is associated to the presence of three zones (see Figure 5.6): i) the cargo terminal (red region); ii) the car terminal (yellow region); and iii) the wet dock

(a)

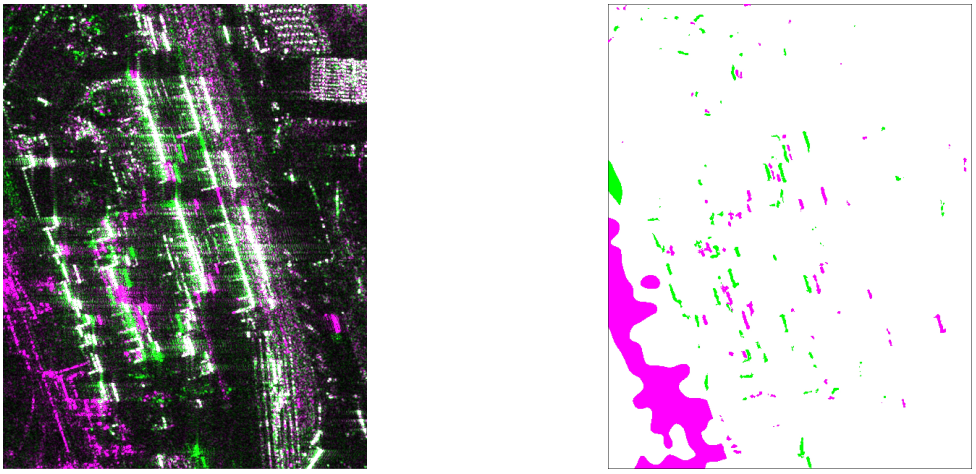


(b)



(c)

Figure 5.6: Port of Livorno (Italy) data set: (a) optical image — GeoEye, Tele Atlas — ©Google — 2011; and (b) RGB multitemporal composition of spotlight COSMOSky-Med images (R:04/24/2010, G:04/23/2010, B:04/24/2010) ©ASI — Agenzia Spaziale Italiana — 2010. All Rights Reserved. (c) CD map obtained with the proposed technique. Pixels that experience an increasing in the value of backscattering are in magenta, pixels that experience a decreasing in the value of backscattering are in green color. Legend for the zones: yellow: container terminal; blue: dock; red: car terminal.



(a) (b)

Figure 5.7: Zoom of cargo terminal area where containers are stacked (red box in left bottom part of Figure 5.6). (a) RGB composition of spotlight COSMO-SkyMed images; and (b) CD map obtained with the proposed method (new containers appear in magenta, removed ones appear in green).

(blue region). In the three mentioned zones we expect different kinds of changes relevant from the application point of view and associated to the movement of: i) containers in the container terminal; ii) cars in car terminals; and iii) cargo ships in the wet dock part. Moreover, we know from meteorological data that between the two acquisition dates a strong storm with wind and rain hit the port. Thus, we expect changes non relevant from the application viewpoint related to the presence of residual water in the scene.

To apply the proposed method, the set $\mathcal{M}_{\text{CD}} = \{\mathbf{M}_{\text{CD}}^0, \ldots, \mathbf{M}_{\text{CD}}^n, \ldots, \mathbf{M}_{\text{CD}}^{N-1}, \}$, $N = 5$, was obtained following the same procedure used for the first data set (see Section 5.3.1). Then, using the available prior information the problem was addressed considering the three zones that make up the port.

Table 5.2: Quantitative parameters associated with changed area retrieved with the proposed approach in the port of Livorno data set.

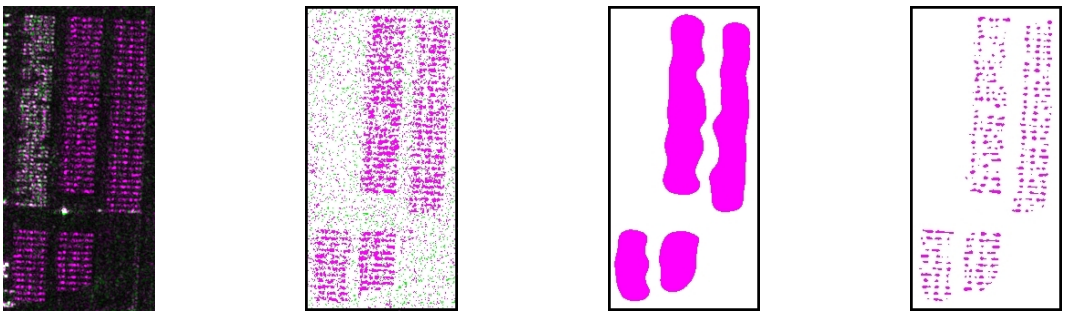
Parameter	Retreived	Reference
Length of cargo ships	Ship1: 117m	Ship1: 116m
	Ship2: 85m	Ship2: 74m
# of cars in the cargo terminal	284	249

In greater detail, the expected changes for each of the three zones were extracted as follows.

Detection of Changes Associated with Movement of Containers in the Container Terminal

for each hot-spot $\mathcal{C}_h^n$ found at the level $n = 4, 3$, bright lines are detected in $\mathbf{X}_1$ and $\mathbf{X}_2$ using the detector with response given by (5.5), with parameters $\omega = 4$, $\omega_1 = 2$, and $\omega_2 = \omega_3 = 4$. For each detected bright line in $\mathbf{X}_1$ and $\mathbf{X}_2$, the CKLD distance was

computed according to (2.30). The CKLD was then thresholded to detect the changes in bright lines location. This allows one to detect the movement of a single container or a stack of containers. An example of container detection is given in Figure5.7 (new containers appear in magenta, removed once in green). As mentioned in Section 5.2.3, containers are often stacked on each other, thus the detection of changes associated with movement of containers cannot fully solved only on the basis of the analysis of bright lines and other issues should be considered. For example, the relation between closed bright lines has to be taken into account [111]. As a matter of fact, a specific analysis should be carried out which is out of the goal of this work.

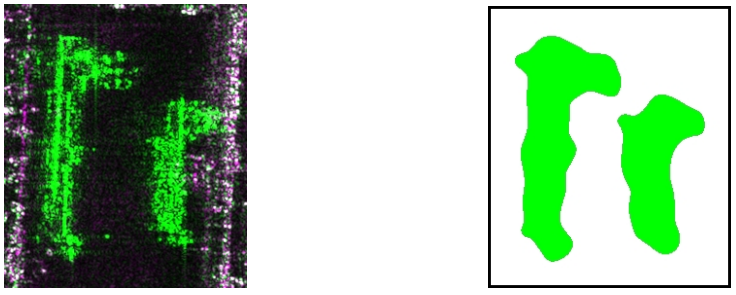


(a) (b) (c) (d)

Figure 5.8: Zoom of cargo terminal area where cars are stacked (red box in the center of Figure 5.6). (a) RGB multiteporal composition of spotlight COSMO-SkyMed images. CD maps obtained by: (b) standard pixel-based thresholding of $\mathbf{X}_{\mathrm{LR}}$; (c) thresholding of $\mathbf{X}^4_{\mathrm{LR}}$; and (d) the proposed technique.

Detection of Changes Associated with Movement of Cars in the Car Terminal

the same procedure used in the first data set was used here. From the qualitative point of view the results involve the same considerations derived for the logistic center and presented in Section 5.3.1 (see Figure 5.8). From a quantitative analysis, 284 new cars were automatically detected, whereas a visual inspection on $\mathbf{X}_2$ resulted in 249 new cars.



(a) (b)

Figure 5.9: Zoom of ships in the wet dock (red box in the top of Figure 5.6). (a) RGB composition of spotlight COSMO-SkyMed images; and (b) CD map obtained with the proposed method.

Detection of Changes Associated with Movement of Cargo Ships in the Wet Dock

due to the expected size of cargo ships, this kind of change was detected using only the information in $\mathbf{X}_{\mathrm{LR}}^4$. Hence, once the hot-spots $\mathcal{C}_h^4$ was derived, to distinguish between changes due to the cargo ship movements from the ones associated to the backscattering of water surface, the normalized ratio of mean values in $\mathcal{C}_h^4$ was used. Among the $\mathcal{C}_h^4$ ($h = 1, \ldots, 166$) hot-spots, two were identified as being compatible with ship movement (see Figure5.9). The two areas have a length of 234 and 170 pixels that, given the pixel spacing and the aspect angle of about 6 degree, correspond to 117 m and 85 m, respectively. The two values were compared with the actual length of the two ships anchor in dock, namely, the oil tanker *Capraia* and the asphalt/bitumen tanker *Bitflower* [115] (see Tab. 5.2). The difference between the estimated and true measures is mainly due to the smoothing effect of the multilevel decomposition. Finally, it is worth noting the two changes detected in the bottom part of the dock. Observing the two original images acquired on April 23rd and April 24th, we can suppose that two cargo ships were docked in the same berth at the two times. In such a situation, the change-detection method can only model the difference between the place of the two cargo ships. However, it cannot trace them back to a simultaneous appearance/disappearance of ships. This is because the frequency of this event, in the Shannon sense, is not in agreement with the sampling frequency of 1 day given by the multitemporal series.

Computational Considerations

It is worth finally mention that the large test scene was processed using a cluster composed by AMD® Opteron™ 6172 processors with 4 GB of RAM per CPU. The image was divided in tiles of 1024×1024 pixel, with an overlap of 32 pixels. The total number of tiles was 12, and each tile was processed by one CPU. The cluster was used also to speed up the feature extraction phase that was done in parallel. Thus, the peak load of the cluster was of 36 CPUs running simultaneously. The total processing time was about 15 minutes, which fulfil the requirement of near-real time processing necessary for surveillance applications, even when large size scene are considered.

5.4 Discussion and Conclusion

In this chapter an approach to change detection in multitemporal VHR SAR images for surveillance applications has been proposed. The approach takes advantage of three concepts: i) the use of multiscale representation for a preliminary detection of areas showing significant changes in backscattering between the two images (hot spots); ii) the exploitation of prior information about typical usage of zones of interest in the area under control; and iii) the definition of features and change detectors optimized for an effective detection of specific changes in each zone of interest.

The method identifies hot spots at different resolution levels according to the multiscale information extracted by using the 2-dimensional stationary Wavelet transform. Each hot spot is then analyzed according to its spatial position within a specific area of interest by using the available prior information. The use of such information allows one

to define the most proper features and change detectors to be used for extracting the expected changes at detailed scale. It is worth noting that the assumption on the availability of the prior information on the usage of different areas (and thus on the expected kinds of change) is reasonable since the method has been developed for high frequency surveillance/monitoring of sensitive areas such as maritime ports, airports, etc.

The approach, which after an initial setup is completely automatic, was tested on two VHR CSK spotlight data sets that show different levels of complexity. The first data set is fairly simple and is related to the freight village "*A. Vespucci*", Livorno (Italy). Here changes are only associated to the movement of cars in the stocking lots. The second data set is complex and is related to the maritime port of Livorno (Italy). Changes due to car, container and ship movements are present. In both cases the proposed method demonstrated to be effective in detecting all the expected kinds of change with a high accuracy. In detail, concerning the complex data set of the maritime port of Livorno, the final change-detection map models with a high geometrical precision both small changes, such as those associated to car movements (which can be automatically counted) and large changes due to cargo ship movements. This is possible because of the hierarchical extraction of hot spots of change and of the specific definition on the basis of the available prior information of feature extraction techniques and change detectors. Indeed, since feature extraction is driven by prior knowledge, the proposed method can effectively detect changes with significantly different properties in terms of shape, modeling and size.

Due to the hierarchical processing and to the implementation on a multi-core cluster we also achieved good performance in terms of computational time. The method converged in 15 minutes on an image of around 55×10^6 pixels running on a cluster architecture.

As a final remark it is important to note that, despite the method has been illustrated on a specific application, it has general validity and can be applied to the surveillance and monitoring of many scenarios.

As future developments of this work we plan to extend the experimental analysis to long time series of images in order to further validate the proposed approach in different conditions toward a possible pre-operational implementation. Moreover, we plan to increase the variety of change detectors currently considered to make it possible the analysis of other scenarios. In particular, we are currently studying the problem of monitoring airport areas.

Chapter 6

A multitemporal detector for the extraction of primitives from SAR images

The availability of Very High Resolution (VHR) Synthetic Aperture Radar (SAR) images regularly acquired by satellites since 2008 has opened new opportunities to study dynamic processes at the resolution of individual objects. At metric resolution objects result inhomogeneous and thus can be identified only by detecting all the primitives that make up them. In this context the development of effective multitemporal primitive detectors becomes very important. Indeed, an efficient combination of the temporal information can both improve the detection of stable primitives and point out the changed primitives. This chapter[1] *presents a novel approach to multitemporal detection of primitives in VHR SAR images that exploits the multitemporal information in order to: i) perform the detection of primitives at a given time; and ii) identify the transitions in the state of primitives between two consecutive observations. This is done by using a compound approach to the detection integrated with the fuzzy logic. Fuzzy logic is used in order to test the semantic consistency among the state of primitives at the different temporal observations and infer the state of a primitive at a given time. The performance of the multitemporal detector is evaluated on a time series of both simulated and real TerraSAR-X images acquired over the city of Lüneburg in Germany. Experimental results confirm the effectiveness of the proposed approach.*

6.1 Introduction

The recent advent of new Synthetic Aperture Radar (SAR) systems, such as TerraSAR-X, TanDEM-X and COSMO-SkyMed (CSK), makes multitemporal data regularly available over the same geographical area with a very high geometrical resolution (VHR). This has opened new opportunities for the analysis of dynamic processes at the resolution of single objects. As in the analysis of low and medium resolution images, some objects may be

[1] Part of this chapter appears in:
[55] C. Marin, F. Bovolo, and L. Bruzzone, "A Novel Multitemporal Detector for the Extraction of Primitive from Very High Resolution SAR Images," in *Geoscience and Remote Sensing, IEEE Transactions on,*, submitted.

directly detected as a single primitive [65]. Nevertheless, in general the high amount of geometrical details present in a VHR SAR image changes completely the perspective of SAR data analysis: objects that are considered homogeneous from a semantic point of view, such as buildings, show a signature that is inhomogeneous at high spatial resolution because of the scattering contributions from sub-objects that make up them. For instance a building signature is made up of: i) a layover area (due to the backscattering contributions coming from the ground, the vertical wall and the roof of the building); ii) a double bounce line (generated by the multiple scattering mechanisms between the ground and the vertical wall); and iii) a shadow area (generated by the occlusion of the sensor due to the building itself) [6,80,86]. Hence in order to properly exploit the amount of geometrical resolution present in the images it is necessary to take into account the backscattering contributions from sub-parts of the investigated object. This can be done by extracting the low-level primitives, which are associated with the sub-parts of an object, and combining them together in order to perform a reliable detection and reconstruction of the object.

There are mainly two factors that modify the performance of object/primitive detection when VHR SAR data with same acquisition parameters are considered: i) the speckle noise; and ii) the environmental conditions during the acquisition. Both these factors vary from acquisition to acquisition and contribute to alter the performance of the detector. In this context, the object/primitive detection may take advantage of the use of multitemporal information. This is a concept that has been already investigated in the context of speckle filtering. In detail, in [37] (which is based on [116]) the authors propose to linearly combine the temporal samples of the multitemporal series so that the output variance is minimal and the mean is preserved. In [117] the homomorphic filtering approach is combined with a discrete cosine transform applied to the logarithmic transformation of the time series allowing a filtering of the noise in the frequency domain. The approach presented in [118] detects homogeneous regions in a 3-D neighborhood of the stacked images which are then filtered using the Kuan filter. These three multitemporal speckle filtering techniques were reviewed and compared in [34]. Few works have been presented for solving the primitive detection problem in SAR images by considering the multitemporal information. In [119] an approach to the detection of roads using multitemporal SAR data has been presented. The approach firstly detects the roads, which are defined as dark linear primitives, independently on each of the images; then it combines the results by applying a fusion technique. Several fusion techniques have been tested leading to the conclusion that there is not a unique optimal fusion technique. The choice of the technique depends on the tolerance of the user to missed and false alarms.

Most of the multi-temporal approaches to both speckle filtering and primitive detection do not explicitly take into account the possibility that pixels completely change over time. Thus, pixels having the same spatial coordinates in the time series are involved in the despecking or detection process. This is a valid procedure for the cases of stable objects (i.e., objects that do not change over time). Whereas it leads to significant errors when unstable objects are present (i.e., objects that have a state transition over time). On the other hand in the literature several works have been presented that investigate the possibility to detect changes between pairs of multitemporal SAR images. These works aim at identifying

the areas of the images associated with a change in the intensity between two different observations acquired on the same geographical area. The change detection concept has been used recently in [120], where the authors propose a multitemporal filtering approach that makes use of a change detection technique. In detail, the method determines the changed and unchanged pixels by means of an exhaustive cross test based on the coefficient of variation. On the basis of this information an adaptive spatio-temporal filtering is performed such that a temporal mean is calculated if at a given time a pixel is unchanged and no filtering is performed if the pixel is changed. Nevertheless, the approach is limited to filtering and does not consider the detection of primitives.

In this chapter we propose an approach to multitemporal detection of primitives in VHR SAR images that exploits the multitemporal information in order to: i) perform the detection of primitives at a given time; and ii) identify the transitions in the state of primitives between consecutive observations (i.e., detect changes). The approach requires the definition of two main parts: one is the design of the detectors that extract primitives/objects at a single time (monotemporal information); the other is the strategy to implement to fully benefit from the multitemporal information. In this work we assume the use of any of the primitive detectors available in the literature and we focus on the definition of a methodology for the effective exploitation of the multitemporal information. This is achieved by proposing a novel compound approach to the detection integrated with the fuzzy logic. In detail, the compound approach is derived in the framework of multitemporal Bayesian decision rule for minimum error. This rule was introduced in the cascade formulation in [121] and in the compound formulation in [122], for addressing problems related to multitemporal image classification. Differently from these two works, this chapter aims at deriving an unsupervised formulation for the compound detection of primitives in multitemporal VHR SAR images. This is done by using the fuzzy logic, which results to be robust against the intrinsic uncertainty of the information sources and does not require any parametric model of the data distributions. The proposed compound detection analyzes the semantic consistency among the state of primitives at the different temporal observations and infers the state of a primitive at a given time. Even though here the proposed approach is presented for primitive detection in VHR SAR images, it has general validity and can be exploited for primitive detection in high- and medium-resolution SAR images. The method, which after an initial setup driven by the nature of primitive to be detected is totally automatic, has been validated on both simulated and real TerraSAR-X multitemporal scenarios. The obtained results highlight the effectiveness of the proposed approach.

The rest of the chapter is organized as follows. Section 6.2 introduces the proposed approach to compound detection of primitives. Section 6.3 describes how the compound approach is integrated in the proposed fuzzy inference system in order to cope with the problems of the probability estimation. Section 6.4 presents the experimental results on both simulated SAR images and challenging real scenarios of urban multitemporal bright line detection. Finally, section 6.5 draws the conclusion and points out the future works.

6.2 Proposed Bayesian Compound Detection Approach

Let us consider two intensity VHR SAR images $\mathbf{X}_1$ and $\mathbf{X}_2$ acquired with the same parameters on the same geographical area at different times. Let $x_{\tau,j}$ be the $j-th$ pixel of the feature that describe an object (or primitive) at the time τ, with $\tau = \{1, 2\}$. Examples of primitives may be lines, regions with specific statistical proprieties, etc. In this work we only consider a single primitive. Let $\Omega = \{\omega_p, \omega_{np}\}$ be the set of possible semantic classes at time τ associated to the considered primitive. The two classes, ω_p and ω_{np}, correspond to the presence and the absence of the considered primitive, respectively. In the framework of the Bayesian rule for minimum error the compound detection problem can be defined by identifying the best (in terms of Bayesian decision theory) pair of classes to be assigned to each pair of features $(x_{1,j}, x_{2,j})$ by considering the temporal dependency between them as follows [122]:

$$(x_{1,j}, x_{2,j}) \in (\omega_n, \omega_m) \quad \text{if} \quad \omega_m, \omega_n = \underset{\omega_i \in \Omega, \omega_k \in \Omega}{\arg\max} \{P(\omega_i, \omega_k | x_{1,j}, x_{2,j})\} \tag{6.1}$$

where $P(\omega_i, \omega_k | x_{1,j}, x_{2,j})$ is the joint posterior probability of the class pair (ω_i, ω_k) given the pair of features $(x_{1,j}, x_{2,j})$. The estimation of the statistical quantities involved in (6.1) is a complex task due to the difficulty of properly modeling the multitemporal dependence in an unsupervised manner. Here, according to the literature [5, 121, 122], we assume class-conditional independence in the time domain to simplify the estimation of the joint conditional posterior probabilities. Under this assumption, (6.1) can be rewritten as:

$$(x_{1,j}, x_{2,j}) \in (\omega_n, \omega_m) \quad \text{if} \quad \omega_m, \omega_n = \underset{\omega_i \in \Omega, \omega_k \in \Omega}{\arg\max} \{p(x_{1,j} | \omega_i) p(x_{2,j} | \omega_k) P(\omega_i, \omega_k)\} \tag{6.2}$$

where $p(x_{1,j}|\omega_i)$ and $p(x_{2,j}|\omega_k)$ are the single-time class-conditional probability density functions representing the presence or absence of a primitive given features $x_{\tau,j}$, and $P(\omega_i, \omega_k)$ is the joint prior probability of having class ω_i at time $\tau = 1$ (i.e., $x_{1,j} \in \omega_i$) and ω_k at time $\tau = 2$ (i.e., $x_{2,j} \in \omega_k$). It is worth noting that two types of detection are achieved by (6.2): i) the detection of the temporal transition of each pair of pixels $(x_{1,j}, x_{2,j})$ from the class ω_i to ω_k with $(\omega_i, \omega_k) \in \Omega$; and ii) the primitive detection on each of the two images $x_{\tau,j}$, which is done by exploiting the information contained in both the images $\mathbf{X}_1$ and $\mathbf{X}_2$. In order to perform such compound primitive detection all the terms involved in equation (6.2) must be estimated. In the following the estimation of the class-conditional probability density function and the joint conditional prior probabilities using unsupervised statistical estimators are discussed.

The single-time class-conditional probability density functions are estimated from the probability density functions (PDF) associated to each feature $x_{\tau,j}$. In detail, by denoting with $p(x_{\tau,j})$ the PDFs associated to $x_{\tau,j}$ we can write:

$$p(x_{\tau,j}) = P_\tau(\omega_p) p(x_{\tau,j} | \omega_p) + P_\tau(\omega_{np}) p(x_{\tau,j} | \omega_{np}) \tag{6.3}$$

where $P_\tau(\omega_p)$ and $P_\tau(\omega_{np})$ denote the a-priori probabilities associated to the presence and absence of a primitive at time τ. Thus by solving the mixture-estimation problem described in (6.3) it is possible to derive the single-time class-conditional probability density functions needed for the compound detection in (6.2).

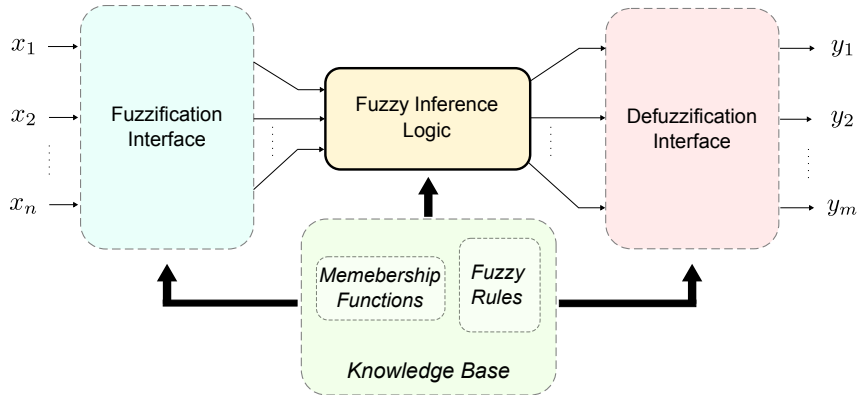


Figure 6.1: Conceptual architecture of a general fuzzy inference system (FIS)

The joint prior probabilities are related to the transitions of the state of the primitives from time $\tau = 1$ to $\tau = 2$. As there are two classes per time i.e., ω_p and ω_{np}, and two time instances to be considered, four class transitions are possible: the absence of class transitions (ω_p, ω_p) and $(\omega_{np}, \omega_{np})$, the transition from primitive to non-primitive (ω_p, ω_{np}), and the transition from non-primitive to primitive (ω_{np}, ω_p). Depending on the nature of the primitive taken into account, these transitions can be described considering an arbitrary index of change $x_{c,j}$. By focusing on the detection of a bright primitive i.e., whose presence is defined by a bright return, and using the filtered version of the log-ratio operator at a the optimal resolution scale that represent the investigated primitive [65], the transitions (ω_p, ω_p) and $(\omega_{np}, \omega_{np})$ are represented by no-change i.e., $x_{c,j} \in \xi_u$, the transition (ω_{np}, ω_p) is represented by an increase in the value of backscattering i.e., $x_{c,j} \in \xi^+$, and the transition (ω_p, ω_{np}) is represented by a decrease in the value of backscattering i.e., $x_{c,j} \in \xi^-$. Therefore the joint prior probabilities of (6.2) are related to the conditional probability density functions of each of the semantic situations $p(x_{c,j}|\xi_u)$, $p(x_{c,j}|\xi^+)$ and $p(x_{c,j}|\xi^-)$ through the distribution of $p(x_{c,j})$

$$p(x_{c,j}) = P(\xi^-)p(x_{c,j}|\xi^-) + P(\xi_u)p(x_{c,j}|\xi_u) + P(\xi^+)p(x_{c,j}|\xi^+) \tag{6.4}$$

The estimation of the mixtures (6.3) and (6.4) can be addressed in different ways [42]. Nonetheless, parametric and nonparametric approaches to estimation may fail at identifying the conditional probability density functions within a reasonable error bound if the right assumptions are not verified. Therefore, in order to cope with the uncertainty of the information sources in a more flexible way in the next section the compound detection approach is developed in the framework of the fuzzy logic.

6.3 Proposed Compound Detector based on a Fuzzy Inference System

The Bayesian formulation of the multitemporal primitive detection described in the previous section requires a robust statistical estimation process. However, the use of either parametric or nonparametric estimators may be affected by approximated models and

thus generate unreliable results. Motivated by the robustness of fuzzy theory to cope with uncertain information sources, we propose to design a fuzzy inference system (FIS) that implements the compound detection mechanism presented in the previous section. Fuzzy logic has been effectively applied to a wide variety of both classical engineering problems such as automotive systems [123–125], and also to remote sensing problems such as classification [126–128]. Its success is mainly due to the fact that fuzzy logic is tolerant to imprecise data, is flexible, and is conceptually easy to be understood. The fuzzy theory was introduced by Zadeh in [90] and provides the framework that can be used to convert expert knowledge into an automatic inference strategy. The most commonly used fuzzy inference technique is the so-called Mamdani method [129]. Fuzzy Inference System (FIS) can be seen as a non-linear mapper between inputs and outputs. The mapping is accomplished by a set of fuzzy rules, called fuzzy conditional statements, which are a set of IF-THEN rules defined by the knowledge available about the problem to be solved. In detail, a fuzzy conditional statement assumes the form "IF a set of conditions are satisfied THEN a set of consequences can be inferred". The if-part of the rule is called antecedent or premise, while the then-part is called the consequence or conclusion. In this statement, antecedents and consequences assume fuzzy values by means of membership functions (MFs), which provide a convenient way for expressing the domain knowledge. The basic structure of the proposed FIS consists of four conceptual components: i) a fuzzification interface, which transforms crisp inputs into fuzzy inputs (this block is not always present as in the presented case); ii) a fuzzy inference logic, which performs the inference procedure; iii) a defuzzification interface, which converts the conclusions of the inference mechanism into crisp outputs; and iv) a knowledge base, which defines the fuzzy rules and the MFs to be applied in order to perform the inference. The architecture of a fuzzy inference system is shown in Fig. 6.1.

The starting point for building a FIS is the definition of the fuzzy rule set. In this chapter we propose to translate the Bayesian compound detection strategy described in section 6.2 into fuzzy conditional statements in order to control the output of a FIS. The inputs of the FIS, which correspond to the statistical variables related to $x_{1,j}$, $x_{2,j}$ and $x_{c,j}$ in the Bayesian compound formulation, are called linguistic variables in fuzzy jargon [123] and are denoted with $\tilde{x}_{1,j}$, $\tilde{x}_{2,j}$ and $\tilde{x}_{c,j}$. In detail, $\tilde{x}_{1,j}$ and $\tilde{x}_{2,j}$, which are defined in the universe of discourse $\mathcal{U}$, are characterized by the term set $\mathbf{T}(\cdot) = \{\tilde{\omega}_p, \tilde{\omega}_{np}\}$, where $\tilde{\omega}_p$ = "primitive" and $\tilde{\omega}_{np}$ = "non-primitive". Whereas $\tilde{x}_{c,j}$, defined in the universe of discourse $\mathcal{T}$, is characterized by the term set $\mathbf{T}(\cdot) = \{\tilde{\xi}_u, \tilde{\xi}^+, \tilde{\xi}^-\}$ where $\tilde{\xi}_u$ = "no-change", $\tilde{\xi}^-$ = "decrease in backscattering" and $\tilde{\xi}^+$ = "increase in backscattering". By calling $\tilde{z}_{c,j}$ the output of the compound inference system, defined in the universe of discourse $\mathcal{V}$, the Bayesian rule for compound detection can be expressed as a set of fuzzy conditional rules as follows

$$
\begin{aligned}
&R^1 : \text{if } \tilde{x}_{1,j} \text{ is } \tilde{\omega}_p \text{ and } \tilde{x}_{2,j} \text{ is } \tilde{\omega}_p \text{ and } \tilde{x}_{c,j} \text{ is } \tilde{\xi}_u \text{ then } \tilde{z}_{c,j} \text{ is } (\tilde{\omega}_p, \tilde{\omega}_p).\\
&R^2 : \text{if } \tilde{x}_{1,j} \text{ is } \tilde{\omega}_{np} \text{ and } \tilde{x}_{2,j} \text{ is } \tilde{\omega}_{np} \text{ and } \tilde{x}_{c,j} \text{ is } \tilde{\xi}_u \text{ then } \tilde{z}_{c,j} \text{ is } (\tilde{\omega}_{np}, \tilde{\omega}_{np}).\\
&R^3 : \text{if } \tilde{x}_{1,j} \text{ is } \tilde{\omega}_p \text{ and } \tilde{x}_{2,j} \text{ is } \tilde{\omega}_{np} \text{ and } \tilde{x}_{c,j} \text{ is } \tilde{\xi}^- \text{ then } \tilde{z}_{c,j} \text{ is } (\tilde{\omega}_p, \tilde{\omega}_{np}).\\
&R^4 : \text{if } \tilde{x}_{1,j} \text{ is } \tilde{\omega}_{np} \text{ and } \tilde{x}_{2,j} \text{ is } \tilde{\omega}_p \text{ and } \tilde{x}_{c,j} \text{ is } \tilde{\xi}^+ \text{ then } \tilde{z}_{c,j} \text{ is } (\tilde{\omega}_{np}, \tilde{\omega}_p).
\end{aligned}
\tag{6.5}
$$

where the output term set is $\mathbf{T}(\cdot) = \{(\tilde{\omega}_p, \tilde{\omega}_p), (\tilde{\omega}_p, \tilde{\omega}_{np}), (\tilde{\omega}_{np}, \tilde{\omega}_p), (\tilde{\omega}_{np}, \tilde{\omega}_{np})\}$. It is worth noting that each of these rules is expressed in natural language. For example, the first rule can be read as: "if the primitive feature $\tilde{x}_{1,j}$ calculated at time $\tau = 1$ is a primitive and if the primitive feature $\tilde{x}_{2,j}$ calculated at time $\tau = 2$ is primitive and the index of change between the two observations $\tilde{x}_{c,j}$ is no change then the pair of pixels $(\tilde{x}_{1,j}, \tilde{x}_{2,j})$ is associated to the detection of the pair (primitive, primitive)". The same is valid for the others rules in (6.5).

The four rules do not cover all the possible situations. Indeed having 2 inputs with 2 possible states and 1 input with 3 possible states results in $2^2 \cdot 3 = 12$ possible rules. Nonetheless the rules expressed before are the only four likely rules that one can infer. Indeed situations such as for example "if the primitive feature $\tilde{x}_{1,j}$ calculated at time $\tau = 1$ is not a primitive and if the primitive feature $\tilde{x}_{2,j}$ calculated at time $\tau = 2$ is not a primitive and the index of change between the two observations $\tilde{x}_{c,j}$ is increase in backscattering" cannot have a logical output. Nonetheless, as we will see later in the chapter these situations will result in an output value equal to zero, which can be managed accordingly.

The fuzzy rules introduced by (6.5) are a linguistic description of the proposed compound primitive detection. Their interpretation is performed by the fuzzy inference logic unit (see yellow block of Fig. 6.1). This requires four steps, which are exemplified in Fig. 6.2: (A) the fuzzy quantification of inputs, (B) the aggregation of the antecedents, (C) the implication from the antecedents to the consequences, and (D) the combination of the consequences across the rules. These steps are described in detail in the following subsection together with the defuzzification interface (E). For sake of notation each rule of (6.5) is recall in its canonical form for each step. Specifically, let l be the number of rules with $l = 1, 2, 3, 4$. The rule-base (6.5) can be expressed in the canonical form as:

$$R^{(l)}: \text{ if } x_1 \text{ is } A_1^l \text{ and } x_2 \text{ is } A_2^l \text{ and } x_3 \text{ is } A_3^l \text{ then } z \text{ is } B^l \tag{6.6}$$

where $x = x_1, x_2, x_3$, defined in $\mathcal{U}_i$, and A_i^l, with $i = 1, 2, 3$, are the input linguistic variables and values, respectively. y, defined in $\mathcal{V}$, and B^l are output linguistic variables and values, respectively. As one can notice all the rules in (6.5) can be obtained by substituting the inputs and outputs variables and values with the appropriate terms.

6.3.1 Fuzzy Quantification of Inputs

The first step of the fuzzy inference logic is the quantification of the input linguistic variables and their uncertainty. This is done by using membership functions (MFs). In Fig. 6.2, which represents the proposed fuzzy inference diagram for the compound detection of the primitive, each of the linguistic variables $\tilde{x}_{1,j}$, $\tilde{x}_{2,j}$ and $\tilde{x}_{c,j}$ is evaluated by MFs in order to quantify its possibility to assume a linguistic value among $\tilde{\omega}_p, \tilde{\omega}_{np}, \tilde{\xi}^+, \tilde{\xi}^-$ and $\tilde{\xi}_u$. The choice of grades of membership is based on the available information (knowledge base) of the system and is described using different functions with different shapes. For example, if the statement "$\tilde{x}_{1,j}$ is $\tilde{\omega}_p$" is considered, it is possible to say that for certain values of the domain $\mathcal{U}$, $\tilde{x}_{1,j}$ is surly primitive whereas for other it is surly non-primitive. This is quantified by using a sigmoid function that maps in a continuous manner and with a given slope the grade of membership of the values of $\tilde{x}_{1,j}$ to the linguistic value $\tilde{\omega}_p$. In

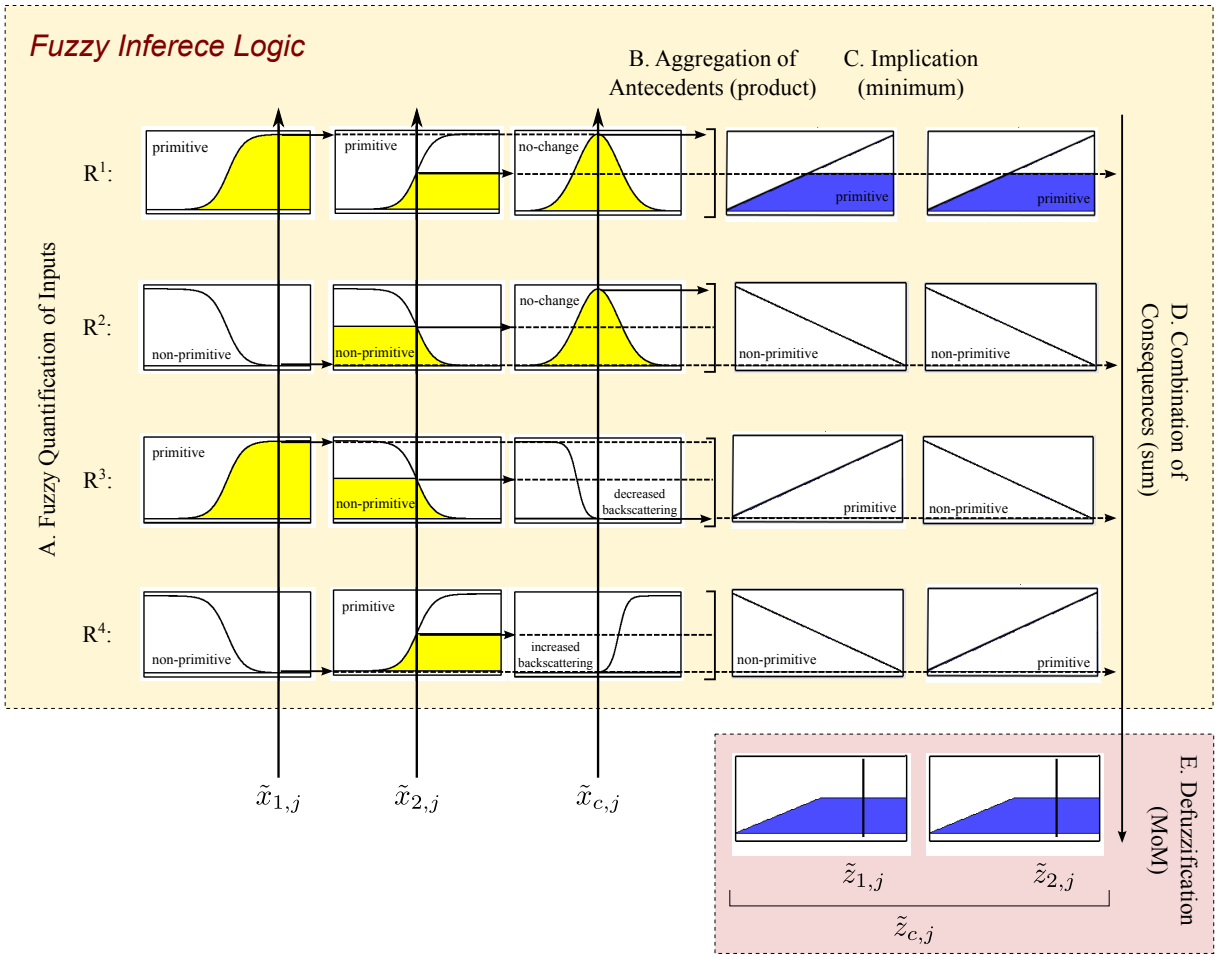


Figure 6.2: Proposed fuzzy inference diagram for the compound detection of the primitive.

detail, by taking into account the canonical form (6.6), a sigmoid function that maps the grade of membership of linguist variable x_i to a linguistic value A_i^l is defined as:

$$\mu_{A_i^l}(x_i) = \frac{1}{1 + \exp[-a_i^l(x_i - b_i^l)]} \tag{6.7}$$

where the constant a_i^l tunes the slope of the function and the constant b_i^l locates the center of the function. Depending on the sign of the parameter a_i^l, the sigmoid membership function is open to the right or to the left. This allows us to describe complementary logic concepts such as x_i is "primitive" and x_i is "non-primitive". The domain of the MF coincides with the universe of discourse of the linguistic variables and the codomain is defined in the range $[0, 1]$. Following the same rationale, in this work the sigmoid function is used to describe the membership of $\tilde{x}_{\tau,j}$ to $\tilde{\omega}_p$ and $\tilde{\omega}_{np}$, and the membership of $\tilde{x}_{c,j}$ to $\tilde{\xi}^+$ and $\tilde{\xi}^-$. Examples of sigmoid MFs for $\tilde{\omega}_p$, $\tilde{\omega}_{np}$ are depicted in Fig. 6.3a, whereas sigmoid MFs for $\tilde{\xi}^-$, $\tilde{\xi}^+$ are depicted in Fig. 6.3b. These MFs allow the definition of the linguistic

values in the rule set (6.5) as fuzzy sets as follows

$$\tilde{\omega}_p = \left\{\tilde{x}_{\tau,j}, \mu_{\tilde{\omega}_p}(\tilde{x}_{\tau,j}) | \tilde{x}_{\tau,j} \in \mathcal{U}\right\} \tag{6.8}$$

$$\tilde{\omega}_{np} = \left\{\tilde{x}_{\tau,j}, \mu_{\tilde{\omega}_{np}}(\tilde{x}_{\tau,j}) | \tilde{x}_{\tau,j} \in \mathcal{U}\right\} \tag{6.9}$$

$$\tilde{\xi}^{+} = \left\{\tilde{x}_{c,j}, \mu_{\tilde{\xi}^{+}}(\tilde{x}_{c,j}) | \tilde{x}_{c,j} \in \mathcal{T}\right\} \tag{6.10}$$

$$\tilde{\xi}^{-} = \left\{\tilde{x}_{c,j}, \mu_{\tilde{\xi}^{-}}(\tilde{x}_{c,j}) | \tilde{x}_{c,j} \in \mathcal{T}\right\} \tag{6.11}$$

In order to map the membership of $\tilde{x}_{c,j}$ to $\tilde{\xi}_u$ a generalized bell-shaped function is used. This because the log-ratio operator concentrate the unchanged values of $\tilde{x}_{c,j}$ around the zero. The generalized bell-shaped membership function is defined as:

$$\mu_{A_i^l}(x_i) = \frac{1}{1 + \left|\frac{x_i - e_i^l}{c_i^l}\right|^{2d_i^l}} \tag{6.12}$$

where the parameter e_i^l locates the center of the curve, c_i^l is the spread of the MF and d_i^l its shape. An example of generalized bell-shaped MF for $\tilde{\omega}_u$ is given in Fig. 6.3b. This MF allows the definition of $\tilde{\xi}_u$ as follows:

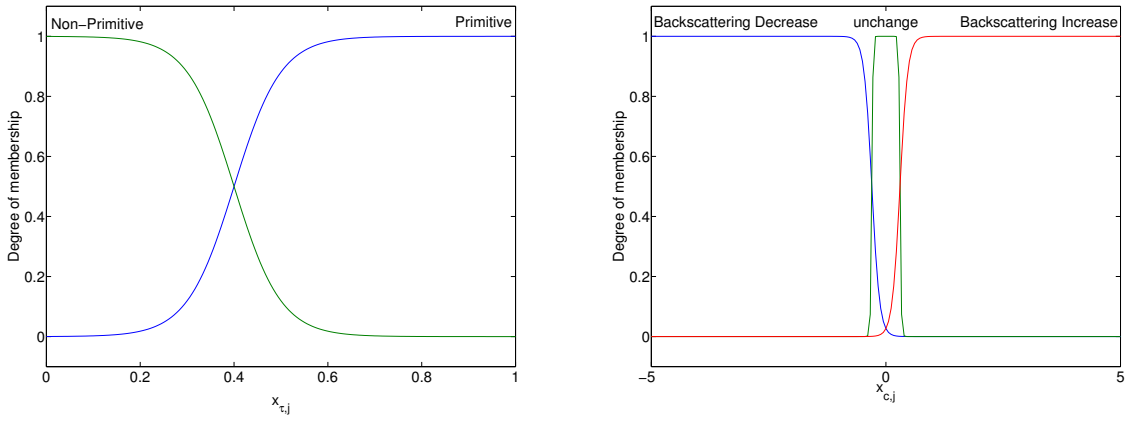
$$\tilde{\xi}_u = \left\{\tilde{x}_{c,j}, \mu_{\tilde{\xi}_u}(\tilde{x}_{c,j}) | \tilde{x}_{c,j} \in \mathcal{T}\right\} \tag{6.13}$$

One can notice that the correct choice of the MFs plays a fundamental role. In particular, if an input is disturbed by noise, the associated MF should be sufficiently wide to reduce the sensitivity to noise. For all the MFs used in this work the spread of the MFs is defined by the parameters a_i^l and c_i^l, which tune the sensitivity to noise of the system. Moreover, since a FIS should always be able to infer a proper value for every possible value of the inputs, the union of the supports of the MFs of a given rule has to cover the entire universe of discourse of the input linguistic variables. This is guaranteed by choosing the element $x_i = x_i^c$ at which $\mu_{A_i^l}(x_i^c) = 0.5$, called crossover point, as the intersection point between MFs representing different linguistic values (see Fig. 6.3). In the extreme case two dominant rules are activated by an equal belief of 0.5. This has the effect to create a relationship between the parameters of the MFs that cover a given universe of discourse. Therefore, for the considered MFs, five parameters have to be tuned by the users. For shake of notation we call with a and b the parameters related to the monotemporal information, and with c, d and e the parameters related to the multitemporal information.

6.3.2 Connection of the Inputs

The second step is the connection of the input MFs for each fuzzy rule (see Fig. 6.2). The interaction is derived by applying fuzzy logic operators. If we consider the canonical IF-THEN rule defined in (6.6), the fuzzy AND between the inputs is a fuzzy set in the space $\mathcal{U}_1 \times \mathcal{U}_2 \times \mathcal{U}_3$ with a MF defined as follows [123]:

$$\mu_A^l(x) = \mu_{A_1^l \cap A_2^l \cap A_3^l}(x_1, x_2, x_3) = t\left[\mu_{A_1^l}(x_1), \mu_{A_2^l}(x_2), \mu_{A_3^l}(x_3)\right] \tag{6.14}$$



(a) (b)

Figure 6.3: Example of membership functions for inputs a) $x_{\tau,j}$ and b) $x_{c,j}$. In detail: a) sigmoid MFs with parameters $a_i^l = 20$ and $b_i^l = 0.35$ for primitive ($\tilde{\omega}_p$) and parameters $a_i^l = -20$ and $b_i^l = 0.35$ for non-primitive ($\tilde{\omega}_{np}$); b) generalized bell MF with parameters $c_i^l = 0.4$, $d_i^l = 15$ and $e_i^l = 0$ for the unchanged pixels ($\tilde{\xi}_u$), sigmoid MF with parameters $a_i^l = 15$ and $b_i^l = 0.8$ for the increase in the backscattering ($\tilde{\xi}^+$), and sigmoid MF with parameters $a_i^l = -15$ and $b_i^l = 0.8$ for the decrease in the backscattering ($\tilde{\xi}^-$). As one can notice some of the parameters are the same (or the opposite) resulting in a total of five parameters to be tuned.

where t is any t-norm operator [125], with $t : [0,1] \times [0,1] \rightarrow [0,1]$. In this work we use the algebraic product as t-norm, which is defined as follows

$$\mu_A^l(x) = \mu_{A_1^l \cap A_2^l \cap A_3^l}(x_1, x_2, x_3) = \mu_{A_1^l}(x_1)\mu_{A_2^l}(x_2)\mu_{A_3^l}(x_3) = \prod_{i=1}^{3} \mu_{A_i^l}(x_i) \tag{6.15}$$

Another common t-norm operator is the minimum. Nonetheless, the product preserves the contribution of each input variable and maintains a similarity with the original Bayesian framework. Thus the antecedents are reduced to a single value included between $[0,1]$. From this value it is possible to interpret the consequences of each fuzzy IF-THEN rule.

6.3.3 Implications

Once the antecedents are reduced it is necessary to evaluate for each rule how much this aggregated value is close to the consequence. For instance, if the first rule R^1 depicted in Fig. 6.2 is considered, the FIS has to evaluate how much the aggregated value of 0.5 represents the output $z_{c,j}$ is $(\tilde{\omega}_p, \tilde{\omega}_p)$. This is done by means of the so called Generalized Modus Ponens (GMP) reasoning. GMP is formulated for the proposed FIS using the fuzzy theory as follows [124]

$$\mu_{B_l'}(z) = \sup_{x \in \mathcal{U}} t[\mu_{A_l'}(x), \mu_{A^l \rightarrow B^l}(x, z)] \tag{6.16}$$

Eq. (6.16) says that given two fuzzy propositions: i) x is A_l' and ii) if x is A_l then z is B_l, it is possible to infer a new proposition: z is B_l' such that closer the A_l' to A_l, the closer

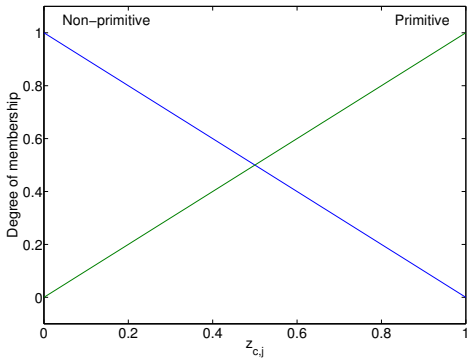


Figure 6.4: Example of triangular membership function with parameters $f = 0$, $h = 1$ and $g = 1$ (primitive) and parameters $f = 0$, $h = 1$ and $g = 0$ (non-primitive).

the B'_l to B_l. Thus, in order to infer the implication i) the output linguistic variables have to be described by MFs; and ii) the implication methods has to be defined. In this work we use triangular functions to describe in a linear way the grade of memberships between $\tilde{\omega}_p, \tilde{\omega}_{np}$ and the output linguistic variables. Generally speaking a triangular curve that maps the grade of membership of the linguistic variable z, defined in the universe of discourse $\mathcal{V}$, to a linguistic value B, is defined as

$$\mu_B^l(z) = \begin{cases} 0, & z < f^l \\ \frac{z-f^l}{g^l-f^l}, & f^l \leq z < g^l \\ \frac{h^l-z}{h^l-f^l}, & g^l \leq z < h^l \\ 0, & z \geq h^l \end{cases} \tag{6.17}$$

or more compactly by

$$\mu_B^l(z) = \max\left(\min\left(\frac{z-f^l}{g-f^l}, \frac{h^l-z}{h^l-f^l}\right), 0\right) \tag{6.18}$$

The parameters f^l and h^l locate the feet of the triangle and the parameter g^l locates the peak. For the considered problem these parameters can be set such that the maximum degree of membership for primitive and non-primitive is in 1 and 0, respectively, and the cross-over point is in 0.5 (see Fig. 6.4).

The implication method is defined exploiting fuzzy logic operators. By using different operators it is possible to obtain different results [124]. Considering crisp input and using the minimum implication (6.16) reduces to

$$\mu_{B'_l}(y) = \min\left[\mu_{A^l}(x^*), \mu_{B^l}(y)\right] \tag{6.19}$$

where x^* is some point in $\mathcal{U}$. This means that if the antecedent is only partially true, then the output fuzzy set is truncated according. In the example of Fig. 6.2 one can notice that the implication clips the output MF in accordance with the considered antecedents values.

At this point the decision about the pair of classes can be taken by selecting the rule that gives the highest grade of membership (R^1 in the case reported in Fig. 6.2). If all the rules return a grade of membership equal to zero (or lower than a given threshold $T_{\min}$ defined by the user), the result is not reliable and has to be rejected and treated accordingly. In addition, the FIS gives the possibility to infer the value $z_{\tau,j}$ in the domain of the monotemporal primitive features at a given time τ, which takes into account the multitemporal information. This is done in the steps described in the following.

6.3.4 Aggregation of Outputs

Since the fuzzy base defined by (6.5) consists of four rules, all the consequences have to be aggregated in order to infer the combined implication. As the rules are a set of independent conditional statements, the union of the fuzzy consequences can be used

$$\mu_{B'}(z) = \bigcup_{l=1}^{4} \mu_{B'_l}(z) = s\left[\mu_{B^1}(z), \cdots, \mu_{B^M}(z)\right] \tag{6.20}$$

where s is any s-norm operator [125]. In this work, inspired by the Bayesian formulation, the maximum operator to infer the aggregated implication for the multitemporal classification is taken into account. In detail, the maximum combination is defined as

$$\mu_{B'}(z) = \sum_{l=1}^{4} \mu_{B^l}(z) \tag{6.21}$$

The result of the combination of each rule is used in the next step in order to find the crisp value that represents the feature value at a given time τ considering the multitemporal information.

6.3.5 Defuzzification

The aim of this step is to map each aggregate output MFs into the space of the feature that describe the primitive (see the red block in Fig. 6.2). The best procedure to do this can vary from case to case. Several strategies have been described in the literature [125]. In this work the mean of maximum (MoM) method is used for the defuzzification of the fuzzy compound detector. Conceptually, the MoM defuzzifier chooses the value z^* as the mean point in which $\mu_{B'}(z)$ reaches its maximum values as depicted in Fig. 6.2. With refer to the canonical form for each rule the MoM is calculated as follows. Let us define m as the set of points at which $\mu_{B'}(z)$ achieves its maximum:

$$m(B') = \{z \in \mathcal{V} | \mu_{B'}(z) = \sup_{z \in \mathcal{V}} \mu_{B'}(z)\} \tag{6.22}$$

From $m(B')$ the MoM defuzzified value z^* can be derived as

$$z^* = \frac{\int_{m(B')} z \, \mathrm{d}z}{\int_{m(B')} \mathrm{d}z} \tag{6.23}$$

Table 6.1: Algorithm of the proposed FIS.

Initialization: define cross-over points and domains of the input MFs;
foreach *pixels of the multitemporal images* **do**
  foreach *rule* R^l **do**
    Aggregate the antecedents according to $\mu_A^l(x) = \prod_{i=1}^{3} \mu_{A_i^l}(x_i)$ (6.15);
    Imply the conclusions according to $\mu_{B_l'}(z) = \min\left[\mu_{A^l}(x^*), \mu_{B^l}(z)\right]$ (6.16);
  end
  if $\arg\max_l \mu_{B_l'} > 0$ **then**
    Assign the class pair (ω_k^l, ω_k^l) to the pair of features $(x_{1,j}, x_{2,j})$;
  end
  Combine the M fuzzy rules according to $\mu_{B'}(z) = \sum_{l=1}^{4} \mu_{B^l}(z)$ (6.21);
  Infer the crisp output $z^* = \frac{\int_{m(B')} z \, \mathrm{d}z}{\int_{m(B')} \mathrm{d}z}$ (6.23) for each output $z_{\tau,j}$;
end

The inferred crisp value can be used in all the applications that use the primitive features such as the classification. In table 6.1 the algorithm that implements the proposed FIS is reported.

6.4 Experimental Results

In this section we illustrate the performance of the proposed approach to primitive detection in multitemporal SAR images. In order to assess the effectiveness of the proposed approach, experiments were carried out on both a series of simulated SAR images and a series of real multitemporal TerraSAR-X images acquired over the city of Lüneburg, Germany during one year.

6.4.1 Design of Experiments and System Setup

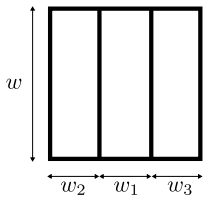


Figure 6.5: Window used by the line detector.

In this sub-section we introduce the primitive detector used in the experiments. Note that the proposed multitemporal detector approach is general and can be used with any detectors. The considered primitives are bright linear primitives. Bright lines are common primitives especially when urban areas are illuminated by a SAR sensor with an adequate resolution. They are generated by multiple reflections of the electromagnetic waves with the structured objects on the ground [130]. Several line detectors have been used to identify bright lines such as the Tupin detector [131], the Steger detector [132],

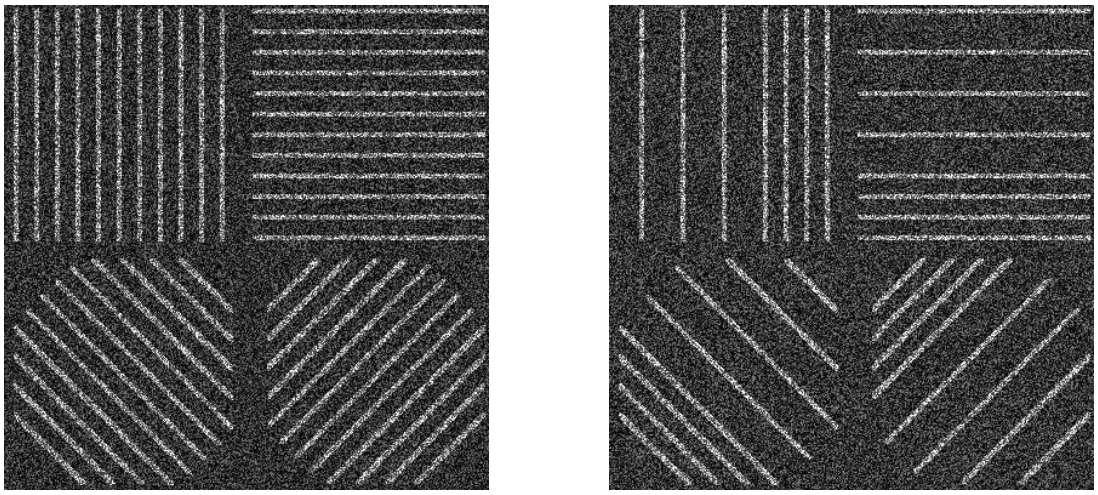
the morphological detector [119], and the Hough transform [133]. All these detectors are defined for a monotemporal SAR image. In this work bright lines are extracted using the line detector proposed in [131]. This is a consolidated approach that has been used in the literature for applications that range from road detection to building detection. The detector consists of the fusion of two statistical and geometrical line detectors, which are applied using a sliding window procedure. For the sake of simplicity, in this work we consider only one of these two detectors, namely the ratio line detector. Its response is the minimum response of the ratio edge detector, firstly described in [134], applied to both sides of a central linear structure

$$r_L = \min(r_E^{(1,2)}, r_E^{(1,3)}) \tag{6.24}$$

where the superscript 1 denotes the central region, superscripts 2 and 3 the two lateral regions (see Fig. 6.5), and $r_E^{(t,u)}$, with $(t,u) \in \{(1,2),(1,3)\}$, is constructed in order to extract only bright edge as follows

$$r_E^{(t,u)} = \begin{cases} 1 - \min\left(\frac{\sigma_t}{\sigma_u}, \frac{\sigma_u}{\sigma_t}\right) & \text{if } \sigma_1 \geq \sigma_2, \sigma_3 \\ 0 & \text{Otherwise.} \end{cases} \tag{6.25}$$

where σ_t is the empirical mean of the region $t = 1$ having $W_t = w \times w_t$ pixels with amplitude A_p^t, i.e., $\sigma_t = (\frac{1}{W_t}) \sum_{p \in t} A_p^t$. The same definition is applied to σ_u (with $u = 2, 3$) i.e., $\sigma_u = (\frac{1}{W_u}) \sum_{p \in u} A_p^u$. In order to detect bright lines with different directions the central linear structure is rotating in 16 different directions over the central pixel and the best response retained. The response of the line detector is high when the detector falls across to a line with a width comparable to the width of the central structure, whereas it is small when the detector encounters smooth areas.



(a) (b)

Figure 6.6: Simulated SAR images. (a) and (b) are the multitemporal images used to represent the scenario of mixed stable and unstable primitives. The sequence of images is corrupted by timely independent speckle noise.

The results are analyzed considering stable primitives (which may slightly vary because

of speckle or different acquisition conditions) and unstable primitives (which change their state). The results of the proposed method are compared with those obtained by the bright line detector applied at each observation independently. Moreover, the performance obtained after applying the multitemporal filtering in its recursive form introduced in [37] is shown for both the proposed method and the ratio line detector.

6.4.2 Simulated Data Set

In order to assess the performance of the proposed approach a set of simulated images has been created. A series of $\tau = 1, \ldots, 14$ images with size 400×400 representing bright lines on black background has been generated. In detail, it represents a scenario where 16 new lines are disappearing/appearing at each new observation in an interleaved manner. The lines have a length that varies from 62 to 190 pixels and a width included in a range of 2 to 4 pixels. These images, which ideally represent the backscattering coefficient σ_τ, are corrupted by speckle noise. In the case of spotlight VHR SAR acquisitions the assumption of distributed scatterer is not verified and therefore the fully developed speckle model may be not valid [135]. Nonetheless, as no accepted speckle statistical models are available for VHR SAR images, in first approximation we corrupt all the images with multiplicative spatially uncorrelated speckle noise under the hypothesis of fully developed speckle. In detail the intensity image $\mathbf{X}_\tau$ has been related to the underlying backscattering coefficient σ_τ as [17]

$$\mathbf{X}_\tau = \sigma_\tau n_\tau \tag{6.26}$$

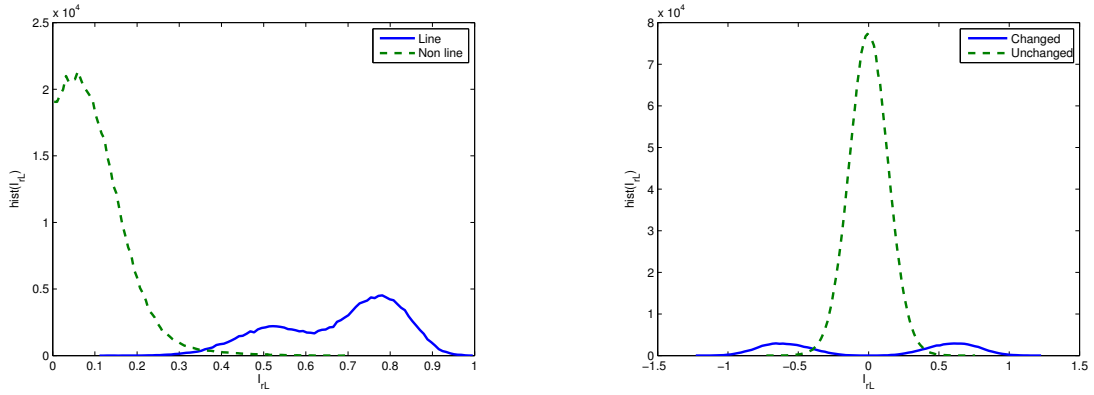
with n_τ being a single-independent random variable whose probability density function (PDF) for an L-looks intensity data is described as

$$P(n_\tau) = \frac{L^L n_\tau^{L-1}}{\Gamma(L)} \exp(-L n_\tau) \tag{6.27}$$

Examples of images obtained with $L = 1$ are reported in Fig. 6.6. It is worth noting that in this simulation speckle varies randomly so each realization represents a scene that has identical statistic but different detail structure.

In order to apply the proposed FIS for multitemporal detection of primitives at time τ, the bright line detector described in the previous section has been used. The window size was set with length of 10 pixels and width of the central structure of 3 pixels. The index of change has been derived using the second level of the log-ratio Wavelet decomposition as described in section 6.2. The obtained image $\mathbf{X}^2_{\mathrm{LR}}$ has a resolution comparable to the size of the bright lines to be detected. Hence changes smaller than this resolution tends to disappear and only changes of a given size are fully preserved, allowing a better detection of the changed bright lines.

Fig. 6.7a reports the mean histogram for the pixels of line and non-line in the simulated images. As one can notice, the distribution of the class ω_p cannot be parametrized using standard statistical models since it is a mixture of different contributions due to the different contrasts measured by the ratio line detector inside the sliding windows [131]. Also nonparametric density estimators may fail if the distribution is not continuous or the bandwidth parameter is not selected in a proper way. As described in Sec. 6.11d, in



(a) (b)

Figure 6.7: Histogram of the two classes line (solid blue) and non line (dashed green) considering the labeled pixels of all the 14 images (simulated data set).

order to evaluate the presence or absence of a line, the proposed FIS needs the definition of the cross-over point a and slope b. For a fixed value of b, the missed alarms decrease by decreasing the value of a, while the false alarms increase. Vice-versa by increasing the value of the a the false alarms decrease, while the missed alarms increase. Therefore, as for the definition of the threshold of the bright line detector presented in [131], the crossover point a may be deducted as a compromise between false alarm and a minimum detectable contrast of the lines. Nonetheless, the slopes b of the MFs, which define the flexibility to cope with uncertain situations, define a new grade of freedom making the selection less strict. Indeed by assign a small value to b the missed alarms are reduced, while the false alarms increase, and vice-versa. A comprehensive test confirms that the values $a = 0.35$ and $b = 20$ lead to the minimum error for the considered data set. By analyzing Fig. 6.7a it is possible to localize the intersection point between the histograms of the two classes at the value of 0.35. By selecting a value of 20 the support of the MFs are equal to $[0, 0.5]$ for the non-line and $[0.2, 1]$ for the line, in accordance with the histogram reported in Fig. 6.7a.

Fig. 6.7b reports the histogram of the classes unchanged (ξ_u), increase (ξ^+) and decrease (ξ^-) of backscattering for $\mathbf{X}^2_{\mathrm{LR}}$. As demonstrated in [55], by exploiting a split-based thresholding technique the, it is possible to derive the two points of intersection among the three distributions i.e., ± 0.4. These points are used as cross-over points a and c for the two sigmoidal MFs defined by (6.7) and the generalized bell MF defined by (6.12). Whereas, the slopes b and d of these MFs were defined equal to 15 defining a support for ξ_u of $[-0.5, 0.5]$, for ξ^+ of $[0.5, +\infty]$ and for ξ^- of $[-0.5, -\infty]$. A comprehensive test confirm that these are the best values in terms of reduction of the overall error.

Performance on the Detection of Primitives at a Given Time

In the first battery of experiments the accuracy of the primitive detection at a given time is evaluated. For each primitive detection method, the best threshold at each time τ has been computed considering the Receiver Operating Characteristic (ROC). The average number

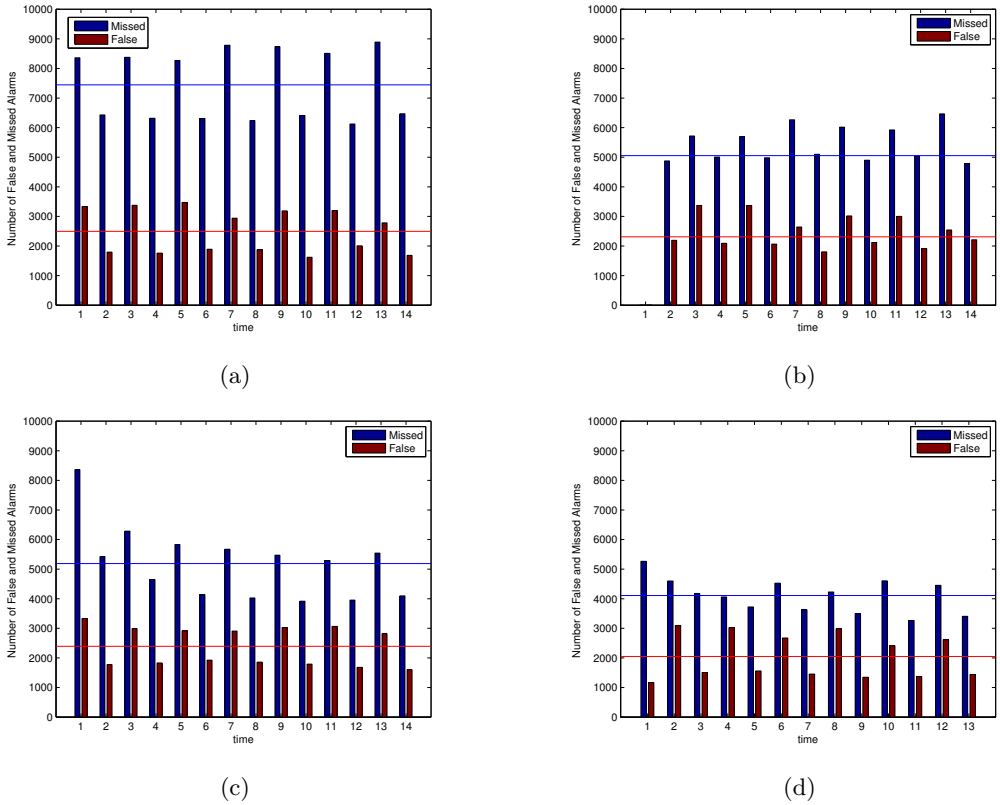


Figure 6.8: Number of Missed and False alarm pixels for: a) Ratio line detector applied independently to each time; b) proposed multitemporal detector applied considering two consecutive times; c) Ratio line detector after the recursive application of the temporal de-speckling proposed in [37]; d) Proposed multitemporal detector applied considering two consecutive times after the multitemporal de-speckling proposed in [37]. Average results are reported in solid lines (simulated data set).

of missed and false alarm pixels for the 14 observations are reported in Table 6.2. From the comparison of these results it is possible to derive different observations. First, the proposed method performs better than the ratio bright line detector applied independently to each image obtaining a reduction of the overall error of 2252 pixels. Moreover, from the comparison of Figg. 6.8(a) and (b) one can notice that the proposed approach is more stable than the bright line detector applied at each single date. This because the proposed method exploits the temporal information in addition to the monotemporal information.

Table 6.2: Average False Alarms, Missed Alarms and Overall Errors for the Bright Line Detection. Results Obtained without De-Speckling (No Despeck) and with De-Speckling (Despeck) Are Given (Simulated Data Set).

	Missed Alarms		False Alarms		Overall Errors	
Approach	No despeck	Despeck	No despeck	Despeck	No despeck	Despeck
Mono-temporal Bright Line Detection	7444	5189	2494	2393	9938	7582
Proposed Approach	4933	4109	2316	2049	7249	6158

Finally we can observe that the pre-processing using a multitemporal de-speckling filter can improve the detection. It is interesting to note that the overall error of the proposed method without pre-processing is lower than the overall error obtained by applying the bright line detectors independently after the multitemporal de-speckling (i.e., 7515 against 7249 pixels).

Performance on the Detection in the Transitions of Primitive State

The capability of the proposed method to identify the transitions of the state of primitives is evaluated by calculating the confusion matrix. In this matrix the true transitions are given on the columns and the estimated transitions detected by the proposed approach on the rows. The terms on the diagonal of the matrix report the correctly recognized transitions while the other terms give the errors. Average results for the 14 images are reported in Table 6.3.In detail, Table 6.3.a reports the confusion matrix obtained by the post-detection comparison and Table 6.3.b the confusion matrix obtained by the proposed method. In both cases no pre-processing was performed in the experiments. The proposed method produces good results i.e., an overall accuracy of 93.05% and a Kappa coefficient of 0.8015, which are higher than the accuracies obtained with the post-detection comparison i.e., 89.95% and 0.7179. It is possible to note that the main source of confusion for the method is given by the transitions $(\tilde{\omega}_p, \tilde{\omega}_{np})$ and $(\tilde{\omega}_{np}, \tilde{\omega}_p)$. This is explained by the fact that the exploited index of change performs as a trade-off between geometrical detail preservation and speckle filtering. Since we used a resolution for the index close to the size of the lines the errors are localized along the border pixels of the lines.

Table 6.3: Confusion Matrix of Primitives State Transitions Obtained by: (a) the Post-Detection Comparison; and (b) the Proposed Method (Simulated Data Set).

		True Transitions			
		(ω_p, ω_p)	(ω_p, ω_{np})	(ω_{np}, ω_p)	$(\omega_{np}, \omega_{np})$
Estimated Transitions	(ω_p, ω_p)	13157	59	50	670
	(ω_p, ω_{np})	1902	4507	15	2237
	(ω_{np}, ω_p)	1893	18	3865	2199
	$(\omega_{np}, \omega_{np})$	3152	1478	1265	112105
Overall Accuracy		89.95			
Kappa		0.7179			

(a)

		True Transitions			
		(ω_p, ω_p)	(ω_p, ω_{np})	(ω_{np}, ω_p)	$(\omega_{np}, \omega_{np})$
Estimated Transitions	(ω_p, ω_p)	14633	14	41	514
	(ω_p, ω_{np})	1072	4748	15	1150
	(ω_{np}, ω_p)	1069	8	4772	1143
	$(\omega_{np}, \omega_{np})$	2212	725	708	113209
Overall Accuracy		93.05			
Kappa		0.8015			

(b)

(a) (b)

Figure 6.9: Lüneburg data-set: (a) optical image [85]; and (b) mean intensity of the 14 images. Viewing direction from the left.

6.4.3 Real Data-Set: Lüneburg

A stack of 14 high resolution spotlight 300-MHz TerraSAR-X data of the town of Lüneburg, situated in Lower Saxony, Germany has been considered in order to evaluate the proposed method. The images were acquired between December 2010 and September 2011 with an almost constant revisiting time of 22 days (see Fig. 6.10). The parameters are the same for all the acquisitions with a nominal incidence angle of 41 degree, a HH polarization and a nominal resolution of 0.6 m in slant range and 1.1 m in the azimuth direction. All the acquisitions have been taken at around 5:30 AM Coordinated Universal Time (UTC). The weather conditions were variable from snow, to rain and dry. Fig. 6.9b shows the mean intensity of the 14 images for the considered portion of the full scene made up of 2048×2048 pixels. Fig. 6.9a shows an optical image of the same area.

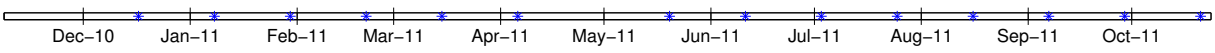


Figure 6.10: Temporal distribution of the TerraSAR-X data acquired on Lüneburg, Germany.

In this scene, bright lines are mainly generated by the multiple reflections of the electromagnetic waves with the several corners present in the structured home-made objects. As one can notice bright lines have different length and width. In order to extract all of them the ratio line detector should be applied with different values for width and length. In this work we report the results of the extraction of lines with a minimum length of 10 pixels and a width of 3 pixels (similar results have been obtained for different ratio line detector parameters). Some of these lines are stable and present a contrast that renders them easily detectable by the line ratio detector. Others have a contrast that is variable.

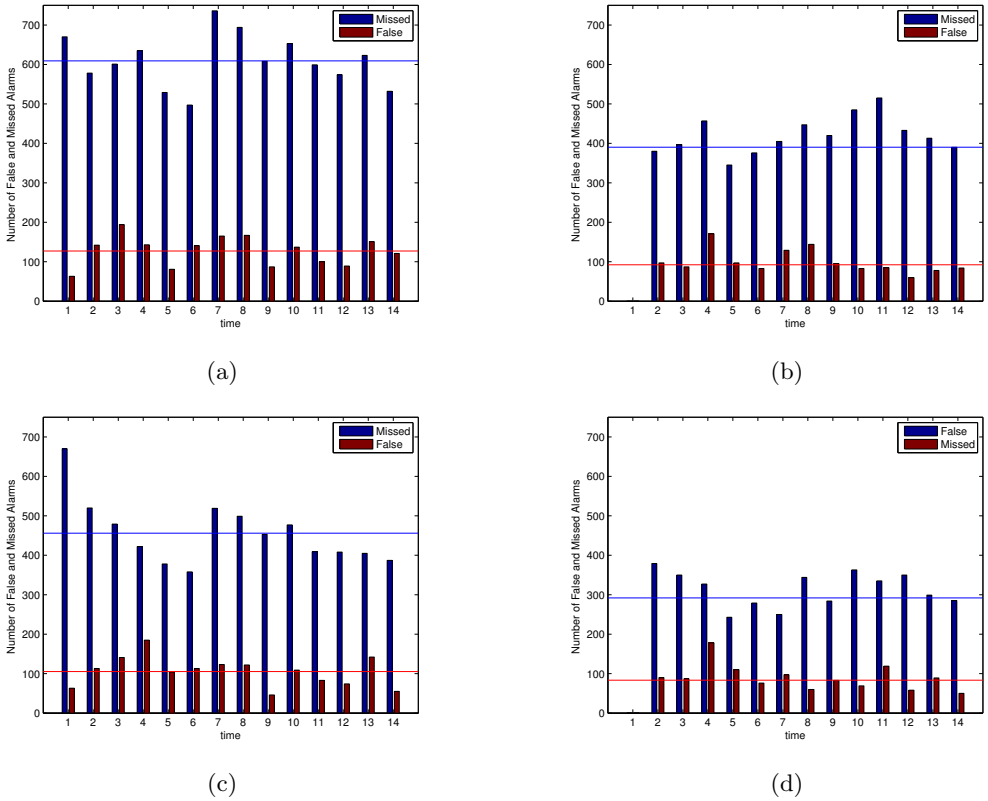


Figure 6.11: Number of Missed and False alarm pixels for: a) Ratio line detector applied independently to each time; b) Proposed multitemporal detector applied considering two consecutive times; c) Ratio line detector after the recursive application of the temporal de-speckling proposed in [37]; d) Proposed multitemporal detector applied considering two consecutive times after the multitemporal de-speckling proposed in [37]. Average results are reported in solid lines (real data-set).

This behavior is probably due to the different weather conditions. Some others bright lines change between the different acquisitions. This is the case of the water treatment plant in the upper part of the image. Here a skimmer arm is rotating inside each of the four circular clarifier basins generating a line that changes positions acquisition to acquisition. This is a very good test target for the proposed approach. Due to the temporal de-correlation, we expect that speckle affects differently each acquisition. In the following the performance of the proposed approach are analyzed quantitatively by considering a reference map generated by photo-interpretation. The five parameters needed for the proper working of the proposed FIS have been set in order to reduce the missed and false alarms. Due to the similarity with the simulated data set illustrated in the previous section only the slopes have been modified in order to better cope with the more uncertain source of information and set to 8.

Table 6.4: Average False Alarms, Missed Alarms and Overall Errors for the Bright line detection. Results Obtained without De-Speckling (No Despeck) and with De-Speckling (Despeck) Are Given (Real data-set).

Approach	Missed Alarms		False Alarms		Overall Errors	
	No despeck	Despeck	No despeck	Despeck	No despeck	Despeck
Mono-temporal Bright Line Detection	609	456	127	105	736	561
Proposed Approach	390	292	92	83	482	375

Performance on the Detection of Primitives at a Given Time

As done for the simulated data set, the best threshold at each time τ has been firstly computed considering the Receiver Operating Characteristic (ROC) for the two methods. The average numbers of missed and false alarm pixels for the 14 observations are reported in Table 6.4. From the comparison of these results it is possible to confirm the observation done in the previous section. First the proposed method performs better than the ratio bright line detector applied independently to each observation (the overall error is halve) and it is more stable (see Fig. 6.11(a) and (b)). Moreover, the pre-processing using a multitemporal filtering improves also in this case the detection. The outcomes of this experiment confirm the results obtained on the simulated data set.

Table 6.5: Confusion Matrix of Primitives State Transitions Obtained by: (a) the Post-Detection Comparison; and (b) the Proposed Method (Real Data-Set).

		True Transitions			
		(ω_p, ω_p)	(ω_p, ω_{np})	(ω_{np}, ω_p)	$(\omega_{np}, \omega_{np})$
Estimated Transitions	(ω_p, ω_p)	5014	110	108	10
	(ω_p, ω_{np})	398	255	0	93
	(ω_{np}, ω_p)	291	0	261	89
	$(\omega_{np}, \omega_{np})$	348	0	1	5283
Overall Accuracy		88.19			
Kappa		0.7964			

(a)

		True Transitions			
		(ω_p, ω_p)	(ω_p, ω_{np})	(ω_{np}, ω_p)	$(\omega_{np}, \omega_{np})$
Estimated Transitions	(ω_p, ω_p)	5546	103	153	56
	(ω_p, ω_{np})	126	262	0	41
	(ω_{np}, ω_p)	126	0	216	57
	$(\omega_{np}, \omega_{np})$	253	0	1	5321
Overall Accuracy		92.52			
Kappa		0.8664			

(b)

Performance on the Detection in the Transitions of Primitive State

Also in this case the capability of the proposed method to identify the transitions of the state of primitives is evaluated by calculating the confusion matrix. Average results for the 14 observations are given in Table 6.5. In detail, Table 6.5.a reports the confusion matrix obtained by the post-detection comparison and Table 6.5.b the confusion matrix obtained by the proposed method. In both cases no pre-processing was performed. As one can notice, the proposed method obtained an overall accuracy 4.33% higher than the post-detection comparison i.e., 92.52% against 88.19% and a Kappa coefficient of 0.86 against 0.79. This confirms the validity of the proposed compound detector also on this last case.

6.5 Conclusion

In this chapter a novel approach to multitemporal detection of primitives in VHR SAR images has been presented. The approach exploits the multitemporal information in order to: i) perform the detection of primitives at a given time; and ii) identify the transitions in the state of primitives between two consecutive observations. This is done by introducing a compound detection approach based on the Bayesian framework. Due to the difficulty to estimate the statistical terms involved in the decision rule in an unsupervised and parametric way, the proposed compound detection mechanism is integrated with the fuzzy logic. Fuzzy logic is used in order to test the semantic consistency among the state of primitives at the different temporal observations and infer the state of a primitive at a given time. This allows the compound detection to avoid the problems of the unsupervised statistical estimation. Note that, even though the proposed approach is presented for primitive detection in VHR SAR images it has general validity and can be exploited for primitive detection in high- and medium-resolution SAR images.

The effectiveness of the proposed approach has been illustrated on both simulated and real VHR TerraSAR-X images by using a detector for bright line identification. When compared to the independent application of the detector at each single time image, the proposed method exhibited significantly higher accuracy and a better stability on both data sets. In the context of the detection of primitive transitions, the proposed method has been compared with the results obtained by a post-detection technique, showing a higher accuracy in both the considered data sets.

With regards to the computation load, the only time consuming step is the derivation of the primitive response using the considered bright line detector. It takes around 5 minutes to elaborate a portion of 4 million pixels of a TerraSAR-X image using a computer with 4GB RAM and four 3.10 GHz CPUs. The time necessary for the inference process is in the order of 55 seconds rendering the overall computational demand reasonable.

As a future development we are investigating an extension of the proposed approach to the analysis of long time series of images. This is particularly important given the forthcoming availability of long image time series acquired by SAR satellite systems (e.g., the ESA Sentinel 1 mission). Moreover, a formulation of the proposed method for solving problems of multi-source primitive detection is under investigation.

Acknowledgment

The authors would like to thank the German Aerospace Center (DLR) for providing the TerraSAR-X images of Lüneburg which were acquired in the framework of the International Society for Photogrammetry and Remote Sensing (ISPRS) working group VII/2 and TerraSAR-X Project LAN0634. This research was supported by the Italian Ministry of the Education, Research and University (MIUR) under the project "A novel integrated system for the analysis of very high spatial and spectral resolution remote sensing data".

Chapter 7

Conclusion

This chapter draws the general conclusions on the methods for change detection in VHR multitemporal SAR images developed in the thesis and presents an outlook of the possible future developments.

7.1 Summary and discussion

In this thesis we analysed and modelled the main properties of multitemporal VHR SAR remote sensing images in order to develop effective unsupervised change-detection methods. In particular, we derived: i) an approach to building change detection; ii) an approach to rapid and accurate damage detection in built-up areas combining stripmap and spotlight images; iii) a hierarchical approach to change detection for surveillance applications; iv) a multitemporal detector for the extraction of primitives. For each topic an analysis of the state of the art was conducted, and the limitations of literature techniques highlighted. Starting from this analysis, novel solutions were theoretically developed, implemented and finally applied to real remote sensing data in order to assess their effectiveness.

With respect to the detection of changed building in chapter 3 a novel approach that takes advantage from a theoretical modeling of the backscattering mechanisms of new and fully destroyed buildings in multitemporal VHR SAR images has been introduced. In detail, new and destroyed buildings are identified by a pattern made up of an area of both increase and decrease of backscattering with spatial properties and alignment depending on the proprieties of the buildings. In order to extract the changes associated with increase and decrease of backscattering, the proposed approach makes use of a multiscale representation of the multitemporal information allowing a detection of changes at the optimal building scale. This information is used to identify the candidates to be changed buildings. The building candidates are analyzed in order to properly detect the new or destroyed building by means of four fuzzy rules. The fuzzy rules are formulated by taking into account the proposed multitemporal changed building model. The aggregated membership resulting from the application of the fuzzy rules makes it possible to identify the class of each building candidate (i.e., new/destroyed building or general change with size comparable to the building size but not related to new or destroyed building). The use of both this fuzzy rule set and the intrinsic reliability obtained when working on the

comparison of two images rather than on each single image has demonstrated its robustness when applied to two different real challenging scenarios related to the detection of demolished and new buildings.

In chapter 4 the two SAR acquisition modes called StripMap (SM) and SpotLight (SL) provided by the new generation of SAR satellite missions has been exploited to detect damages after catastrophic events. The two complementary SAR modes have been used to perform different duties during the emergency response phase. In detail SM images, which are characterized by a wide coverage (around 1500 Km^2) and a high resolution (3 to 5 m), have been exploited to detect the urban changed areas (hot-spots) spread around the large observed territory. This has been done using a multiscale strategy based on the Wavelet transform. The obtained hot-spots are then used to drive the selection from the archive of pre-crisis spotlight SAR images, which are characterized by a moderate coverage (around 100 Km^2) and a very high resolution (1 m), and the acquisition of new post-crisis VHR SAR images. These images are used to perform change detection at high spatial resolution with the goal to detect collapsed buildings. This analysis is conducted by taking advantage of the theoretical model of new/destroyed building and the detection method illustrated in chapter 3 by exploiting the Curvelet transform (instead of Wavelet) in order to better preserve the linear details that characterize urban structures in VHR SAR images. The proposed method demonstrated to be effective in quickly detecting the most affected urban areas and accurately identify the demolished buildings. This was possible because of both the multiscale representation of the multitemporal information and the high robustness of the proposed building change detection method to deviate from the ideal model considered for changed buildings.

In chapter 5 the high-complexity backscattering behaviors due to wide diversity of man-made structures typical of a home-made scenarios has been addressed. The proposed approach identifies changed areas (hot-spots) at different resolution levels according to a multiscale technique based on the 2-dimensional stationary Wavelet transform. Each hot-spot is analyzed according to its spatial position within a specific area of interest by using the available prior information. The use of such information allows one to define the best change detectors to be used for extracting the expected changes at fine scale. It is worth noting that the assumption on the availability of the prior information on the usage of different areas (and thus on the expected kinds of change) is reasonable since the method has been developed for high frequency surveillance/monitoring of sensitive areas such as maritime ports, airports, and so on. The proposed method, applied in two different real scenarios, has demonstrated to be effective in detecting all the expected kinds of change with a high geometrical accuracy despite the changes were associated to objects with significantly different properties in terms of shape, modeling and size. This is possible because of the hierarchical extraction of hot spots of change and of the specific definition on the basis of the available prior information of feature extraction techniques and change detectors.

In chapter 6 a novel approach that exploits the multitemporal information in order to: i) perform the detection of primitives at a given time (e.g., lines, object with specific shape); and ii) identify the transitions in the state of primitives between two consecutive observations has been presented. This is done by introducing a compound detection ap-

proach based on the Bayesian framework. Nonetheless, due to the difficulty in estimating the statistical terms in an unsupervised way, the proposed compound detection mechanism is integrated with the fuzzy logic. Fuzzy logic is used in order to test the semantic consistency among the state of primitives at the different temporal observations and infer the state of a primitive at a given time. This allows the compound detection to avoid the problems of the unsupervised statistical estimation. Even though the proposed approach is presented for primitive detection in VHR SAR images it has general validity and can be exploited for primitive detection in high- and medium-resolution SAR images. The effectiveness of the proposed approach has been illustrated on both simulated and real VHR SAR images by using a detector for bright line identification. When compared to the independent application of the detector at single time image, the proposed method exhibits significantly higher accuracy and a better stability on both data sets. In the context of the detection of primitive transitions, the proposed method has been compared with the results obtained by a post-detection technique, showing a higher accuracy in both the considered data sets.

Each chapter contributes to improve the state of the art on the analysis of multitemporal VHR SAR images, facing and solving relevant problems related to their analysis.

7.2 Concluding remarks and future developments

In this research activity we developed approaches that improve the capability of identifying changes in multitemporal VHR SAR images. In detail, we address the issues related to the correct exploitation of the very high geometrical content considering real challenging change detection scenarios. In this context some issues remain open and need to be addressed in future developments. We identify (among the others) the future developments reported in the following.

With the launch of the second generation of COSMO-SkyMed constellation there will be the availability of multitemporal high resolution polarimetric data. The polarimetric information allows the separation among different scattering mechanisms (e.g., between volumetric and multiple scattering). Therefore the possibility to exploit this information in order to detect partially damaged buildings should be investigated.

With the passage of time the archive of high and very high resolution SAR images of the current and incoming missions will become richer and richer. In order to properly exploit this extremely high amount of data, the proposed methods should be extended in order to cope with the challenging problems of long-time series e.g., high amount of data to be managed and processed.

The possibility of acquiring data with different incidence angles and with different orbit nodes (i.e., ascending and descending) may be also exploited to solve the effects of interaction and occlusions of neighboring objects with respect the changed target of interest e.g., collapsed buildings.

The cogent time requirements of real emergency scenarios may be addressed by the integration of the proposed approach for rapid and accurate damage detection in an interoperability framework that is able to use the data coming from the different SAR missions currently available allowing a considerable reduction of the response time.

New primitive/object detectors to identify different kind of change should be developed taking into account the works already presented in the field of target detection. This may contribute to a significant improvement of the performance of change detection approaches for surveillance applications.

List of Publications

International Journals

[J1] C. Marin, F. Bovolo, L. Bruzzone, *"A Hierarchical Approach to Change Detection in Very High Resolution SAR Images for Surveillance Applications"* in IEEE Transactions on Geoscience and Remote Sensing, v. 51, n. 4 part I (2013), p. 2042-2054.

[J2] F. Bovolo, C. Marin, L. Bruzzone, *"Building Change Detection in Multitemporal Very High Resolution SAR Images"* in IEEE Transactions on Geoscience and Remote Sensing, v. 53, n. 5 (2015), p. 2664-2682.

[J3] C. Marin, F. Bovolo, L. Bruzzone, *"A Novel Multitemporal Detector for the Extraction of Primitive from Very High Resolution SAR Images "* in IEEE Transactions on Geoscience and Remote Sensing, (in review).

Book Chapters

[B1] L. Bruzzone, C. Marin, F. Bovolo, *"Damage Detection in Built-up Areas Using SAR Images."* In: Beer M., Patelli E., Kougioumtzoglou I., Au I. (Ed.) Encyclopedia of Earthquake Engineering: SpringerReference (www.springerreference.com). Springer-Verlag Berlin Heidelberg.

International Conferences

[C1] C. Marin, F. Bovolo, L. Bruzzone, *"Rapid and Accurate Damage Detection in Built-up Areas Combining Stripmap and Spotlight SAR Images"* in IEEE International Geoscience and Remote Sensing Symposium (IGARSS), 13-18 July 2014.

[C2] C. Marin, F. Bovolo, L. Bruzzone, *"A novel multitemporal detector for primitive extraction from VHR SAR images"* in IEEE International Geoscience and Remote Sensing Symposium (IGARSS), 21-27 July 2013.

[C3] C. Marin, F. Bovolo, L. Bruzzone, *"Detection of changed buildings in multitemporal Very High Resolution SAR images"* in IEEE International Geoscience and Remote Sensing Symposium (IGARSS), 21-27 July 2013.

[C4] F. Bovolo, C. Marin, L. Bruzzone, *"A novel approach to building change detection in very high resolution SAR images"* in SPIE Conference on Image and Signal Processing for Remote Sensing XVII, SPIE 2012, Edinburgh, United Kingdom, 24-27 September 2012.

[C5] F. Bovolo, C. Marin, L. Bruzzone, *"A novel hierarchical approach to change detection with very high resolution SAR images for surveillance applications"* in IEEE International Geoscience and Remote Sensing Symposium (IGARSS), 22-27 July 2012.

[C6] S. Serpico, L. Bruzzone, G. Corsini, W. Emery, P. Gamba, A. Garzelli, G. Mercier, J. Zerubia, N. Acito, B. Aiazzi, F. Bovolo, F. Dell'acqua, M. De Martino, M. Diani, V. Krylov, G. Lisini, C. Marin, G. Moser, A. Voisin, C. Zoppetti, *"Development and Validation of Multitemporal Image Analysis Methodologies for Multirisk Monitoring of Critical Structures and Infrastructures"* in IEEE Internatioal Geoscience and Remote Sensing Symposium (IGARSS), Munich, Germany, 22-27 July 2012.

[C7] F. Bovolo, L. Bruzzone, C. Marin, *"A Multilevel Approach to Change Detection for Port Surveillance with Very High Resolution SAR Images"* in IEEE Sixth International Workshop on the Analysis of Multi-Temporal Remote Sensing Images, Trento, Italy, 12-14 July 2011.

Bibliography

[1] S. Liang, *Advances in Land Remote Sensing: System, Modeling, Inversion and Application*. Springer Netherlands, 2008. 1

[2] G. Joseph, *Fundamentals of Remote Sensing*. Universities Press, 2005. [Online]. Available: https://books.google.it/books?id=peUFLCy4zLYC 1

[3] M. K. Ridd and J. Liu, "A comparison of four algorithms for change detection in an urban environment," *Remote Sensing of Environment*, vol. 63, no. 2, pp. 95 – 100, 1998. 1

[4] T. Hame, I. Heiler, and J. San Miguel-Ayanz, "An unsupervised change detection and recognition system for forestry," *International Journal of Remote Sensing*, vol. 19, no. 6, pp. 1079–1099, 1998. 1

[5] L. Bruzzone and S. Serpico, "An iterative technique for the detection of land-cover transitions in multitemporal remote-sensing images," *IEEE Trans. Geosci. Remote Sens.*, vol. 35, no. 4, pp. 858 –867, jul 1997. 1, 94

[6] D. Brunner, G. Lemoine, and L. Bruzzone, "Earthquake damage assessment of buildings using VHR optical and SAR imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 5, pp. 2403 –2420, may 2010. 1, 31, 32, 35, 46, 59, 75, 92

[7] F. Dell'Acqua and P. Gamba, "Remote sensing and earthquake damage assessment: Experiences, limits, and perspectives," *Proceedings of the IEEE*, vol. 100, no. 10, pp. 2876–2890, Oct 2012. 2, 58, 62, 63

[8] S. Quegan, T. Le Toan, J. Yu, F. Ribbes, and N. Floury, "Multitemporal ERS SAR analysis applied to forest mapping," *IEEE Trans. Geosci. Remote Sens.*, vol. 38, no. 2, pp. 741 –753, mar 2000. 2, 17

[9] T. Le Toan, F. Ribbes, L.-F. Wang, N. Floury, K.-H. Ding, J. A. Kong, M. Fujita, and T. Kurosu, "Rice crop mapping and monitoring using ERS-1 data based on experiment and modeling results," *IEEE Trans. Geosci. Remote Sens.*, vol. 35, no. 1, pp. 41 –56, jan 1997. 2

[10] D. Mason, R. Speck, B. Devereux, G.-P. Schumann, J. Neal, and P. Bates, "Flood detection in urban areas using TerraSAR-X," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 2, pp. 882 –894, feb. 2010. 2, 75

[11] R. Werninghaus and S. Buckreuss, "The TerraSAR-X Mission and System Design," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 2, pp. 606 –614, feb. 2010. 2

[12] G. Krieger, A. Moreira, H. Fiedler, I. Hajnsek, M. Werner, M. Younis, and M. Zink, "Tandem-x: A satellite formation for high-resolution sar interferometry," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 45, no. 11, pp. 3317–3341, Nov 2007. 2

[13] F. Caltagirone, G. De Luca, F. Covello, G. Marano, G. Angino, and M. Piemontese, "Status, results, potentiality and evolution of COSMO-SkyMed, the Italian Earth Observation constellation for risk management and security," in *Geoscience and Remote Sensing Symposium (IGARSS), 2010 IEEE International*, july 2010, pp. 4393 –4396. 2

[14] Y. Bazi, L. Bruzzone, and F. Melgani, "An unsupervised approach based on the generalized Gaussian model to automatic change detection in multitemporal SAR images," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 4, pp. 874 – 887, april 2005. 3, 16, 17, 18, 26

[15] J. Inglada and G. Mercier, "A New Statistical Similarity Measure for Change Detection in Multitemporal SAR Images and Its Extension to Multiscale Change Analysis," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 5, pp. 1432 –1445, may 2007. 3, 20, 22, 75

[16] J. Curlander and R. McDonough, *Synthetic Aperture Radar: Systems and Signal Processing*, ser. Wiley Series in Remote Sensing and Image Processing. Wiley, 1991. 11, 14

[17] C. Oliver and S. Quegan, *Understanding Synthetic Aperture Radar Images*, ser. Artech House remote sensing library. Artech House, 1998. 12, 105

[18] ——, *Understanding Synthetic Aperture Radar Images*, ser. SciTech radar and defense series. SciTech Publ., 2004. 14, 32

[19] ceos and esa, "The earth observation handbook." [Online]. Available: http://www.eohandbook.com/ 14

[20] A. Singh, "Digital change detection techniques using remotely-sensed data," *International Journal of Remote Sensing*, vol. 10, no. 6, pp. 989–1003, 1989. [Online]. Available: http://www.tandfonline.com/doi/abs/10.1080/01431168908903939 15, 16

[21] C. Chen, *Image Processing for Remote Sensing*. CRC Press, 2007. 15, 16

[22] L. Bruzzone and F. Bovolo, "A novel framework for the design of change-detection systems for very-high-resolution remote sensing images," *Proceedings of the IEEE*, vol. 101, no. 3, pp. 609–630, 2013. 15, 30, 57

[23] G. Camps-Valls, L. Gomez-Chova, J. Munoz-Mari, J. Rojo-Alvarez, and M. Martinez-Ramon, "Kernel-based framework for multitemporal and multisource

remote sensing data classification and change detection," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 46, no. 6, pp. 1822–1835, June 2008. 16

[24] G. Moser and S. Serpico, "Generalized minimum-error thresholding for unsupervised change detection from sar amplitude imagery," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 44, no. 10, pp. 2972–2982, Oct 2006. 16, 17, 26

[25] F. Bovolo and L. Bruzzone, "A split-based approach to unsupervised change detection in large-size multitemporal images: Application to tsunami-damage assessment," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 6, pp. 1658 –1670, june 2007. 16, 17, 26, 62, 78

[26] E. Rignot and J. van Zyl, "Change detection techniques for ERS-1 SAR data," *IEEE Trans. Geosci. Remote Sens.*, vol. 31, no. 4, pp. 896 –906, jul 1993. 16, 18, 19, 25

[27] A. Atto, E. Trouve, Y. Berthoumieu, and G. Mercier, "Multidate divergence matrices for the analysis of sar image time series," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 51, no. 4, pp. 1922–1938, April 2013. 16

[28] P. Lombardo and T. Pellizzeri, "Maximum likelihood signal processing techniques to detect a step pattern of change in multitemporal sar images," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 40, no. 4, pp. 853–870, Apr 2002. 16

[29] F. Bujor, E. Trouve, L. Valet, J.-M. Nicolas, and J.-P. Rudant, "Application of log-cumulants to the detection of spatiotemporal discontinuities in multitemporal sar images," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 42, no. 10, pp. 2073–2084, Oct 2004. 16

[30] R. Dekker, "Texture analysis and classification of ERS SAR images for map updating of urban areas in The Netherlands," *IEEE Trans. Geosci. Remote Sens.*, vol. 41, no. 9, pp. 1950–1958, 2003. 16, 30

[31] M. Matsuoka and F. Yamazaki, "Use of satellite SAR intensity imagery for detecting building areas damaged due to earthquakes," *Earthquake Spectra*, vol. 20, pp. 975–994, 2004. 16, 31, 59

[32] L. G. Brown, "A survey of image registration techniques," *ACM Comput. Surv.*, vol. 24, pp. 325–376, December 1992. [Online]. Available: http://doi.acm.org/10.1145/146370.146374 17

[33] R. Touzi, "A review of speckle filtering in the context of estimation theory," *IEEE Trans. Geosci. Remote Sens.*, vol. 40, no. 11, pp. 2392 – 2404, nov 2002. 17

[34] E. Trouve, Y. Chambenoit, N. Classeau, and P. Bolon, "Statistical and operational performance assessment of multitemporal SAR image filtering," *IEEE Trans. Geosci. Remote Sens.*, vol. 41, no. 11, pp. 2519–2530, Nov 2003. 17, 92

[35] A. Lopes, E. Nezry, R. Touzi, and H. Laur, "Structure detection and statistical adaptive speckle filtering in SAR images," *International Journal of Remote Sensing*, vol. 14, pp. 1735–1758, 1993. 17, 81

[36] A. Lopes, R. Touzi, and E. Nezry, "Adaptive speckle filters and scene heterogeneity," *IEEE Trans. Geosci. Remote Sens.*, vol. 28, no. 6, pp. 992 –1000, nov 1990. 17, 18

[37] S. Quegan and J. J. Yu, "Filtering of multichannel SAR images," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 39, no. 11, pp. 2373–2379, Nov 2001. 17, 92, 105, 107, 110

[38] X. Su, C.-A. Deledalle, F. Tupin, and H. Sun, "Two-step multitemporal nonlocal means for synthetic aperture radar images," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 52, no. 10, pp. 6181–6196, Oct 2014. 17, 27, 28

[39] F. Bovolo and L. Bruzzone, "A detail-preserving scale-driven approach to change detection in multitemporal SAR images," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 12, pp. 2963 – 2972, dec. 2005. 20, 23, 62, 75, 77

[40] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York, NY, USA: Wiley-Interscience, 1991. 21

[41] S. Kullback and R. A. Leibler, "On information and sufficiency," *Annals of Mathematical Statistics*, vol. 22, pp. 49–86, 1951. 21

[42] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern Classification (2nd Edition)*. Wiley-Interscience, 2000. 22, 26, 95

[43] H. Xie, L. Pierce, and F. Ulaby, "SAR speckle reduction using wavelet denoising and Markov random field modeling," *IEEE Trans. Geosci. Remote Sens.*, vol. 40, no. 10, pp. 2196 – 2212, oct 2002. 23

[44] F. Argenti and L. Alparone, "Speckle removal from SAR images in the undecimated wavelet domain," *IEEE Trans. Geosci. Remote Sens.*, vol. 40, no. 11, pp. 2363 – 2374, nov 2002. 23

[45] L. Bruzzone and L. Carlin, "A multilevel context-based system for classification of very high spatial resolution images," *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 9, pp. 2587 –2600, sept. 2006. 23

[46] T. Fung and E. LeDrew, "The determination of optimal threshold levels for change detection using various accuracy indices," *Photogrammetric Engineering and Remote Sensing*, vol. 54, no. 10, pp. 1449 – 1454, 1988. 25

[47] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *Journal of the Royal Statistical Society, Series B*, vol. 39, no. 1, pp. 1–38, 1977. 26, 27, 78

[48] Y. Bazi, L. Bruzzone, and F. Melgani, "Image thresholding based on the EM algorithm and the generalized Gaussian distribution," *Pattern Recogn.*, vol. 40, no. 2, pp. 619–634, Feb. 2007. 26, 27, 78

[49] L. Bruzzone and D. Prieto, "Automatic analysis of the difference image for unsupervised change detection," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 38, no. 3, pp. 1171–1182, May 2000. 26

[50] T. Kasetkasem and P. Varshney, "An image change detection algorithm based on markov random field models," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 40, no. 8, pp. 1815–1823, Aug 2002. 26

[51] G. Quin, B. Pinel-Puyssegur, J.-M. Nicolas, and P. Loreaux, "Mimosa: An automatic change detection method for sar time series," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 52, no. 9, pp. 5349–5363, Sept 2014. 27, 28

[52] B. Aiazzi, L. Alparone, S. Baronti, A. Garzelli, and C. Zoppetti, "Nonparametric change detection in multitemporal sar images based on mean-shift clustering," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 51, no. 4, pp. 2022–2031, April 2013. 27, 28

[53] F. Baselice, G. Ferraioli, and V. Pascazio, "Markovian change detection of urban areas using very high resolution complex sar images," *Geoscience and Remote Sensing Letters, IEEE*, vol. 11, no. 5, pp. 995–999, May 2014. 27, 28

[54] A. Schmitt, B. Wessel, and A. Roth, "An innovative curvelet-only-based approach for automated change detection in multi-temporal sar imagery," *Remote Sensing*, vol. 6, no. 3, pp. 2435–2462, 2014. 27, 28

[55] C. Marin, F. Bovolo, and L. Bruzzone, "Building change detection in multitemporal very high resolution SAR images," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. PP, no. 99, pp. 1–19, 2014. 29, 91, 106

[56] P. Gamba, "Human settlements: A global challenge for EO data processing and interpretation," *Proceedings of the IEEE*, vol. 101, no. 3, pp. 570–581, 2013. 30

[57] F. Henderson and Z.-G. Xia, "SAR applications in human settlement detection, population estimation and urban land use pattern analysis: a status report," *IEEE Trans. Geosci. Remote Sens.*, vol. 35, no. 1, pp. 79–85, 1997. 30

[58] P. Du, S. Liu, P. Gamba, K. Tan, and J. Xia, "Fusion of difference images for change detection over urban areas," *IEEE J. Sel. Topics Appl. Earth Observations Remote Sens.*, vol. 5, no. 4, pp. 1076–1086, 2012. 30

[59] T. Balz, "SAR simulation of urban areas: Techniques and applications," in *Radar Remote Sensing of Urban Areas*, ser. Remote Sensing and Digital Image Processing, U. Soergel, Ed. Springer Netherlands, 2010, vol. 15, pp. 215–231. 30

[60] U. Soergel, U. Thoennessen, A. Brenner, and U. Stilla, "High-resolution sar data: new opportunities and challenges for the analysis of urban areas," *Radar, Sonar and Navigation, IEE Proceedings* -, vol. 153, no. 3, pp. 294–300, 2006. 30, 59

[61] A. Brenner and L. Roessing, "Radar imaging of urban areas by means of very high-resolution sar and interferometric sar," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 10, pp. 2971–2982, 2008. 30, 59

[62] Y. Dong, B. Forster, and C. Ticehurst, "Radar backscatter analysis for urban environments," *International Journal of Remote Sensing*, vol. 18, no. 6, pp. 1351–1364, 1997. 30, 34, 59

[63] G. Franceschetti, A. Iodice, and D. Riccio, "A canonical problem in electromagnetic backscattering from buildings," *IEEE Trans. Geosci. Remote Sens.*, vol. 40, no. 8, pp. 1787 – 1801, aug 2002. 30, 34, 59

[64] Z.-G. Xia and F. Henderson, "Understanding the relationships between radar response patterns and the bio- and geophysical parameters of urban areas," *IEEE Trans. Geosci. Remote Sens.*, vol. 35, no. 1, pp. 93–101, 1997. 30, 32, 59

[65] F. Bovolo, C. Marin, and L. Bruzzone, "A hierarchical approach to change detection in very high resolution sar images for surveillance applications," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 51, no. 4, pp. 2042–2054, April 2013. 30, 73, 92, 95

[66] M. Chini, N. Pierdicca, and W. Emery, "Exploiting SAR and VHR optical images to quantify damage caused by the 2003 Bam earthquake," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 1, pp. 145–152, 2009. 31, 59

[67] D. Ehrlich, H. Guo, K. Molch, J. Ma, and M. Pesaresi, "Identifying damage caused by the 2008 Wenchuan earthquake from VHR remote sensing data," *International Journal of Digital Earth*, vol. 2, no. 4, pp. 309–326, 2009. 31, 59

[68] T. Balz and M. Liao, "Building-damage detection using post-seismic high-resolution SAR satellite data," *International Journal of Remote Sensing*, vol. 31, no. 13, pp. 3369–3391, 2010. 31, 59

[69] R. Dekker, "High-resolution radar damage assessment after the earthquake in Haiti on 12 january 2010," *IEEE J. Sel. Topics Appl. Earth Observations Remote Sens.*, vol. 4, no. 4, pp. 960 –970, dec. 2011. 31, 59, 62, 63

[70] F. Dell'Acqua, C. Bignami, M. Chini, G. Lisini, D. Polli, and S. Stramondo, "Earthquake damages rapid mapping by satellite remote sensing data: L'Aquila april 6th, 2009 event," *IEEE J. Sel. Topics Appl. Earth Observations Remote Sens.*, vol. 4, no. 4, pp. 935 –943, dec. 2011. 31, 59

[71] Y. Dong, Q. Li, A. Dou, and X. Wang, "Extracting damages caused by the 2008 Ms 8.0 Wenchuan earthquake from SAR remote sensing data," *Journal of Asian Earth Sciences*, vol. 40, no. 4, pp. 907 – 914, 2011, the 2008 Wenchuan Earthquake, China and Active Tectonics of Asia. 31, 59

[72] D. Jin, X. Wang, A. Dou, and Y. Dong, "Post-earthquake building damage assessment in Yushu using airborne SAR imagery," *Earthquake Science*, vol. 24, no. 5, pp. 463–473, 2011. 31, 59

[73] L. Dong and J. Shan, "A comprehensive review of earthquake-induced building damage detection with remote sensing techniques," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 84, no. 0, pp. 85 – 99, 2013. 31, 59

[74] V. Poulain, J. Inglada, M. Spigai, J.-Y. Tourneret, and P. Marthon, "High-resolution optical and SAR image fusion for building database updating," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 8, pp. 2900–2910, 2011. 31, 59

[75] J. Tao, S. Auer, and P. Reinartz, "Detecting changes between a dsm and a high resolution sar image with the support of simulation based separation of urban scenes," in *Synthetic Aperture Radar, 2012. EUSAR. 9th European Conference on*, 2012, pp. 95–98. 31, 59

[76] P. Brett and R. Guida, "Earthquake damage detection in urban areas using curvilinear features," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 9, pp. 4877–4884, 2013. 31, 59

[77] A. Ferro, D. Brunner, L. Bruzzone, and G. Lemoine, "On the relationship between double bounce and the orientation of buildings in vhr sar images," *Geoscience and Remote Sensing Letters, IEEE*, vol. 8, no. 4, pp. 612 –616, july 2011. 31, 34, 79

[78] U. Soergel, U. Thönnessen, and U. Stilla, "Iterative building reconstruction from multi-aspect insar data," in *3-D reconstruction from airborne laserscanner andInSAR data. International Archives of Photogrammetry and Remote Sensing.*, Maas, Vosselmann, and Streilein, Eds., vol. 34, 2003, pp. 186–192. 31, 34

[79] M. Quartulli and M. Datcu, "Stochastic geometrical modeling for built-up area understanding from a single SAR intensity image with meter resolution," *IEEE Trans. Geosci. Remote Sens.*, vol. 42, no. 9, pp. 1996–2003, 2004. 31

[80] A. Thiele, E. Cadario, K. Schulz, U. Thonnessen, and U. Soergel, "Building recognition from multi-aspect high-resolution InSAR data in urban areas," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 11, pp. 3583–3593, 2007. 31, 32, 92

[81] F. Xu and Y.-Q. Jin, "Automatic reconstruction of building objects from multiaspect meter-resolution sar images," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 7, pp. 2336–2353, 2007. 31

[82] R. Guida, A. Iodice, D. Riccio, and U. Stilla, "Model-based interpretation of high-resolution sar images of buildings," *Selected Topics in Applied Earth Observations and Remote Sensing, IEEE Journal of*, vol. 1, no. 2, pp. 107–119, 2008. 31

[83] H. Sportouche, F. Tupin, and L. Denise, "Extraction and three-dimensional reconstruction of isolated buildings in urban scenes from high-resolution optical and SAR spaceborne images," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 10, pp. 3932–3946, 2011. 31, 34

[84] A. Ferro, D. Brunner, and L. Bruzzone, "Automatic detection and reconstruction of building radar footprints from single VHR SAR images," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 2, pp. 935–952, 2013. 31

[85] "Google maps." [Online]. Available: https://maps.google.com 33, 48, 50, 52, 53, 109

[86] A. Bennett and D. Blacknell, "Infrastructure analysis from high resolution SAR and InSAR imagery," in *Remote Sensing and Data Fusion over Urban Areas, 2003. 2nd GRSS/ISPRS Joint Workshop on*, may 2003, pp. 230 – 235. 32, 34, 39, 92

[87] Y. Zhang, H. Chen, C. Ding, and H. Wang, "A novel approach for shadow enhancement in high-resolution SAR images using the height-variant phase compensation algorithm," *IEEE Geosci. Remote Sens. Lett.*, vol. 10, no. 1, pp. 189–193, 2013. 34

[88] J. Groen, R. Hansen, H. Callow, J. Sabel, and T. Sabo, "Shadow enhancement in synthetic aperture sonar using fixed focusing," *Oceanic Engineering, IEEE Journal of*, vol. 34, no. 3, pp. 269–284, 2009. 34

[89] F. Bovolo, C. Marin, and L. Bruzzone, "A novel approach to building change detection in very high resolution SAR images," in *Proc. SPIE*, vol. 8537, 2012, pp. 85 370Y–85 370Y–12. 37

[90] L. Zadeh, "Fuzzy sets," *Information and Control*, vol. 8, pp. 338–353, 1965. 40, 44, 96

[91] P. Heckbert, *Graphics Gems IV*, ser. Graphics gems series. AP Professional, 1994. 41

[92] P. M. Larsen, "Industrial applications of fuzzy logic control," *International Journal of Man-Machine Studies*, vol. 12, no. 1, pp. 3 – 10, 1980. 44

[93] C. B. Barber, D. P. Dobkin, and H. Huhdanpaa, "The quickhull algorithm for convex hulls," *ACM Trans. Math. Softw.*, vol. 22, no. 4, pp. 469–483, Dec. 1996. 45

[94] D. Brunner, K. Schulz, and T. Brehm, "Building damage assessment in decimeter resolution sar imagery: A future perspective," in *Urban Remote Sensing Event (JURSE), 2011 Joint*, 2011, pp. 217–220. 46

[95] S. Kuny, K. Schulz, and H. Hammer, "Signature analysis of destroyed buildings in simulated high resolution SAR data," in *Geoscience and Remote Sensing Symposium (IGARSS), 2013 IEEE International*, 2013. 46

[96] "Geoportale regione abruzzo." [Online]. Available: http://geoportale.regione.abruzzo.it/ 48

[97] "Bing maps." [Online]. Available: www.bing.com/maps/ 50

[98] World Bank, "The great east japan earthquake learning from megadisasters : Knowledge notes, executive summary," in *EAP DRM Knowledge Notes*, 2012. 58

[99] [Online]. Available: http://earthquake.usgs.gov/ 58

[100] [Online]. Available: http://www.disasterscharter.org/ 58

[101] [Online]. Available: http://www.unitar.org/unosat/ 58

[102] [Online]. Available: http://supersites.earthobservations.org/ 58

[103] J.-L. Starck, E. Candes, and D. Donoho, "The curvelet transform for image denoising," *Image Processing, IEEE Transactions on*, vol. 11, no. 6, pp. 670–684, Jun 2002. 63

[104] Q. Liu, "Efficiency analysis of container ports and terminals," Ph.D. dissertation, Univ. College London, 2010. 74

[105] Livorno Port Authority, "http://www.portauthority.li.it." 74, 85

[106] C. S. Regazzoni, G. Vernazza, and G. Fabri, Eds., *Advanced Video-Based Surveillance Systems*. Norwell, MA, USA: Kluwer Academic Publishers, 1998. 74

[107] E. Kozan, "Optimising container transfers at multimodal terminals," *Mathematical and Computer Modelling*, vol. 31, no. 1012, pp. 235 – 243, 2000. 74

[108] F. Bovolo and L. Bruzzone, "An adaptive technique based on similarity measures for change detection in very high resolution sar images," in *Geoscience and Remote Sensing Symposium, 2008. IGARSS 2008. IEEE International*, vol. 3, july 2008, pp. III –158 –III –161. 75

[109] F. Tupin, H. Maitre, J.-F. Mangin, J.-M. Nicolas, and E. Pechersky, "Detection of linear features in sar images: application to road network extraction," *IEEE Trans. Geosci. Remote Sens.*, vol. 36, no. 2, pp. 434 –453, mar 1998. 79, 80

[110] A. Ferro, D. Brunner, and L. Bruzzone, "Building detection and radar footprint reconstruction from single VHR SAR images," in *Geoscience and Remote Sensing Symposium (IGARSS), 2010 IEEE International*, july 2010, pp. 292 –295. 79

[111] T. Johnsen, "Change detection and detailed analysis of stacking configuration of containers in terrasar-x sar images," in *Radar Conference, 2010 IEEE*, may 2010, pp. 609 –614. 81, 88

[112] G. Palubinskas and H. Runge, "Radar signatures of a passenger car," *Geoscience and Remote Sensing Letters, IEEE*, vol. 4, no. 4, pp. 644 –648, oct. 2007. 81

[113] D. J. Crisp, "The state-of-the-art in ship detection in synthetic aperture radar imagery," *Information Sciences*, no. DSTO-RR-0272, p. 115, 2004. 82

[114] G. M. Amdahl, "Validity of the single processor approach to achieving large scale computing capabilities," in *Proceedings of the April 18-20, 1967, spring joint computer conference*, ser. AFIPS '67 (Spring). New York, NY, USA: ACM, 1967, pp. 483–485. 83

[115] Livorno Ship Info Service, "http://www.shipinfo.it." 89

[116] J. Bruniquel and A. Lopes, "Multi-variate optimal speckle reduction in SAR imagery," *International Journal of Remote Sensing*, vol. 18, no. 3, pp. 603–627, 1997. 92

[117] D. Coltuc, E. Trouve, F. Bujor, N. Classeau, and J. Rudant, "Time-space filtering of multitemporal SAR images," in *Geoscience and Remote Sensing Symposium, 2000. Proceedings. IGARSS 2000. IEEE 2000 International*, vol. 7, 2000, pp. 2909–2911 vol.7. 92

[118] M. Ciuc, P. Bolon, E. Trouvé, V. Buzuloiu, and J.-P. Rudant, "Adaptive-neighborhood speckle removal in multitemporal synthetic aperture radar images," *Appl. Opt.*, vol. 40, no. 32, pp. 5954–5966, Nov 2001. 92

[119] J. Chanussot, G. Mauris, and P. Lambert, "Fuzzy fusion techniques for linear features detection in multitemporal SAR images," *IEEE Trans. Geosci. Remote Sens.*, vol. 37, no. 3, pp. 1292–1305, May 1999. 92, 104

[120] T. Le, A. Atto, E. Trouve, and J.-M. Nicolas, "Adaptive multitemporal SAR image filtering based on the change detection matrix," *IEEE Geosci. Remote Sens. Lett.*, vol. PP, no. 99, pp. 1–5, 2014. 93

[121] Swain, "Bayesian classification in a time-varying environment," *Systems, Man and Cybernetics, IEEE Transactions on*, vol. 8, no. 12, pp. 879–883, Dec 1978. 93, 94

[122] L. Bruzzone, D. Prieto, and S. Serpico, "A neural-statistical approach to multitemporal and multisource remote-sensing image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 37, no. 3, pp. 1350 –1359, may 1999. 93, 94

[123] C. Lee, "Fuzzy logic in control systems: fuzzy logic controller. I," *Systems, Man and Cybernetics, IEEE Transactions on*, vol. 20, no. 2, pp. 404–418, Mar 1990. 96, 99

[124] ——, "Fuzzy logic in control systems: fuzzy logic controller. II," *Systems, Man and Cybernetics, IEEE Transactions on*, vol. 20, no. 2, pp. 419–435, Mar 1990. 96, 100, 101

[125] L.-X. Wang, *A Course in Fuzzy Systems and Control*. Upper Saddle River, NJ, USA: Prentice-Hall, Inc., 1997. 96, 100, 102

[126] F. Wang, "A knowledge-based vision system for detecting land changes at urban fringes," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 31, no. 1, pp. 136–145, Jan 1993. 96

[127] J. Desachy, L. Roux, and E.-H. Zahzah, "Numeric and symbolic data fusion: A soft computing approach to remote sensing images analysis," *Pattern Recognition Letters*, vol. 17, no. 13, pp. 1361 – 1378, 1996, special Issue on Non-conventional Pattern Analysis in Remote Sensing. 96

[128] G. L. Mota, R. Q. Feitosa, H. L. Coutinho, C.-E. Liedtke, S. Mller, K. Pakzad, and M. S. Meirelles, "Multitemporal fuzzy classification model based on class transition possibilities," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 62, no. 3, pp. 186 – 200, 2007. 96

[129] E. Mamdani and S. Assilian, "An experiment in linguistic synthesis with a fuzzy logic controller," *International Journal of Man-Machine Studies*, vol. 7, no. 1, pp. 1 – 13, 1975. 96

[130] S. Auer and S. Gernhardt, "Linear signatures in urban SAR images – partly misinterpreted?" *Geoscience and Remote Sensing Letters, IEEE*, vol. 11, no. 10, pp. 1762–1766, Oct 2014. 103

[131] F. Tupin, H. Maitre, J.-F. Mangin, J.-M. Nicolas, and E. Pechersky, "Detection of linear features in SAR images: application to road network extraction," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 36, no. 2, pp. 434–453, Mar 1998. 103, 104, 105, 106

[132] C. Steger, "An unbiased detector of curvilinear structures," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 20, no. 2, pp. 113–125, Feb 1998. 103

[133] R. O. Duda and P. E. Hart, "Use of the hough transformation to detect lines and curves in pictures," *Commun. ACM*, vol. 15, no. 1, pp. 11–15, Jan. 1972. [Online]. Available: http://doi.acm.org/10.1145/361237.361242 104

[134] R. Touzi, A. Lopes, and P. Bousquet, "A statistical and geometrical edge detector for SAR images," *IEEE Trans. Geosci. Remote Sens.*, vol. 26, no. 6, pp. 764 –773, nov 1988. 104

[135] F. Argenti, A. Lapini, T. Bianchi, and L. Alparone, "A tutorial on speckle reduction in synthetic aperture radar images," *Geoscience and Remote Sensing Magazine, IEEE*, vol. 1, no. 3, pp. 6–35, Sept 2013. 105