Title: A Hierarchical Approach to Change Detection in Very High Resolution SAR Images for Surveillance Applications

This paper appears in: Geoscience and Remote Sensing, IEEE Transactions on

Date of Publication: April 2013

Author(s): Bovolo, F. ; Marin, C. ; Bruzzone, L.

Volume: 51 , Issue: 4

Page(s): 2042 - 2054

DOI: 10.1109/TGRS.2012.2223219

A Hierarchical Approach to Change Detection in Very High Resolution SAR Images for Surveillance Applications

Francesca Bovolo, *Member, IEEE,* Carlo Marin, *Student Member, IEEE,* and Lorenzo Bruzzone, *Fellow, IEEE*

***Abstract*—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 paper 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.**

***Index Terms*—Multitemporal images, Change detection, Very high geometrical resolution images, Synthetic aperture radar, Image processing, Remote sensing.**

I. Introduction

CHANGE detection is a process of primary importance for a large number of applications, including urban planning [1], natural resources monitoring [2], agricultural surveys [3], natural hazard prevention and monitoring [4]. In this context, optical sensors have been extensively exploited and several automatic and unsupervised change-detection methods have been developed. Unlike optical sensors, Synthetic Aperture Radar (SAR) systems have been less exploited. 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 and geometry distortions, which make any automatic analysis difficult. Despite the presence of speckle noise, the use of SAR sensors in change detection is highly attractive from the operational viewpoint. In fact, SAR has the advantage over optical sensors of being insensitive to atmospheric and sunlight conditions. This makes it possible to ensure the data acquisition on an area of interest in advance according to end-user requirements (e.g., seasonal and agricultural calendars [5], [6]) or during a crisis event (e.g., hurricanes, floods [7]).

In the past few years, we have observed an increased availability of Very High Resolution (VHR) SAR images. This is due to a new generation of satellites, such as TerraSAR-X [8] and COSMO-SkyMed (CSK®) [9], which are able to acquire images with a resolution of up to 1 meter. 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. The significant amount of geometrical details present in a VHR image changes completely the perspective of SAR data analysis: objects that are considered homogeneous from a semantic point of view, such us buildings, show a signature that is inhomogeneous at high spatial resolution because of the scattering contributions from sub-objects that compose 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 [4], [10], [11]. Hence, in order to properly detect changes in the objects present in the scene it is necessary to take into account backscattering contributions from sub-objects of the investigated object. An illustrative example of the use of this concept for the detection of building radar footprint in single date VHR SAR images is given in [10]. The rationale behind this approach is that each building can be modeled by means of interrelated characteristics that depend on both the considered object and the acquisition geometry. The low-level features of a building (i.e., layover, double-bounce, and shadowing effects) are combined using a production net in order to derive more structured primitives, which can be associated with a semantic meaning that allows a reliable detection and reconstruction of the building.

The short repetition interval guaranteed by the 4 satellites of the CSK® constellation (i.e., less than 12 hours in emergency situation) opens new attractive opportunities for both the scientific and the user communities. In this scenario it is possible to develop change-detection (CD) methods for solving specific surveillance and monitoring problems of sites of general interest, such as ports, airports, industrial sites, etc. [12]. Nevertheless, the combination of short repetition

F. Bovolo, C. Marin, and L. Bruzzone are with the Department of Information Engineering and Computer Science, University of Trento Via Sommarive, 14 I-38123, Povo, Trento, Italy, (e-mail: lorenzo.bruzzone@ing.unitn.it).

Manuscript received Month Day, Year; revised Month Day, Year.

interval and very high geometrical resolution leads to some challenging issues that should be addressed. Unsupervised change-detection techniques developed for medium/high resolution SAR generally compare pixel-by-pixel (e.g., according to ratio or log-ratio operators) two SAR images acquired on the same geographical area at different times by assuming that they are similar to each other except for the presence of changes occurred on the ground. This approach is not suitable to be applied to VHR SAR images due to: i) the intrinsic complexity of the electromagnetic backscattering mechanisms that increases with the complexity of the imaged objects at metric resolution; and ii) possible differences in acquisition conditions of multitemporal images (e.g., view angle, moisture content on the ground). Thus the same object may show different backscattering behaviors at two dates even though it did not experience any change, leading to a large amount of false alarms if a pixel-by-pixel comparison between multitemporal images is performed. Moreover, depending on the application, some changes may be of interest to the end-users, whereas others may not. Accordingly, there is a need to change the way of approaching the problem with respect to sensors having resolutions in the range of 10–30 meters. It is no longer possible to analyze an image on the basis of single pixels or also of small neighborhood systems [13], [14]. Instead it is necessary to explicitly model the semantic meaning of objects. This can be achieved by exploiting both the intrinsic SAR image properties (e.g., acquisition geometry and scattering mechanisms) and all available prior information on the condsidered application.

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 [4], [7]. However, they do not take explicitly into account the multiscale proprieties of the expected changes. The main idea behind multiscale strategies is that there should be an optimal representation level for each ground object. For example, in the multiscale analysis of a very high resolution image (e.g., 1 meter resolution), 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 [15]. In the recent literature, two papers have made use of this concept for the detection of changes in SAR images having resolution in the range 10-30 meters [16], [17]. In [16] the authors proposed a scale-driven approach to change detection that aims at both preserving the geometrical details and filtering out the speckle noise in the homogeneous regions. In [17] a cumulant-version of the *Kullback-Leibler* distance is used as a statistical measure for detecting the changes, and the concept of the Multiscale Change Profiles (MCP) is introduced. MCPs are indicators used to choose the proper window size that guarantees at the same time both enough samples for computing the KL-distance in a reliable manner and preservation of geometrical details. This rationale was extended in [18] by computing the KL-distance between adaptive regions homogeneous both in space and time (i.e., multitemporal parcels) [19]. Even though this model conceptually overcomes the two contrasting requirements mentioned before, and better fits with the high geometrical proprieties of VHR images, it poses the problem on how to generate the parcels in an effective way. Despite their nice property of being multiscale, the mentioned 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). 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.

In this work the method is designed and illustrated on the specific application of surveillance of freight traffic. The specific application is introduced in order to better illustrate the method. However the approach is general and can be easily used in other applications dealing with multitemporal VHR images for surveillance purposes. As it will be explained in detail later, in such a scenario it is reasonable to assume that prior information is available.

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 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 paper is organized into four sections. Section II introduces the problem of change detection for port surveillance applications with VHR SAR images by describing the usual structure of a port. This description gives a better understanding of the rationale behind the proposed approach

(a) (b)

Fig. 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).

for the use of the prior information. Section III describes the proposed approach to change detection. Section IV presents the data sets, illustrates the tuning of the proposed method on the considered application, and shows the experimental results. Section V draws the conclusions of this work.

II. Background and Motivations

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. In detail we will focus 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 [20]–[22]. 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 [20], [23]. 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 Fig. 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 Fig. 1b which was acquired on the same geographical area of Fig. 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 feature(s) to enhance relevant changes and reduce the effects of non-relevant ones. These features 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 features 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

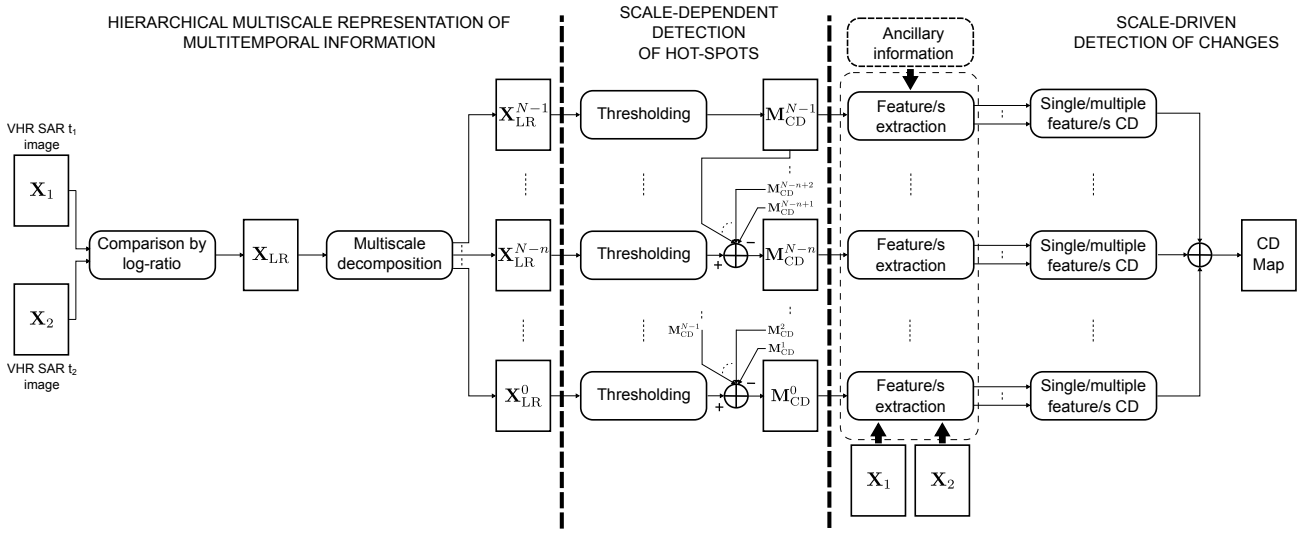


Fig. 2. Conceptual flow of the proposed approach for monitoring and surveillance applications.

of interest such as airports, industrial areas, logistic centers and so on.

The proposed method makes use of the prior information about the zones of interest in the investigated site in order to model, extract and exploit the semantic meaning of backscattering changes at the different resolution levels.

III. 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. Fig. 2 shows the block scheme of the proposed approach.

A. 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 [16]. 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. III-C).

The multiscale representation is obtained as follows. Differences in backscattering are highlighted by means of the log-ratio image $\mathbf{X}_{\mathrm{LR}}$ i.e., $\mathbf{X}_{\mathrm{LR}} = \log \mathbf{X}_2/\mathbf{X}_1$. This is the most common operator for highlighting changes in multitemporal SAR data [14]. 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 $\mathcal{X}_{\mathrm{MS}} = \left\{\mathbf{X}^0_{\mathrm{LR}}, \ldots, \mathbf{X}^n_{\mathrm{LR}}, \ldots, \mathbf{X}^{N-1}_{\mathrm{LR}}\right\}$ 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., $\mathbf{X}^0_{\mathrm{LR}} \equiv \mathbf{X}_{\mathrm{LR}}$. 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)[1]. 2D-SWT applies to the considered signal at each resolution level n appropriate

[1]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.

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 four different images at the next scale. Thus, the image $\mathbf{X}_{\mathrm{LR}}$ is decomposed into four 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 images 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 sub-band $\mathbf{X}_{\mathrm{LR}}^{\mathrm{LL}_n}$ at each scale. Thus, the outputs at a generic resolution level can be expressed analytically as follows:

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

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. 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 sub-band $\mathbf{X}_{\mathrm{LR}}^{\mathrm{LL}_{n+1}}$ the inverse stationary wavelet transform (2D-ISWT) is applied $n+1$ times.

B. 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 [24]. 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. Once the size is defined at each resolution level, splits are sorted according to their probability to contain a significant amount of changed pixels. As log-ratio images are considered, the sorting procedure can be reasonably done according to the value of the variance $\sigma_{s,n}^2$, $s = 1, \ldots, S$ computed on the pixels of each split. The desired set $\mathcal{P}_S$ of splits with the highest probabilities to contain changes is defined by selecting the first elements of the sorted set which fitful the following inequality:

$$\sigma_{s,n}^2 \geq \bar{\sigma}_n^2 + B\sigma_n, \quad s = 1, \ldots, S, \quad n = 1, \ldots, N-1 \tag{2}$$

where $\bar{\sigma}_n^2$ denotes the average variance of splits at level n, σ_n represents the standard deviation of the variance of splits and $B > 0$ is a constant. High values of B will make the selection strict, so that only those splits with high variance will be selected (i.e., the ones with the highest probability of containing changes), vice-versa low values of B will make the selection gentle and more blocks will be chosen. It is important to chose B considering the trade-off between selecting only high-variance splits and having a sufficient number of samples for performing next steps devoted to pixel labeling.

The split-based selection allows defining subset of pixels in which the class of change shows a higher prior probability than in the whole image. At each resolution level the defined pixel sub-set is used to compute the change-detection map. Here the thresholding method described in [25] 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 [25], the generalized-Gaussian model and the well known Expectation-Maximization (EM) algorithm [26] 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 Fig. 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 $\mathcal{M}_{\mathrm{CD}} = \left\{\mathbf{M}_{\mathrm{CD}}^0, \ldots, \mathbf{M}_{\mathrm{CD}}^n, \ldots, \mathbf{M}_{\mathrm{CD}}^{N-1}\right\}$ change-detection maps complementary to each other. Thus each map in $\mathcal{M}_{\mathrm{CD}}$ 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 [24] 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.

C. 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 $\mathcal{M}_{\mathrm{CD}}$ 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 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 II (see Fig. 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.

1) 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 [27], 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.

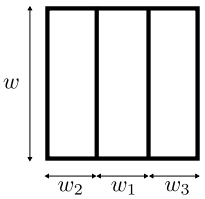


Fig. 3. Window used by the line detector.

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 [28] but modified in order to extract the bright lines instead of the dark ones. This detector was successfully used in [10] 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 Fig. 3 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 Fig. 3:

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

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:

$$r_E^{(t,c)} = \begin{cases} 1 - \min\left(\frac{\mu_t}{\mu_c}, \frac{\mu_c}{\mu_t}\right) & \text{if } \mu_1 \geq \mu_2, \mu_3 \\ 0 & \text{Otherwise.} \end{cases} \tag{4}$$

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 Fig. 3:

$$\rho = \min(\rho_{(1,2)}, \rho_{(1,3)}) \tag{5}$$

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:

$$\rho^2_{(t,c)} = \begin{cases} \frac{1}{1+(M_t+M_c)\frac{M_t \mathrm{LCV}_t^2 \mathrm{CR}^2_{(t,c)} + n_t \mathrm{LCV}_c^2}{M_t M_c (\mathrm{CR}_{(t,c)}-1)^2}} & \text{if } \mu_1 \geq \mu_2, \mu_3 \\ 0 & \text{Otherwise.} \end{cases} \tag{6}$$

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:

$$\Sigma(r_L, \rho) = \frac{r_L \rho}{1 - r_L - \rho + 2 r_L \rho}, \quad \text{with } r_L, \rho \in [0, 1] \tag{7}$$

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 [28]. 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 paper we use the *Kullback-Leibler* (KL) divergence. KL divergence [29] is a measure of the distance among statistical distributions and gives an indication on the difference between the shapes of two probability density functions (pdfs). It is formally defined as follows. Let $X_1(f_L)$ and $X_2(f_L)$ be two random variables associated to the pixels belonging to a feature of bright line f_L in images $\mathbf{X}_1$ and $\mathbf{X}_2$, respectively, and let $f_{X_1(f_L)}(x)$ and $f_{X_2(f_L)}(x)$ be the corresponding probability density functions. Then the KL distance based on bright linear features is defined as:

$$\mathrm{KL}\left(X_1(f_L)|X_2(f_L)\right) = \int \log \frac{f_{X_2(f_L)}(x)}{f_{X_1(f_L)}(x)} f_{X_2(f_L)}(x)\, \mathrm{d}x \tag{8}$$

In order to perform the comparison of multitemporal structures according to (8), pdfs should be known. Nonetheless, assuming that the density to be approximated is not too far from a Gaussian pdf, Inglada *et al.* in [17] demonstrated that it is possible to model the shape of a statistical distribution using the infinite *Edgeworth* series expansion of cumulats truncated at a given order. According to the results provided in [17], the cumulant-based pdf estimation presents a more robust behavior with respect to other parametric models (i.e., Gaussian or Pearson based), even though the hypothesis that the pdf to be approximated is Gaussian does not accurately fit the statistical model of SAR images.

Since the cumulant-based KL divergence (as the KL-divergence) is not symmetric, a symmetric version between two observations $X_1(f_L)$ and $X_2(f_L)$, called cumulant-based KL distance (CKLD), can be defined as:

$$\mathrm{CKLD}(X_1(f_L)|X_2(f_L)) = \mathrm{CKLD}(X_2(f_L)|X_1(f_L)) = \mathrm{KL}_{\text{Edgeworth}}\left(X_1(f_L)|X_2(f_L)\right) + \mathrm{KL}_{\text{Edgeworth}}\left(X_2(f_L)|X_1(f_L)\right) \tag{9}$$

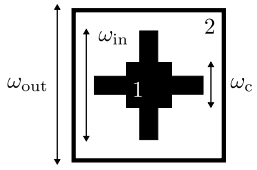


Fig. 4. Window used by the isolated scatterer detector.

The KL distance is then thresholded in order to identify changes in each analyzed double-bounce line of containers.

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 [30] can mitigate the ambiguity on possible staking configurations of containers.

2) 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 [31] 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 [31] 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 [32]. 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. Fig. 4 shows the geometry of the isolated scatterer detector proposed in [32]. 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:

$$r_{\text{PT}} = 1 - \min\left(\frac{\mu_1}{\mu_2}, \frac{\mu_2}{\mu_1}\right) \tag{10}$$

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 $r_{\mathrm{PT}\min}$. 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.

3) 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 [33]. 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:

$$r = 1 - \min\left(\frac{\mu_{\mathcal{C}_h^n}^{\tau_1}}{\mu_{\mathcal{C}_h^n}^{\tau_2}}, \frac{\mu_{\mathcal{C}_h^n}^{\tau_2}}{\mu_{\mathcal{C}_h^n}^{\tau_1}}\right) \tag{11}$$

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.

D. 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* [34]), 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. III-A and Sec. III-B). 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. III-C), 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.

IV. 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).

A. 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

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

Fig. 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.

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. 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:

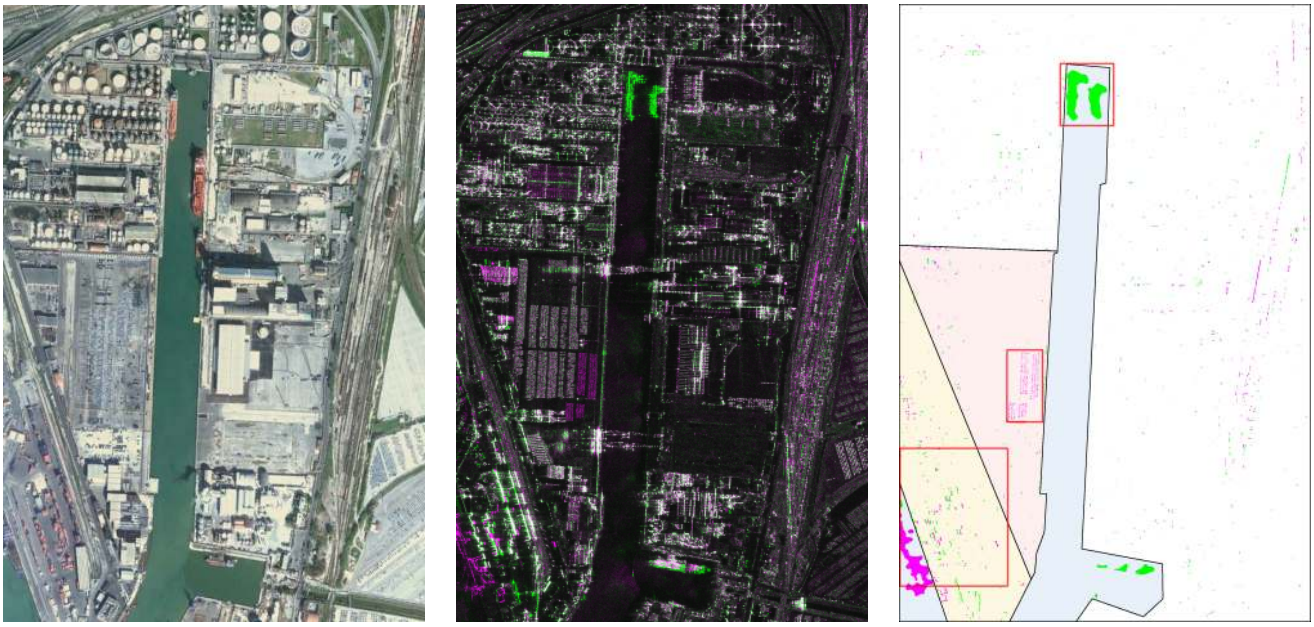
$$\{-0.0105974, 0.0328830, 0.0308414, -0.187035, \\ -0.0279838, 0.630881, 0.714847, 0.230378.\}$$

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 III-B. 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 sufficient 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 (10), with $\omega_{\mathrm{c}} = 3$, $\omega_{\mathrm{in}} = 5$ and $\omega_{\mathrm{out}} = 11$. 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.

TABLE I
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

Since no ground truth is available for the scene considered, the CD map obtained with the proposed method was compared



(a) (b) (c)

Fig. 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.

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 $\mathbf{X}_{\mathrm{LR}}$ (i.e., $\mathbf{M}_{\mathrm{CD}}^{0}$) and (c) the change-detection map obtained by thresholding $\mathbf{X}_{\mathrm{LR}}^{4}$ (i.e., $\mathbf{M}_{\mathrm{CD}}^{4}$). As expected, the change-detection map obtained with a standard pixel-based thresholding of $\mathbf{X}_{\mathrm{LR}}$ is affected by a high number of false alarms due to noisy components. The change-detection map obtained by thresholding $\mathbf{X}_{\mathrm{LR}}^{4}$ (Fig. 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 (Fig. 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 Fig. 5d, respectively), whereas a visual inspection on $\mathbf{X}_1$ and $\mathbf{X}_2$ resulted in 302 new cars and 808 removed (see Table I). This operation was not possible considering the CD map obtained with standard pixel-based thresholding of $\mathbf{X}_{\mathrm{LR}}$.

B. 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 [21]. It is located on the Tyrrhenian Sea in the northwestern 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. 6a shows the optical image corresponding to the same area taken from Google Maps. This is a GeoEye RGB true color composition. Fig. 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 Fig. 6): i) the cargo terminal (red region); ii) the car terminal (yellow region); and iii) the wet dock (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.

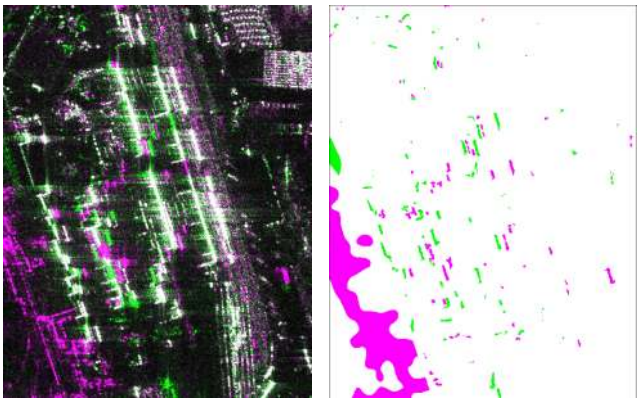
To apply the proposed method, the set $\mathcal{M}_{\mathrm{CD}} = \{\mathbf{M}_{\mathrm{CD}}^{0}, \ldots, \mathbf{M}_{\mathrm{CD}}^{n}, \ldots, \mathbf{M}_{\mathrm{CD}}^{N-1}, \}$, $N = 5$, was obtained following the same procedure used for the first data set (see Section IV-A). Then, using the available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

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

1) 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 (7), with parameters $\omega = 4$, $\omega_1 = 2$, and $\omega_2 = \omega_3 = 4$. For each

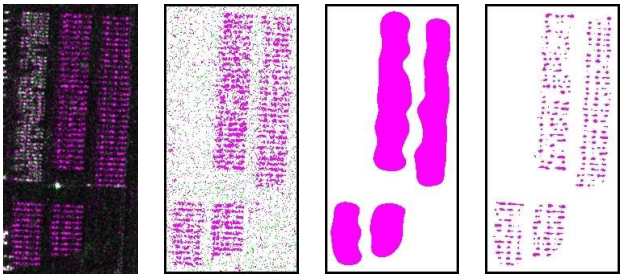
TABLE II
QUANTITATIVE PARAMETERS ASSOCIATED WITH CHANGED AREA RETRIEVED WITH THE PROPOSED APPROACH IN THE PORT OF LIVORNO DATA SET.

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



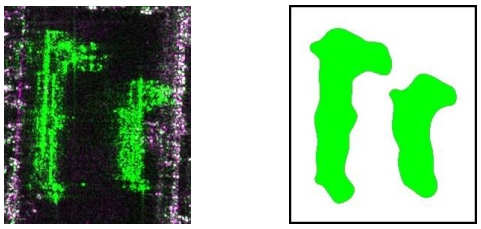
(a) (b)

Fig. 7. Zoom of cargo terminal area where containers are stacked (red box in left bottom part of Fig. 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).



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

Fig. 8. Zoom of cargo terminal area where cars are stacked (red box in the center of Fig. 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 $\mathbf{X}_{\mathrm{LR}}^4$; and (d) the proposed technique.



(a) (b)

Fig. 9. Zoom of ships in the wet dock (red box in the top of Fig. 6). (a) RGB composition of spotlight COSMO-SkyMed images; and (b) CD map obtained with the proposed method.

detected bright line in $\mathbf{X}_1$ and $\mathbf{X}_2$, the CKLD distance was computed according to (9). 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 Fig.7 (new containers appear in magenta, removed once in green). As mentioned in Section III-C1, 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 [30]. As a matter of fact, a specific analysis should be carried out which is out of the goal of this work.

2) 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 IV-A (see Fig. 8). From a quantitative analysis, 284 new cars were automatically detected, whereas a visual inspection on $\mathbf{X}_2$ resulted in 249 new cars.

3) 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 Fig.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* [35] (see Tab. II). 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.

4) 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.

V. Discussion and Conclusion

In this paper 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.

Acknowledgment

This work has been developed within the agreement no. I/041/09/0, project ID-2181, program "Development and validation of multitemporal image analysis methodologies for multirisk monitoring of critical structures and infrastructures" founded by the Italian Space Agency.

References

[1] 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.

[2] 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.

[3] 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.

[4] 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.

[5] 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.

[6] 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.

[7] 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.

[8] 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.

[9] 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.

[10] 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.

[11] 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.

[12] S. Brusch, S. Lehner, T. Fritz, M. Soccorsi, A. Soloviev, and B. van Schie, "Ship surveillance with TerraSAR-X," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 3, pp. 1092 –1103, march 2011.

[13] 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.

[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.

[15] 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.

[16] 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.
[17] 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.
[18] 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.
[19] L. Bruzzone and D. F. Prieto, "An adaptive parcel-based technique for unsupervised change detection," *International Journal of Remote Sensing*, vol. 21, no. 4, pp. 817–822, 2000.
[20] Q. Liu, "Efficiency analysis of container ports and terminals," Ph.D. dissertation, Univ. College London, 2010.
[21] Livorno Port Authority, "http://www.portauthority.li.it."
[22] C. S. Regazzoni, G. Vernazza, and G. Fabri, Eds., *Advanced Video-Based Surveillance Systems*. Norwell, MA, USA: Kluwer Academic Publishers, 1998.
[23] E. Kozan, "Optimising container transfers at multimodal terminals," *Mathematical and Computer Modelling*, vol. 31, no. 1012, pp. 235 – 243, 2000.
[24] 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.
[25] 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] 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.
[27] 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.
[28] 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.
[29] S. Kullback and R. A. Leibler, "On information and sufficiency," *Annals of Mathematical Statistics*, vol. 22, pp. 49–86, 1951.
[30] 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.
[31] 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.
[32] 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.
[33] 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.
[34] 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.
[35] Livorno Ship Info Service, "http://www.shipinfo.it."

Francesca Bovolo (S'05–M'07) received the "Laurea" (B.S.), the "Laurea Specialistica" (M.S.) degrees in telecommunication engineering (*summa cum laude*) and the Ph.D. in Communication and Information Technologies from the University of Trento, Italy, in 2001, 2003 and 2006, respectively. She is presently a research fellow at the Remote Sensing Laboratory of the Department of Information Engineering and Computer Science, University of Trento. Her main research activity is in the area of remote-sensing image processing. In particular, her interests are related to multitemporal remote sensing image analysis and change detection in multispectral and SAR images, and very high resolution images. She conducts research on these topics within the frameworks of several national and international projects. Francesca Bovolo is a referee for the *IEEE Transaction on Geoscience and Remote Sensing; IEEE Geoscience and Remote Sensing Letters; IEEE Transactions on Image Processing; IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing; International Journal of Remote Sensing; Pattern Recognition; Pattern Recognition Letters; Remote Sensing of Environment; Photogrammetric Engineering and Remote Sensing; Photogrammetry and Remote Sensing; IEEE Transactions on Aerospace and Electronic Systems; Sensors*. Francesca Bovolo ranked first place in the Student Prize Paper Competition of the 2006 *IEEE International Geoscience and Remote Sensing Symposium* (Denver, August 2006). Since January 2011 she is an associate editor of the IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing. She is guest editor of the Special Issue on Analysis of Multitemporal Remote Sensing Data of the IEEE Transaction on Geoscience and Remote Sensing. She is the Technical Chair of the Sixth International Workshop on the Analysis of Multi-temporal Remote-Sensing Images (MultiTemp 2011). Since 2006, she has served on the Scientific Committee of the SPIE International Conference on Signal and Image Processing for Remote Sensing. She has served on the Scientific Committee of the *IEEE Fourth and Fifth International Workshop on the Analysis of Multi-Temporal Remote Sensing Images (MultiTemp 2007 and 2009)*.

Carlo Marin (S'12) received the M.Sc. in telecommunication engineering in 2011 from the University of Trento, Italy where he is currently a PhD student in Information and Communication Technologies (ICT). At present, his research interests are related to multitemporal analysis and target detection using very high geometrical resolution SAR images.

Lorenzo Bruzzone (S'95–M'98–SM'03–F'10) received the laurea (M.S.) degree in electronic engineering (*summa cum laude*) and the Ph.D. degree in telecommunications from the University of Genoa, Italy, in 1993 and 1998, respectively. He is currently a Full Professor of telecommunications at the University of Trento, Italy, where he teaches remote sensing, radar, pattern recognition, and electrical communications. Dr. Bruzzone is the founder and the director of the Remote Sensing Laboratory in the Department of Information Engineering and Computer Science, University of Trento. His current research interests are in the areas of remote sensing, radar and SAR, signal processing, and pattern recognition. He promotes and supervises research on these topics within the frameworks of more than 28 national and international projects. He is the author (or coauthor) of 114 scientific publications in referred international journals (76 in IEEE journals), more than 170 papers in conference proceedings, and 16 book chapters. He is editor/co-editor of 10 books/conference proceedings and 1 scientific book. He has served on the Scientific Committees of several international conferences and was invited as keynote speaker in more than 20 international conferences and workshops. He is a member of the Managing Committee of the Italian Inter-University Consortium on Telecommunications. Since 2009 he is a member of the Administrative Committee of the IEEE Geoscience and Remote Sensing Society. Dr. Bruzzone ranked first place in the Student Prize Paper Competition of the 1998 IEEE International Geoscience and Remote Sensing Symposium (Seattle, July 1998). He was a recipient of the Recognition of IEEE Transactions on Geoscience and Remote Sensing (TGRS) Best Reviewers in 1999 and was a Guest Co-Editor of different Special Issues of international journals. In the past years joint papers presented by his students at international symposia and master theses that he supervised have received international and national awards. He was the General Chair/Co-chair of the First, Second and Sixth IEEE International Workshop on the Analysis of Multi-temporal Remote-Sensing Images (MultiTemp), and is currently a member of the Permanent Steering Committee of this series of workshops. Since 2003, he has been the Chair of the SPIE Conference on Image and Signal Processing for Remote Sensing. From 2004 to 2006 he served as an Associated Editor of the IEEE Geoscience and Remote Sensing Letters, and currently is an Associate Editor for the IEEE TGRS and the Canadian Journal of Remote Sensing. Since April 2010 he has been the Editor of the IEEE Geoscience and Remote Sensing Newsletter. In 2008 he has been appointed as a member of the joint NASA/ESA Science Definition Team for the radar instruments for Outer Planet Flagship Missions. Since 2012 he has been appointed Distinguished Speaker of the IEEE Geoscience and Remote Sensing Society. Dr. Bruzzone is member of the Italian Association for Remote Sensing (AIT).